

06\23

I  
C  
I

# ICIRI CBRNE DIARY



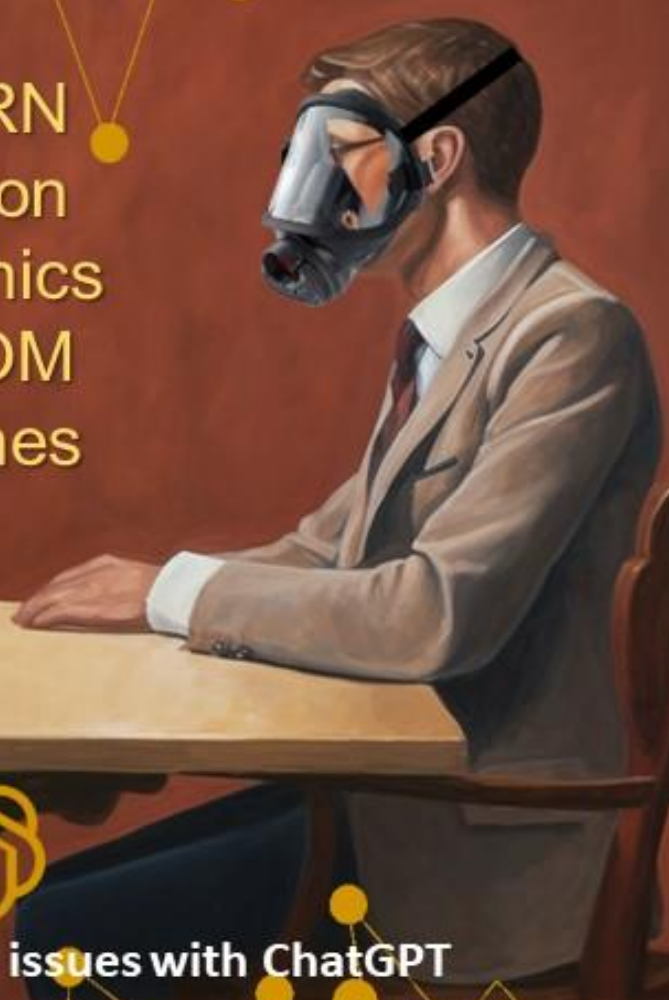
June 2023

*Dedicated to Global  
First Responders*

**PART B**



CBRN  
Decon  
AI ethics  
NEOM  
Drones



The Editor discuss CBRN issues with ChatGPT



Covid **0** patients identified

An International CBRNE Institute publication

ICI  
International  
**CBRNE**  
INSTITUTE



# DIRTY R-NEWS



## Nobel Prize winner Beatrice Fihn discusses abolishing nuclear weapons in the age of the Ukraine war

By **Samantha Powers** (Assistant Campus Editor)

Source: <https://dailynorthwestern.com/2023/05/23/campus/nobel-prize-winner-beatrice-fihn-discusses-abolishing-nuclear-weapons-in-the-age-of-the-ukraine-war/>



May 23 – Beatrice Fihn, former executive director of the Nobel Peace Prize-winning International Campaign to Abolish Nuclear Weapons, gave a talk at Northwestern’s Buffett Institute for Global Affairs on Monday. She spoke on the heightened threat of nuclear war since Russia’s invasion of Ukraine and what the average citizen can do to combat the use of nuclear weapons.

Fihn said other methods of mass destruction, like biological warfare, have been delegitimized on a global scale — and she argued the same can be done for nuclear weapons through action from banks, businesses and citizens.

“This is something we can all do,” Fihn said. “Regular people can speak about the unspeakable, center humans in these conversations and expose the realities of these weapons.”

She said when talking about the destruction associated with nuclear warfare, it’s important to elevate communities that have been impacted the most.

Fihn said when wealthy countries practice nuclear testing, they often do so in impoverished and disempowered countries. Racism can come into play because some lives are seen as less valuable than others, she said.

Fihn added nuclear warfare is also a feminist issue, citing that women face disproportionate health impacts from nuclear fallout compared to men.

It’s important to highlight the human toll of nuclear warfare, she said.

“The supporters of these weapons of mass destruction hope to frame the issue of nuclear weapons ... in abstract and sterile terms to avoid exposing what they are actually talking about,” Fihn said.

Fihn then detailed her work on passing the United Nations Treaty on the Prohibition of Nuclear Weapons, a global nuclear weapons ban that currently has 92 signatories. She noted that none of the nine countries that currently possess nuclear weapons have signed the treaty.



Chicago Ald. Maria Hadden (49th), who sponsored the Chicago City Council's legislation urging President Biden to bring the country "Back from the Brink" on nuclear weapons, attended the talk. The resolution calls on the federal government to renounce the option of preemptive nuclear strikes, end the independent, unchecked power of the president to launch a nuclear attack and cancel plans to enhance the U.S. nuclear arsenal with new weapons.

Hadden said her constituents played a crucial role in her decision to champion the legislation. She said Chicagoans pay more than \$500 million a year in federal taxes that go toward funding nuclear weapons — money that she said would be better spent on safe streets, education and public health.

Hadden added she involved the 49th Ward Youth Advisory Council in the process of passing this legislation, stressing the importance of highlighting young voices. When issues threaten the future of life on the planet, she said it's important to work intergenerationally. "When we talk about any long-term impact from government decision-making, we have to keep young people centered in our eyes," Hadden said.

Fihn said now, in the era of Russia's invasion of Ukraine, people can no longer ignore the threat of nuclear war.

Weinberg student Maelys Bory, a French exchange student, said she attended Monday's talk because of her interest in political science. She said young people like her fellow NU students should have the chance to hear from speakers like Fihn on pressing global issues.

"I think the conflict (in Ukraine) shows the emergence of the issue in the sense that it can happen," Bory said. "It's not a mere risk, it's not a small thing we can just live with and have in the back of our minds."

## "Right of boom" : Meet the experts who respond to nuclear disaster

Source: <https://thebulletin.org/2023/05/right-of-boom-meet-the-experts-who-respond-to-nuclear-disaster/#post-heading>

May 25 – It's a warm spring day in downtown Indianapolis. An emergency operations team is delivering its situation report, one of several coming in from county governments and field teams responding to the chaos after a 10-kiloton nuclear device exploded in the city, just an hour earlier. Radiation, fires, and limited capacity at area hospitals and shelters complicate treatment of the wounded and communication with a panicked public.

One team leader relays a request from a heavily damaged adjacent county to help house survivors of the nuclear blast. Field teams report back on the radiation doses they've received while navigating a pickup truck through the city to respond to people who need assistance.

"We didn't wreck anything," one driver says.

"Show-offs!" yells another.

These reports are part of a three-hour drill I'm observing inside a large conference room at the Hyatt Regency hotel in Indianapolis. The jurisdictions used in the exercise are fictional, with sci-fi county names like Endor, Caprica, and Druidia. The truck missions are



run on a video-game-like simulator, designed by a former Energy Department scientist who, among other things, assessed radiological impacts in Japan during the 2011 Fukushima nuclear disaster.

A field team member runs a simulated rescue mission driving through dangerously contaminated areas during the nuclear detonation response and recovery drill on the first day of the NREP conference.

The team members and radiation experts guiding the proceedings are all gathered for the four-day National Radiological Emergency Preparedness (NREP) conference, an

[annual meeting](#) of emergency management officials from local, state, and federal government agencies, along with representatives of utility companies, health organizations, and vendors with products ranging from expert guidance to Geiger counters. These are the people who try to plan for the aftermath of some of the worst imaginable radiological emergencies. The goal is to make sure that they, and the public, have some idea how to respond when disaster strikes.





The mood around the conference tables is typical of most any professional conference, with a shared travel-weariness offset by occasional jokes amusing only to insiders. After a short break, Bill Irwin, who leads the nuclear detonation drill, tries to get participants' attention to start its next part. "As is the case in real world emergencies," he announces, "they roll on despite the fact that people are trying to get pretzels."

A map of the region immediately affected by the mock 10 kiloton nuclear detonation centered over Indianapolis during the emergency preparedness drill at the 2023 NREP conference. The fictional county jurisdictions of Caprica, Druidia, Endor, and Vega are outlined, along with predicted zones of radiation and physical damage.

Despite the jovial atmosphere in the room, the events being contemplated are deadly serious. They're based on the Federal Emergency Management Agency (FEMA) "[Planning Guidance for Response to a Nuclear Detonation](#)," one of several major scenarios—including the use of a dirty bomb or chemical weapon—that federal, state, and local agencies have developed plans for and trained on in the years since 9/11.

Irwin is a member of a volunteer group of subject matter experts known by the acronym ROSS (Radiological Operations Support Specialists); they attend conferences and trainings like this one to help local officials coordinate their emergency plans during a radiological disaster. The ROSS group was established in 2015 by FEMA's Chemical, Biological, Radiological, and Nuclear Office

after a decade-long assessment identified gaps in the United States' preparedness for use of an "improvised nuclear device." Selected from locations across the country, ROSS experts can be [called on](#) in an actual emergency to support local authorities quickly. Like other ROSS volunteers helping with the proceedings and many of the 300-odd conference attendees, Irwin has a background in health physics, a field focused on the biological effects of radiation, from dental X-rays to nuclear fallout. In his day job, he leads the Radiological and Toxicological Sciences Program at the Vermont Department of Health.

He keeps the drill moving. Standing at a podium in front of a Powerpoint slide with bullet points like "Initial public messaging" and "Initial instructions to first responders," Irwin says, "So, this is one of the sad ones. It talks about how many people within certain areas have died." He's referring to a mock report generated for the drill that includes the immediate estimated casualties of a nuclear blast in the notional city—within the severe and moderate damage zones around ground zero, nearly 90 percent of the population, or about 49,000 people, are killed or injured. Irwin notes it's important to prepare local authorities and first responders for catastrophes of that scale. "This is something that we have to come to grips with, that there were a lot of deaths. It happens in natural disasters too," he says. After all, the 2004 tsunami in the Indian Ocean killed more than 200,000 people. "So don't think just because this is a nuclear detonation, we have to be afraid to talk about what we're doing the [first] day."

The storyline of the 10-kiloton drill dates to FEMA's 2010 guidance documents. The [latest version](#) of the nuclear detonation scenario, published in 2022, outlines proposed response to nuclear explosions as large as 100 kilotons—the yield of most individual warheads [currently deployed by Russia](#). Drilling on the consequences of a nuclear detonation means some of the emergency protocols appropriate during a less explosive accident—say, a reactor meltdown—get tossed out the window. For instance, Irwin notes, emergency workers responding to a nuclear detonation may not be able to avoid radiation exposures that are above—and perhaps significantly above—those allowed in less dire situations.

Most of the work that goes into radiological emergency planning is more mundane than reacting to extreme events like nuclear terrorism or war. The main context of emergency preparedness discussions at the conference involves accidents at nuclear power plants. The conference features three days of



presentations on everything from flying drones that detect radiation at accident sites to former journalists who help officials practice interacting with the media during an emergency.

The conference was first held in 1990 to help state authorities sort out how to implement federal regulations that emerged after the United States' worst nuclear power emergency, the Three Mile Island accident in 1979. Ken Evans began attending three decades ago; most of the people I meet at the conference make sure to mention his name. Evans started as a health physicist in Arkansas in the 1970s, worked for both utility companies and state agencies, and recently retired from his role as a radiological emergencies specialist with the Illinois Emergency Management Agency. When I catch up with him, he's only got a few minutes to talk before he has to leave for a side-meeting with FEMA officials about some updated regulations. A lilting Arkansas accent makes his precise recitation of arcane Nuclear Regulatory Commission (NRC) codes almost pleasant. ("NUREG-0654, you may have heard of that?" he says with a chuckle. "That's the Bible of emergency preparedness.")

Before the Three Mile Island accident, nuclear utilities (known formally as licensees) only had to demonstrate "reasonable assurance" of emergency preparedness onsite—that is, at the nuclear power plant facilities themselves—to be licensed to operate. "Emergency preparedness for offsite was purely voluntary," Evans says. After the accident, the NRC made nuclear plant licensing contingent on preparedness for offsite emergencies as well (based on a 10-mile zone for atmospheric radiation and a 50-mile zone for "ingestion pathways," like contamination of water and agricultural systems).

Evans tells me how emergency planning for nuclear plant emergencies developed since the 1980s in phases: the creation of the new off-site emergency regulations after Three Mile Island; increasing formalization of FEMA's planning guidance in the 1990s (alongside the NREP conference); the emergence of the "all hazards" approach to emergencies in the aftermath of 9/11 and hurricanes Katrina and Rita in 2005; and the evolution of new standards since the tsunami and reactor meltdowns at the Fukushima Daiichi nuclear power plant in 2011.

In the beginning, Evans says, offsite emergency planners weren't exactly sure how to apply the new regulations, let alone coordinate and standardize preparedness across states and the country. "Here we are, all stumbling around, including the regulators themselves," Evan says, referring to FEMA and the NRC. "I mean, it's one thing to write things out, it's another thing to implement things." So after a decade or so of uncertainty about how exactly to follow the rules, a group of frustrated state officials convened the first radiological emergency preparedness conference in 1990. "It was a chance for everyone to get together, share experiences, try to figure out what I would call basic information. But now we've kind of evolved on that," Evans says. In the years since, the NREP conference has become known as the most important opportunity for radiological emergency preparedness organizers and trainers to discuss changes to federal regulations, share state of the art tools and techniques that can be passed on to local emergency workers, and workshop ideas with counterparts from across the country.

As Evans describes the evolution of the conference, we're distracted by a robot dog walking behind him. One of the vendors is demonstrating its customized Boston Dynamics model, Cerberus QUGV, which has been outfitted with a bank of wireless radiation sensors and dosimeters to assist with risky exploration of irradiated spaces. Turning back to me, Evans says, "We didn't have any robots at the early conferences ... but the conference has grown in attendees and scope!"



Cerberus QUGV is a Boston Dynamics robot customized by RADeCO, a Connecticut-based company that specializes in air sampling and detection systems for the energy and defense industries. The robot dog is equipped with radiation detectors and a wireless system to transmit data from risky locations to remote operators.

At an afternoon session providing an overview of emergency management, Todd Smith of the NRC's Office of Nuclear Security and Incident Response explains that planners should focus on the most likely incidents, but can extend preparedness to less likely

accidents. "You're an emergency manager—you have to balance the resources of your community, right?" Smith contrasts the likelihood of radiological emergencies with wildfire, flood, and hurricane probabilities,



## ICI C<sup>2</sup>BRNE DIARY – June 2023

and turns to a slide that puts the probability of certain kinds of nuclear reactor failure at 1 in 10 million years, or in some cases, he says, “beyond the lifetime of the universe.”

“All the uncertainty, all the unknowns. That’s why we’re doing what we’re doing,” Smith says.

Because most of the protocols (and funding) around radiological emergencies emerge from a regulation framework that governs nuclear power plants, accidents involving commercial power are the main focus of radiological preparedness in the United States. But the people attending the conference also prepare for a variety of other radiological situations, like lost radioactive sources, or rocket launches with radioactive payloads (like the plutonium-heated Perseverance Mars rover), and, occasionally, nuclear detonations. Although those incidents are rare compared to natural disasters and other industrial accidents, the conference is taking place just weeks after several reports of radioactive items being lost and recovered and during an ongoing leak of tritium at the Monticello nuclear plant in Minnesota.

On an elevator during a lunch break, an attendee from Minnesota tells another how she noticed her own farm pictured in a slide at one of the sessions. “I said, ‘That’s my chicken coop!’” Her farm had been used as the site for an exercise on ingestion pathways, which helped emergency planners practice taking samples, but also gave them some unexpected insight that could be important in a radiological emergency: Cows, like cats, use their tongues to groom each other, which means if contaminated material lands on their hair, it can end up inside their bodies too.

Tornadoes roll across the Midwest while I’m standing at a booth with representatives of IPAWS, the FEMA system that pushes cell phone notifications about weather alerts, child abductions, or even [incoming missiles](#). As I head down to the next session I run into Kelly Van Buren, Emergency Services Coordinator at the San Luis Obispo County Office of Emergency Services (which handles offsite emergency planning for California’s only remaining nuclear plant, Diablo Canyon), who will lead management of the NREP conference next year. She pauses to show me a photo of a semi-truck that has flipped over in the storm and is blocking the I-65 highway north of Indianapolis. Then she races off.

Moments later I’m not surprised to see an alert on the conference’s smartphone app, warning anyone traveling northbound about the tornadoes and the accident. This is a conference of emergency preparedness experts, after all.



A diversity of dosimeters and other radiation detectors was on offer from multiple vendors at the 2023 NREP conference, alongside drones and other equipment for radiological emergencies.

Communication is a primary theme of the conference; about one in five sessions are explicitly devoted to improving public messaging around emergencies. At an interactive spokesperson training, Eric Singer warms up the couple of dozen people gathered for his session while his colleague sets up a tripod with a camera (or as Singer calls it, an ENG—“electronic news gathering device”).

Singer is a former broadcast news anchor who now works as an emergency management communication specialist with the Risk and Crisis Communication Program team from Argonne National Laboratory. The team provides hands-on media training for emergency responders across the country. Singer’s presentation provides tips on best practices for appearing on camera, ranging from how to keep one’s eyes on a camera lens to avoiding jargon and acronyms and staying on message. One slide is titled “What to say when you can’t say anything.” The idea is to learn “how to say ‘no comment’ in a different way,” Singer explains, but without missing the opportunity to educate and inform the public.



Focus instead on “how the stew is made,” Singer says, like outlining the steps authorities and scientists are taking to gather more information and respond. Some of this advice might seem to enter spin doctor territory, but the point isn’t to obfuscate. Uncertainty and confusion can exacerbate dangerous situations, or make a minor problem worse by exaggerating hazards and causing panic during an emergency. So the goal of the exercise is connecting with the public to get clear information across. Singer’s next session is in the same room and goes even deeper. It’s called “Empathy isn’t for Weaklings.”

Communicating clearly is a thread that runs through nearly all the workshops I attend at the conference, even the more technical ones. At a day-long introduction to the Federal Radiological Monitoring and Assessment Center (FRMAC), Brian Hunt details the various software systems the center’s teams use for collecting and analyzing radiation data. A tool built by Sandia National Laboratories to track data requests from other agencies looks a bit like the popular project management app Trello, but it’s currently being used to support FRMAC’s tracking of atmospheric conditions during the war in Ukraine.

Making sure new information is vetted and accessible before it’s used to communicate updates to the public is an important part of FRMAC’s work in emergencies. “Most of what we do is not intended for the public, because it’s not couched in terms that people will understand,” Hunt says. “The general public doesn’t know anything about radiation, and you’re going to have to couch it properly, so that the proper messages get through.”

The phrase “right of boom” is used repeatedly during the four days of the conference, referring to what is done to respond after some event (the boom) creates a radiological emergency. Everything done to prevent those events from happening is “left of boom”—that might include reactor and power plant safety, but also nonproliferation, arms control, and even nuclear deterrence. The emphasis at NREP is almost entirely on what’s right of boom—for the most part, this is a conference of health physicists and emergency managers, not nuclear reactor designers or defense experts.

But some of the conference-goers find it problematic that these areas remain so separated. Arnold Bogis spent his early career working on issues like nuclear non-proliferation (left of boom) at think tanks like Harvard University’s Belfer Center for Science and International Affairs. Today he works for the Association of State and Territorial Health Officials, where he aided a CDC-funded project developing guidance for use of cytokines to treat severe radiation exposure after a nuclear explosion (right of boom). This is his second year at the NREP conference, but he’s been involved with the radiological emergency crowd for a while, traveling as far as Guam to help teach local emergency workers how to deal with radiation.

Based on his experience with both left and right of boom experts, Bogis feels a lack of communication between them has real implications. He cites the crisis at the Zaporizhzhia nuclear power plant in Ukraine as one example of something both communities should be interested in. “I haven’t heard anything about that plant here at all, not one mention. ... No one is talking about this ongoing experiment in nuclear safety,” he says, and he’s largely right—the most I hear about it is a comment during a slideshow of a tour to Chernobyl on the last morning of the conference.

“You rarely see the FEMA people over at nonproliferation conferences,” he says. “I find it pretty unlikely ... that anybody at this conference, even probably the federal folks, has ever read anything out of Arms Control Today.”

With more interaction between the fields, those working to prevent use of nuclear weapons could incorporate the realities of emergency planning into their research and their advocacy. Likewise, emergency planners would benefit from deeper understanding of nuclear defense and nuclear energy debates, especially when it comes to public communication. So Bogis would like to see experts from think tanks like the Nuclear Threat Initiative and the Carnegie Endowment for International Peace attending, and even presenting, at future NREP conferences. “Come and give a talk here,” he says, “I bet people will be fascinated.”

The felicitously named Martin Vigil participated in the 10-kiloton nuclear detonation drill with Bogis and is also concerned that not enough is being done to prevent the nuclear disasters. When I introduce myself after the drill, he tells me he has his own homemade Doomsday Clock on his mantle at home, and he points out that he’s wearing a shirt embroidered with his own version of the old Civil Defense logo.

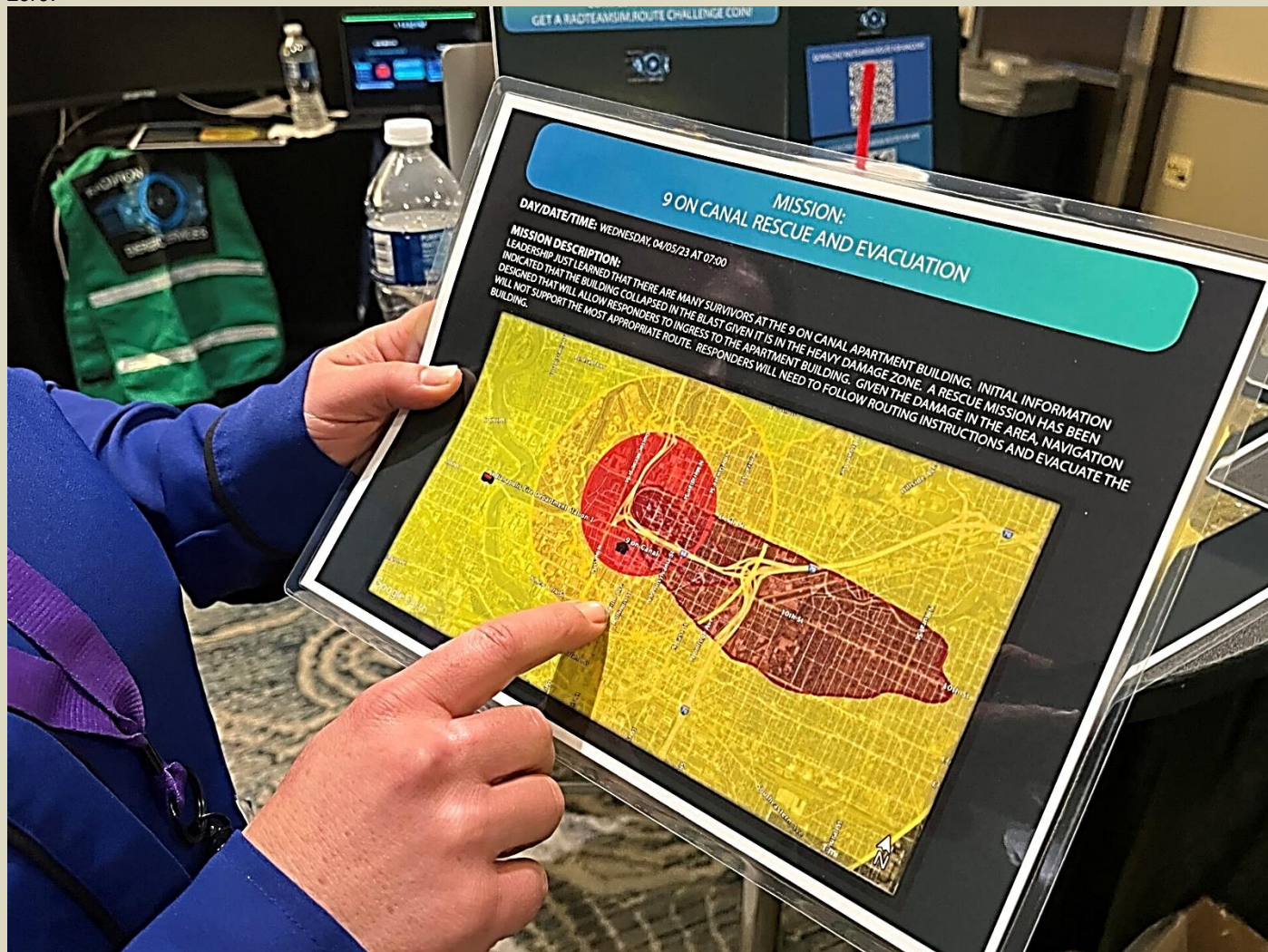
Today Vigil is the director of the Santa Fe County Office of Emergency Management, which among other things helps secure transports out of Los Alamos. (“We don’t know what’s in there,” he says.) Vigil is concerned about global nuclear tensions and broken arrows (close calls with nuclear warheads). He’s also very worried that the country isn’t prepared for a nuclear detonation. “Nationwide, there’s, I think, 1,800 burn beds and less for pediatric specialties,” he says. “Some of the modeling for just a 10-kiloton detonation, you know, you’re up in the 80,000-to-100,000 range [of burn victims]. How do we even begin to manage that?”

Most people I speak with at the conference are concerned about rising nuclear tensions worldwide, in light of Russia’s invasion of Ukraine and the dangers surrounding the Zaporizhzhia nuclear plant. But few tell me they see a heightened threat of nuclear war changing much about how they prepare for emergencies. Rae Walker, an emergency preparedness coordinator in the Radiation Control Program at the Texas Department of State Health Services and the outgoing chair of the NREP conference steering committee, says her own impression is that “the tempo is different” between the concerns of those who focus on weapons of mass destruction and those who focus on emergency preparedness with nuclear plants. One of the ROSS volunteers puts it another way: If they’re able to





respond to a major nuclear catastrophe, they'll follow their plans and training; if not, well, it's probably better to be closer to ground zero.



Wendy Renno explains one of the training missions from the nuclear detonation drill that used the RadTeamSim.Route driving simulator she designed. Renno previously worked for the Department of Energy's Federal Radiological Monitoring and Assessment Center, including during its response to the 2011 Fukushima Daiichi nuclear plant emergency.

As support for nuclear power has waned since the 1980s and plant development in the United States has stalled, the nuclear industry has suffered significant declines in the size of its workforce. It's a problem not just for the nuclear plants, but also for emergency preparedness; fewer people with decades of knowledge and experience—the Ken Evanses of the world—are available to mentor younger planners. Walker notes that retirement of older professionals (including “Navy nukes,” as Van Buren calls those who got into the nuclear field through their work on nuclear submarines) has been an increasing concern, and that the pandemic appears to have accelerated the process.

But the conference may be able to help with the aging-out problem. Walker says that her committee has only recently started gathering more detailed demographic data about conference registrants, but that this year's conference does skew younger than it has before. (The youngest person attending appears to be a 19-year-old emergency responder from St. Lucie, Florida.) Perhaps more important, many of this year's attendees are new to the conference itself. During the opening plenary session, Walker asks any newcomers to stand, and a good third of the crowd gets out of their seats.

One of the youngest members of the conference steering committee is Jack Wiley. He's in his 20s and studied nuclear engineering as a way to work on green energy, a common motivation for others his age who are concerned about climate change. He switched his attention to health physics and now works in the Radiological Emergency Preparedness section of the Washington State Department of Health. He's been working there four years and has already seen a dozen retirements in that time. He tells me there's an especially big gap between those in their late 40s or older, and people his age who are early in their careers. “They don't know the



stories behind what happened. They might read it in a textbook, but have no context for it,” Wiley says. “I’m trying to bridge that gap by, you know, being present and asking those things and hearing those stories firsthand. ... Really trying to extract out as much of that information before people retire.” I ask whether he’s noticed any concern or changes in preparedness given the situation with the Zaporizhzhia plant in Ukraine. Besides getting new background radiation readings in case something disastrous does happen, he hasn’t noticed his colleagues taking extended precautions. But paying attention is important. “I like historical things happening,” he says. “Not things like this. But it’s interesting to me to see the reaction of the [radiological emergency preparedness] community based on historical events of this scale.”

A much less dramatic historical event is part of NRC Commissioner David Wright’s keynote address at the conference. He notes that the new Vogtle 3 reactor in Georgia was hooked up to the grid just that weekend, 14 years after construction began, and the first time a new reactor has generated commercial power in the US since 2016. There’s no cheering among those assembled, but no heckling either. Few of the people I speak with express strong opinions about nuclear energy one way or another, so I don’t get a good sense of how many think this emergency preparedness work is worth the effort, to keep aging nuclear plants in business. Do the energy benefits of nuclear power justify living with the uncertainty and risks that these emergency experts must constantly prepare for year after year?

New developments may change the calculus. In January, the NRC certified a small modular nuclear reactor design for the first time. The potential impact of an industry-wide shift to small modular reactors comes up now and then in a few sessions of the conference, but most still see that as a long way off. One thing that no one doubts though, is that the new regulations that accompany the new reactors will probably change how offsite emergency planning is done. According to Evans, it’s possible that with the new reactors, the NRC’s required emergency planning zones will be reduced—potentially eliminating the need for the offsite planning that’s been in place since the Three Mile Island accident over four decades ago. But the broader need for radiological emergency preparedness won’t go away anytime soon. It will take many years to get new reactors online; it also takes years to decommission them. And as the participants in the first day’s drill established, there are other dire possibilities to be prepared for.

Thomas Gaulkin is the *Bulletin of the Atomic Scientist’s* multimedia editor.

## A Nuclear Collision Course in South Asia

The Budding Arms Race Among China, India, and Pakistan

By Andrew F. Krepinevich, Jr.

Source: <https://www.foreignaffairs.com/china/nuclear-collision-course-south-asia>

May 26 – In the summer of 2021, the world learned that China was dramatically expanding its nuclear arsenal. Satellite imagery showed Beijing building as many as 300 new ballistic missile silos. The Pentagon now projects that China’s stockpile of nuclear weapons, which had for years rested in the low hundreds, could spike to 1,500 warheads by 2035, confirming suspicions that Beijing has decided to join Russia and the United States in the front rank of nuclear powers.

Security experts are only beginning to sort through the implications of China’s nuclear breakout. They would do well to consider Ashley Tellis’s new book, *Striking Asymmetries*, which assesses the implications of Beijing’s actions from the vantage point of the rivalries between South Asia’s three [nuclear](#) powers: China, India, and Pakistan. In a work that should be required reading for senior political and military leaders, Tellis presents a compelling case why this tripolar nuclear system, which has for decades remained remarkably stable, may be on the verge of becoming far more dangerous.

Tellis draws upon decades of experience in South Asian [security](#) affairs, unique access to senior policymakers and military leaders in the three rivals’ defense establishments, and a remarkable ability to make seemingly abstract technical concepts readily understood by those with even a passing interest in the subject matter. The result is the most comprehensive, informed, and accessible assessment to date of this nuclear rivalry—and one that cannot be ignored.

### The race is on

[China](#) and Pakistan have a long and close relationship, in part built around their mutual view of India as a rival. India finds itself sandwiched between these two often hostile powers. Yet despite a history of wars and persistent low-grade conflict between India and its two rivals, a general war has been averted since India and Pakistan became nuclear powers a quarter century ago. Moreover, the three countries have not found themselves caught up in a nuclear arms race. Until recently, they viewed their nuclear weapons



primarily as political instruments, not as tools for actual warfighting. All three adopted a “minimum deterrent” nuclear posture, maintaining the lowest number of nuclear weapons necessary to inflict unacceptable damage to their adversaries’ key cities even after suffering a nuclear attack.

In keeping with this strategy, the three Asian rivals avoided maintaining a significant portion of their arsenals on high alert. Instead, they stored their weapons in caves, in deep underground facilities, or in other concealed locations. Rejecting American and Russian notions that “retaliation delayed is retaliation denied,” the three countries, especially China and India, forswore the need for a swift response to a nuclear attack. To be sure, they would respond eventually—in days, weeks, or even months—but they did not accept the imperative of immediacy. As a result, these countries have avoided making heavy investments in early warning systems while retaining centralized control over their arsenals.

But the prospects for sustaining this era of minimum deterrence appear increasingly shaky. The tripolar rivalry has not been locked in amber: Tellis describes strongly held beliefs among top security officials in China, [India](#), and Pakistan that their nuclear postures are inadequate. Led by China and Pakistan, with India following in their wake, the three rivals are now on a course that will result in a dramatic expansion of their nuclear arsenals, even if Russia and the United States pursue substantial cuts to theirs.

### Two against one

At the core of Tellis’s assessment are the differences—“asymmetries”—driving the tripolar rivalry. One fundamental difference is that China and [Pakistan](#) are revisionist powers seeking to alter the existing order, while India remains content with the status quo. China possesses the most formidable nuclear arsenal of the three, followed by Pakistan, with India trailing.

There is also an asymmetry in the three powers’ strategic focus. Pakistani security officials are obsessed with India, while India’s focus is overwhelmingly on China. China’s sights, however, have shifted beyond regional to global rivalries, principally with the [United States](#). It is this competition with Washington that is driving Beijing’s nuclear breakout. For China, India’s deterrent is rapidly assuming a peripheral role, similar to that played by China in American nuclear planning during the Cold War.

Beijing’s support for Pakistan’s nuclear weapons program, which includes providing Islamabad with blueprints for a bomb and fissile material, has further complicated India’s position. Pakistan’s leaders are looking to abandon minimum deterrence in favor of “full-spectrum deterrence,” where their nuclear forces cover multiple contingencies in the event of war with India. There are three central factors spurring Pakistani officials to adopt this more aggressive posture. First, Islamabad is aware that its conventional forces are weaker than India’s and believes it has no alternative but to employ, if need be, its nuclear forces to offset this asymmetry. Second, given that India is far larger than Pakistan, Islamabad believes it must be able to inflict greater destruction on India in a retaliatory strike than India will inflict on it. This requires Pakistan to maintain a larger nuclear arsenal to target India’s population and economic hubs in the event of war. Third, Pakistan also hopes that its nuclear forces prevent India from undertaking large-scale military action against it in response to Islamabad’s ongoing support for militant groups in the disputed region of Kashmir.

Tellis shows that accomplishing full-spectrum deterrence will require Pakistan to expand its arsenal substantially. For instance, he notes that stopping a major advance of Indian conventional forces into Pakistani territory would require scores of so-called tactical nuclear weapons, weapons that Islamabad currently lacks.

### A fragile peace

Although Tellis argues that Beijing’s and Islamabad’s nuclear provocations do not automatically portend growing instability in the region, the evidence he presents suggests otherwise. He finds that Beijing’s growing arsenal will not necessarily place India’s security at greater risk—but describes a set of highly plausible Chinese actions that, in combination with a superpower-sized arsenal, risk undermining India’s confidence in its own nuclear deterrent.

To begin with, Beijing is seeking the capability to launch nuclear reprisals far more quickly than ever before. This requires China to maintain a portion of its force on heightened alert, which may not have posed a threat to India when China possessed a few hundred weapons. But if Beijing placed a significant percentage of its [expanded arsenal](#) of 1,000 or more warheads on high alert, the strategic ground would shift considerably. India would now face a neighbor capable of launching a large-scale attack with little or no warning. India’s ability to withstand a nuclear strike and retain the capacity to inflict catastrophic destruction in response is closely tied to the security of its underground nuclear storage sites. China currently lacks the ability to destroy them—even assuming it knows their locations. That could change, however, once China’s arsenal has more than 1,000 warheads, especially if China improves the accuracy of its weapons. Such a development, combined with Beijing’s adoption of increased alert levels for its nuclear forces, would set alarm bells ringing in New Delhi; Indian officials could conclude that China has the capacity to disarm India’s nuclear weapons arsenal. China may also enhance its air and missile defenses, making matters even more precarious for India. These defenses would minimize the threat posed by any “broken-back” Indian nuclear retaliation—in other words, an attack that uses whatever weapons survive a disarming Chinese strike. But New Delhi would surely know that employing the remnants of its arsenal to retaliate against China would leave it vulnerable to Pakistani nuclear blackmail. Put simply, India would risk being left with no credible nuclear



deterrent to resist coercion by Islamabad. Tellis is correct to note that China's development of these capabilities is not assured. Yet during Beijing's decades-old conventional military buildup, it has sought to match every significant U.S. capability, including stealth fighters, military satellite constellations, aircraft carriers, and cyberweaponry. Tellis recognizes that even if China creates such a set of capabilities, it must still know the location of India's storage sites in order to target them—and have high confidence that its [intelligence](#) is accurate and comprehensive. This uncertainty could restrain Beijing. But at the same time, New Delhi may not feel comfortable simply trusting that its nuclear sites have not yet been unearthed by Chinese intelligence or presuming that Chinese leaders are wary of taking big risks.

### New Delhi's dilemma

How might India respond to China's and Pakistan's nuclear provocations? Tellis points out that India is not without options—but that each path has its pitfalls. First, he shows that if India wanted to, it could easily match China weapon for weapon. Yet he believes New Delhi would prefer to maintain its minimum deterrent strategy, emphasizing its ability to inflict severe damage on its adversaries' cities. This stems in no small part from the expense India would incur by following Beijing in its quest to match America's nuclear arsenal. Still, Tellis acknowledges that India's arsenal will have to expand its nuclear holdings to possess the warheads needed to inflict unacceptable damage on both China and Pakistan. And as India increases its arsenal, Pakistan is sure to do the same—completing the regional chain reaction triggered by China's nuclear expansion.

Tellis rejects the “more of the same” option of expanding India's underground storage facilities, showing persuasively that it would prove costlier to accomplish than it would for China to simply expand the number of weapons needed to destroy them. Rather, he argues, India's solution is to be found in stealth and mobility. This could be achieved by creating a nuclear ballistic missile submarine force and by shifting more of India's arsenal to mobile road and rail missile launchers.

As for China's air and missile defenses, Tellis points out that India might address the problem by deploying penetration aid decoys on its missiles. These decoys are designed to present themselves as actual warheads to missile defense radars, thereby inducing the defender to expend precious interceptor missiles engaging false targets. This would offset, if only partially, New Delhi's need to expand its nuclear arsenal.

Yet even if India were to pursue these actions, it would still face significant challenges. The [threat](#) of a Chinese preemptive strike may compel India to develop an effective early warning system to enable it to reduce its arsenal's vulnerability by sending its weapons out to sea and flushing its land-based missiles from their silos. New Delhi would also have to establish a new command-and-control system to direct the actions of its nuclear submarines. Yet while India is in the process of constructing nuclear-powered ballistic submarines, it still has a long way to go in building a significant force and overcoming the technological hurdles necessary to create a credible seaborne nuclear deterrent. Tellis notes that among these challenges, New Delhi is experiencing problems with its naval nuclear reactor designs. Then there are India's nuclear weapons. New Delhi has only conducted a handful of nuclear tests—not enough to validate its thermonuclear designs to offer high confidence that these weapons will perform as designed. Its most reliable weapon has a yield of 12 kilotons, whereas China's weapons have yields as much as 100 times greater. Addressing these shortfalls may require India to resume testing—and risk incurring sanctions from the United States and other nations.

Tellis hints at a tantalizing solution to India's problems. The United States could provide India with a reliable thermonuclear weapon design. The trilateral security pact among Australia, the United Kingdom, and the United States that is known as [AUKUS](#), which will assist Australia in acquiring nuclear-powered submarines, could be expanded to include India. Might the Americans also share their nuclear reactor designs with New Delhi? But for this to happen, India, which has kept the United States at arm's length practically since its birth, would have to finally and firmly close ranks with the leading Indo-Pacific democracies and formally forsake the nonaligned strategic autonomy it has long enshrined at the heart of its foreign policy.

[Andrew F. Krepinevich, Jr.](#), is a Senior Fellow at the Hudson Institute and an Adjunct Senior Fellow at the Center for a New American Security.

## Unmonitored networks put US nuclear arsenal at risk, GAO finds

By Edward Graham

Source: <https://www.nextgov.com/cybersecurity/2023/05/unmonitored-networks-put-us-nuclear-arsenal-risk-gao-finds/386910/>

May 30 – The Energy Department needs to take additional steps to prevent insider threats to the nation's nuclear arsenal — including working to identify the total number of classified networks across the department to fully monitor users' activity — according to a recent report from the Government Accountability Office. GAO's [report](#) — released on May 24 — reviewed the effectiveness of Energy's Insider Threat Program, one of the department's risk mitigation initiatives that is designed “to further protect



against insider threats from employees, contractors and trusted visitors.” The study was requested in a House [report](#) accompanying the fiscal year 2022 National Defense Authorization Act. Despite the program being established in 2014, GAO said that multiple independent assessments conducted in the intervening years found that Energy “has not implemented seven required measures for its Insider Threat Program, even after independent reviewers made nearly 50 findings and recommendations to help DOE fully implement its program.” Four of these “unmet minimum standards” were previously identified in a March 2022 memo sent from the Office of the Director of National Intelligence to the Energy Secretary, while the remaining three “were found to be unmet through DOE’s Office of Enterprise Assessments’ review of DOE’s Insider Threat Program in 2021.” GAO identified continuing concerns — first conveyed by ODNI — about Energy’s efforts related to “monitoring user activity on all classified networks.” The report noted that “minimum standards require that insider threat programs include the technical capability to monitor user activity on all classified networks,” but that the department’s Insider Threat Analysis and Referral Center “has not met full user activity monitoring coverage requirements on all classified networks.” While GAO said the department “has processes for addressing concerns on unmonitored classified networks should an event be detected by other means,” it noted that Energy officials “have not identified the total number of DOE’s stand-alone classified networks, which leaves them unaware of the extent to which the Insider Threat Program falls short of minimum standards for user activity monitoring.” The watchdog also highlighted Energy’s continued failure to produce an annual progress report on its various threat mitigation programs since 2017, which is meant to document “annual accomplishments, resources allocated, insider threat risks to the agency, recommendations and goals for program improvement and major impediments or challenges.” Energy officials told GAO that an annual report had not been completed since last decade “because the program decided to wait until independent assessments of the Insider Threat Program were completed,” and because “program staff did not have access to classified materials while working remotely during the COVID-19 pandemic, which contributed to some of the delay in annual reporting.” In its last annual report, Energy “reported experiencing about 250 unclassified insider threat-related security incidents in 2017,” with the department considering “about 100 of those incidents to be serious.” Most of these incidents “were unintentional,” according to the watchdog, and included “sending classified information over unclassified systems, leaving security areas unattended and not properly protecting classified information.” Other unmet standards GAO identified in Energy’s Insider Threat Program included inconsistent insider threat awareness training for employees, the department’s inability to validate the completion of training “for all cleared employees and contractors,” the lack of “a formalized independent assessment element” for oversight compliance reviews and no established procedures for personnel accessing sensitive or protected data. The report also cited a failure to ensure staff associated with the program “were fully trained on legal issues, response actions, handling of data and records, civil liberties, privacy and investigative referral requirements.” The watchdog faulted Energy for dividing “significant responsibilities” of the program between two offices, noting that “the program’s senior official resides within the security office, while operational control for insider threat incident analysis and response resides within the Office of Counterintelligence — a part of the organization with its own line of reporting to the Secretary of Energy.” In addition to not fully integrating the program within one office, GAO said Energy “has not identified and assessed the human, financial and technical resources needed to fully implement its Insider Threat Program.” “For example, DOE’s budget does not include dedicated funding for its contractor-run nuclear weapons production and research sites to carry out their responsibilities for implementing the program,” the report said. “Unless DOE identifies and assesses the resources needed to support the Insider Threat Program, it will be unable to fully ensure that components are equipped to respond to insider threat concerns, potentially creating vulnerabilities in the program.”

GAO made seven recommendations to Energy, including calling for the department “to track and report on actions it takes to address reviewers’ findings and recommendations, to establish a process to better integrate program responsibilities and to assess resource needs for the program.” Energy agreed with all of the watchdog’s recommendations.

[Edward Graham](#) is a technology reporter for Nextgov/FCW, where he reports on national security technologies and policies, the Department of Veterans Affairs and election security issues. He received his master’s degree in Journalism and Public Affairs from American University in 2016.

## A book, nuclear weapons, means and ends

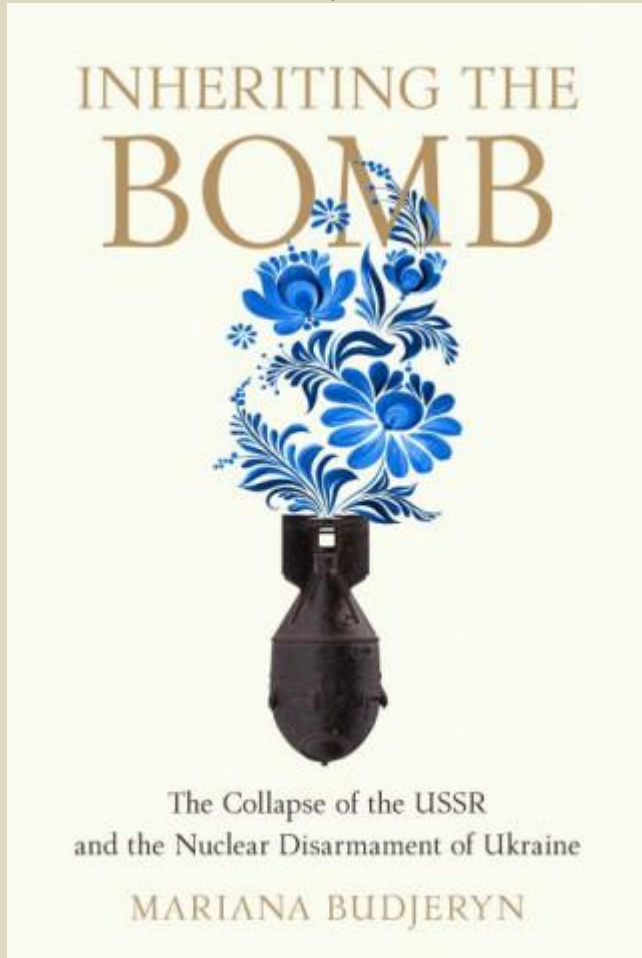
By John Erath

Source: <https://armscontrolcenter.org/a-book-nuclear-weapons-means-and-ends/>

One of the benefits of working at the Center is the opportunity to collaborate with some of the top experts in the field, people who bring knowledge, insight and a variety of different viewpoints to discussions of international security. Earlier this year, I was privileged to chair a panel including one such expert, [Mariana Budjervyn](#), who has recently published a new book, [Inheriting the Bomb](#), a study of Ukraine’s nuclear legacy and the process by which it gave up the nuclear weapons inherited from the Soviet Union. The



story of Ukraine's progress toward disarmament is intimately connected with Ukraine's realization of its regained status as an independent state and the way it defined its nationhood. In researching the book, Mariana was able to gain access to previously



unavailable sources and recently declassified records — as well as interviews with some of those involved — to tell the story in a more complete way than has been previously related.

With Ukraine under attack from its larger neighbor, this book assumes a greater importance. Some commentators have suggested that Ukraine was wrong to give up its nuclear weapons, and that possession of such weapons might have been enough to deter Moscow from invading. Mariana demonstrates conclusively that this was never an option. Ukraine lacked much of the necessary command and control and support infrastructures and was not in a position financially to build them. Additionally, the international community, on whom Ukraine depended for assistance, was strongly against adding nuclear weapons states. Perhaps most importantly, the legacy of the Chernobyl accident and first-hand experience with the effects of radiation led to reluctance to rely on anything nuclear for national security.

One aspect of the book which I had not previously considered was the crucial role of the United States in Ukraine's denuclearization process. U.S. [Cooperative Threat Reduction \(CTR\)](#) was a key facilitator and enabled the destruction of thousands of nuclear weapons throughout the former USSR. Even more important was Washington's insistence on a non-nuclear Ukraine in the process of managing the dissolution of the Soviet Union. Although this was undoubtedly the right, and really the only possible, outcome, it was the way that U.S. leaders went about getting there that possibly could have been done better. Thinking about the process as it was described in the book, it seems as though there was a major error.

Both the Bush and Clinton administrations chose to define the problem to be solved as denuclearizing Ukraine. In reality, this should have

been a means to an end, one element in a broader strategy to make the Eurasian region more stable following the demise of the Soviet empire. By treating Ukraine's ratification of the [START treaty](#) and accession to the [NPT](#) as a non-nuclear state in return for a rehashing of existing security assurances as the goal, U.S. policy did not address the other challenge to security in the former Soviet space: Russian revanchism, or determination to dominate what had been their empire. This was demonstrated early on by Russian attempts to influence Ukrainian politics and by the encouragement of separatist elements in Crimea and the Donbas, a playbook implemented with varying degrees of success in the Baltics, Georgia, and Moldova.

By making the nuclear issue the primary focus of Ukraine policy, Washington may have inadvertently undermined its own non-proliferation goals. The message seemed clear that America cared only about weapons of mass destruction, so that defining the problem as getting Ukraine to give up its nuclear weapons may have contributed to the perceived importance of such weapons. The decades following the Cold War saw this played out again and again, in North Korea, Iraq, Libya, South Asia and most recently, Iran. The question of how to manage Ukraine's denuclearization also brought about a rare communion of U.S. and Russian interests. Both large countries prioritized having Russia as the sole inheritor of the Soviet nuclear capability, and both brought pressure on Ukraine to achieve the goal. In the end, though, Russia let the United States do most of the diplomatic heavy lifting, effectively getting what it wanted for little more than symbolic concessions. The most important of these, the 1994 Budapest Memorandum, in which the nuclear states "guaranteed" Ukrainian sovereignty, proved easy to discard in 2014 when the time came to seize Ukrainian territory. It is not possible to state that had Washington identified the real issue — the peaceful succession of independent states to the USSR — correctly and molded its policy to use denuclearization as a means to that end, Russia would not have invaded in 2014 or 2022. As Mariana writes, "Nuclear decisions are multicausal." Keeping Ukraine as an appendage, or ersatz colony, of Russia is perceived in Moscow as too vital an interest to forego. However, the current situation shows that subordinating all other issues to the question of how to induce Ukraine to give up its nuclear weapons was ultimately unsuccessful in terms of averting Russian aggression and did not lead to a secure Ukraine able to deter revanchism. Arms control remains a tool, when done correctly a highly effective one, to improve international security, but becomes less effective when treated as an end in itself. The Clinton



administration scored an important success in helping Ukraine divest its nuclear weapons, given that keeping them would have been all but impossible. It needed to do more in helping Ukraine build its own security.

For this reason, I hope that Mariana's book becomes required reading for students of arms control and policymakers alike. As many are starting to question the future of arms control, it is good to be reminded that it continues to have an important role as part of broader diplomatic efforts and should be an element of future security arrangements. As we look to the eventual end of the current war, it is well to consider what contribution arms control can have in building stability.

## Prove It Before You Use It: Nuclear Retaliation Under Uncertainty

By Johnathan Falcone, Jonathan Rodriguez Cefalu, Michael Kneeshaw, and Maarten Bos

Source: <https://warontherocks.com/2023/06/prove-it-before-you-use-it-nuclear-retaliation-under-uncertainty/>

June 01 – It is [2028](#), and the United States Space Force's [early warning radar modernization](#) is complete. Technical Sergeant Jack Nichols works at [Buckley Space Force Base](#) operating systems that detect and assess ballistic missile threats against the United States and Canada. Since arriving at the Colorado base, Nichols has experienced his share of false alarms. However, these are no ordinary false alarms; the system Nichols watches provides early warning that the United States is under ballistic missile attack. While these existential alerts would distress most, he maintains an “old school” validation protocol: He evaluates the warning against his sensor's input settings and raw data output, resolving any concerns.

But today, the warning that flashed across his screen was different. Recent modernization efforts introduced [next-generation sensors and machine learning-powered tools](#) to manage the increased flow of information. These purported improvements made the raw data inaccessible to Tech Sgt. Nichols. The system had identified an incoming missile, but he couldn't help but wonder: What if this was a [mistake](#)? What if the system had been [hacked](#) or had malfunctioned? And, just as unsettling, what if the newly implemented algorithm had made a decision based on [flawed or biased data](#)?

To some extent, his concerns do not matter. His training dictates that he has [less than two minutes](#) to evaluate and report the warning. This expediency ensures the president maintains the option to launch a retaliatory nuclear strike before an adversary's weapon — if a first-strike weapon is, in fact, inbound — strikes the American homeland. Nichols understood that the president's decision to retaliate requires balancing the inherent limitations of early warning accuracy with the concern that presidential control may be lost if the warning turns out to be true. But, he wondered, could the pressure from this uncertainty be alleviated if the president could issue a delayed order?

### A New Nuclear Era

Russian President Vladimir Putin's [threats](#) to deploy tactical nuclear weapons in Ukraine as well as North Korea's provocative [ballistic missile testing](#) have [renewed concerns](#) about the possibility of nuclear escalation. Meanwhile, China's burgeoning [submarine-launched](#) deterrent capability and Iran's [rebuilding](#) of its nuclear capability have provided additional reasons for concern.

This unease is exacerbated by the advanced offensive capabilities in cyberspace demonstrated by these [same actors](#). U.S. adversaries, such as Russia and China, have targeted [critical national infrastructure](#), including [electrical grids](#) and [nuclear facilities](#). Perhaps what is most destabilizing is that these adversaries are [incentivized to hide their capabilities](#) until they are ready to be used, so the true extent of the [cyber-nuclear threat](#) is unknown.

Given this security environment, the Biden administration continues the push [to modernize](#) the nation's nuclear deterrent. This modernization effort includes investing in the capacity and hardening of the [nuclear command, control, and communications](#) architecture. Furthermore, it potentially entails the integration of [machine learning systems](#) and other emerging technologies — despite [objections from experts](#) writing in these virtual pages — as outlined in the [Nuclear Posture Review](#).

However, as the hypothetical vignette in our introduction illustrates, modernizing equipment and systems may not be enough to achieve the administration's goals of [“non-use and to reduce the risk of a nuclear war.”](#) President Biden — and any future U.S. leader — still [retains](#) the “launch-under-attack” option. In this approach, when early warning sensor data indicates a [“medium or high confidence”](#) of a threat, the White House is alerted, and the president and their advisors convene. At this emergency conference, the president will be briefed their options and decide whether or not to launch nuclear weapons, even if the warning's legitimacy is not [conclusively](#) determined.

This approach is a remnant of the Cold War. We argue it is inadequate in today's strategic landscape, given the proliferation of nuclear weapons and cyber capabilities, as well as the [technical limitations](#) and [human biases](#) associated with the use of automated and machine learning systems. Instead, we argue that this administration should break from its predecessors and adopt a “decide-under-attack” posture. This action would shift the retaliation posture from a time-constrained decision in the fog of war to deliberate action based on evidence of an attack.



## Cold War Posture Endures

In the 1970s, the United States [was concerned](#) that the Soviet Union could launch a surprise attack using thousands of land-based missiles against then-vulnerable [Minuteman missiles](#) and command and control nodes. The concern was that after this attack, the United States would be unable to retaliate with nuclear weapons. To deter this threat and maximize response options, a [launch-under-attack posture](#) was adopted in 1979. Under this posture, Minuteman missiles were required to [launch within 30 minutes](#) of receiving reliable warning that the United States was under attack. Later, in the 1980s, submarine-launched ballistic missiles [were also configured](#) to this posture.

This policy was extended even after the fall of the Soviet Union. Planners determined that an effective counterstrike required, at minimum, a [five-minute launch sequence](#). This left the remaining 25 minutes for satellite and radar detection, operator assessment, communication to the president, and a nuclear-use decision. These time constraints encouraged successive U.S. administrations to maintain the launch-under-attack policy.

However, [simulations by Massachusetts Institute of Technology researchers](#) have demonstrated that hundreds of silo-based Minuteman missiles would likely survive a first strike. In fact, according to recent [analysis](#) published on this platform, the United States would maintain “more warheads per retaliatory target than before the Russian strike,” weakening the primary rationale for the posture. This somewhat puzzling result is due to the [survivability of U.S. silo-based missiles](#) and the fact that there will be fewer [military targets](#) remaining, since many Russian missile silos will be empty after a first strike.

## “Launch-Under-Attack” in a Complex World

A launch-under-attack posture exposes the United States to an increased risk of [accidental or mistaken launch](#) in the modern nuclear era. To be available as an option, launch-under-attack relies on accurate warning data and a viable launch capability. The follow-on corollary is that to be effective in its deterrence role, adversaries must believe that a first strike would be detected and retaliatory weapons would be employed. Underpinning these capabilities is the nuclear command, control, and communications architecture. But unlike during much of the Cold War, modernized [command and control systems](#) are more reliant on computers and thus are susceptible to cyber exploitation. This is a significant risk when combined with an outdated retaliatory option, as it impacts incentives for preemptive or retaliatory nuclear launch decision-making.

Two cyber risks are routinely discussed in policy circles. First, [critical hardware](#) and [software components](#) may be compromised in supply chains. Adversaries can introduce malware or malicious code to [digital and automation components](#) to infiltrate both networked and non-networked elements of communications systems. If U.S. nuclear systems were compromised by a supply chain attack, it could either undermine the national command authority’s confidence in its second-strike capability or, from the adversary’s perspective, reduce the risk of a retaliatory strike.

The second cyber risk is [spoofing](#), which involves the [injection of false data](#) into key computer-mediated systems. [Spoofing](#) can take two forms in early warning systems: hiding actual inbound missiles or creating fake signals of inbound missiles. The former is more likely to originate from a nuclear peer in an effort to further compress Washington’s decision-making window by obfuscating early warning data in hopes of increasing the effectiveness of a first strike. The latter, on the other hand, is more likely to be injected by a non-peer or terrorist group aiming to manipulate global perceptions of American brinkmanship or trigger [catalytic nuclear war](#) between two or more powers.

During a crisis, cyber vulnerabilities can increase the risk of a preemptive strike or a mistaken launch. This is because cyber attacks can disrupt critical systems, which can reduce trust in early warning and second-strike capabilities. Additionally, such attacks can create confusion and make it difficult to distinguish between a genuine attack and a false alarm, potentially resulting in a mistaken launch from the side that thinks it is under nuclear attack. The launch-under-attack posture exacerbates this problem because it requires a decision to be made. Even if the president opts for nonretaliatory measures, this is still a deliberate choice amidst the prevailing uncertainty.

The rationale for this posture has also been challenged by proliferation, which has driven increased demands on technical systems. When the launch-under-attack posture was first implemented, there were only two major nuclear powers. This is not the case today. The 2022 Nuclear Posture Review [recognizes](#) both Russia and China as major nuclear powers and strategic competitors. In the absence of strategic intelligence suggesting an imminent first strike, the already short decision timeline is further compressed by the need to collate early warning system detection with ever-growing sets of radar and intelligence data. Command, control, and communications systems — particularly early warning system components — that are modernized and integrate machine learning will help [alleviate some of this information-induced pressure](#). However, technical limitations and [human biases](#) introduce additional risks.

Fundamental to machine learning systems are the data and algorithms that train the system. Data, which is used to train algorithms, can be [poisoned](#) or [biased](#), while the algorithms themselves [may produce results](#) of [indeterminable quality](#). Moreover, training machine learning systems based on infrequent occurrences is challenging. In the case of implementing machine learning tools for early warning systems,





the infrequency of missile launches poses a unique challenge for training these systems. And in the absence of real-world data, simulations will be used to generate the necessary data sets. Effective simulation data will rely on intelligence about adversary delivery capabilities. Inaccurate intelligence risks creating [bias](#) in the system's training, and there may be insufficient opportunities to validate the models using real-world events.

Accurately assessing nuclear capabilities is a [challenge](#) because intelligence is [fallible](#), and [open source data](#) reveals only so much. But these assessments and the follow-on technical challenges may be more pronounced in a scenario where the primary nuclear threat is temporarily a non-peer, [like North Korea](#). If a machine learning system is overtrained on particular data, it can make [inaccurate predictions](#) when presented with new information. For example, if early warning systems are overtrained on data from known Russian and Chinese capabilities, the model may misclassify sensor data from a new North Korean capability. More generally, proliferation — to include both [new states](#) developing nuclear weapons and existing powers [expanding](#) capabilities — generates greater uncertainty in model outputs. This uncertainty may make it more difficult for decision-makers to assess a threat.

Compounding these technical weaknesses is an [operator's tendency to overestimate](#) the system's accuracy, particularly as operators are further removed from the original data. For instance, when an operator interprets radar data, they will determine whether a missile is there or not. When an algorithm performs this interpretation, it may simply output whether an attack is in progress or not. Again, because actual events are infrequent, a system will frequently and correctly evaluate “no attack,” convincing operators and decision-makers that the system is more accurate than it is. This can lead to an overconfidence dubbed [automation bias](#), and it is especially prevalent in military settings due to [training](#) and organizational trust. The human-machine interaction at the operator level, combined with the launch-under-attack option for the president, are conducive to facilitating a positive launch decision, even without certainty of a threat.

### Building Resiliency Through Policy

The [Swiss cheese model](#) of accident causation is a risk management tool used in a variety of industries. The model uses a slice of cheese to represent individual safeguards. Each safeguard has inherent weaknesses, which are portrayed by the holes in each slice. In the visual analogy, by stacking multiple slices of cheese together, the likelihood of an unwanted outcome is reduced. Ideally, enough cheese slices are stacked so that the holes do not align, and threats are thwarted.

In the U.S. nuclear architecture, multiple safeguards are stacked to prevent weaknesses in each component from aligning. However, the launch-under-attack posture creates an opportunity for system weaknesses to align by creating incentives to overly trust early warning systems, which is where the nuclear-use decision chain begins. Even as just an option, the president will face a [“premium on haste in a crisis”](#) to launch from a high confidence warning, or otherwise face the strategic and political repercussions of indecision. Thus, the posture's mere availability paradoxically constrains the president's decision-making process, which is informed by vulnerable machine-produced data in a time-compressed, high-stress environment.

### *Decide-Under-Attack in the Electronic Environment*

The current retaliatory posture must consider two factors: first, the inherent and increasing vulnerability of systems that inform decision-making, and second, the fundamental importance of presidential control in U.S. nuclear policy. It is crucial for a retaliatory posture to ensure the availability of weapons and command and control from the use decision to execution.

Retired Adm. James Winnefeld, former commander of North American Aerospace Defense Command, proposed an [approach](#) that better balances deterrence and safety. This posture, called “decide-under-attack,” introduces a delayed response option to reduce the time pressure inflicted by launch-under-attack.

Ultimately, an attack warning will prove to be real or false. But the president will decide whether to launch weapons or not without knowing if it is the former or latter. Among the four possible scenarios, two outcomes must be avoided. The first would be that the president fails to launch when an attack warning is real. The second would be an irretrievable retaliatory strike even though the warning is false. The cyber- and system-based vulnerabilities highlight the uncertainty inherent in the information that feeds this decision-making process. And due to the induced time constraints, a launch-under-attack posture increases the likelihood of these unwanted outcomes. Decide-under-attack improves upon launch-under-attack by allowing the president to opt for a delayed response. This option extends the reach of command and control and reduces the pressure caused by uncertainty and time constraints. Upon receiving a warning, the president can choose to order specific or all components of the nuclear triad to execute a delayed attack. For example, the president may decide to ready the submarine- and land-launched components while keeping the long-range bombers grounded to minimize the potential for escalation if the warning proves false.

In a scenario where the president has a higher degree of confidence that the warning is real and is concerned about the survivability of the land and sea components, they may also order the strategic aircraft to take flight. Even if it is a real warning and the president becomes incapacitated (or communications are lost), [weapons would be available](#) and the command and control concept would be intact, enabling a retaliatory strike.



However, if the warning proves false, the president can cancel the strike. The risk of a premature decision is reduced because the president knows that the order could still be carried out even in the event of their death or disrupted communications. Decide-under-attack effectively addresses the risk of mistaken launch in today's posture by pivoting the retaliation decision from time-constrained to proof-based. Furthermore, the proposed posture serves as a deterrent to adversaries with cyber capabilities. A strategic adversary could launch a real strike and use cyber-based tactics to induce additional uncertainty. This heightened uncertainty may overwhelm the president, making it difficult to initiate a retaliatory response. Consequently, this situation may create incentives for adversaries to launch a first strike. However, if adversaries believe that a delayed retaliatory response is likely, the incentive to launch such a cyber-nuclear attack is reduced. Other actors, namely terrorists with [cyber capabilities](#), may try to provoke a preemptive launch by [fabricating a false signal](#). The decide-under-attack posture addresses this by delaying the response until there is greater evidence, such as additional sensor correlation or confirmation of weapons impact. A potential weakness of this approach would be if an adversary could convincingly deliver a false signal across multiple systems to provoke a launch order and then disrupt communications. However, the time delay, combined with the availability of alternative communications methods (since the warning was false), adds layers of resilience to prevent a mistaken launch. Moreover, this approach accounts for system and human biases that could potentially lead to actions based on a false warning, which have been the [sources](#) of near-accidental or mistaken launches. As such, the decide-under-attack option builds resiliency by expanding the decision space. That space can be used to recall an order, modify an order to achieve a proportional response, or validate the inbound weapon's origin. This posture not only increases the credibility of Washington's retaliatory capability but also accounts for false nuclear alarms caused by anything from equipment malfunction and algorithmic error to deliberate spoofing and human fallibility.

### Conclusion

Returning to the scene at Buckley Space Force Base: Tech Sgt. Nichols stared at the warning on his console. For a moment he wondered if this was his Colonel Petrov moment, a Soviet officer credited with "[saving the world](#)" when he deliberately failed to act on an erroneous report of an incoming American strike. But unlike the dilemma facing the Soviet colonel, Nichols knew that modern nuclear brinkmanship was more complex than ever before, with many different nuclear actors and the constant threat of terrorism. And although he knew that advanced systems were imperfect, who was he to question the machine? Fortunately, Tech Sgt. Nichols had been briefed on a new launch policy. The president had abandoned the old launch-under-attack posture for a decide-under-attack approach. This meant that before any nuclear exchange began, the retaliatory decision would give greater weight to proof than "time to impact" of an inbound threat. He was assured that he could report the notification and then take additional time to verify its origin, validity, and accuracy without fear that it would be too late to alter his original report. This renewed his confidence in the systems, both machine and human, that are responsible for the world's safety.

**Johnathan Falcone** is an active-duty U.S. Navy officer currently serving as a chief engineer in the Littoral Combat Ship program. He was awarded the 2022 Alfred Thayer Mahan Literary Award by the Navy League of the United States and is a graduate of the Princeton School of Public and International Affairs and Yale University.

**Jonathan Rodriguez Cefalu** is the founder and Chairman of Preamble, Inc., a company on a mission to provide ethical guardrails for AI systems. Jonathan holds a computer science degree, with honors, from Stanford University. He created the Snapchat Spectacles augmented reality glasses when his first startup Vergence Labs was acquired by Snap Inc. in 2014.

**Michael Kneeshaw** is a bioinformatics scientist and researcher with a focus on machine learning and simulations. He is currently leading the development of a wargame simulator called SIMC4, which is special-built for simulating catalytic nuclear war scenarios. The project is funded by the Preamble Windfall Foundation, a 501(c)(3).

**Maarten Bos** is a quantitative experimental behavioral researcher, with expertise in decision science, persuasion, and human-technology interaction. He has worked in academia and industry research laboratories, and his work has been published in journals including Science, Psychological Science, and the Review of Economic Studies. His work has been covered by the Wall Street Journal, Harvard Business Review, NPR, and the New York Times. Maarten received his Ph.D. in the Netherlands and postdoc training at Harvard Business School.

## Finland's nuclear catacombs nearly ready to house waste

Source: <https://techxplore.com/news/2023-06-finland-nuclear-catacombs-ready-house.html>

June 06 – Deep within the bedrock in western Finland, flashing lights from heavy machinery cast shadows in eerie tunnels that lead to a tomb that must remain unperturbed for millennia. There is a reason the catacomb on the lush island of Olkiluoto was buried 400 meters (1,312 feet) below ground: It will house the world's first spent nuclear fuel disposal facility.



The project, which aims to solve the age-old quagmire of what to do with [nuclear waste](#), is nearing completion in a region that is already home to Europe's largest nuclear reactor.



The catacomb will solve the problem of long-term storage for spent nuclear fuel from Finland's nuclear reactors.

The Onkalo repository is designed to house a total of 6,500 metric tons of uranium, covering the [spent fuel](#) produced by Finland's five nuclear reactors during their lifetime.

Low-carbon [nuclear power](#) has been hailed as a climate-friendly energy source as the world struggles to combat global warming.

Environmentalists, however, worry about their safety and warn that disposing [radioactive waste](#) carries huge risks.

Around 400,000 tons of used fuel have been discharged from reactors worldwide so far, the World Nuclear Association estimates, most of which is stored in temporary storage facilities near power plants.

But the long-term storage problem has cast a shadow on nuclear projects.

Onkalo's operator Posiva has developed its solution together with the Swedish Nuclear Fuel and Waste Management Company (SKB). The plan is to encapsulate the used uranium in thick copper shells and bury the up to five-meter-long canisters in the bedrock, before sealing the tunnel with a massive, wedge-shaped and steel-reinforced concrete plug.

Construction began at Onkalo—Finnish for "hollow"—in 2004, and final trials will be conducted in early 2024.

"Once we have established that we are ready, we can then move on to the actual operational phase in the mid-2020s," geologist Johanna Hansen, Onkalo's research and development coordinator, told AFP.

### Final resting place

The immense amount of time it takes for radiation to decrease to safe levels poses significant challenges.

It can take around 250,000 years for the high-level [toxic waste](#) to reach the radiation levels of a naturally occurring uranium deposit.

Put in context, 250,000 years ago Europe was in the midst of an Ice Age and was home to Neanderthals and mammoths, while Homo sapiens had only recently emerged in Africa.

In other words, before the [waste](#) no longer poses a threat to the environment or people, the landscape of Olkiluoto island may have changed dramatically.

Posiva insists their canisters are designed to withstand even major changes taking place on Earth.





Nuclear waste is stored in holes covered by lids such as these

But some experts have cast doubt on the Finnish-Swedish disposal method. "It is not a solution, it's a reduction of risk," Jan Haverkamp, nuclear energy expert at Greenpeace, told AFP. Haverkamp claimed that Posiva is giving "no real attention" to the discourse around corrosion of the copper canister in Sweden. For example, researchers at the Swedish Royal Institute of Technology (KTH) have questioned the plan's safety by highlighting greater risks of copper corrosion than previously acknowledged, which could result in leakage. "The case for not producing more waste is still standing as strong as it was standing 20 years ago," he added. Posiva and SKB have refuted KTH's experimental studies, and the Finnish Radiation and Nuclear Safety Authority has concluded that the plans meet the requirements of nuclear and radiation safety.

### 'Leading the way'

Allison Macfarlane, professor at the University of British Columbia, told AFP that while no plan is "100 percent safe," Onkalo is "certainly the solution that's received the most research."

The alternative, she added, would be "leaving the waste where it is above ground without a plan indefinitely."

If the nuclear waste is left above ground and our society falls apart "it will definitely get into the environment and affect humans," she explained. She believes that in a reasonably selected location, the waste "will remain safe for thousands and tens of thousands of years." "I think it is the internationally agreed-upon solution to the problem of nuclear waste," Macfarlane said.

Locations for potential nuclear waste repositories have been selected in Sweden, France and Switzerland, with Canada close to making a decision. "Now that we are leading the way, others are following and doing the same," Posiva's Hansen said.

But many projects have faced resistance and locations have been difficult to agree on, such as France's Cigeo project in the country's northeast.

In April, Finland's next-generation Olkiluoto 3 nuclear reactor, the largest in Europe, went into regular production. A poll published the same month by the trade association Finnish Energy showed record-high support for nuclear energy in Finland, with 68 percent of Finns in favor.

Macfarlane said "Sweden and Finland have shown us that the technical challenges can be surmounted." "Other challenges that need to be addressed are the political ones."



## Hinge Points: An Inside Look at North Korea's Nuclear Program

By Siegfried S. Hecker

Stanford University Press | 2023 | Reviewed by Sharon Squassoni

Source: <https://www.armscontrol.org/act/2023-06/book-reviews/hinge-points-inside-look-north-koreas-nuclear-program>

The 70th anniversary of the alliance between South Korea and the United States this spring provided a much-needed opportunity to underline the two countries' solidarity and commitment to peace and security in Northeast Asia. North Korea's unrelenting missile tests and China's awkward support for Russian sabotage of the so-called international order undoubtedly have made South Korean President Yoon Suk Yeol nervous. His own plan to rein in North Korea, the boldly named Audacious Plan, was rejected summarily last year by Kim Yo Jong, sister of North Korean leader Kim Jong Un.

Anyone hoping for creative new ideas on reducing the threat from North Korea at the White House summit in April between Yoon and U.S. President Joe Biden was disappointed. That, however, was not really the point. The Washington Declaration that the two leaders announced in the Rose Garden on April 26 was all about ally reassurance. For months, South Korean policy elites had been taking the temperature in Washington on nuclear options to improve regional security. Between two disastrous and, one hopes, impossible options—the return of U.S. nuclear weapons to South Korea or the development of an indigenous South Korean nuclear arsenal—lay a less offensive approach to create a NATO-like nuclear-sharing arrangement.

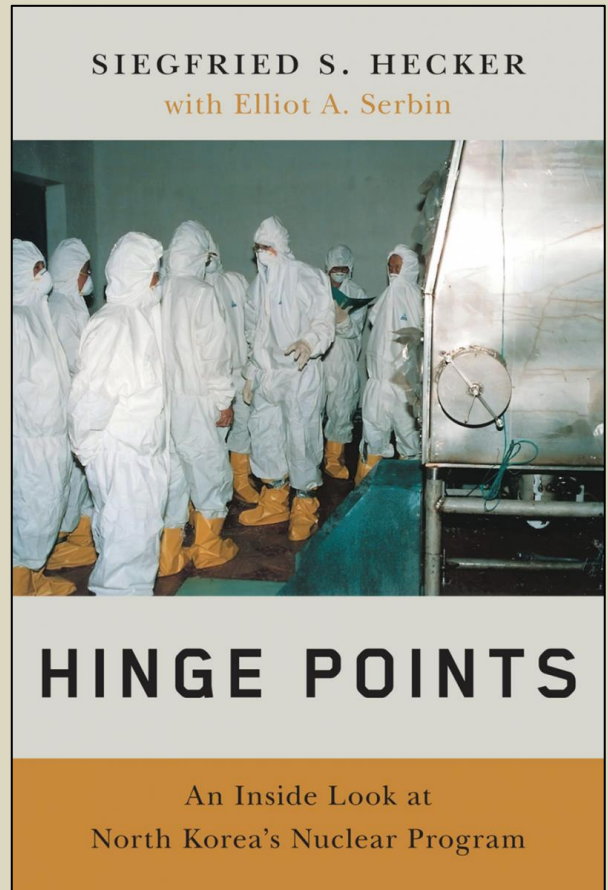
It is this third option that unwisely forms the basis of the declaration. Although there is no mention of NATO in the text, the announced creation of a Nuclear Consultative Group is clearly a nod to NATO's Nuclear Planning Group. The problem is that NATO nuclear-sharing arrangements involve weapons that are stationed on European soil, whereas South Korea has not hosted U.S. nuclear weapons on its soil since 1991, when the United States withdrew nonstrategic nuclear weapons in response to the collapse of the Soviet Union.

The declaration refers specifically to South Korean "conventional support to U.S. nuclear operations in a contingency" and to improved "combined exercises and training activities on the application of nuclear deterrence on the Korean peninsula." Yoon, in a speech at Harvard University later in the week, called the declaration "an inevitable choice" and suggested that the new bilateral arrangement would be more effective than NATO nuclear weapon-sharing agreements.

Despite this, it is more than likely that South Korea walked away with less in the declaration than it had hoped. The declaration specified that the "U.S. commitment to extended deterrence to [South Korea] is backed by the full range of U.S. capabilities, including nuclear." Explicit mention of U.S. nuclear capabilities now seems necessary, given the flurry of protests in Asia when the Biden administration seemed to waver last year about whether it would respond to Russian nuclear use with nuclear weapons. On a positive note, South Korea reiterated its commitment to the nuclear Nonproliferation Treaty (NPT), even if it did not extol the security virtues of forswearing nuclear weapons.

Yoon's postsummit remarks hinted at his frustrations. He confidently told the audience that, of course, South Korea could make nuclear weapons if it so chose. "However, nuclear weapons are not just a matter of technology. There are complex politics and economics and political and economic equations related to nuclear weapons," he said. "There are various values and interests that must be given up when possessing nuclear weapons." Yoon clearly understands that his choice was between his own nuclear weapons and the U.S.-South Korean alliance, but it is damning that the solution is to renuclearize extended deterrence when the United States has committed repeatedly to reducing reliance on nuclear weapons.

The message that only nuclear weapons bring security is not one that will be lost on Kim Jong Un. Although Yoon may be forgoing his own nuclear weapons now in favor of closer planning with Washington on U.S. nuclear deterrence, the extended deterrence equation always fails to eliminate nuclear weapons as a reasonable recourse.



To longtime observers, the Biden-Yoon summit is the latest chapter in decades of ineffectual policies to reduce the nuclear threat from North Korea. Isolating, berating, and belittling North Korea, in combination with ostentatious displays of U.S. nuclear might, have rarely caused the hermit kingdom to cease its provocative actions. There is no reason why South Korea and the United States would believe the latest actions outlined in the Washington Declaration would be effective. They are a temporary measure aimed more at dampening South Korea's domestic debate rather than encouraging North Korean cooperation.

Discerning why and how North Korea has made choices to roll back, freeze, or accelerate its nuclear arsenal development is particularly tough. Thankfully, John Lewis, a Stanford University professor who played a leading role in facilitating dialogue with North Korea, recruited Siegfried Hecker, a former director of Los Alamos National Laboratory, to travel multiple times to North Korea to engage in discussions with technical and political officials and to visit key nuclear sites. Hecker's first trip took place in 2004, and he traveled nearly every year to North Korea until 2010. His experiences and analysis are captured engagingly in this book. An unwilling recruit at first, Hecker had planned to spend his retirement on topics on which he was more expert, collaborating as best he could with Russian scientists in cooperative threat reduction initiatives and building ties with Chinese nuclear weapons scientists to improve nuclear security. He certainly had his share of "nuclear tourism" in far-flung places, including Russia's nuclear test site in Semipalatinsk, Kazakhstan, where he discovered that the Russians had left intact nuclear devices in testing shafts. Inordinately modest, Hecker has produced a masterful analysis not just of the policy debates, decisions, and mistakes that have produced a standoff with North Korea, but also of the technical dilemmas faced in North Korea's nuclear weapons program and in the United States' own expectations and analysis of the North's achievements.

Hecker argues that there were at least six hinge points in the last 20 years when outcomes could have been vastly different if different choices had been made. His underlying assumption is that the North Koreans have always pursued a dual-track approach: developing nuclear weapons for security and negotiating them away for economic and security gains. This analysis assumes that, with the right incentives, cooperation is possible.

The first hinge point occurred in October 2002, when the Bush administration torpedoed the 1994 Agreed Framework, a deal to provide light-water reactors and heavy fuel oil in exchange for North Korea's shutdown of its plutonium-production reactor and reprocessing plant. Rather than working through the agreement to resolve compliance issues, the Bush administration effectively shattered the deal, triggering North Korea's withdrawal from the NPT. The second hinge point occurred in September 2005 when the Bush administration undermined the joint statement agreed during the six-party talks, and a third fateful decision came when the Obama administration walked away from the Leap Day deal in response to a North Korean satellite launch.

In January 2015, the Obama administration failed to take North Korea's proposed nuclear testing moratorium seriously, after which the North Koreans conducted three more nuclear tests. In February 2019, President Donald Trump literally walked away from the Hanoi summit with Kim because the two sides failed to agree on what might be captured in the definition of North Korea's Yongbyon nuclear facility. There are many more missteps in U.S. policymaking, but Hecker identifies these as the major ones and highlights how the North Korean nuclear weapons program benefited in their aftermath. At one point, he suggests that, "History will not be kind to Washington." A metallurgist by training, Hecker was perhaps the perfect interlocutor to discuss Pyongyang's then-plutonium-based nuclear weapons program. As a career government scientist working on nuclear weapons, however, he also was aware of the security risks of his travel to North Korea. Surprised at first that he was given permission to travel, he meticulously documented his discussions and briefed government officials and nongovernmental experts before and after his trips.

This is the fundamental value of such Track 2 meetings, which is to gather and impart information to governments when direct, official meetings between governments are difficult if not impossible to convene. One of the more interesting elements of the book is Hecker's interactions with Chinese nuclear weapons scientists on the margins of those North Korea trips; it clearly must have been invaluable to compare notes with foreign scientists who likely had close connections to North Korea's nuclear weapons program. Overall, the reader cannot fail to be impressed by Hecker's devotion long after retirement to public service.

For consumers of policy memoirs, Hinge Points may seem disarmingly candid. Referring to the chief U.S. demand of North Korea, Hecker suggests at one point that "no one quite knew what 'denuclearization' even meant." Unfortunately, this is probably still true today, despite the reams of official papers devoted to denuclearization road maps.

Unlike other physical scientists who have turned their hands to policy, Hecker understands the complexities without trying to reduce them to simple solutions. His frustration at the lack of progress with North Korea is palpable and a refreshing contrast to the tired cynicism of experts in Washington who believe that Pyongyang is a hopeless case yet refuse to adopt different approaches. One can only hope Hecker has another chance to visit North Korea to brighten the prospects for diplomacy.

Nuclear expert **Sharon Squassoni** is a research professor at The George Washington University and co-chair of the Science and Security Board and ex officio member of the Governing Board at the Bulletin of the Atomic Scientists.



## Ukrainian dam is destroyed; nuclear plant lives in a ‘grace period’

By François Diaz-Maurin

Source: <https://thebulletin.org/2023/06/ukrainian-dam-is-destroyed-nuclear-plant-lives-in-a-grace-period/>



Cooling spray ponds at the Zaporizhzhia nuclear power plant. (Undated photo: Energoatom)

June 06 – In the early hours of Tuesday, June 6, video footage [circulated](#) of a destroyed dam in southern Ukraine with large swaths of water flowing through. The Kakhovka dam—located about 70 kilometers upstream of the city of Kherson—is a critical piece of infrastructure, hosting a hydroelectric power plant and managing a reservoir that supplies water for drinking, irrigation, and cooling of the upstream six-reactor Zaporizhzhia nuclear power plant—Europe’s largest.

It was unclear on Tuesday what or who caused the breach in the dam, which is under Russian control, although it was hard not to blame Russia given the timing of the attack, which happened one day after Ukraine [reportedly launched](#) its long-awaited spring counteroffensive. Both countries denied responsibility and have blamed each other throughout the day. Ukraine said Russia was responsible for the explosion of an engine room of the hydroelectric plant, in part to prevent Ukrainian troops from crossing the Dnipro River downstream, while Russia said Ukrainian forces conducted a sabotage attack. Russia’s defense minister Sergei Shoigu made the [acrobatic suggestion](#) that because Ukraine wanted to transfer some military units and equipment from Kherson to other parts of the front to help with its counteroffensive, making the river wider downstream would make it easier to defend Kherson with fewer forces.

A third scenario [being advanced](#) on Tuesday was that the dam might have suffered from a structural failure after the water level of the Kakhovka reservoir had reached a 30-year high, leading it to be at beyond-design storage capacity since May. No evidence of any of those scenarios had emerged on Tuesday night, although Ukraine’s intelligence directorate [reportedly said](#) it had collected evidence that the hydroelectric plant had been destroyed using explosives, and US intelligence was also [“leaning towards”](#) Russia’s responsibility. In October 2022, Ukraine and Russia traded accusations of preparing attacks against the dam, with Ukraine [warning](#) that Russia had mined the dam and Russia [responding](#) the following day with the opposite suggestion.

The destruction of the dam caused immediate life-threatening flooding and [evacuation](#) of thousands of people living downstream of the dam along the Ukrainian-controlled right bank of the Dnipro River. Early



satellite imagery [was showing](#) large areas being flooded a few hours only after the breach. While the water was quickly rising to dangerous levels downstream, the water level in the upstream Kakhovka Reservoir was dropping, which could have severe nuclear safety implications for the nearby plant.



Map of the Kakhovka reservoir in southern Ukraine. The Zaporizhzhia nuclear plant is located about 130 kilometers upstream of the Kakhovka dam on the Dnipro river. (Map: Thomas Gaulkin / Google Earth)

The Kakhovka Reservoir serves as the Zaporizhzhia plant's [ultimate heat sink](#), an essential safety function of removing the radioactive decay heat generated by the fuel inside the shutdown reactors and spent fuel pools. The plant has a cooling pond that pumps its water from the Kakhovka Reservoir. According to the director general of the International Atomic Energy Agency (IAEA), Rafael Mariano Grossi, the water level of the Kakhovka Reservoir [was dropping](#) on Tuesday at a rate of 5 centimeters per hour, adding that "water in the reservoir was at around 16.4 meters at 8 am. If [it] drops below 12.7 meters, then it can no longer be pumped." This would theoretically leave operators with about three days to pump as much water as possible to fill up the pond. But local Ukrainian military officials [estimated](#) that the water level was dropping at the much higher rate of about 15 centimeters per hour; leaving only 24 hours for the operators to do so.

Commenting on Twitter, Edwin Lyman, a nuclear safety expert at the Union of Concerned Scientists, [described](#) the situation at the Zaporizhzhia nuclear power plant as a "slow-motion disaster." "The impact on the plant is something we are going to see unfold over time," Lyman further explained to the *Bulletin*. "There is a grace period to address this problem, but it's not infinite."

In a statement, the American Nuclear Society was [more reassuring](#): "Even if the water drops in the breached reservoir to pre-dammed levels for the Dnipro River, the nuclear plant has mobile pumping units that can be used to access water from alternative sources. The plant also has special floating water intakes which allow the facility to draw water when the reservoir is at low levels." In his statement, Grossi said that the pond was sufficient to provide water for cooling for several months, although he added the IAEA will confirm this information.

Although the plant has not been generating electricity for several months and therefore requires less cooling than if its reactors were operating, before the dam's breach on Tuesday, one reactor was apparently still in "hot shutdown" to produce steam for onsite purposes. "That means keeping an elevated pressure and temperature, [which] raises some questions about the stability of that reactor if something else happens," Lyman explained. "It's not a normal mode of operating a nuclear reactor. And, because it's not a normal mode, normal emergency procedures may not be available if something goes wrong."

Since February 2022, the Russia-Ukraine war has had [devastating consequences](#) on Ukraine's water resources and infrastructure. But with the destruction of the Kakhovka dam on Tuesday, another safety





layer vanished. In his statement, Grossi said that there was “no immediate risk to the safety of the plant.” But that is “assuming nothing else happens,” Lyman told the *Bulletin*. “The plant is stable for now, but it is becoming increasingly more vulnerable.” Grossi conceded that “it is vital that this cooling pond remains intact.”  
 “Nothing must be done to potentially undermine its integrity.”

**François Diaz-Maurin** is the associate editor for nuclear affairs at the Bulletin of the Atomic Scientists. Previously, Diaz-Maurin was a MacArthur Foundation Nuclear Security Visiting Scholar at the Center for International Security and Cooperation (CISAC), Stanford University, and a European Commission’s Marie Skłodowska-Curie Fellow. He has been a scientific advisor to members of the European Parliament on nuclear issues, and he is a founding member of the Emerging Leaders in Environmental and Energy Policy network (ELEEP) of the Atlantic Council, Washington D.C. and the Ecologic Institute, Berlin. Prior to joining academia, Diaz-Maurin spent four years as a research engineer in the nuclear industry in Paris, France and Boston, MA. There, he worked on the safety design of new reactors and of a treatment plant to vitrify Hanford’s tank waste from WWII and Cold War nuclear weapons production. Diaz-Maurin received multi-disciplinary training in civil engineering (B.Sc./M.Sc., University of Rennes 1, 2004/2007, both with distinction), environmental and sustainability sciences (Ph.D., Universitat Autònoma de Barcelona, 2013, summa cum laude and “Extraordinary Ph.D.” Award), and nuclear materials, geochemistry of radionuclides and nuclear security (postdoctoral training, Stanford University, 2017–2019).

## Hamish de Bretton-Gordon: We are now dangerously close to nuclear war

Source: <https://www.msn.com/en-ca/news/world/hamish-de-bretton-gordon-we-are-now-dangerously-close-to-nuclear-war/ar-AA1ciP5y>

June 08 – The blowing of the dam at Ukraine’s Kakhovka hydroelectric power plant was, quite simply, an act of terror by a terrorist state. With his army failing and his air force stuck in its hangars, it would appear Vladimir Putin is prepared to do almost anything to cling on to the Russian-occupied areas of Ukraine and his throne in the Kremlin. This is another war crime to add to the growing list, a list that includes the unlawful deportation of children — something that has led the International Criminal Court to issue a warrant for his arrest.



The Russian military motivation behind the blast is clear and not unexpected. The vast area to the west of the dam is a “tank highway” to Crimea, and Putin knows his demoralized forces are likely to collapse in the face of Challenger and Leopard tanks charging towards them. The flooding will likely block this axis for many weeks. The ecological and agricultural damage alone will be legion, and with no power coming out of the Kakhovka plant or the Zaporizhzhia nuclear power plant, Ukraine is going to be short on electricity for a while.

This type of terrorism is not new and is to be predicted from the tyrant. I had the honour to be one of the Peshmerga’s chemical weapons advisers in the fight with another terrorist state: ISIL. In 2017, as the Islamic State was falling in Iraq, it blew up the Al Mishraq sulphur mine south of Mosul. From a tactical perspective this had the same effect as blowing the dam at Kakhovka. The 400,000 tonnes of very toxic sulphur dioxide went across the route of the advancing Iraqi army’s direct approach to Mosul and delayed it for several days, allowing ISIL to dig deeper into the city. At one point the toxic cloud was heading to the Kurdistan capital, Erbil, with over one million people in mortal danger. Thankfully the “gods” intervened, and the poison dissipated in the higher atmosphere. When you have no limits or concern for civilian casualties like ISIL and Putin, sadly virtually nothing is off limits.

But the Ukrainians are canny. No doubt the Ukrainian high command will have strategized for such an eventuality and will have numerous lines and methods of attack planned. At the early stages it also looks as though the Russian plan may have backfired, with Russian troops defending this sector scabbling for high ground and the water needed for Crimea disappearing into the Black Sea.

However, the long-term intended consequence may well be the weaponization of the Zaporizhzhia power plant. For months now there have been warnings that Putin would think nothing of triggering some kind of incident there if he felt it would give Russia a tactical advantage. Some have dismissed them. But if the West shrugs now, we take a major step closer to such an incident becoming a horrifying reality.

With no power and no water at this huge nuclear power plant, the chance of the meltdown of reactors and spent nuclear fuel starts to become plausible. Putin has threatened the West with nuclear weapons since the beginning of this war, but even if this is a hollow threat, the power plant could still be used as an improvised nuclear weapon, with plausible deniability. It is uncertain what contamination would ensue or where it would go, but it would be a global humanitarian and environmental disaster.

The “Special Military Operation” is now in its death throes, with a rampant, confident, well-trained and equipped Ukraine army on the march. It is becoming clearer that the Russian military will hit a speed bump conventionally, and with a leader and army commanders with no morals or scruples we must brace



ourselves for further unconventional violence. The attack on schools and hospitals was a portent of the evil of Putin's regime, and the blowing of the dam is another move of truly terrible intent.

Whatever it takes, we must ensure Ukraine prevails as quickly as possible, as there is still the spectre of escalation to chemical, biological and — terrifyingly — nuclear war.

[Hamish de Bretton-Gordon](#) is a former commander of U.K. and NATO CBRN (Chemical, Biological, Radiological, Nuclear) forces.

**EDITOR'S COMMENT:** In my humble opinion, "experts" should urge the international military and political communities towards efforts leading to peace instead of covert threats for a CBRN war if Ukraine does not defeat Russia – logic dictates that this is almost impossible. I dare to add that experts might desire a CBRN war mainly because they are not going to be on the front line – too many others to do that! Some like the ongoing "proxy war". Let us be serious. Let logic prevail. Please!

## Training for Nuclear Incidents and Preparing WMD Responses

By Justin S. Griffin

Source: <https://www.homelandsecuritynewswire.com/dr20230609-training-for-nuclear-incidents-and-preparing-wmd-responses>

June 09 – On a warm spring day in Albuquerque, more than 270 people from across the Labs came through Steve Schiff Auditorium for the first Weapons of Mass Destruction Counterterrorism and Incident Response Showcase.

Attendees learned about myriad opportunities to serve as a responder for the Nuclear Emergency Support Team and participate in training drills and exercises in response to incidents involving WMDs and radiological materials. A call was also made to collaborate on research and development to keep pace with rapidly evolving threats.

Representatives from [Sandia's](#) Joint Technical Operations Team, the Consequence Management Program, Accident Response Group, Radiological Assistance Program, Stabilization Program and many more were on-hand to educate on their important role as part of Sandia's Nuclear Counterterrorism and Incident Response program. The program includes the Nuclear Emergency Support Team mission, as well as numerous technology development efforts and training programs that comprise Sandia's nuclear incident response resources.

As explained by Mike Enghauser, Radiological Triage skillset lead, "Radiological material can end up in almost any location or any place and take on almost any shape and form." The incident response program focus on detecting, evaluating, responding to and mitigating threats is not purely hypothetical work. "It is not practice. It is not an exercise. It is real life stuff," he said.

Visitors to the showcase were also shown virtual reality and haptics technology demonstrations and sat down for lunch from a local food truck.

Art Shanks, senior manager in WMD Counterterrorism and Response, was thrilled with the success of the showcase. "I hope this event brought broader awareness across the Labs' leadership for those already supporting the mission so their management — and management all the way up the Sandia chain — will understand the importance and significance of the Nuclear Counterterrorism and Incident Response mission. I also hope that staff from around the Labs that are interested in this very important mission area learned about opportunities where they could help support as either a responder or as part of the cutting edge research that supports the programs," Art said.

### Highlighting a National Asset

Before the event, a tour included Deputy Labs Director David Gibson and multiple senior leadership team members. Associate Labs Director Justine Johannes kicked off the day by speaking of the vital efforts of the incident response program and its impact on the Nuclear Emergency Support Team mission. "[The Nuclear Counterterrorism and Incident Response program] serves a national mission by drawing on capabilities from across every division at the Laboratories," Justine said, further characterizing the program as a "combination of operational readiness enabled by research and development innovation and technology development."

"I was very excited to see how engaged the leadership was during their visit," Art said. "I also heard from many people, including several from the senior leadership team, about how helpful this event was to give visibility to all the different elements and help connect how all the assets are involved, including the impacts on current real-world events. For a program that often works quietly in the background with a high level of operational security, it was great to give a little visibility to those supporting the mission."

Also in attendance for the tour was Mike Peters, Albuquerque deputy director of NNSA's Office of Nuclear Incident Response. He praised the efforts of the Nuclear Counterterrorism and Incident Response program and its support of the Nuclear Emergency Support Team for his office, which relies upon responders across Sandia to be among the first to respond to a nuclear incident.



The Nuclear Emergency Support Team roster currently includes 210 Sandians from the New Mexico and California sites, spanning every division at the Labs. Jay Tilden, the NNSA deputy under secretary for counterterrorism and counterproliferation, has said that the “diversity of [the team’s] missions” and its scientific capabilities “set the unit apart as a national asset.”

Lindsay Klennert, Sandia’s executive chief of staff, said that while the incident response mission is already well-regarded by the senior leadership team, often in association with response activities and exercises, the showcase “provided an opportunity for Sandia’s leadership to see the breadth of the program — the number of divisions that support the mission, the space utilized, the variety technical expertise required and impressive capabilities developed by Sandia to enable response.

“Everyone walked away with an appreciation for how the program reaches across the Labs to push knowledge and capability development to best respond to the unknown at a moment’s notice,” she said.

### Ready When We Are Needed

Nuclear Emergency Support Team responders come from a variety of backgrounds but share one important commonality: the passion to serve and use their training to make the nation safer.

Kevin Hart, Radiological Assistance Program manager, explained his love for his job as “the ability to take all the technical training that I’ve had and get out in the field, carry backpacks and work with law enforcement, civil support teams and the general public to help protect the nation from nuclear terrorism or radiological terrorism activities.”

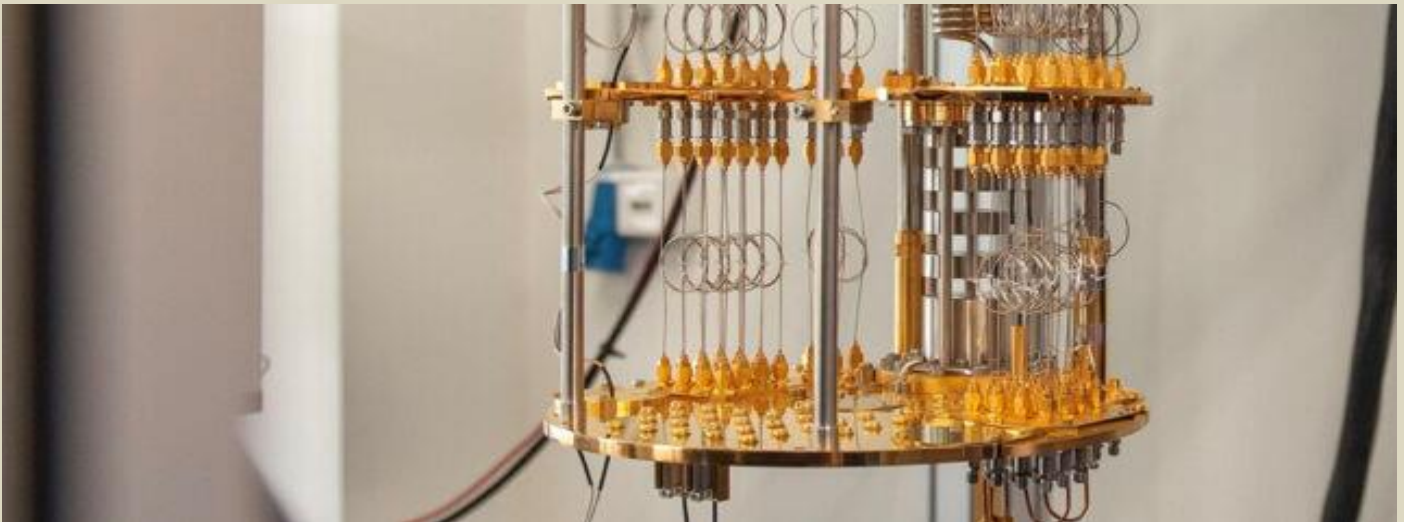
Sean Fournier, a nuclear engineer, echoed that sentiment. “It is exciting work knowing that you are playing a key role in helping to keep the nation safe. “We practice and build technologies and devote our careers to build a program that we hope to never have to use. But it is very important that we are ready when we are needed,” Sean said.

The WMD Counterterrorism and Incident Response Showcase brought those already devoted to the mission together while extending an invitation to those interested in engaging. “I am very excited by how effectively this event helped those already supporting the program to connect how all the different elements of the program work together,” Art said. “I also hope those who are interested in becoming more involved were able to make contacts and identify some ways to become involved, whether by becoming a responder or as part of the research and development efforts that support the program.”

[Justin Griffin](#) is a Program Communications Specialist at Sandia National Lab.

## New Device Detects Radiation at a Trillionth of The Usual Scale

Source: <https://www.sciencealert.com/new-device-detects-radiation-at-a-trillionth-of-the-usual-scale>



Part of the cryogenic measurement setup. ([Bluefors](#))

June 12 – A team of researchers has accurately measured power at a scale that’s a trillion times smaller than what’s possible with standard instruments. It means microwave radiation can be more precisely assessed in quantum physics experiments. Being able to measure power at ultra-low levels is helpful for scientists building quantum systems – systems that are incredibly small in scale and usually [incredibly cold](#) in terms of their temperature. Now we can take those measurements with much greater accuracy. For example, the new



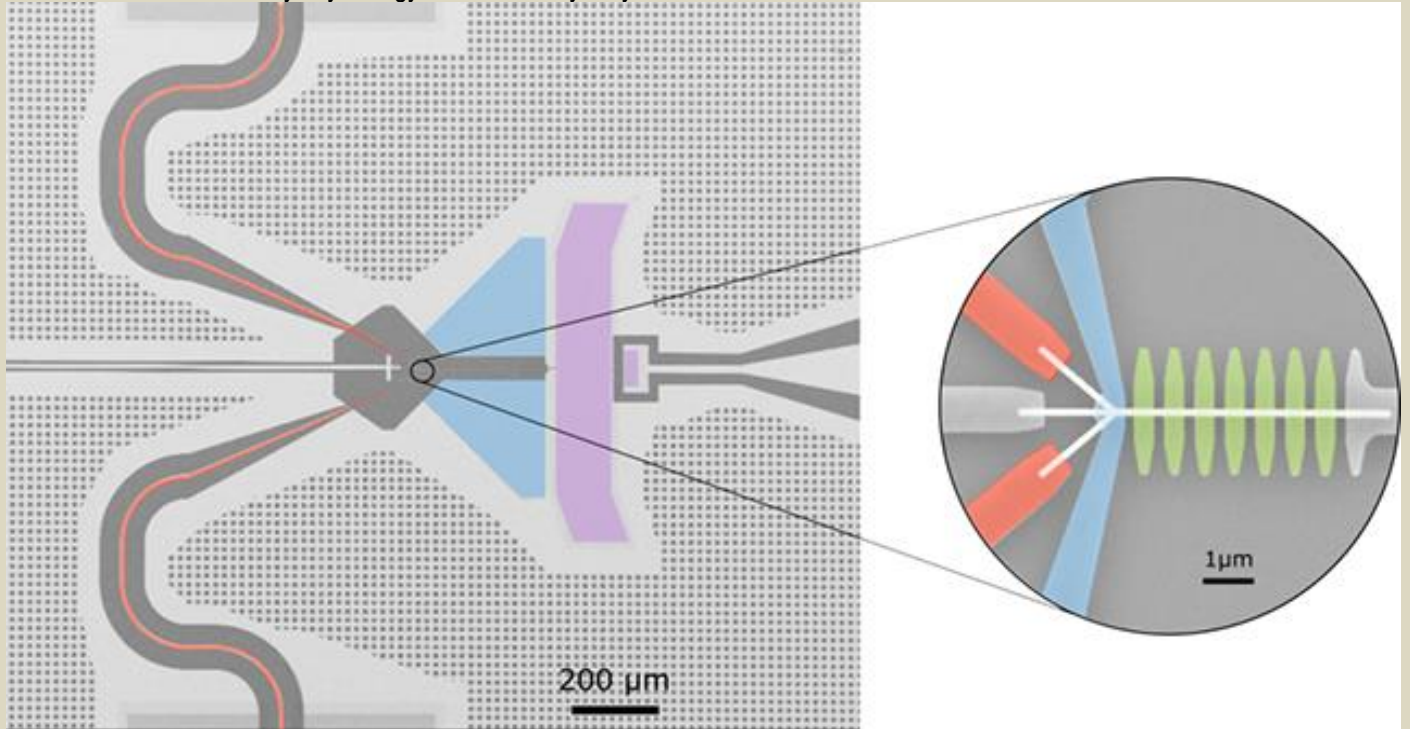
system could be used to better prepare and calibrate qubits – particles at the center of [quantum computers](#) that take the place of classic bits – to ensure they're operating as intended and that the readings they produce are correct.

"Commercial power sensors typically measure power at the scale of one milliwatt," [says](#) Russell Lake, a senior scientist at the Bluefors quantum technology company in Finland.

"This bolometer does that accurately and reliably at 1 femtowatt or below. That's a trillion times less power than used in typical power calibrations."

In quantum experiments, energy is measured using a special thermometer called a [bolometer](#). It tracks temperature through a small strip of material – usually a metal or semiconductor – that changes its electrical resistance as it absorbs energy.

Researchers added a heater with a known current and voltage in the new system. By knowing precisely how much heat was put in, the scientists detected very tiny energy shifts made by very weak microwaves.



A diagram of the power sensor on a silicon chip. ([Jean-Philippe Girard/Aalto University](#))

Part of the reason quantum physics is so challenging is that quantum systems are very fragile, and they can be broken or interfered with by the smallest of disturbances, including the tools we use to try to measure them. One of the ways that the new approach can help is by detecting those disturbances.

"For accurate results, the measurement lines used to control qubits should be at very low temperatures, void of any thermal photons and excess radiation," [says](#) quantum physicist Mikko Möttönen from Aalto University in Finland.

"Now, with this bolometer, we can actually measure that radiation temperature without interference from the qubit circuitry."

The new setup is known as a nanobolometer, and early tests on weak microwaves passing through a radio frequency transmission line showed that the instrument could precisely record changes in power.

This work builds on [previous research](#) into creating a bolometer capable of measuring a qubit's energy state. The approach is scalable and doesn't use much energy while eliminating any potential interference for the qubit.

Bolometers can be used in a wide variety of scenarios, including as part of [deep space telescopes](#), but if they can be practically put to use on qubits, then it means we're another step closer to fully realized [quantum computing](#) systems.

"Measuring microwaves happens in wireless communications, radar technology, and many other fields," [adds](#) Lake. "They have their ways of performing accurate measurements, but there was no way to do the same when measuring very weak microwave signals for quantum technology."

"The bolometer is an advanced diagnostic instrument that has been missing from the quantum technology toolbox until now."

●► The research has been published in the [Review of Scientific Instruments](#).



## China could potentially have as many ICBMs as U.S. or Russia by turn of decade: SIPRI

Source: <https://www.thehindu.com/news/national/china-could-potentially-have-as-many-icbms-as-us-or-russia-by-turn-of-decade-sipri/article66961196.ece>

June 13 – The size of China’s nuclear arsenal increased from 350 warheads in January 2022 to 410 in January 2023, and it is expected to keep growing, according to estimations by Stockholm International Peace Research Institute (SIPRI).

“Depending on how it decides to structure its forces, China could potentially have at least as many intercontinental ballistic missiles (ICBMs) as either the U.S. or Russia by the turn of the decade, the Stockholm-based think tank said in its annual assessment released on Monday said. The key finding in SIPRI Yearbook 2023 is that the number of operational nuclear weapons started to rise as countries’ long-term force modernisation and expansion plans progressed. “China has started a significant expansion of its nuclear arsenal. It is increasingly difficult to square this trend with China’s declared aim of having only the minimum nuclear forces needed to maintain its national security,” Hans M. Kristensen, Associate Senior Fellow with SIPRI’s Weapons of Mass Destruction Programme, said in a statement issued by SIPRI.

### India’s arsenal expanding

India and Pakistan appear to be expanding their nuclear arsenal. Both countries introduced and continued to develop new types of nuclear delivery system in 2022, the assessment noted.

“While Pakistan remains the main focus of India’s nuclear deterrent, India appears to be placing growing emphasis on longer-range weapons, including those capable of reaching targets across China.” According to SIPRI estimates, the **nuclear arsenals of India grew from 160 in 2022 to 164 in 2023 and that of Pakistan from 165 to 170.** India, which has a

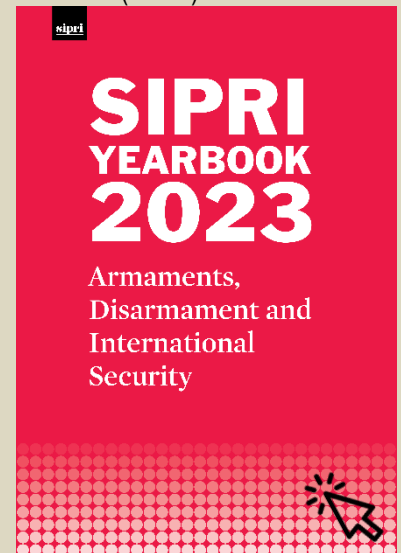
no-first-use policy for use of nuclear weapons and has completed its nuclear triad with the fielding of two ballistic missile nuclear submarines, is in the process of upgrading its ballistic missiles. While a submarine-launched intermediate-range ballistic missile is under development, a new generation ballistic missile, ‘Agni Prime’, with a range between 1,000 km-2,000 km, is close to being inducted, which will replace older Agni missiles in this range. India has also inducted Agni-5, which has a range of over 5,000 km. The nine nuclear-armed states —the United States, Russia, the United Kingdom, France, China, India, Pakistan, the Democratic People’s Republic of Korea (North Korea) and Israel—continue to modernise their nuclear arsenals and several deployed new nuclear-armed or nuclear-capable weapon systems in 2022, SIPRI said. **Of the total global inventory of an estimated 12, 512 warheads in January 2023, about 9,576 were in military stockpiles for potential use — 86 more than in January 2022.** Russia and the U.S. together possess almost 90% of all nuclear weapons. The sizes of their respective nuclear arsenals (useable warheads) seem to have remained relatively stable in 2022, although transparency regarding nuclear forces declined in both countries in the wake of Russia’s invasion of Ukraine in February 2022, the statement added. In the wake of the war in Ukraine, U.S. suspended its bilateral strategic stability dialogue with Russia, and in February 2023, Russia announced it was suspending its participation in the 2010 Treaty on Measures for the Further Reduction and Limitation of Strategic Offensive Arms (New START) — the last remaining nuclear arms control treaty limiting Russian and U.S. strategic nuclear forces. Talks about a follow-on treaty to New START, which expires in 2026, were also suspended. Nevertheless, by SIPRI’s assessment, deployed strategic nuclear forces of both countries remained within the New START limits as of January 2023, the statement added.

## Iran Can Produce Enough Weapon-Grade Uranium for a Nuclear Weapon in 12 Days

By David Albright, Sarah Burkhard, Spencer Faragasso, and Andrea Stricker

Source: <https://www.homelandsecuritynewswire.com/dr20230612-iran-can-produce-enough-weapongrade-uranium-for-a-nuclear-weapon-in-12-days>

June 12 – A new [report](#) from the [Institute for Science and International Security](#) summarizes and assesses information in the International Atomic Energy Agency’s (IAEA’s) quarterly report for 31 May 2023, *Verification and monitoring in the Islamic Republic of Iran in light of United Nations Security Council resolution 2231 (2015)*, including Iran’s compliance with the Joint Comprehensive Plan of Action (JCPOA).



## Findings

- Iran can now break out and produce enough weapon-grade enriched uranium for a nuclear weapon in 12 days, using only three or four of its advanced centrifuge cascades and little more than one-third of its existing stock of 60 percent enriched uranium. This breakout could be difficult for the IAEA to detect promptly, if Iran took steps to delay inspectors' access.
- Using more of its remaining stock of 60 percent enriched uranium in the same three or four cascades and much of its stock of near 20 percent enriched uranium in the vast bulk of its production-scale cascades, Iran could produce enough weapon-grade uranium (WGU) for an additional four nuclear weapons within the first month of a breakout.
- In the second month, using its further remaining stocks of 20 and 60 percent material and part of its stock of less than 5 percent low enriched uranium (LEU), Iran could produce enough WGU for another two weapons. Using its residual stock of less than 5 percent low-enriched uranium (LEU), Iran could produce enough WGU for an eighth weapon by the end of the third month.
- In summary, Iran could produce enough WGU for five nuclear weapons in one month, seven in two months, and a total of eight in three months.
- Iran's stockpile of 60 percent highly enriched uranium (HEU) was 114.1 kg (Uranium mass, or U mass) or 168.8 kg uranium hexafluoride mass (hex mass) as of May 13. With a monthly average production rate of 9 kg (U mass) per month, Iran could amass enough 60 percent HEU for three nuclear weapons by mid-June.
- Iran is producing 60 percent HEU from 5 percent LEU feed in advanced centrifuge cascades at the above-ground Pilot Fuel Enrichment Plant (PFEP) and the below-ground Fordow Fuel Enrichment Plant (FFEP); the latter includes an IR-6 centrifuge cascade that is easily modifiable to change operations. This cascade was at the center of an IAEA-detected undeclared mode of operation in January 2023, interconnected with another IR-6 cascade to produce HEU, and subsequently, the IAEA detected the presence of near-84 percent HEU particles at the cascade's product sampling point.
- The IAEA assessed that Iran's explanation for its production of 83.7 percent enriched uranium — which Iran claimed was the result of "unintended fluctuations" — was "not inconsistent" with "additional information and supporting operational data" Iran provided during meetings with the IAEA and "had no further questions on the matter at that stage."
- The IAEA did not provide any details of its findings, or report whether it checked if any of Iran's information was falsified, but notes that it did investigate the information's consistency. The IAEA has used the more mathematical double negative "not inconsistent," rather than "consistent," in the past to express a finding for which it has a lower confidence level because of insufficient basis to confirm information made available by the state. As such, the IAEA may revisit a "not inconsistent" finding. Nonetheless, this episode has left an indelible impression that Iran can quickly and perhaps secretly produce 90 percent enriched uranium, if it chooses to do so.
- The IAEA has been seeking increased access and intensification of verification activities at the FFEP. In its separate report on Iran's compliance with the Nuclear Non-Proliferation Treaty (NPT), the IAEA reported that it installed enrichment monitoring devices (EMD) at both the FFEP and at the PFEP to "monitor the enrichment level of the HEU being produced by Iran." These monitors are not JCPOA-related but are installed pursuant to Iran's comprehensive safeguards agreement (CSA) with the agency. IAEA Director General Rafael Grossi confirmed in a press conference that the EMD data will notify the IAEA of "another oscillation or otherwise" in the enrichment level in "real-time."
- Iran continues to keep the majority (60 percent) of its stock of 60 percent HEU at the Esfahan site, where it maintains a capability to make enriched uranium metal.
- As of May 13, 2023, Iran had an IAEA-estimated stock of 470.9 kg of 20 percent enriched uranium (U mass and in the form of UF<sub>6</sub>), equivalent to 696.6 kg (hex mass). Iran also had a stock of 38.8 kg (U mass) of 20 percent uranium in other chemical forms.
- The average production rate of 20 percent enriched uranium at the FFEP remained steady at 12.2 kg (U mass) or 18 kg (hex mass) per month.
- At the Natanz Fuel Enrichment Plant (FEP), Iran has a total of 36 cascades of IR-1 centrifuges, 21 cascades of IR-2m centrifuges, four cascades of IR-4 centrifuges, and three cascades of IR-6 centrifuges installed. It did not install any additional centrifuge cascades at the FEP during this reporting period, although it has additional eight IR-4 centrifuge cascades planned and the installation of one IR-4 cascade was ongoing.
- Iran did not install any additional advanced centrifuge cascades at the FFEP, where it is currently operating six IR-1 centrifuge cascades and two IR-6 centrifuge cascades, although it plans to install up to 14 additional IR-6 centrifuge cascades.
- Iran installed one IR-4 centrifuge cascade at the PFEP since the last reporting period. It is unclear why Iran has not installed additional centrifuge cascades per its announced plans, especially following a spike in advanced centrifuge deployment from August 2022 to February 2023. It is



unclear whether this means Iran produced fewer centrifuges than expected, implying possible manufacturing difficulties, or is keeping newly produced machines in storage instead.

- Iran's current, total operating enrichment capability is estimated to be about 19,100 separative work units (SWU) per year, a slight increase over the end of the last reporting period, largely due to the added cascade at the PFEP. As of this reporting period, Iran was not yet using its fully installed enrichment capacity at the FEP.
- Iran's stockpile of near 5 percent LEU was 1340.2 kg (U mass) or 1982.5 kg (hex mass). Average production of near 5 percent LEU at the FEP decreased, but despite large amounts of 5 percent LEU being used as feedstock for 20 and 60 percent uranium production, Iran's overall near 5 percent LEU stock increased slightly.
- Despite the increase during this reporting period in the amount of uranium enriched between two and five percent, Iran has not prioritized stockpiling this material over the past two years. In addition, it has not made planned progress on the Enriched Uranium Powder Plant, a key civil facility to convert less than five percent enriched uranium hexafluoride into a uranium oxide powder for use in nuclear power reactor fuel. These two choices are at odds with Iran's contention that its primary goal is to accumulate 4-5 percent enriched uranium for use in nuclear power reactor fuel. Instead, Iran has used this stock extensively to produce near 20 percent and 60 percent enriched uranium, far beyond any of Iran's civilian needs.
- Iran's overall reported stockpile of enriched uranium increased by 983.7 kg (U mass), largely due to an increase in uranium enriched to less than 2 percent.
- The IAEA reported unsatisfactory progress by Iran on resolving a discrepancy in Iran's natural uranium inventory at the Uranium Conversion Facility (UCF). It reported a shortfall in Iran's declaration, which may indicate that Iran mixed into the UCF inventory undeclared uranium it used in the past at the Lavisian-Shian site during its early-2000s nuclear weapons program.
- The IAEA reports that Iran provided an updated design information questionnaire (DIQ) for the Arak reactor indicating that it is reorienting the reactor consistent with the conceptual design set out in the JCPOA. The reactor will have a power of 20 MW(th). Iran informed the IAEA that it expects to commission the reactor in 2023 and start operations in 2024, although the reported status of the reactor construction raises doubts about whether these milestones will be reached.
- The IAEA underscores that "for two and a quarter years Iran has not provided updated declarations and the Agency has not been able to conduct any complementary access under the Additional Protocol to any sites and locations in Iran."
- The IAEA reports that in line with an IAEA/Iran Joint Statement from March 2023, in May, "the Agency installed surveillance cameras at workshops in Esfahan where centrifuge rotor tubes and bellows are manufactured." However, despite reinstalling surveillance cameras, Iran is still not turning over the video footage to the IAEA, nor has Iran provided past footage from nuclear-related sites from February 2021 to June 2022. Moreover, there remain other Iranian centrifuge manufacturing and assembly facilities where Iran has still not permitted the IAEA to reinstall cameras.
- The IAEA urges Iran to hasten its cooperation. It writes, "The process of implementing the activities set out in the Joint Statement needs to be sustained and uninterrupted in order that all of the commitments contained therein are fulfilled. In addition, the Agency has informed Iran that for these activities to be effective the Agency needs to re-establish a satisfactory understanding of Iran's inventory of centrifuge rotor tubes and bellows, including those in assembled centrifuges."
- The absence of monitoring and surveillance equipment, particularly since June 2022, has caused the IAEA to doubt its ability to ascertain whether Iran has diverted or may divert advanced centrifuges. A risk is that Iran could accumulate a secret stock of advanced centrifuges, deployable in the future at a clandestine enrichment plant or during a breakout at declared sites. Another risk is that Iran will establish additional centrifuge manufacturing sites unknown to the IAEA. Iran has proven its ability of moving manufacturing equipment to new, undeclared sites, further complicating any future verification effort and contributing to uncertainty about where Iran manufactures centrifuges.
- The IAEA concluded previously that "Iran's decision to remove all of the Agency's equipment previously installed in Iran for surveillance and monitoring activities in relation to the JCPOA has [had] detrimental implications for the Agency's ability to provide assurance of the peaceful nature of Iran's nuclear program."
- Concern about Iran's installation of advanced centrifuges at an undeclared site increases as its 60 percent HEU stocks grow. Such a scenario is becoming more worrisome and viable, since a relatively small number of advanced centrifuge cascades would suffice for the rapid enrichment of the 60 percent enriched material to weapon-grade. This hybrid strategy involves the diversion of safeguarded HEU and the secret manufacture and deployment of only three or four cascades of advanced centrifuges. With greater uncertainty about the number of advanced centrifuges Iran is making, there is a greater chance of Iran hiding away the requisite number of advanced centrifuges to realize this scenario.
- Combined with Iran's refusal to resolve outstanding safeguards violations, the IAEA has a significantly reduced ability to monitor Iran's complex and growing nuclear program, which notably has unresolved nuclear weapons dimensions. The IAEA's ability to detect diversion of



nuclear materials, equipment, and other capabilities to undeclared facilities remains greatly diminished.

David Albright is President and Founder of the Institute for Science and International Security.

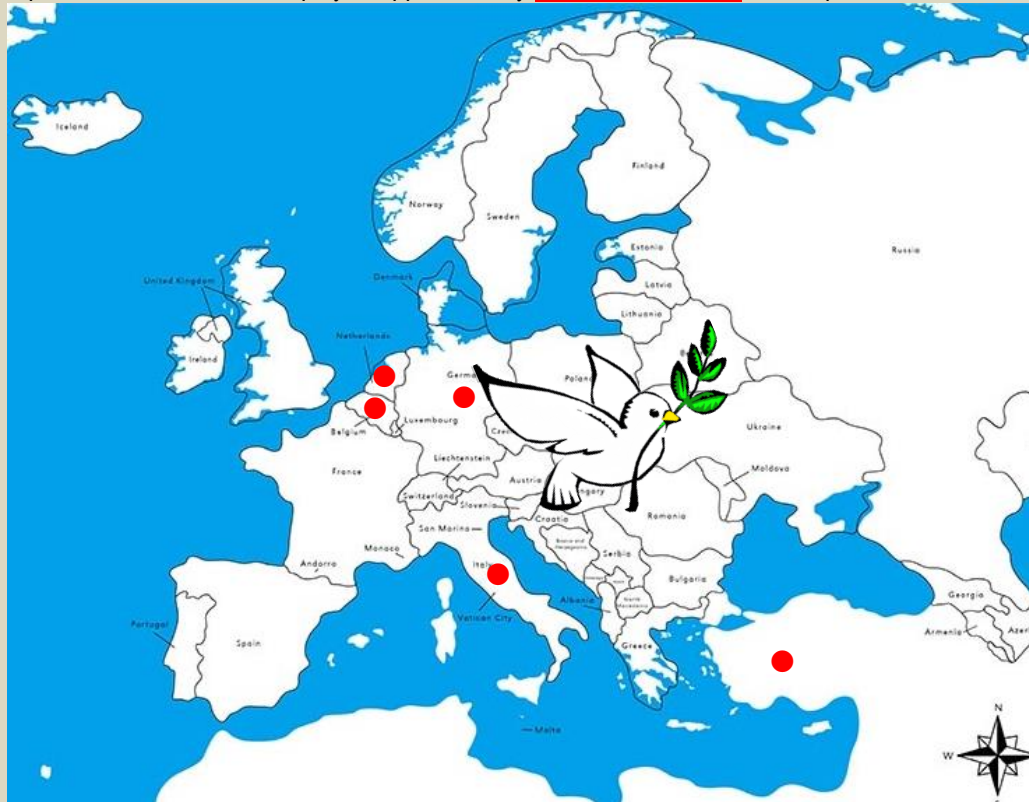
Sarah Burkhard is Research Associate, and Spencer Faragasso is Research Fellow, at the Institute for Science and International Security.

## ICAN: U.S. secretly deploy nuclear weapons in Europe

Source: [https://www.tvr.by/eng/news/v\\_mire/ican\\_ssha\\_tayno\\_razmestili\\_yadernoe\\_oruzhie\\_v\\_evrope/](https://www.tvr.by/eng/news/v_mire/ican_ssha_tayno_razmestili_yadernoe_oruzhie_v_evrope/)



June 12 – The International Campaign to Abolish Nuclear Weapons (ICAN) policy and research coordinator Alicia Sanders-Zacre reported that the US has deployed approximately **150 nuclear shells** at European air bases without making official statements. She



specified that the shells are deployed in **the Netherlands, Turkey, Italy, Germany and Belgium**, as BeITA reports.

Independent experts estimate that approximately 150 nuclear warheads are deployed in U.S. air bases in these countries, said Sanders-Zacre during a briefing with journalists of the Association of Journalists ACANU, accredited by the UN. She also expressed serious concern to the International Campaign to Abolish Nuclear Weapons (ICAN) about the deployment of this type of weapon in any European country, and also recalled the Cuban Missile Crisis and added that such actions could be a potential starting point and increase the risk of nuclear war. According to Sanders-Zacre,

European countries, which demand that the whole world be transparent about nuclear weapons, have shown clear duplicity in this situation. She stressed that it is important to pay attention to the tools available to us, and to ask questions about what the bombs are and where they are located. ICAN is an international public organization founded in 2007. Its primary mission is to support and implement the Nuclear Weapons Ban Treaty. For its efforts to achieve nuclear disarmament, the organization was awarded the Nobel Peace Prize in 2017.

**EDITOR'S COMMENT:** A new word for Armageddon lovers: **PEACE**

## One nuclear-armed Poseidon torpedo could decimate a coastal city. Russia wants 30 of them.

By Silky Kaur

Source: <https://thebulletin.org/2023/06/one-nuclear-armed-poseidon-torpedo-could-decimate-a-coastal-city-russia-wants-30-of-them/>

June 14 – Imagine finding yourself on a serene seashore, only to realize that a catastrophic tsunami is imminent. However, this is no ordinary tsunami; this one is radioactive.

In January 2023, the Russian news agency TASS [reported](#) that Russia had produced the first set of nuclear-powered, very long range, nuclear-armed torpedoes known as “Poseidon.” Strategic experts are

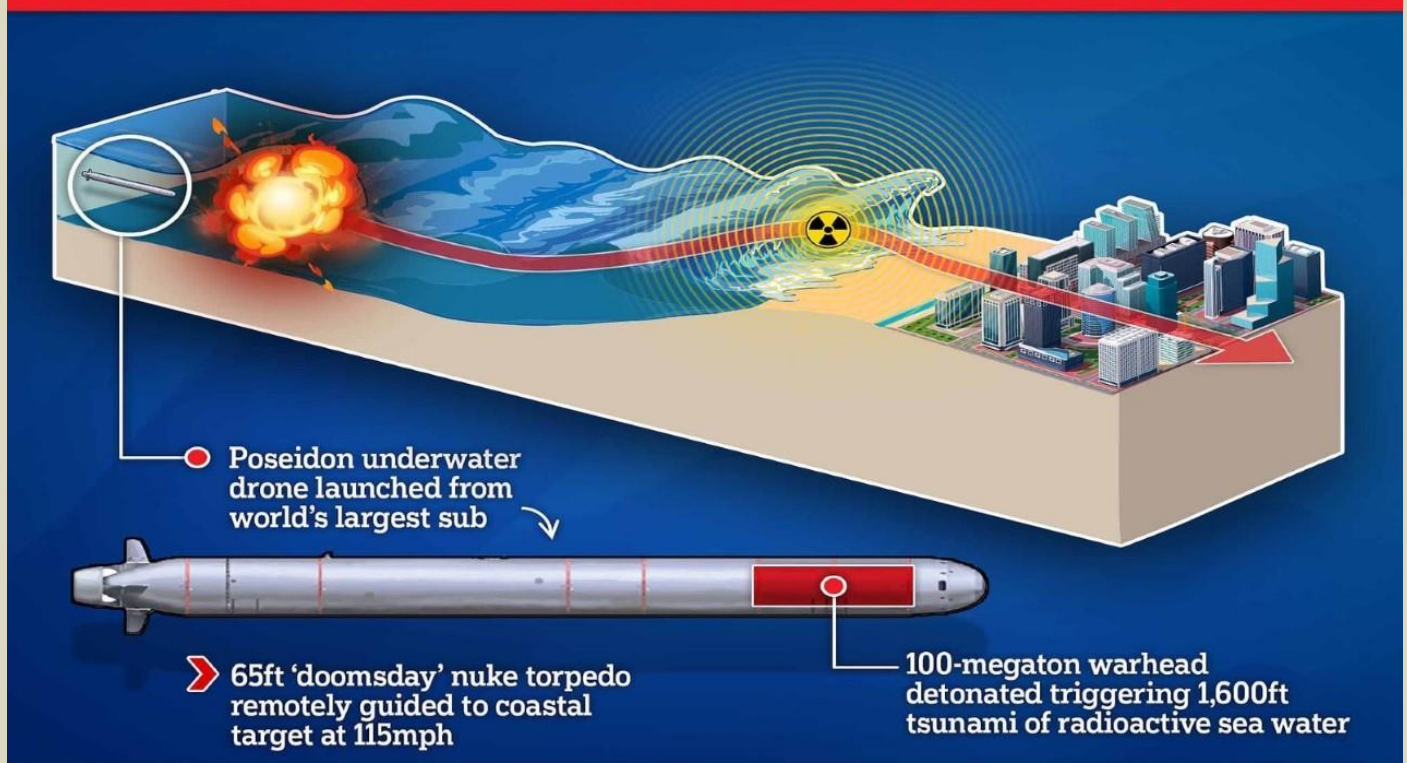




warning that the Poseidon torpedo would have the potential to devastate a coastal city, cause radioactive floods, and result in millions of deaths. Over the past few years, tabloid news outlets [have painted](#) a hauntingly vivid picture of a towering, 1,000-foot-tall radioactive tsunami violently crashing onto British shores, pulverizing everything in its path, and transforming whole cities into barren, lifeless lands.



## POSEIDON: RUSSIA'S 'DOOMSDAY' NUCLEAR TORPEDO



### Is Poseidon even real?

In ancient Greek mythology, Poseidon was revered as the god of the oceans, protector of sailors and those who worked or traveled at sea. In the current context, "Poseidon" takes on a different meaning, with a destructive weapon capable of causing catastrophic events. Russia's Poseidon—also known in the United States as Canyon, Ocean Multipurpose System, and Status-6—was first revealed by the Russian Navy in 2015 and reportedly [tested](#) for the first time in November 2016. The torpedo—a nuclear-powered



underwater drone equipped with nuclear weapons—is designed to be launched from submarines. The first Poseidon is expected to be deployed by 2027. There is buzz that this new weapon is poised to shake the landscape of modern warfare. And this may owe less to its borrowed name than to some unmatched features—at least on paper.

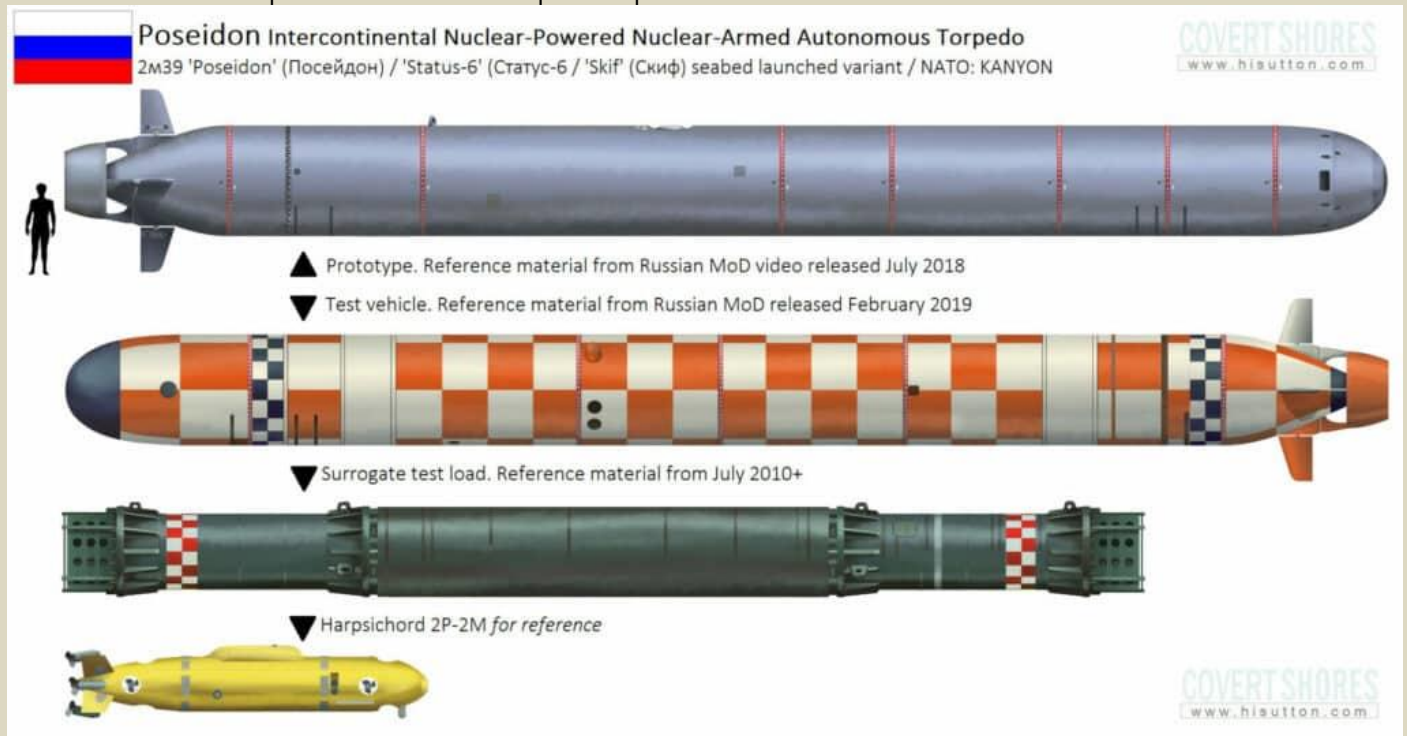
Powered by a compact nuclear reactor, it is believed the Poseidon could travel at [unprecedented speeds](#) of 100 knots (185 kilometers per hour), have a range of approximately 10,000 kilometers, and operate at depths of up to 1,000 meters. Designed to evade detection by acoustic tracking devices and other traps, the Poseidon has a diameter of approximately 1.6 to two meters. Particularly riveting is the torpedo's devastating payload: a nuclear warhead with a [likely yield](#) of at least several megatons (with early reports suggesting it could yield up to 100 megatons). For comparison, Russia's Tsar Bomba—the most powerful nuclear weapon ever tested—had an estimated blast yield of about 50 megatons.

While some reports [claim](#) that Russia's Poseidon may exist only as a propaganda scheme, experts generally agree that the system is “[very real](#)” and has received significant resources from the Russian armed forces, although many details remain unknown. [Intelligence reports](#) have suggested that Poseidon has undergone many trials, evidenced by the fact that some submarines have been modified and some are being specially built to accommodate for the larger and heavier Poseidon. For instance, the [Sarov](#) submarine [is believed](#) to have been modified to test Poseidon prototypes. According to TASS, the Russian Navy intends to purchase at least [30 Poseidon torpedoes](#) and deploy them on four submarines.

### Poseidon's origin and roles

The idea of torpedoes fitted with a nuclear warhead was first conceived in the 1950s, when the Soviet Union began two separate programs to develop submarine-launched nuclear torpedoes—the T-5 and the T-15. These programs were part of a broader strategy to expand the Soviet Union's nuclear capabilities and gain an edge in a nuclear crisis.

In October 1962, during the Cuban Missile Crisis, a Soviet submarine armed with a T-5 nuclear-tipped torpedo [came close](#) to launching it against US forces because its commander, Valentin Savitsky, believed that a US-Soviet Union war had already started. It is only because the submarine's deputy commander, [Vasili Arkhipov](#), convinced other top officers that launching the torpedo would be a fatal mistake that a potential nuclear catastrophe was prevented.



Prototype of Russia's Poseidon nuclear-powered/nuclear-armed torpedo next to the silhouette of a human being and other Russian underwater weaponry. (Image courtesy of HI Sutton/Covert Shores – hisutton.com)

The Poseidon torpedo, however, will have no human [onboard](#) to make critical decisions after it has been launched. It will be controlled through a combination of remote communications and onboard automation. These forms of guidance may lead to problems that include hacking by third parties, loss of control because of technical malfunctions, and environmentally caused accidents that may lead to wrong signaling and thus inadvertent escalation.





The Belgorod—Russia’s largest submarine—reportedly will be capable of carrying up to six Poseidon torpedoes. A second Poseidon-capable submarine—the Khabarovsk—reportedly is under development. (Image: Russian Navy)

The Poseidon can be viewed as an evolution in Russia’s nuclear deterrence strategy. Russia currently possesses the [Dead Hand](#), an automated nuclear weapons command system also known as the Perimeter. Established during the Cold War, the Perimeter was created to ensure that, in a crisis, the Soviet Union could respond to a nuclear first strike, even if Russia’s armed forces were destroyed and all its leadership was eliminated. But with the technological advances of US ballistic missile defense systems, the Russian Perimeter system, which concerns intercontinental ballistic missiles only, has lost its preeminence.

Russian authorities describe the Poseidon torpedo as a multi-purpose system, suggesting that it could serve several roles.

It is widely speculated that the Poseidon may have been developed as a reaction to advances in US ballistic missile defense capabilities. In March 2018, Russian President Putin [stated](#) that Poseidon and other advanced weapons were developed because of the demise of the 1972 US-Soviet Anti-Ballistic Missile Treaty, which Putin said was “the cornerstone of the international security system.” After the United States pulled out of the treaty in 2002 to build its new national missile defense system, Russia started working on improving their military equipment and weapons to maintain a strong nuclear deterrent capability against US assets.

Besides being touted as a weapon that can circumvent US ballistic missile defenses for use [against](#) aircraft carrier groups and coastal targets, Poseidon also reflects Russia’s larger nuclear strategy and doctrine. After the Cold War ended, the United States shifted its focus to de-emphasizing the role of nuclear weapons in its security strategy. But Russia continued to greatly rely on nuclear weapons and even [adopted](#) a nuclear doctrine of “escalate to de-escalate,” which consists of using nuclear coercion and messaging tactics to achieve strategic goals. The development and deployment of new nuclear weapons like the Poseidon are considered to contribute to Russia’s broader strategy.

#### Effect on strategic stability

Advances in artificial intelligence and the use of autonomous platforms for nuclear delivery such as unmanned aerial vehicles, unmanned underwater vehicles, and hypersonic glide vehicles have [raised concerns](#) over potential negative impacts on strategic stability. Weaker nuclear-armed countries would remain unable to keep up with progress in AI technologies, which also pose new risks by reducing the



direct human control over nuclear weapons use. In this context, the deployment by Russia of its nuclear-capable, autonomous Poseidon torpedo could be seen to also threatening strategic stability.

However, many experts and US officials—including former US Defense Secretary James Mattis—believe that the addition of the Poseidon torpedo to Russia’s nuclear arsenal [would not alter](#) the existing [strategic nuclear balance](#) with the United States because Russia already possesses the capability to launch nuclear-armed missiles at US cities. The introduction of Poseidon would not significantly increase the existing threat and, even if used, its impact would be limited to coastal areas.

Instead, many see Poseidon primarily as a [psychological weapon](#) intended for nuclear signaling rather than for its actual use. Some [even speculate](#) that the Poseidon will remain at the prototype stage and serve mainly political objectives, such as renewed talks with the United States on ballistic missile defense systems. The weapon’s primary intent may be in the uncertainty, speculation, and fear it can provoke.

### How to respond?

Automated weapons like Poseidon, especially if controlled by artificial intelligence, undoubtedly add complexity to decision-making in a crisis. It effectively transforms AI into an active participant during times of strategic adversity. This has sparked a heated debate on whether these systems should be fully automated or not, as automation brings significant risks of misjudging the intentions of an opponent. To better understand the challenges posed by unmanned automated nuclear vehicles like Poseidon, further research is imperative, which requires gathering more information about the weapon itself.

While ensuring that research in automation only aims to reduce any risks associated with accidents or inadvertent use, it is equally important for countries to continue engaging in arms control and risk reduction talks whenever possible. Such dialogues may play a crucial role in providing insight into each other’s nuclear forces, helping to alleviate uncertainties about possible new weapon systems, including Poseidon.

[Silky Kaur](#) holds a PhD from Jawaharlal Nehru University. She has previously served as an associate fellow at the Centre for Air Power Studies in New Delhi.

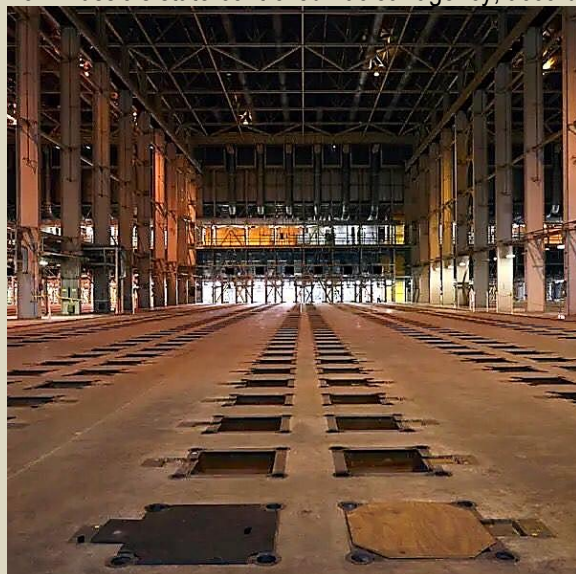
## US Still Pays Russia **\$1 Billion a Year** for Nuclear Energy

Source: <https://themessenger.com/news/u-s-still-pays-russia-1-billion-a-year-for-nuclear-energy>

June 15 – Even as the United States pulled out of purchasing Russian fossil fuels amid start of the war in Ukraine, over two years later, American companies are still buying roughly \$1 billion worth of enriched uranium a year from Russia’s state-controlled nuclear agency, according to [a New York Times report](#).



**ROSATOM**



Among the last of America’s major economic ties to Russia, the payments for cheap nuclear fuel flow to subsidiaries of Rosatom, which is linked closely to Russia’s military, according to the newspaper.

The transactions supply the fuel that produces more than half of the U.S.’s emissions-free energy, with demand only expected to grow in the coming years. Approximately a third of enriched uranium used in the U.S. comes from Russia, which is the world’s cheapest supplier.

But U.S. independence from Russia in the nuclear market appears to be years away.

[Inside the American Centrifuge Plant in Piketon, Ohio, with thousands of “waiting” holes for centrifuges. Photo: New York Times](#)

No American company enriches uranium - a necessary process that creates the needed components to fuel nuclear power plants, the newspaper reported.

“We cannot be held hostage by nations that don’t have our values, but that’s

what has happened,” West Virginia Democratic Sen. Joe Manchin III, who chairs the Senate’s energy committee, told the Times. (Manchin is also [a sponsor of a bill](#) to help boost the nation’s ability to enrich uranium through federal subsidies.)

In Piketon, Ohio, the American Centrifuge Plant has 16 centrifuges that can enrich uranium - and space for thousands more, the *Times* reported.





But right now, the facility lies dormant and the company operating it said it could take more than 10 years for it to produce enriched uranium at a scale similar to Rosatom.

"It's inexplicable that over a year after Russia invaded Ukraine, the Biden administration does not appear to have a plan to end this dependence," James Krellenstein, the director of the clean energy consulting firm GHS Climate, that recently released [a white paper](#) about the issue, told the *Times*.

"We could eliminate almost all of America's dependence on Russian enrichment by finishing the centrifuge plant in Ohio." According to the newspaper, estimates from the Department of Energy say the U.S. will have to more than double its nuclear power capacity to meet the nation's goals for emission reductions.

This week, the department released a draft of a request for proposals intended to boost U.S. uranium enrichment.

Kathryn Huff, assistant secretary for nuclear energy, said the draft is an important step toward ending "reliance on Russia."

## Queering nuclear weapons: How LGBTQ+ inclusion strengthens security and reshapes disarmament

By Louis Reitmann, and Sneha Nair

Source: <https://thebulletin.org/2023/06/queering-nuclear-weapons-how-lgbtq-inclusion-strengthens-security-and-reshapes-disarmament/>



June 15 – "They should not allow mentally ill people near weapons of mass destruction." That was one of dozens of derogatory [tweets](#) that the Vienna Center for Disarmament and Non-Proliferation received in response to a December 2022 [panel discussion](#) on LGBTQ+ identity in the nuclear weapons space. Most of these tweets were purely hateful, written by trolls. But some respondents explained their opposition, saying that talk about queerness was inserting a non-issue and "[derailing](#)" discussions of nuclear weapons. All showed a keen determination to misunderstand the purpose of the event.

While the event received an outpouring of vocal and wide-reaching support from some of the best-known figures in the nuclear field, the disparaging tweets illustrated the common belief that queer identity has no relevance for nuclear policy, and that examining the relationship between queerness and nuclear policy is intended to push a social agenda rather than to address substantive issues.

During this Pride Month, we would like *Bulletin* readers to understand that the visible representation and meaningful participation of queer people matters for nuclear policy outcomes. **Discrimination against queer people can undermine nuclear security and increase nuclear risk.** And queer theory can help change how nuclear practitioners, experts, and the public think about nuclear weapons.



### It's about people

Equity and inclusion for queer people is not just a box-ticking exercise in ethics and social justice; it is also essential for creating effective nuclear policy. **Studies in psychology and behavioral science show that diverse teams [examine](#) assumptions and evidence more carefully, [make](#) fewer errors, [discuss](#) issues more constructively, and better [exchange](#) new ideas and knowledge.**

When the stakes of making best-informed decisions are as high as they are with nuclear weapons, governments cannot afford to lose out on the human capital and innovation potential of queer people.

**REALLY?**



Informed by their life experiences, queer people have specific skills to offer that are valuable in a policy and diplomacy context. LGBTQ+ people often must navigate being different from those around them; develop the ability to listen and empathize; and mobilize the skill and perseverance to make themselves heard.

Diversity and inclusion are especially important for the policy community dealing with arsenal development and nuclear posture. Women familiar with this “nuclear priesthood” describe it as “male-dominated and unwelcoming.” Homogenous groups like this are prone to groupthink and hostile to critical examination of baseline assumptions about how adversaries construct and identify nuclear threats and risks. For nuclear weapons policy, this has meant the perpetuation of theories like deterrence and crisis stability, which have contributed to increasing nuclear arsenals and a growing risk of nuclear use.

Such workplace cultures also create enormous psychological stress for minority staff, including queer people, who spend lots of time and energy adapting to role expectations, rather than focusing on bringing their full, authentic potential to the policy-making process. This truth is reflected in the personal experiences of queer officials working on nuclear weapons issues. Richard Johnson, US Deputy Assistant Secretary of Defense for Nuclear and Countering Weapons of Mass Destruction Policy, [recounts](#) feeling inhibited to speak up and contribute in a workplace culture dominated by traditional masculinity that treated homosexuality as a risk factor.

### Exclusion creates nuclear security risks

Exclusion and unfair treatment of queer individuals and other minorities by a homogenous, cis-heteronormative community of practitioners also creates vulnerabilities in nuclear decision making. Cis-heteronormativity is the automatic assumption that someone is heterosexual and identifies with the sex assigned to them at birth. It creates the idea that being heterosexual and cisgender is normal and natural, whereas being queer or trans is a deviation.

Being LGBTQ+ has historically been considered a security risk. Akin to the “Red Scare” anti-communism movement, the “[Lavender Scare](#)” was a campaign persecuting and dismissing gay and lesbian federal employees. The linking of homophobia and national security concerns seems to stem from sensationalized case studies of defections of US intelligence specialists to the Soviet Union during the Cold War. This [legacy](#) of queerness being considered a security risk is still pervasive in the nuclear field.

Even where laws protecting individuals from discrimination on the basis of sexual identity exist, security practitioners can decide whether a queer person is suitable for a job, on the basis of whether or not an individual is fully “out” and the risks hiring managers believe this poses for the national or nuclear security enterprise. An often-cited reason for excluding queer people is that they could be [blackmailed](#) into becoming insider threats, if the risk of having their sexuality outed is leveraged against them. However, such concerns are based on outdated cultural attitudes toward the LGBTQ+ community, rather than evidence that queer people pose a security risk. Despite setbacks, public [acceptance](#) of the queer community is rising [globally](#), and the supposed links between espionage and homosexuality have been [unfounded](#). However, nuclear facilities still have a reputation for being unwelcoming toward queer people and have [failed to investigate](#) allegations of homophobia and harassment. In part, this is due to the lack of diversity in the nuclear field. Homogenous organizations run a higher [risk](#) of isolating queer employees, leaving them vulnerable to pressure. Employees in the majority can feel threatened by those they perceive as “different” and exclude them due to discomfort, rather than any legitimate risk factors. Nuclear security practice needs to refrain from treating an individual’s behavior or identity as a risk and focus instead on identifying misbehaviors that indicate malicious intent.

By failing to create a welcoming workplace at nuclear facilities—whether military or civilian—practitioners risk reducing the effectiveness of an organization’s nuclear security culture. A [report](#) examining sexual harassment in the US National Nuclear Security Administration found that such harassment “can produce harmful psychological, physical, occupational, and economic effects on harassed employees. It can also affect the environment in which they work and lead to decreased organizational performance and productivity and increased employee turnover. In national security settings, sexual harassment can undermine an organization’s core values, cohesion, and readiness, as well as public goodwill.” These risks of sexual harassment are also relevant in the context of homophobia and LGBTQ+ exclusion from the nuclear field, and they demonstrate the security vulnerabilities that nuclear facilities face when failing to address discrimination against queer people.

Including a wider range of perspectives in nuclear decision making creates a more comprehensive definition of [who or what constitutes a “threat” to nuclear security](#). An example of this is the [threat](#) posed by some white supremacist groups with plans to acquire nuclear weapons or material, which can go undetected when a white-majority workforce does not perceive these groups and their ideological motivation as a relevant threat to their nuclear security mission. Individuals targeted by these kinds of groups—including women, people of color, and the LGBTQ+ community—are more likely to identify these types of behaviors and attitudes as security risks and can play a crucial role in identifying a potential insider threat.

### Queer theory: changing the narrative

Queer identity is also relevant for the nuclear field because it informs theories that aim to change how officials, experts, and the public think about nuclear weapons. Queer theory is a field of study, closely related to feminist theory, that examines sex- and gender-based norms. It shines a light on the harm done



by nuclear weapons through uranium mining, nuclear tests, and the tax money spent on nuclear weapons ([\\$60 billion](#) annually in the United States) instead of on education, infrastructure, and welfare. The queer lens prioritizes the rights and well-being of people over the abstract idea of national security, and it challenges the mainstream understanding of nuclear weapons—questioning whether they truly deter nuclear war, stabilize geopolitics, and reduce the likelihood of conventional war. Queer theory asks: Who created these ideas? How are they being upheld? Whose interests do they serve? And whose experiences are being excluded?

Queer theory also identifies how the nuclear weapons discourse is gendered: Nuclear deterrence is associated with “rationality” and “security,” while disarmament and justice for nuclear weapon victims are coded as “emotion” and a lack of understanding of the “real” mechanics of security. The Greenham Common Women’s Peace Camp, a 19-year protest against the storage of US nuclear missiles in the United Kingdom, called attention to the gendered nature of nuclear weapons. The camp’s inhabitants—many lesbian—recognized that the same male-dominated power structures underpinned the oppression of women and nuclear armament. Their protests, often involving feminine-coded symbols like pictures of children, defined nuclear weapons by the existential threat they pose, instead of the protection they supposedly offer. From the queer perspective, the allegation of “derailing” substantive discussions through a non-traditional perspective on nuclear weapons is itself an attempt to exclude marginalized voices and reinforce the idea that nuclear weapons are a domain only for “serious” and “rational” (i.e., male) actors.

Queer theory is also about [rejecting](#) binary choices and zero-sum thinking, such as the tenet that nuclear deterrence creates security and disarmament creates vulnerability. It identifies the assumptions and interests these ideas are built on—and imagines alternatives that serve a broader range of interests, including those of the invisible and resource-stripped.

Indeed, queer theory helps us not only see the bad of a world with nuclear weapons, but also imagine the good of a world without them. It envisions using the resources freed up by nuclear disarmament to build structures that tangibly increase people’s safety and well-being through healthcare, social housing, etc. In this scenario, the more than \$100 billion that nuclear-armed states spend on nuclear weapons every year could be used to address the climate crisis, which could [kill](#) up to 83 million people by 2100.

Queer theory is rooted in the lived experiences of queer people. Because of the rejection they face in cis-heteronormative society, many find validation and purpose outside traditional models for community, career, and family. And yet, they continue to feel society’s pressure to adapt, hide, and edit themselves. This strengthens queer people’s facility to question dominant ideas of what is right, important, or common-sense, and to ask who decides that.

It also creates a sensitivity to the struggles of other marginalized groups. Tully Starr, an Australian activist, [explains](#): “Being queer, coming out, and experiencing discrimination has been a catalyst for questioning the status quo... [Queer people] gravitate towards like-minded individuals, and each person/community presents an opportunity to expand our ideas of what is possible.”

Finally, queer theory informs the struggle for nuclear justice and disarmament. For example, queer artist and writer [Jessie Boylan](#) highlights the harm done by nuclear weapons by documenting the social and environmental consequences of nuclear testing in Australia as part of the [Atomic Photographers Guild](#). Queer theory helps to shift the perception of nuclear weapons as instruments for security by telling the hidden stories of displacement, illness, and trauma caused by their production and testing.

### The time to do better is now

As the nuclear field continues to reflect on its legacy of exclusion and homogeneity during this Pride Month, we as stakeholders, decision makers, and advocates for change should also realize our privilege of being able to openly discuss the challenges facing the LGBTQ+ community without fear of criminalization, retribution, or death. Participating as oneself in the nuclear field is a right that should be extended to all.

However, including the LGBTQ+ community in the nuclear field is far more than a social issue campaign. It is up to allies, people in power, and the institutions they serve to vocalize their support for LGBTQ+ inclusion, not just because it is the right thing to do, but also because queer people add value to nuclear weapons policy and discourse. Decision makers should look to LGBTQ+ inclusion for better nuclear policy outcomes, and build environments in which queer people can bring their specific skills and lived experiences to bear without fear. Arguments to the contrary are as stagnant and outdated as those who voice them.

**Louis Reitmann** is a research associate at the Vienna Center for Disarmament and Non-Proliferation (VCDNP), focusing on export controls; nuclear disarmament; and diversity, equity, and inclusion in the nuclear field. He is a board member of the Emerging Voices Network, organized by the British American Security Information Council (BASIC). Previously, he served as an export control officer at Imperial College London and worked with the European Union’s Special Envoy for Disarmament and Non-Proliferation in Brussels. Louis holds an MSc in International Relations from the London School of Economics.

**Sneha Nair** is a research analyst with the Nuclear Security Program and Partnerships in Proliferation Prevention Program at the Stimson Center and Coordinator for the International Nuclear Security Forum, where she focuses on nuclear security; insider threats; chemical, biological, radiological, and nuclear disinformation; and domestic violent extremist threats to national security and critical infrastructure. Before joining the Stimson Center, she worked at the Comprehensive Nuclear-Test-Ban Treaty Organization and



the Organisation for the Prohibition of Chemical Weapons. She has a master's degree in geography and international relations from the University of St Andrews in the United Kingdom.

**EDITOR'S COMMENT:** Unbelievable! I just read that the way you are having sex affects your risk perception and makes you a better analyst! Damn! For being straight and too old to change!

## Did Pakistan Just Overhaul Its Nuclear Doctrine?

By Sitara Noor

Source: <https://foreignpolicy.com/2023/06/19/pakistan-india-nuclear-weapons-zero-range-cold-start-doctrine/>

June 19 – As Pakistan celebrated the 25th anniversary of its first nuclear tests last month, it also appeared to share more details than usual about its current nuclear posture.

Speaking at a seminar [hosted](#) by the Institute of Strategic Studies Islamabad on May 24, retired Lt. Gen. Khalid Kidwai [provided](#) new details of Pakistan's otherwise ambiguous nuclear doctrine.

During his address, he reiterated some of the long-standing positions of Pakistan's nuclear policy, but also expanded the scope of this doctrine by explaining what he called vertical and horizontal dimensions. Most striking was his statement that "vertically the spectrum encapsulates adequate range coverage from 0 meters to 2,750 kilometers [about 1,700 miles] as well as nuclear weapons destructive yields at three tiers—strategic, operational, and tactical."

Reducing the minimum range to 0 meters is unprecedented and, if implemented, points to a major shift in Islamabad's nuclear policy thinking. Talk of zero-range weapons suggests that Pakistan is either going to develop artillery shells as the United States, Soviet Union, and United Kingdom did during the Cold War, or it could be a hint that Pakistan could possibly lay nuclear land mines across the India-Pakistan border to deter Indian advances.

[Sitara Noor](#) is a research fellow in the Project on Managing the Atom at Harvard University's Belfer Center for Science and International Affairs.

## 5 detained in plot to export radioactive caesium-137 abroad

Source: <https://en.mehrnews.com/news/202351/5-detained-in-plot-to-export-raradioactiv-caesium-137-abroad>

June 23 – Five people were detained for trying to export a kilogram of radioactive cesium-137 to use it to the detriment of Russian interests during a special operation, the Russian Federal Security Service (FSB) said.

The Russian Federal Security Service said that those suspected of trying to smuggle cesium-137 out of the country to discredit Russia had been arrested and are confessing. A case was opened on the illegal handling and smuggling of radioactive substances, Sputnik reported.

"The Russian Federation's Federal Security Service, with the support of the Russian Interior Ministry, has detained five members of an organized crime group acting under the coordination of a citizen of Ukraine... The criminals, acting in the interests of a foreign customer, took steps to acquire **1kg of the cesium-137 isotope for \$3.5 million to smuggle the radioactive substance from the Russian Federation** to somewhere abroad for use to the detriment of Russian interests during the special military operation," the statement said.



The radioisotope cesium-137 is used in gamma-fetoscscopy, in inspection and control complexes, in measuring equipment, for radiation sterilization of food, medicines, and drugs, and in radiotherapy for the treatment of malignant tumors.

At the same time, radioactive cesium is one of the most important so-called dose-forming radionuclides - products of uranium and plutonium fission. It poses a great danger if ingested by humans.

According to experts, cesium-137 can be used to make a "dirty bomb".





ICI  
International  
**CBRNE**  
INSTITUTE



# EXPLOSIVE NEWS

## Cyprus's Thin Green Line

By Mr. David Oliver (Former Deputy Editor of CBNW)

Source: <https://nct-cbnw.com/cypruss-thin-green-line/>



June 03 - Almost 50 years since the island of Cyprus has been divided – following Turkish invasion – into Turkish Northern Cyprus and the Greek Cypriot Republic of Cyprus. A demilitarized zone called the Green Line divides the two, a buffer zone, which, according to UN estimates, remains contaminated with over 7,000 mines.

In 1974, a Greek military coup, which aimed to unite the island with mainland Greece, led to a Turkish invasion and the division of the island between Turkish Northern Cyprus and the Greek Cypriot Republic of Cyprus. Cyprus remains divided to this day by what has become known as the Green Line.

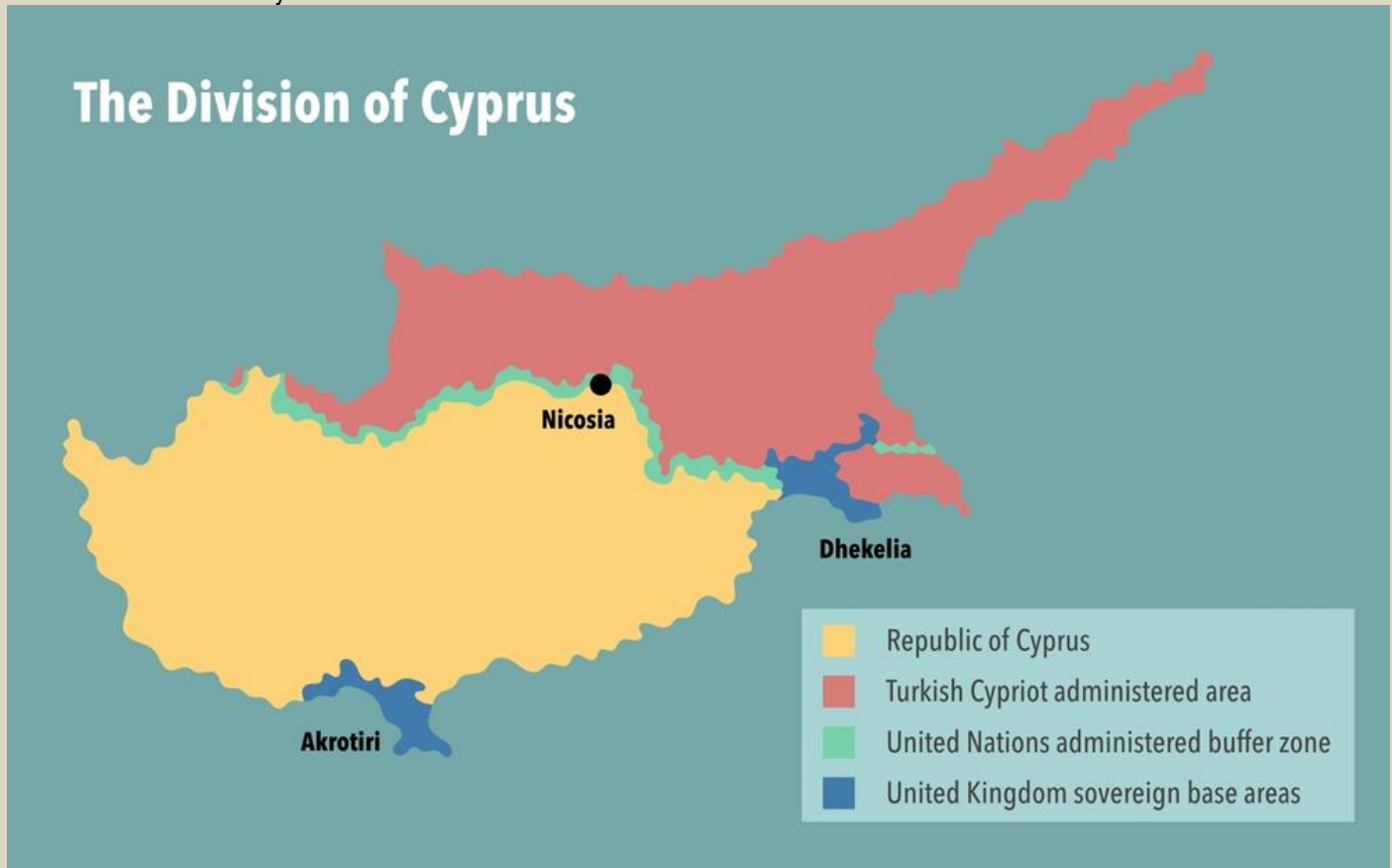
The Green Line today in Cyprus is not simply a line drawn through Nicosia, but a buffer zone between two states, the Turkish Cypriot north and Greek Cypriot south. It is described as a demilitarized zone where its northern and southern limits are the lines where the belligerents stood following the ceasefire of 16 August 1974, as recorded by the United Nations Peacekeeping Force in Cyprus (UNFICYP). This zone including the Kokkina enclave is over 180 km long, is 7.4 km at its widest and 3.3 m at its narrowest point. The buffer zone is home to some 10,000 people and there are several villages and farms located within. UNFICYP is headquartered from the Blue Beret Camp next to the abandoned airport of Nicosia. The mission has a strength of 1,000, including British soldiers who serve on Operation Tosca, a six-month UN tour of duty during which they are responsible for maintaining the integrity of the Buffer Zone. The mission's mandate was last renewed on 30 January 2023 and extended until 31 January 2024.

UNMAS assistance is focused on the removal of mines and ERW that create barriers between the two communities and threaten peace. Its operations contribute to confidence building measures and enable progress towards the implementation of a final settlement agreement.

UNFICYP data indicates that there are 47 suspected hazardous areas in Cyprus that may be contaminated with mines and/or ERW, amounting to approximately 1.7 million square meters of land. Each area requires



survey to determine the need for clearance. To prepare for future surveys and clearance, UNMAS conducts planning and coordination; data analysis and management of information on minefields and suspected hazardous areas; and conducts assessments. The activities help ensure that survey and clearance can get underway to release safe land to communities as soon as access is authorised by the sides.



To achieve this goal, UNMAS conducts advocacy and outreach activities to raise awareness and promote a mine-free Cyprus, provides training in mine/ERW awareness for UNFICYP personnel to help increase force protection and the implementation of its activities; supports planning and coordination of UN activities; and implements the establishment or maintenance of minefield marking, as needed. UNMAS has also provided mine/ERW awareness training to 15 UN Police Officers and “training of trainers” to 26 UNFICYP officers to help raise their awareness of the threat of mines/ERW and how to operate safely in Cyprus.

Numerous experienced international mine clearance specialists have been deployed to Cyprus. In 2015 a twenty-strong team of Cambodian peacekeepers from the United Nations Interim Force in Lebanon (UNIFIL), deployed with UNFICYP as part of an inter-mission co-operation agreement.

The UN Development Programme (UNDP) established the UN Mine Action Centre in Cyprus (UNMACC) with funding from the European Union in 2003. Preparations began for mine clearance activities in the buffer zone, supported by UNFICYP. G4S Ordnance Management, a leading organization working for environmental remediation of landmines and unexploded ordnance, had worked under the supervision of UNMACC since 2004 and completed its activities at the buffer zone excluding the four minefields that had not been opened to demining process. UN officials said three of the four minefields belonged to Greek Cypriots and one was under the control of the Turkish Cypriot side. Teams of G4S Ordnance Management had cleared the buffer zone of nearly 27,000 mines in six years. 18,000 were anti-personnel mines and 9,000 were anti-tank mines. In July 2016, UNMAS became an integral component of UNFICYP, providing dedicated expertise in mine action planning and coordination, operational demining capacity, and data management of mine action information. The objectives of UNMAS in Cyprus operations are: to reduce the threat of landmines and explosive remnants of war (ERW); to contribute to confidence-building measures; to facilitate a return to normal living conditions; and to ensure safer freedom of movement for UNFICYP, UN partners, and communities. UNMAS also assists the Committee on Missing Persons to facilitate safe access to areas it conducts activities and technical guidance to UNFICYP for small arms ammunition storage.

●► [Read the full article at the source's URL.](#)





## US military considered making a bomb to turn enemy soldiers gay

Source: <https://www.thepinknews.com/2016/05/05/us-military-considered-making-a-bomb-to-turn-enemy-soldiers-gay/>

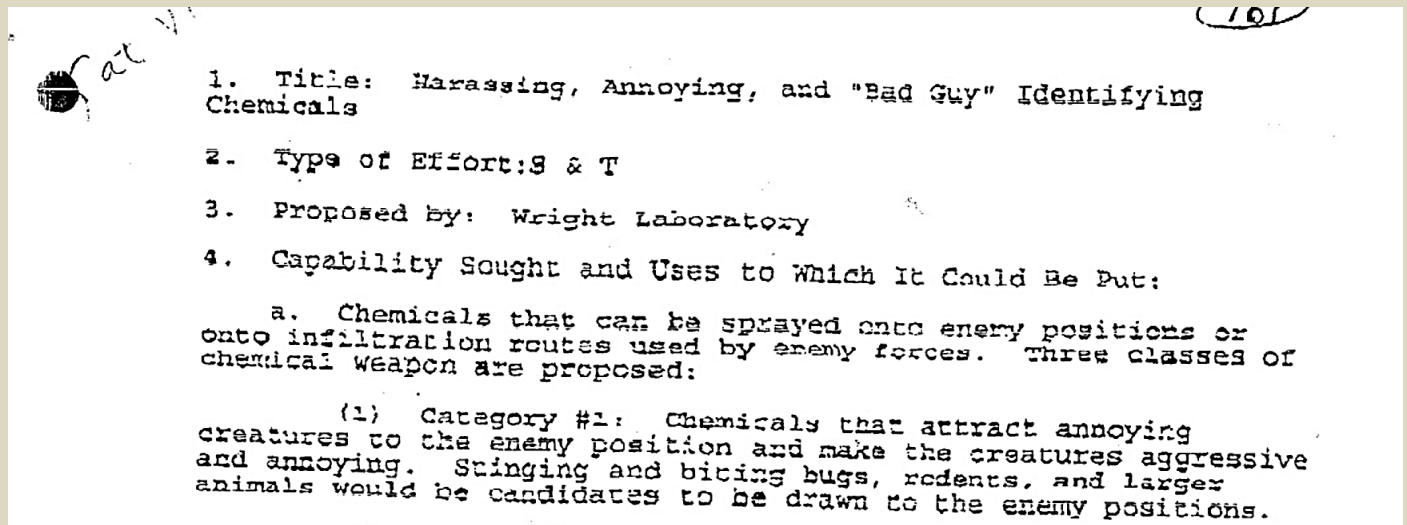
**May 2016** – The US Military once seriously looked into a suggested weapon that would make people gay.

The news comes from unearthed 1994 documents – written the same year that Bill Clinton signed ‘Don’t Ask Don’t Tell’, which banned soldiers from being openly gay.

The proposals came from the US Air Force Wright Laboratory in Dayton, Ohio, which was seeking \$7.5m in funding to research a number of ideas for chemical weapons.

Some of the ideas suggested are beyond comical – for instance, a chemical that would attract bees to enemy soldiers so they get stung – but one in particular caught our eye.

[The document](#) suggests that it could develop a powerful aphrodisiac to spray on people, in order to make enemy soldiers have sex with each other.



However, to be fair they did note some downsides... describing the plan as “distasteful”.

Complete with several spelling mistakes, it reads: “Category #3: Chemicals that effect human behavior so that discipline and morale in enemy units is adversely affected.

“One distasteful but completely non-lethal example would be strong aphrodisiacs, especially if the chemical also caused homosexual behavior.

“Another example would be a chemical that made personnel very sensitive to sunlight.”

The document laments: Chemicals that affected human behaviours in ways that would be disruptive to unit morale and effectiveness would need to be created.

“Chemicals that attract bees and cause them to sting would need to be identified or created.”

However, make sure you don’t drop the gay bomb and the bees bomb in the same place... or you might end up with gay bees.

**EDITOR’S COMMENT:** There is no smoke without a fire! Reconsider now with the aid of AI ...

## EOD: refining skillsets, cultivating fortitude

Source: <https://www.dvidshub.net/news/447021/eod-refining-skillsets-cultivating-fortitude>

June 12 – Explosives have been used for centuries, developing over the past few decades, becoming more complex and aggressive. To protect its bases and assets, the U.S. Air Force trains an elite team of preeminent tactical and technical explosives experts—Explosive Ordnance Disposal.

The 21st Civil Engineer Squadron EOD at Peterson Space Force Base, Colorado remains ready to support any of the nine mission sets of EOD, which include unexploded ordnance, counter improvised explosive device, weapons of mass destruction and nuclear response.

To come face-to-face with these events, the EOD developed their expertise through a rigorous curriculum of physical training, academic instruction and continuous hands-on practice with explosives.



## ICI C<sup>2</sup>BRNE DIARY – June 2023

Even after completing over a year of training, it is not over when EOD members arrive at their first duty station. There are over 300 training tasks and events to complete in order to be fully qualified and certified as an EOD team member—followed by continuous recertification to sustain peak mental and physical performance.



U.S. Air Force Tech. Sgt. David Thomas, an EOD team leader at the 21st CES EOD, compared the level of training to a firefighter. “Most people are generally running away from a burning building. These guys are running in,” Thomas said. “They’re not doing that blindly, just like we’re not doing this blindly. There’s a lot of training—very extensive training.”

This training prepares them for real-world operations that save lives. These operations often involve working with other agencies to synchronize response efforts.

In 2023, the 21st CES EOD finetuned this collaboration with FBI response team and local law enforcement in a joint training at the U.S. Air Force Academy—one of five military installations that the 21st CES EOD supports in the Front Range region, in addition to the Colorado Springs bomb squad.

“If they are to come across anything military related, a UXO for instance, they’re going to reach out to us, because that’s our expertise,” Thomas said.

Cheyenne Mountain Space Force Station, one of the five installations that the 21st CES EOD supports, is truly unique to any other EOD shop. The base is encased in a mountain, which offers additional challenges to EOD technicians.

“When you’re in there, you are essentially inside of a big cave,” Thomas said. “That very much changes the explosive effects, the damage that would be caused and how the energy is going to be transferred throughout all the different mediums there.”

Another one of the nine core mission areas that EOD supports in the Air Force is the Very Important Persons Protective Support mission set.

Most recently, the 21st CES EOD supported the Secret Service to protect President Joe Biden as he flew into Peterson SFB to speak at the 2023 USAFA graduation ceremony. They provided two direct support teams to secure Air Force One’s landing and assigned members to a dedicated standby team during the graduation.

The 21st CES EOD area of responsibility is not limited to Colorado though. In early 2023, two members of the 21st CES EOD team completed a nine-day U.S. Secret Service mission to Ghana, Africa, to conduct operations in support of the Vice President Kamala Harris.



## ICI C<sup>2</sup>BRNE DIARY – June 2023

“I think it was a great experience getting to go on this tasking to Ghana,” said Senior Airman Steven Nguyen, one of the EOD technicians assigned to this mission. “I got to work with the local fire department, the local police and the local military. I would’ve never had an opportunity like this if I wasn’t in EOD.”

However, this was not the first time the unit supported the vice president, as they also deployed in late 2021 for Vice President Harris’s trip to New York.

Looking back at the last few decades, specifically during the war in Afghanistan, a critical response capability for EOD was the counter IED mission set.

“The IED was the biggest threat during the Afghan War,” said Staff Sgt. Nicholas Logie, an EOD team leader for the 21st CES EOD. “A lot of EOD techs lost their lives over them.”

Airman 1st Class Matthew Seidler, an EOD technician from the 21st CES EOD, died from an IED explosion in Afghanistan in 2012.

“One of his team members was a good friend of mine that I went to school with,” said Master Sgt. Daniel Esselstrom, the 21st CES EOD noncommissioned officer in charge.

The EOD community is relatively small compared to other career fields, especially in the Air Force. This creates an additional bond between EOD members.

“It’s a very big deal anytime we lose anyone, especially when it’s an EOD tag,” Thomas said. “It’s like it was one of my brothers or sisters that it happened to.”

Despite the risk, the 21st CES EOD members said that they must trust their fellow EOD technicians with their lives. This creates a bond that some describe as family.

“When your coworkers are solely responsible for your safety, I think that that kind of forces a stronger bond between people,” Thomas said. Just as the threats facing EOD develop and evolve, so do the members of the 21st CES EOD. They adapt to whatever scenario they encounter, ready to respond to any threat.

Multiple members of the 21st CES EOD agreed on one thing, when the time comes for action, they are ready.



**When stepped on a landmine, it does not make a soft “click” and explode after the pressure is lifted as shown in the movies. Instead, It explodes the moment you step on it.**



## Common explosives and their uses

Source: <https://www.ukfrs.com/guidance/search/common-explosives-and-their-uses>

### Nitroglycerine

A very powerful and extremely sensitive liquid explosive that is usually mixed with other inert materials to form propellant, dynamites and blasting gelatines. The liquid is particularly sensitive to heat, flame, shock, oxygen or UV radiation and is not transported or used in its pure form as it is too dangerous to handle. Some types of explosives containing nitroglycerine such as dynamite, release nitroglycerine liquid when in contact with water. This can present particular dangers during firefighting or where the explosives are involved in flooding. Nitroglycerine is also transported and used either as a dilute solution in alcohol or as a mixture with a solid dilutant for pharmaceutical purposes. Nitroglycerine is highly toxic in both vapour and liquid form and is readily absorbed through the skin.

### Trinitrotoluene (TNT)

A comparatively insensitive explosive mostly used for military purposes. It is a stable solid substance that is generally safe to handle. Small, unconfined quantities may burn quietly but larger quantities, particularly under confinement, are likely to burn to detonation. On decomposition, toxic fumes are given off. TNT is moderately toxic by ingestion and can be absorbed through the skin.

### Nitrocellulose (guncotton, pyrocotton, nitrocotton)

Nitrocellulose is a white or cream coloured fibrous material produced in a range of forms. Its properties depend on the amount of nitrogen in it. Types with a nitrogen level above 12.6% are generally used in explosives manufacture. Those below that are used for other purposes. Nitrocellulose is sensitive to impact and friction and easy to ignite by flame when dry and is therefore normally transported either wetted with water or alcohol or it is plasticised. Explosives-grade nitrocellulose can present a fire or explosion hazard depending on its level of dryness and whether it is confined. Non-explosives-grade nitrocellulose will normally burn and is often classified as a class 4.1 flammable solid. Its principal uses are as an ingredient in propellant mixtures for ammunition and rocket motors, and in nitroglycerine explosives. Non-explosive grades of cellulose nitrate are used with other substances in the manufacture of paints and lacquers. Small arms propellant powders are usually made from nitrocellulose (single base) or nitro-cellulose with nitroglycerine (double base) and are usually in the form of a free-flowing granular material. For the most part, they burn fiercely but under confinement many types can burn to detonation. As a consequence, the hazard can be either a fireball or a mass explosion.

### Primary explosives such as fulminates, azides and styphnates

These are initiating explosives that are highly sensitive and dangerous, particularly when they contain heavy metals such as lead, silver or mercury. They explode violently and are commonly used in initiators such as detonators and cap compositions. These types of explosive are infrequently transported and are only carried when wetted with water or other suitable liquid to minimise their sensitivity to impact and friction. Transport packages are carefully designed to minimise the risk of wetting agents leaking and the explosive consequently drying out. Some forms of these substances, such as styphnic acid or potassium azide, have low explosive properties. However, they can react with other substances such as copper or lead salts to become highly dangerous forms. This can be a serious problem where these substances have been spilled and allowed to get in contact with copper, lead or other reactive species. Because the sensitive forms tend not to dissolve in water, a hazard could rapidly develop.

### Gunpowder/black powder/black blasting powder

These are generally made from a mixture of charcoal, sulphur and potassium nitrate and are frequently described as low explosives. They are very sensitive to ignition from sparks, heat and friction. They burn violently even when loose and uncompressed and when confined may explode. They also release volumes of smoke on burning or exploding. Black powder is used as a blasting explosive as well as an ingredient in some types of sporting cartridges, fireworks and pyrotechnics.

### Oxidisers

These substances are not explosives in their own right but can explode in a fire under certain conditions. Two types commonly used in explosives are worthy of comment.

- Nitrates are used in explosives such as gunpowder, emulsion and slurry-blasting explosives and pyrotechnic substances. The most common are potassium, sodium, barium and ammonium nitrates. If mixed with a fuel these compounds can burn fiercely or explode. Wood, if impregnated with oxidisers, can burn fiercely, which can be a potential problem on vehicles with wooden floors or in wooden storage buildings where oxidisers have been transported or stored. Care must be



taken in warehouses to ensure that oxidisers are segregated from other dangerous goods. Ammonium nitrate is used as a fertiliser and is commonly found in farming areas. It is usually supplied as a porous prill (pellet) that can readily absorb liquids. It forms the basis of ammonium nitrate and fuel oil (ANFO) explosives often used by terrorist organisations. It is possible that other fuels such as sawdust and metal powders could also form explosive mixtures with ammonium nitrate. There have been a small number of reports throughout the world of nitrates exploding without other chemicals being present, when confined or subject to severe heat. In recognition of this danger, the size of stack of ammonium nitrate in storage is normally carefully controlled

- Chlorates, mostly in the form of potassium chlorate, are mainly used in pyrotechnic compositions. As with nitrates, chlorates can decompose violently in a fire, particularly if mixed with fuels. As a general rule, chlorates are more aggressive and less stable in a fire than nitrates

### Emulsion and slurry explosives

Emulsion and slurry explosives are relatively new types of explosives that are mixtures of nitrates and other substances, often in a water-based system. They are now replacing nitroglycerine-based explosives for many uses, particularly in quarrying. A small proportion of emulsion and slurry explosives are produced as pre-packed, ready to use products. Most of this type of explosives, however, is manufactured at the scene (shot hole), immediately prior to use. This mixing is carried out on a specially designed mixer truck that carries the necessary ingredients for producing a fully active explosive.

### Peroxide explosives (improvised or home-made explosives)

Peroxide explosives are extremely dangerous. Only small quantities are needed to cause serious injury or explosions, the constituents are readily available from almost any high street and they are easy to make, with instructions for their manufacture easily available on the internet

Emergency services may encounter them at many routine incidents (e.g. domestic property fires)

The main constituents are:

- Hydrogen peroxide (e.g. hair dye, etc.)
- Acid (e.g. battery acid, brick cleaner, citric acid used in brewing, etc.)
- Acetone (e.g. nail varnish remover, etc.)
- Hexamine (e.g. camping stove fuel tablets etc.)

It is important for fire and rescue service responders to know and remember these main constituents because identifying the possibility of home-made explosives being present in the early stages of an apparently routine incident can save lives.

Peroxide explosives can vary greatly in appearance. Pure substances form a white powder, but they may also be granular in texture like sugar, or even form a sticky goo. Responders should not rely on physical appearance alone to identify this hazardous material. It is more important to recognise the raw constituents along with other indicators such as mixing jars and containers or a fridge or freezer to keep the substance cool.

Common types of peroxide explosives are:

- Triacetone triperoxide (TATP) is an explosive that has appeared as a terrorist weapon in the Middle East. TATP has been used by suicide bombers in [Israel](#), and was chosen as a detonator in 2001 by the thwarted shoe bomber Richard Reid. It is one of the most sensitive explosives known and is extremely sensitive to impact, temperature change and friction
- Hexamethylene triperoxide diamine (HMTD) is another peroxide-type explosive that is less sensitive than TATP but still dangerous

## Landmine Use in Ukraine

Source: <https://www.hrw.org/news/2023/06/13/landmine-use-ukraine>

June 13 – This background briefing surveys the types of landmines used in Ukraine since the full-scale Russian invasion began on February 24, 2022.<sup>[1]</sup> Antipersonnel mines explode by the presence, proximity, or contact of a person, and can kill and injure people long after armed conflicts end. The 1997 Mine Ban Treaty prohibits antipersonnel mines, but not antivehicle mines or command-detonated (remote-controlled) mines, which are also surveyed in this briefing.<sup>[2]</sup> Other international agreements weakly restrict anti-vehicle mine use, and international efforts to strengthen existing protections have stalled.<sup>[3]</sup> Visuals and accounts from deminers working in Ukraine indicate that the use of antipersonnel and anti-vehicle mines in the armed conflict is resulting in a large, dispersed, and complex level of contamination that will threaten Ukrainian civilians and hinder recovery efforts for years to come.





## Findings

Ukraine is severely contaminated with landmines and explosive remnants of war (ERW) from the armed conflict that began in 2014 and surged with Russia's full-scale invasion in February 2022.<sup>[4]</sup>

Landmines have been documented in 11 of Ukraine's 27 regions: Chernihivska, Dnipropetrovska, Donetska, Kharkivska, Khersonska, Kyivska, Luhanska, Mykolaivska, Odeska, Sumska, and Zaporizka.

### Antipersonnel Mines

- Russian forces are known to have used at least 13 types of antipersonnel mines since February 2022.
- Human Rights Watch documented Ukrainian forces repeatedly used rocket-delivered PFM antipersonnel blast mines in attacks on and around the city of Iziom during the summer months of 2022 when Russian forces controlled the city and its environs.
  - Ukrainian Deputy Defense Minister Oleksandr Polishchuk responded by letter to Human Rights Watch's questions about Ukrainian forces' use of PFM antipersonnel mines, saying that Ukrainian authorities cannot comment on the types of weapons used during the armed conflict "before the end of the war and the restoration of our sovereignty and territorial integrity."<sup>[5]</sup>
  - On January 31, 2023, the Ministry of Foreign Affairs stated that Human Rights Watch's findings "will be duly studied by the competent authorities of Ukraine."<sup>[6]</sup>
  - The use of antipersonnel mines by Ukraine is a violation of the 1997 Mine Ban Treaty, to which Ukraine is party.
- Russian forces also emplaced numerous victim-activated booby-traps as they retreated from positions they had taken, occupied, or fortified during the 2022 invasion. Booby-traps can function as antipersonnel mines when a person unintentionally activates the fuze and detonates the device.

### Anti-Vehicle Mines

- Both Russian and Ukrainian forces have used at least 13 types of anti-vehicle mines (also called anti-tank mines).
- Hand-emplaced TM-62 series anti-vehicle mines appear to be the type most frequently deployed.
- States supplying Ukraine with weapons have transferred seven types of anti-vehicle mines. Those states include Estonia, France, Germany, Sweden, UK, and the US.

### General Observations

- Several landmine delivery methods have been documented since February 2022: hand-emplaced, mechanically laid, scattered by truck-mounted projectors, and remotely delivered by rockets.
- Both Russia and Ukraine stockpile landmines inherited from the former Soviet Union.
- Russia continues to produce both antipersonnel and anti-vehicle mines; several new, previously unseen Russian landmines have been identified in use for the first time, including some antipersonnel mines produced as late as 2021.
- Landmine casualties have been reported in the conflict in Ukraine, including by Human Rights Watch, and new mine use can also be seen in the denial of access to homes, infrastructure, transportation routes, and agricultural lands used by civilians.<sup>[7]</sup>
- Landmines in fields and on rural paths and roads are harming agricultural production.

## Recommendations

All parties to the armed conflict in Ukraine should ensure that no antipersonnel mines are used by any actor and destroy any antipersonnel mines that they have seized or otherwise acquired. In addition:

1. Ukraine should ensure that its forces respect their obligations under the Mine Ban Treaty in all circumstances. The Ukrainian government should confirm to states parties that remaining stockpiles of PFM antipersonnel mines currently in storage awaiting destruction cannot be used in combat operations.
2. Russia should fully abide by the prohibitions and restrictions of the Convention on Conventional Weapons Amended Protocol II, including the detectability and reliability of any mines it uses in this conflict, as well as marking, monitoring, and ensuring the effective exclusion of civilians from mined areas.<sup>[8]</sup>
3. All parties to the Mine Ban Treaty supporting Russia's war effort, such as Belarus, and several states parties that are supplying Ukraine with weapons, need to ensure that these activities do not violate the prohibition on assisting, encouraging or inducing a non-party to engage in activities prohibited by the treaty.<sup>[9]</sup> This means avoiding participating in the planning for use of antipersonnel mines; not committing to rules of engagement that permit use of antipersonnel mines; not accepting orders to use, request others to use, or train others to use antipersonnel mines, not



knowingly deriving military benefit from the use of the weapon by others; and not providing security, storage, transportation or transit for antipersonnel mines.

4. Ukraine should continue efforts to identify victims, collect comprehensive data about victims and their needs, inform them of their rights and availability of services, and provide appropriate and timely compensation, assistance, and support to those injured and the families of those killed by mines, including medical care, psychosocial support, prosthetics where appropriate and ongoing rehabilitation needs, as well as to other landmine victims in Ukraine. Donor countries should support the Ukrainian government's efforts to identify victims, provide medical and other assistance, destroy stockpiled mines, and clear mine-contaminated areas.
5. In territory occupied by its forces and in line with its obligations under the Geneva Conventions, Russia should provide explosive ordnance risk education for the population, compensation to those injured and the families of those killed by mines, medical care for mine victims including prosthetics where appropriate and other rehabilitation needs.

## Methodology

This background briefing draws on information collected by Human Rights Watch in Ukraine, including from witnesses to landmine use, victims of landmines and their families, first responders, doctors, and Ukrainian deminers. It also relies on several other sources, including:

- Verified videos and photos, obtained directly or via social media, showing landmines or evidence of their use.
- Information shared with Human Rights Watch by international and national demining personnel, armament research specialists, domestic and international media, military analysts, and human rights organizations;
- Displays of landmines cleared by government authorities and deminers;
- Munitions technical reference guides maintained by the Geneva International Centre for Humanitarian Demining (GICHD), Collective Awareness to UXO (CAT-UXO.com), Fenix Insight, Jane's Information Group, Armament Research Services (ARES), and Russia Military Parade. Of particular utility for ordnance identification is the GICHD's "Explosive Ordnance Guide for Ukraine," second-edition guide.[\[10\]](#)
- *Landmine Monitor Report* profiles of Russia and Ukraine.[\[11\]](#)

This document updates a June 2022 briefing paper by Human Rights Watch.[\[12\]](#) It should be regarded as an interim overview because active combat operations limit access by independent observers to battlefield areas. It is not currently possible to survey mine contamination systematically or determine the full number of casualties caused by landmine use given ongoing hostilities. Some of the data presented here is taken from government sources or was posted to social media by third parties and should be treated with care, particularly in attributing responsibility for landmines use and its compliance with treaty obligations and international humanitarian law.

Other types of mines reported in the Ukraine conflict by various sources and media outlets have not been included in this interim accounting for various reasons. Some have dubious sourcing, visual materials that appear to be staged, or they appear to show mines from displays and inert models of mines and ERW used for recognition or risk education training. This background briefing also does not address sea mines or anti-landing mines such as the PDM-1M.

Human Rights Watch welcomes corrections and further information.

### Types of Landmines Used in Ukraine since 2022

Human Rights Watch has identified at least 13 types of antipersonnel mines and 13 types of anti-vehicle mines used in the current conflict in Ukraine through June 2023 (see tables below). The 2022 background briefing identified at least seven types of antipersonnel mines and six types of anti-vehicle mines.

All manner of landmine delivery methods have been used except for air-delivered mines: hand-emplaced, mechanically laid, scattered by truck-mounted projectors and remotely delivered by rocket.

Factory markings on the landmines used by Russia show that they were manufactured in the Soviet era and subsequently in Russia; some antipersonnel mines were produced by Russia as recently as 2021.

Ukraine inherited a significant stockpile of landmines from the former Soviet Union but destroyed most of them.[\[13\]](#) It has also acquired several types of anti-vehicle mines from allied countries.

The 1997 Mine Ban Treaty prohibits antipersonnel mines but not anti-vehicle mines or command-detonated (remote-controlled) mines. Other international agreements, notably the 1996 Amended Protocol II of the Convention on Conventional Weapons (CCW), provides weak restrictions on the use of anti-vehicle mines.[\[14\]](#)

### Antipersonnel Landmines used in Ukraine

Russian forces have used at least 13 types of antipersonnel mines since February 24, 2022. In September, October, and December 2022, Human Rights Watch spoke with Ukrainian deminers who were involved in clearance operations in the Kharkivska region, including in Iziium and in parts of the



Khersonska region, following the retreat from those areas of Russian forces. They identified numerous types of antipersonnel mines they had found and neutralized in recently retaken areas, all of which are known to be in Russian stockpiles, including OZM-72 bounding fragmentation mines and PMN-series blast mines (both PMN-2 and PMN-4). [15]

Russian forces have also emplaced victim-activated booby-traps at positions that it has taken, occupied, or fortified over the past 16 months. Deminers told Human Rights Watch that they have cleared and destroyed multiple victim-activated booby traps from areas that were formerly under Russian control. The booby traps were constructed with various types of hand grenades equipped with tripwires, including F-1, RGD-5, and RGN-type grenades. Booby-traps can function as antipersonnel mines when the fuze that is used is activated unintentionally by a person.

Ukrainian forces repeatedly used rocket-delivered PFM-series antipersonnel blast mines in attacks on and around the city of Iziom in 2022 when Russian forces controlled the city and its environs. The use of antipersonnel mines by Ukraine is a violation of the 1997 Mine Ban Treaty.

Additionally, the United States has transferred command-detonated M18A1 Claymore directional fragmentation munitions to Ukraine, but without their victim-activated trip-wire fuzes. [16]

#### Antipersonnel Landmines in Ukraine since February 24, 2022

Name	Origin	Type	Initiation	Notes
<a href="#">MOB</a>	Russia	Fragmentation	Multiple options	A modern hand-emplaced directional multi-purpose mine that is emplaced either in a command-detonated or victim-activated manner. When used in victim-activated mode with a mechanical pull, tension release, or seismic fuze, they are prohibited by the Mine Ban Treaty. This mine is only used by Russia.
<a href="#">MON-50</a>	USSR/Russia	Fragmentation	Tripwire/command	The MON-series of hand-emplaced directional multi-purpose antipersonnel mines can be used either in a command-detonated or victim-activated manner. [17] When used in victim-activated mode with a mechanical pull, tension release, or seismic fuze, they are prohibited by the Mine Ban Treaty.
<a href="#">MON-90</a>	USSR/Russia	Fragmentation	Tripwire/command	
<a href="#">MON-100</a>	USSR/Russia	Fragmentation	Tripwire/command	
<a href="#">MON-200</a>	USSR/Russia	Fragmentation	Tripwire/command	
<a href="#">OZM-72</a> [18]	USSR/Russia	Fragmentation	Tripwire/command	A multi-purpose bounding munition emplaced either in a command-detonated or victim-activated manner. When used in victim-activated mode with a mechanical pull, tension release, or seismic fuze, they are prohibited by the Mine Ban Treaty.
<a href="#">PFM-1/</a> <a href="#">PFM-1S</a>	USSR	Blast	Pressure, some SD	Uniquely shaped and constructed, this plastic-cased mine can be scattered by mine-laying rockets and dispensers mounted on trucks or helicopters. It contains 37 grams of a liquid high explosive. Both Russia and Ukraine stockpile this type.
<a href="#">PMN-2</a> [19]	USSR/Russia	Blast	Pressure	A circular, plastic-cased mine. Ukraine destroyed its stockpile of this type in 2003.
<a href="#">PMN-4</a> [20]	Russia	Blast	Pressure	Modern circular, plastic-cased mine produced by Russia. First publicly displayed by Russia in 1993, it has never been stockpiled by Ukraine.
<a href="#">POM-2/</a> <a href="#">POM-2R</a> [21]	USSR/Russia	Fragmentation	Tripwire, SD	A metal-case bounding mine delivered by helicopter, ground-fired rockets, or other means. POM-2 and POM-2R mines are stockpiled by Russia, Ukraine destroyed its stocks of this mine in 2018.



Name	Origin	Type	Initiation	Notes
<a href="#">POM-3</a>	Russia	Fragmentation	Seismic	Used only by Russia, POM-3 mines were first publicly displayed during annual military exercises in 2021. <a href="#">[22]</a> The POM-3 is scattered by rockets or truck-mounted launchers. Ukraine does not possess the POM-3 mine or its delivery system. Markings on an expended delivery canister pictured with POM-3 mines that failed to deploy properly indicate it was produced in 2021. <a href="#">[23]</a>

Note: SD=self-destruct

### Anti-vehicle Landmines used in Ukraine

Both Russian and Ukrainian forces have used at least 13 types of anti-vehicle mines (also called anti-tank mines). Agricultural production is being negatively affected due to the use of landmines in fields and on rural paths and roads.

The hand or mechanically emplaced TM-62 series anti-vehicle blast mines equipped with an MVCh-62 pressure activated fuze appears to be the most common type of anti-vehicle mine used since Russia's full-scale invasion. These mines are often buried but are also seen laid on top of the ground.

Several countries supplying Ukraine with military assistance have transferred anti-vehicle mines to Ukraine, including Estonia, France, Germany, Sweden, the United Kingdom, and the United States.

### Anti-vehicle Landmines in Ukraine since February 24, 2022

Name	Origin	Type	Initiation	Notes
<a href="#">PTKM-1R</a> <a href="#">[24]</a>	Russia	Shaped-charge, top-attack	Seismic sensor	Produced and used only by Russia, the mine canister is hand-emplaced and its sensors detect the seismic and thermal signatures of vehicles to distances of 200 to 300 meters. Once a target is detected, a sensor-fuzed submunition is ejected into the air and scans for a target. Once a target is located a shaped charge is fired downwards onto the top of the target.
<a href="#">PTM-1/ PTM-1G</a> <a href="#">[25]</a>	USSR/Russia	Blast	Pressure, SD	A plastic-bodied, rectangular mine scattered by helicopter or rocket artillery (Grad/Uragan). Both Russia and Ukraine stockpile this type of mine
<a href="#">PTM-3</a> <a href="#">[26]</a>	USSR/Russia	Shaped-charge	Magnetic Influence, SD	Rectangular, metal-cased mine scattered from individual tubes or truck-mounted, helicopter, and rocket dispensers. Only Russia stockpiles this type of mine.
<a href="#">PTM-4M</a>	Russia	Shaped-charge	Magnetic influence, SD/SDA	Modern metal-cased, rectangular mine scattered from individual tube or truck, helicopter, and rocket mounted dispensers. This mine has not been documented previously and the marking on the disperser indicates production in 2021. <a href="#">[27]</a>
<a href="#">TM-62M</a>	USSR/Russia	Blast	Pressure	Large, metal-cased circular mine that is either hand emplaced or mechanically laid. In addition to the commonly seen MVCh-62 pressure fuze, it is compatible with many other types of fuzes. Both Russia and Ukraine possess this type.
<a href="#">TM-62P3</a>	USSR/Russia	Blast	Pressure	Large, plastic-cased circular mine that is either hand emplaced or mechanically laid. In addition to the commonly seen MVCh-62 pressure fuze, it is compatible with many other types of fuzes. Both Russia and Ukraine possess this type.



Name	Origin	Type	Initiation	Notes
<a href="#">TM-83[28]</a>	USSR/Russia	Shaped-charge	Off-route, seismic, infra-red	Emplaced above the ground, this cylindrical metal-cased mine possessed by both Russia and Ukraine.
<a href="#">DM-22</a> , <a href="#">PARM-2[29]</a>	Germany	HE/AT	Off-route, SDA. Infra-red sensor, command, crush wire	Mounted above the ground on a tripod, this mine fires a high explosive/anti-tank (HE/AT) fin-stabilized projectile into its target.
DM-31, <a href="#">FFV-028[30]</a>	Germany, Sweden	Shaped charge	Magnetic Influence, SDA, some SD	A modern, metal-cased mine that is either hand-emplaced or mechanically laid.
<a href="#">DM-1399</a> , <a href="#">AT-2[31]</a>	Germany	Shaped charge	Magnetic Influence, scratch/contact wire, SD, SDA	A total of 28 of these mines are delivered by 227mm rocket. Using a small parachute to facilitate deployment, this landmine sits on the surface of the ground on five fold-out legs.
<a href="#">HPD-2A2[32]</a>	France	Shaped charge	Magnetic influence, SD, SDA	A modern, mechanically laid mine with a sensitive initiating sensor that could detonate the mine prematurely.
<a href="#">M/56[33]</a>	Denmark	Pressure	Pressure, tilt rod	A large, non-metal cased blast mine that is a copy of the French Model 1951/52 blast landmine.
PK-14[34]	Estonia	Command detonated	Off-route, shaped charge	Also called M14, this item is emplaced above the ground.
<a href="#">RAAMS[35]</a>	USA	Shaped charge	Magnetic Influence, SD, SDA	Delivered by 155mm artillery projectile, each containing nine mines. The projectiles in this system have long or short duration self-destruct times: M718/M718A1 (RAAM-L) SD time over 24 hours; M741/M741A1 (RAAM-S) SD time under 24 hours.

Note: SD=self-destruct; SDA=self-deactivate; RAAMS=Remote Anti-Armor Mine System

Many types of anti-vehicle mines, including some on this list, have devices, features, and characteristics designed to deter and attack military clearance efforts (called countermine) during combat. Some could have an antihandling device that causes the mine to explode from an unintentional or innocent act of a person. Others feature a fuzing mechanism that is sensitive enough to be activated and trigger the mine by the unintentional act of a person. Certain types of antivehicle mines rely on a tripwire, breakwire, or tilt rod as their sole initiating mechanism.[\[36\]](#)

No matter its label or design intent, there is widespread agreement that any explosive device that is capable of being detonated by the unintentional act of a person is an antipersonnel mine and thus prohibited under the Mine Ban Treaty. This is the view of a majority of the treaty's 164 states parties and also shared by the International Committee of the Red Cross.[\[37\]](#)

### Impact of Landmines on Ukrainian Civilians

Ukraine's contamination from landmines and explosive remnants of war (ERW) dates to World War II, worsened during armed conflict that began in 2014, and has expanded greatly since Russia's full-scale invasion in February 2022. [\[38\]](#) In the current conflict, landmines have been used in at least 11 regions of Ukraine: Chernihivska, Dnipropetrovska, Donetsk, Kharkivska, Khersonska, Kyivska, Luhanska, Mykolaivska, Odeska, Sumska, and Zaporizka.

Currently, there is no systematic reporting on the number of mine incidents and victims, nor is any disaggregated casualty information available publicly. Ukrainian Prime Minister Denys Shmyhal stated in April 2023 that about 174,000 square kilometers of Ukrainian land (larger than the total landmass of Greece) is now contaminated with explosive objects, which killed 226 people in the month of March 2023 alone in territory controlled by Ukraine.[\[39\]](#)

Three civilians were reportedly seriously injured by Russian landmines in three separate regions on March 17, 2023: a 57-year-old man in Kherson who was collecting firewood, the driver of a tractor in Chernihiv, and a 66-year-old woman walking in the forest in Izium, Kharkiv. [\[40\]](#)

In 2022, Ukraine fired numerous 9M27K3 mine-laying rockets carrying PFM antipersonnel mines into at least nine different Russian-occupied areas or near Russian military facilities in and around Izium city. Human Rights Watch researchers verified 11 civilian casualties from these mines, of which four people lost their foot or lower leg. One, a 77-year-old man, died some days after he was injured by a mine,



although other factors beyond the mine-inflicted injuries may have contributed to his death. Local healthcare workers told investigators they had treated about 50 people for injuries from antipersonnel mines.

Other landmine casualties have been reported in the conflict in Ukraine, but many likely go unrecorded, especially if the victim was alone at the time of the incident. The impact of landmine use can also be seen in preventing civilians from accessing homes, infrastructure, transportation routes, and agricultural lands.

### Russia's Response

Russia has not joined the Mine Ban Treaty. In November 2020, it told the UN General Assembly that it “shares the goals of the treaty and supports a world free of mines,” but views antipersonnel mines “as an effective way of ensuring the security of Russia's borders.”<sup>[41]</sup>

Russia is party to the prohibitions and restrictions on mines, booby-traps, and other devices found in 1996 Amended Protocol II of the UN Convention on Conventional Weapons (CCW).<sup>[42]</sup> However, it has not reported any measures that it has taken to comply with Amended Protocol II (AP II) in its invasion of Ukraine, such as “feasible precautions” or steps necessary to ensure the effective exclusion of civilians from the areas such as fencing, signs, warnings, and monitoring, as required by article 5.2 of AP II.

Russia is also bound by prohibitions against deliberate, indiscriminate and disproportionate attacks against civilians found in the First Additional Protocol to the Geneva Conventions of 1949 (AP I)<sup>[43]</sup> and customary international humanitarian law.<sup>[44]</sup> Belarus, which has supported Russia's full-scale invasion of Ukraine, acceded to the Mine Ban Treaty on September 3, 2003.<sup>[45]</sup>

### Ukraine's Response

Ukraine signed the Mine Ban Treaty on February 24, 1999 and ratified the treaty on December 27, 2005. Ukraine is also party to CCW Amended Protocol II but their obligations on antipersonnel mines are superseded by the stricter prohibitions contained in the Mine Ban Treaty.

Ukraine inherited a substantial stockpile of antipersonnel mines after the breakup of the Soviet Union. It destroyed more than 3.4 million antipersonnel mines between 1999 and 2020, including PFM mines. In 2021, Ukraine reported to the UN secretary-general that 3.3 million stockpiled PFM mines still need to be destroyed. According to Ukrainian officials, the only type of antipersonnel mines remaining in Ukraine's stockpiles are PFM mines contained in 9M27K3 220mm rockets.

Human Rights Watch wrote to Ukrainian authorities on November 3, 2022 to present a summary of findings of its research into Ukraine's apparent use of PFM-series antipersonnel mines in and around Izium during 2022. Deputy Defense Minister Oleksandr Polishchuk replied on November 24 and stated that Ukrainian authorities cannot comment on the types of weapons used during the armed conflict “before the end of the war and the restoration of our sovereignty and territorial integrity.” The letter stated, “Ukraine is a reliable member of the international community, and it fully commits to all international obligations in the sphere of mine usage. This includes the non-use of anti-personnel mines in the war.”<sup>[46]</sup>

The deputy minister also wrote: “[D]uring the course of the war, Ukraine's Armed Forces strictly adhered to the standards of IHL [international humanitarian law] and provisions of the international conventions to which Ukraine is a party. This includes, inter alia, the Convention on the Prohibition of the Use, Stockpiling, Production... (hereinafter the Ottawa Convention), as well as the Convention on Certain Conventional Weapons, which are considered excessively injurious or whose effects are indiscriminate.”<sup>[47]</sup>

On January 31, 2023, in response to Human Rights Watch's report about Ukraine's PFM-series antipersonnel mine use, the Ministry of Foreign Affairs stated that the findings “will be duly studied by the competent authorities of Ukraine.”<sup>[48]</sup> The government has not provided any subsequent updates, though it is expected to address the matter, including steps it has taken, at the next intersessional meeting of the Mine Ban Treaty on June 19-21 in Geneva.<sup>[49]</sup>

### International Response

All parties to the conflict in Ukraine are bound by treaties that prohibit or regulate landmines in addition to the general laws of war. The 1997 Mine Ban Treaty comprehensively prohibits all types of victim-activated explosive devices, regardless of the technical features and regardless of the mine's predicted longevity, delivery method, or type of manufacture (improvised or factory-made).

The 20th Meeting of States Parties to the Mine Ban Treaty in November 2022 “condemned the use of anti-personnel mines anywhere, at any time, and by any actor.” Since March 28, 2022, Ukraine and at least 40 other countries have condemned or expressed concern at Russia's use of antipersonnel landmines in Ukraine: Albania, Austria, Australia, Belgium, Bosnia and Herzegovina, Bulgaria, Canada, Colombia, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Georgia, Greece, Guatemala, Hungary, Ireland, Italy, Japan, Latvia, Lithuania, Luxembourg, Malta, Netherlands, New Zealand, North Macedonia, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, South Korea, Spain, Sweden, Switzerland, United Kingdom, United States, and the European Union.

Landmine use in Ukraine has also been condemned by successive Mine Ban Treaty presidents as well as the special envoy for universalization.<sup>[50]</sup> On February 3, 2023, the president of the Mine Ban Treaty,



Ambassador Thomas Göbel of Germany, expressed concern at the alleged use of antipersonnel mines by members of the Ukrainian armed forces as reported by Human Rights Watch. In a statement, he said that he will work “together with the Convention’s Committee on Cooperative Compliance” to “use the established procedures to seek clarification of the allegations.” He said, “We are confident that we can continue to fully rely on Ukraine’s cooperation in this respect.”<sup>[51]</sup> The International Campaign to Ban Landmines, 1997 Nobel Peace Laureate, has condemned the use of antipersonnel landmines in Ukraine, as has its member organizations, including its chair, Human Rights Watch.<sup>[52]</sup>



ICI  
International  
**CBRNE**  
INSTITUTE



# CYBER NEWS





## Cybersecurity researchers gain control of ESA nanosatellite in an ethical hacking exercise

Source: <https://interestingengineering.com/culture/hackers-gain-control-esa-nanosatellite>

Apr 28 – A team of four cybersecurity researchers from the multinational technology company Thales was successful in hacking into a nanosatellite belonging to the European Space Agency (ESA). The attempt was carried out as part of ESA's cybersecurity exercise as part of its CYSAT conference.

With countries opening up [space to private players](#), there is a surge in the number of satellites orbiting the planet. A hacked satellite is a potential concern for governments around the world and it can be used to transmit sensitive information or even be weaponized. The ESA introduced the Hack CYSAT challenge, the first of its kind in the world to understand the potential impact of a real cyberattack.



### How hackers gained control of a Nanosatellite

The satellite made available for this challenge was the OPS-SAT demonstration nanosatellite that was launched in 2019. According to a press release from Thales, the team of cybersecurity researchers accessed the satellite's onboard system and "used standard access rights to gain control of its application environment."

The intrusion allowed the hackers to gain access to the satellite's global positioning system attitude control system as well as its onboard camera. The researchers also exploited several vulnerabilities in the satellite which allowed them to enter malicious code into the satellite's system.

Doing so, the researchers compromise the data the satellite was sending back to Earth, especially by modifying the images captured by the onboard camera. In addition to this, the hackers could also mask selected geographical areas visible in the satellite imagery to simulate the hiding of activities therein.

The ESA remained in control of the satellite during the test and also returned it to normal operation later on, so there isn't a nanosatellite spiraling out of control in orbit as of now.

"This unprecedented exercise was a chance to raise awareness of potential flaws and vulnerabilities so that they can be remediated more effectively, and to adapt current and future solutions to improve the cyber resilience of satellites and space programs in general, including both ground segments and orbital systems," said Pierre-Yves Jolivet, VP Cyber Solutions at Thales in a press release.

While the vulnerabilities of the ESA satellite are worrying, those in the commercial satellites are a greater cause of concern. Last year, *Interesting Engineering* reported how a hacker built a \$25 tool to [hack into SpaceX's Starlink](#) system which has a constellation of nearly 3,600 satellites in low-Earth orbit

A [Bloomberg](#) report last month stated that Russia managed to hack into several mainstream satellite internet systems in February last year. Around the same time, hacker group Anonymous claimed that it had [hacked into Russian spy satellites](#) in response to its invasion of Ukraine, a charge Russia denied.

## Is your car safe from a cyber attack?

By Lindsay James

Source: <https://eandt.theiet.org/content/articles/2023/05/is-your-car-safe-from-a-cyber-attack/>

May 17 – There's no doubt about it, our cars are smarter and more connected than ever before. While this opens a world of possibility, there's also a major downside: they are far more vulnerable to cyber attacks. What are automakers doing about it?

In January 2022, 19-year-old David Colombo from Dinkelsbühl, Germany, announced via Twitter that he had been able to hack at least 25 Tesla vehicles in 13 countries and partially take them over.

"So, I now have full remote control of over 25 Teslas in 13 countries and there seems to be no way to find the owners and report it to them," he tweeted. Luckily, Colombo's intentions were good. As the founder of



cyber-security firm Colombo Technology, he used his actions simply to demonstrate the security flaw of the third-party software that Tesla was using, and to warn automakers the world over about the danger of malicious attacks.

Colombo isn't alone in his mission. Cyber-security firm McAfee demonstrated how it could trick autonomous vehicles (AVs) into speeding over 50mph above the speed limit. And Ubiquitous System Security Lab, along with a series of partners, demonstrated how 'poltergeist' attacks – where attacks are made against the camera-based computer-vision systems found in AVs – can trick self-driving cars' machine-learning systems into ignoring obstacles.

"If we don't find vulnerabilities in the vehicles of tomorrow, threat actors will," Colombo wrote in an article for Medium. "Malicious cyber-attacks on vehicles and/or the automotive ecosystem can have disastrous outcomes affecting not just the public image of automakers and OEMs, but also having a direct impact on human lives, infrastructure, and other aspects. A fleet hacked by a threat actor with malicious intent would be a worst-case scenario that should be avoided at all costs."

The possibilities don't bear thinking about. "The ultimate risk would be a fleet of commercial vehicles or buses could be taken over remotely and turned into missiles," said Tu Le, founder of global innovation and management consultancy firm Sino Auto Insights. "Autonomous systems in military vehicles could be disabled while attacking or while under attack – and made useless. Warships could be taken over and used to attack or ambush unsuspecting entities. This is just to name a few."

However, Mike Ramsey, a research director focused on automotive and smart mobility at analyst firm Gartner, argues that the chances of malicious attacks causing real danger to human lives are slim. But he does believe that ransomware attacks are a very real possibility. "You have to think about things in terms of incentives," he says. "There's not a lot of incentive to do something dangerous. However, threat actors can hack a network and shut down a large number of vehicles or make them non-functional and then demand a ransom. Hackers are very good at figuring out the economics of this. When automakers are faced with a decision of fixing \$25m-worth of cars or paying a ransom of \$1m, there's not really much of a decision to be made."

Elad Robb, head of cyber threat intelligence, AutoThreat at Upstream Security, agrees: "As more autonomous and connected vehicles interact with other vehicles, road signs over networks, mobile applications and charging infrastructure, it will be more lucrative for malicious actors to attack them," he says. "Put simply, adversaries follow the money, so the greater the gain, the more likely they are to try and gain access."

Fortunately, the attacks reported to date, while dangerous, have not carried the impact of their full potential – but they have not been trivial. "In April 2022, an EV charging station in the Isle of Wight was hacked to show inappropriate content, with some EV owners also experiencing high-voltage fault codes, leaving them stranded," Robb says. "Also, in February, a Japanese OEM was forced to shut down 14 manufacturing facilities as a result of a cyber attack."

Meanwhile, Honda recently acknowledged that hackers had found a way to remotely start the engine of some of its models, and unlock doors by taking control of the car's remote keyless entry system.

These aren't isolated incidents. According to Upstream's 2022 Global Automotive Cybersecurity Report, the number of cyber events on cars soared by a massive 225 per cent between 2019 and 2022.

Why such a dramatic uptick? According to Gartner's Ramsey, there are multiple reasons. "Not only are there so many more vehicles on the road today, but the majority of these now have embedded connectivity," he says. "Embedded connectivity means that they are sending and receiving information all the time – and that makes them more vulnerable to cyber attacks."

In fact, the number of connected vehicles on our roads will increase by 134 per cent, from 330 million in 2018 to 775 million this year, according to Juniper Research. And, by 2025, a connected car will produce 25GB of data per hour and up to 500GB if fully autonomous.

"As the automotive industry transforms from individual, siloed vehicles into an interconnected smart mobility ecosystem, it expands the auto-industry ecosystem from vehicles into services," explains Robb. "As vehicles become software-defined and more connected, threats and attacks escalate accordingly. New attack surfaces continuously emerge and are exposed by cyber-security experts."

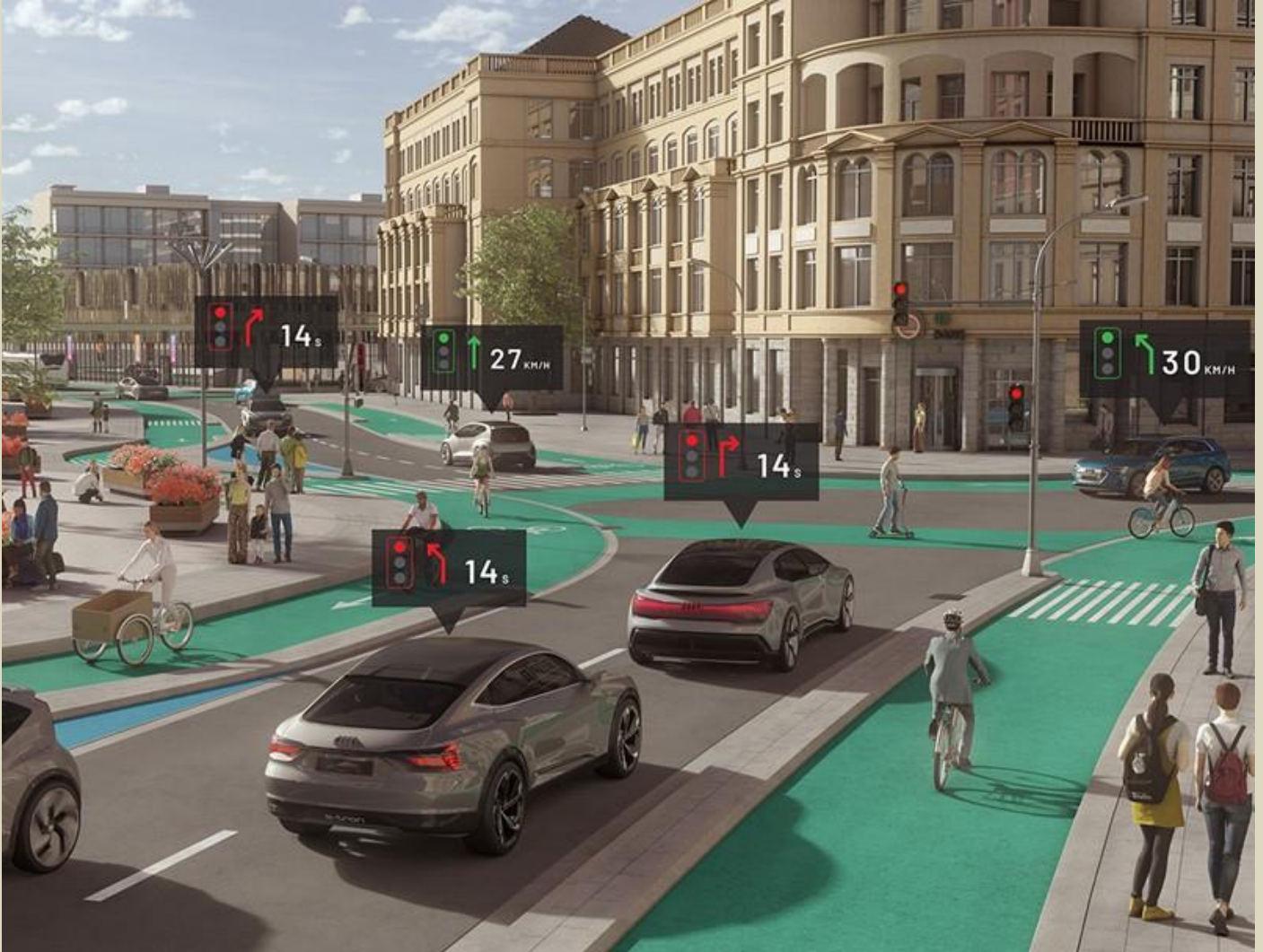
It appears that some automakers are paying attention. "Pioneering automotive players are definitely beginning to recognise that this is a unique use case where traditional IT cyber-security solutions may not fully meet the complex needs of protecting vehicles on the road, and have started implementing purpose-built solutions," says Robb.

Audi, for example, is prioritising the issue. "Audi is taking technical, organisational and process-related measures to ensure automotive security," says Christian Hartmann, a company spokesperson for electric mobility and automated driving. "We are doing this from the development stage, through encryption and authentication of data connections between cars, for example, right through to the backend. We are searching for weak points, and security experts and pen testers check the processes from the start of development to the start of production. We are constantly expanding the security mechanisms to develop new functions around data security, protecting the vehicle against hacking and privacy demands."

Meanwhile, the approach taken by Waymo – the company formerly known as the Google self-driving car project – is to think about cyber security holistically. "At Waymo, we build an autonomous driver, not a car," says Stacy Janes, head of cyber security at the company. "It can be applied to different vehicle types and use cases from passenger cars to delivery vans to big rig trucks. So when we think about cyber security,



that includes the Waymo Driver, the vehicle platform it's going to be applied to, how they interact and communicate with each other and the infrastructure that supports that."



There are threats that are unique to each application of the Waymo Driver. "We use a risk-based approach to identify these threats, evaluate what the main risks are and prioritise what to focus on first, working our way down through the list," Janes says. "Of course, this approach is not unique to the autonomous driving space and has been widely used across many other industries."

Waymo uses layers of security to protect its autonomous driving system – especially its safety-critical functions like steering and braking, and the way it interacts with the base vehicle.

"We also consider the security of our wireless communication," Janes adds. "The Waymo Driver does not rely on a constant connection to operate safely. While on the road, all communications between the operations centres and the vehicles are encrypted, including those between Waymo's operations support staff and riders. The Waymo Driver can communicate with the operations centre to gather more information about road conditions, while the Waymo Driver maintains responsibility for the driving task at all times."

That's not all. Waymo also has diverse mechanisms for noticing anomalous behaviour and internal processes for analysing those occurrences. "Should Waymo become aware that someone has attempted to impair its vehicle's security, it will trigger its company-wide incident response procedure, which involves impact assessment, containment, recovery and remediation," Janes says.

However, Le at Sino Auto Insights believes that this approach is the exception rather than the rule. "Many traditional automotive manufacturers are still trying to learn the basics of software development," he says. "So, generally speaking, they may be aware of the exposure that's created by manufacturing and selling smart, electric vehicles to the public (because they have high-paid lawyers that will tell them), but they're not savvy enough to know how to set up a bulletproof/hackproof infrastructure that's always one step ahead of the bad guys.

"Software hasn't been in their wheelhouse for the last hundred or so years, so most of them are still trying to understand the implications of having vulnerable firmware, control modules, operating systems or



infotainment systems, for example,” he continues. “Also, remember that they may rely on partners for software add-ons, so there’s another vulnerability they’re likely relying on their partner to secure.”

It’s also widely accepted that regulations are lacking. In fact, at the moment, it appears that the move toward connected autonomous vehicles is outpacing the cyber-security measures and regulations that are in place.

“Despite regulatory progress, automotive-specific cyber-security standards have not been fully mandated worldwide,” says Robb.

“With the proliferation of connectivity and software-based services, the attack surfaces on vehicles are rapidly expanding.”

However, change is on the horizon. New regulations are being implemented in Europe with the aim of protecting vehicles from both today’s known threats, and the unknown threats of the future.

“Since July 2022, a common set of cyber-security requirements is mandatory for all new vehicle types, and it will become mandatory for all new vehicles produced from July 2024,” explains Sonya Gospodinova, a spokesperson at the European Commission. “The new regulations on automated vehicles refer to the cyber-security requirements and add specific requirements when needed.” Moreover, she adds, the new NIS2 Directive on cyber security specifically includes motor vehicles among the sectors covered.

When the NIS2 directive enters into force, the manufacture of motor vehicles, trailers and semi-trailers, like other entities in sectors that are dependent on network and information systems and that provide key services to the EU economy and society, will be required to take cyber-security measures and report significant incidents with a view to increasing the overall level of cyber resilience throughout the internal market. However, Gospodinova is quick to recognise that cyber security must be constantly implemented to be future-proof. “Manufacturers have to put in place a cyber-security management system covering the whole lifecycle of the vehicle, from design to decommissioning, including software updates,” she says. “A challenge faced by the industry today is to define a cyber-security approach that allows third parties to develop and offer services for the vehicle users.”

Robb shares this opinion. “In order to mitigate and thwart these attempts to gain access to sensitive and critical systems,” he says, “OEMs should focus on ensuring they have a holistic view of all potential access points – from companion apps to charging stations as well as vehicles already on the road. Having the tools in place to monitor and understand the live state of the vehicle, consumer or application that interacts with the vehicle to detect unusual or malicious activity is key so that OEMs can respond efficiently via a dedicated virtual security operations centre.”

This is where Le believes that the majority of automakers will continue to fall short. “I think generally many CAV manufacturers outsource their cyber security to ‘security’ partners or rely on/assume that their suppliers are taking the necessary steps,” he says. “As long as the supply contract says it’s not their responsibility and they bear none for any hacks, then I think that’s as far as they’ll take it.” Ramsey agrees, adding that, while the new regulations will lead to some action, automakers won’t really make cyber security a priority until something serious happens. “At some point in the not-too-distant future, there’s going to be a major hacker event that will probably wake up the industry and start changing the talk around it,” he says. “I don’t know when that’s going to happen, of course, or the context of it, but it’s almost inevitable. As sad as it sounds, it probably needs to happen in order for the speed of implementation to quicken. If one automaker has a million cars that stop working all of a sudden, and the cost of that is astronomical, that is a pretty big motivator to change.”

## Cyber security

### Three key steps to success

Mike Ramsey, research director for automotive and smart mobility at Gartner, says that pioneering automakers are prioritising cyber security by focusing on three key areas:

Vehicle design – Automakers are having code written in a way that’s not as vulnerable as it has been in the past. Part of that is better standards for code development. Until very recently, code was almost entirely done by suppliers rather than the manufacturers themselves. There wasn’t a lot of standardisation. That’s changing slowly. So the design of the code is being set up so that it has to meet higher-level security standards.

Security at the edge – Pioneers are putting security around everything – in the telecommunications control unit, in the power control units, and even in the computers themselves so that there’s security at the edge, in the design and in the network. As a result, automakers are better equipped to notice when someone is trying to get into the network.

Security operations centres – Leading automakers are setting up security operations centres so that they can better monitor their fleet and ensure they have the ability to respond quickly and effectively if something does go wrong.

## How to protect a car

### Waymo’s self-driving car technology

Secondary compute: the Waymo Driver autonomous driving system has a secondary on-board computer that’s always running in the background. It’s designed to bring the vehicle to a safe stop should it detect a failure of the primary system.



Backup collision detection and avoidance system: multiple backup systems, including independent collision avoidance systems, constantly monitor the road ahead and behind the vehicle for objects such as pedestrians, cyclists and other vehicles. They can slow or stop the car in the rare event that the primary system doesn't respond.

Redundant steering: the steering system features a redundant secondary drive motor system with independent controllers and separate power supplies.

Redundant braking: a full, secondary braking system is designed to bring the vehicle to a safe stop if it should be required.

Backup power systems: independent power sources are provided for each of the critical driving systems. They ensure that the Waymo Driver remains up and running in the event of power failures or circuit interruptions.

Redundant inertial measurement systems for vehicle positioning: this helps the Waymo Driver accurately track its motion along the road. These two systems cross-check each other and assume control from one another if a fault is detected in either system.

- ✓ Cyber security: protecting the driving system from malicious activity is paramount. Waymo has developed a process to identify, prioritise and mitigate cyber-security threats in alignment with industry- and government-defined security best practices.
- ✓ Peripheral vision system: radar.
- ✓ 360° lidar: 360° Vision System.
- ✓ Forward-facing long-range camera: radar.
- ✓ Perimeter lidar: peripheral vision system.
- ✓ Perimeter lidar perimeter vision system.

## What will stop AI from flooding the internet with fake images?

Source: <https://www.vox.com/technology/23746060/ai-generative-fake-images-photoshop-google-microsoft-adobe>

June 03 – On May 22, a fake photo of an explosion at the Pentagon [caused chaos online](#).

Within a matter of minutes of being posted, the realistic-looking image [spread on Twitter and other social media networks](#) after being retweeted by some popular accounts. Reporters asked government officials all the way up to the [White House press office](#) what was going on.

The photo was quickly determined to be a hoax, likely generated by AI. But in the short amount of time it circulated, the fake image had a real impact and even briefly moved [financial markets](#).

This isn't an entirely new problem. Online misinformation has existed since the dawn of the internet, and crudely photoshopped images fooled people long before generative AI became mainstream. But recently, tools like ChatGPT, DALL-E, Midjourney, and even [new AI feature updates](#) to Photoshop have supercharged the issue by making it easier and cheaper to create hyperrealistic fake images, video, and text, at scale. Experts say we can expect to see more fake images like the Pentagon one, especially when they can cause political disruption.

One report by Europol, the European Union's law enforcement agency, [predicted that as much as 90 percent](#) of content on the internet could be created or edited by AI by 2026. Already, [spammy news sites seemingly generated entirely by AI](#) are popping up. The anti-misinformation platform NewsGuard started tracking such sites and found nearly three times as many as they did a few weeks prior.

"We already saw what happened in 2016 when we had the first election with a flooding of disinformation," said Joshua Tucker, a professor and co-director of NYU's Center for Social Media and Politics. "Now we're going to see the other end of this equation."

So what, if anything, should the tech companies that are rapidly developing AI be doing to prevent their tools from being used to bombard the internet with hyperrealistic misinformation?

One novel approach — that some experts say could actually work — is to use metadata, watermarks, and other technical systems to distinguish fake from real. Companies like Google, Adobe, and Microsoft are all supporting some form of labeling of AI in their products. Google, for example, [said at its recent I/O conference that](#), in the coming months, it will attach a written disclosure, similar to a copyright notice, underneath AI-generated results on Google Images. OpenAI's popular image generation technology DALL-E already adds a colorful stripe watermark to the bottom of all images it creates.

"We all have a fundamental right to establish a common objective reality," said Andy Parsons, senior director of Adobe's content authenticity initiative group. "And that starts with knowing what something is and, in cases where it makes sense, who made it or where it came from."

In order to reduce confusion between fake and real images, the content authenticity initiative group developed a tool Adobe is now using called content credentials that tracks when images are edited by AI. The company describes it as a nutrition label: information for digital content that stays with the file wherever



it's published or stored. For example, Photoshop's latest feature, Generative Fill, uses AI to quickly create new content in an existing image, and content credentials can keep track of those changes.

AI-labeling tools like Adobe's are still in their early stages, and by no means should they be considered a silver bullet to the problem of misinformation. It's technically possible to manipulate a watermark or metadata. Plus, not every AI generation system will want to disclose that it's made that way. And as we've learned with the rise of online conspiracy theories in recent years, people will often ignore facts in favor of believing falsehoods that confirm their personal beliefs. But if implemented well — and especially if these labels are seen as more neutral than traditional social media fact-checking — AI disclosures could be one of our only hopes for navigating the increasingly blurry distinction between fake and real media online.

Here is how some of these early AI markup systems could work, what the limitations are, and what users can do to navigate our confusing post-truth internet reality in the meantime.

### The devil is in the metadata

When you look at an image on social media or a search engine today, odds are you don't know where the photo came from — let alone if it was created by AI. But underneath the hood, there's often a form of metadata, or information associated with the digital image file, that tells you basic details, like when and where the photo was taken. Some tech companies are now starting to add specific metadata about AI to their products at the moment of creation, and they're making that information more public in an effort to help users determine the authenticity of what they're looking at.

Google recently said it will start marking up images made by its own new AI systems in the original image files. And when you see an image in Google Search that's made by Google's AI systems, it will say something like "AI-generated with Google" underneath the image. Going a step further, the company announced it's partnering with publishers like Midjourney and stock photography site Shutterstock to let them self-tag their images as AI-generated in Google Search. This way, if you come across a Midjourney image in Google Search, it will say something like "Image self-labeled as AI-generated"

Google Search public liaison Danny Sullivan said that this kind of AI labeling is part of a broader effort to give people more context about images they're seeing.

"If we can show you a helpful label, we're going to want to do that," said Sullivan, "but we're also going to want to try to give you background information that we can determine independent of the label."

That's why Google is also adding an "About this image" feature next to image search results — whether they are AI labeled or not — that you can click and see when the image was first indexed by Google, where it may have first appeared, and where else it's been seen online. The idea is, if you searched for, say, "Pentagon explosion" and saw a bunch of images in the results, you would be able to see a fact-checked news article debunking the piece.

"These tools are really designed to help people understand information literacy more and bake it into the search product itself," said Sullivan.

Other major industry players have also been working on the issue of how to label AI-generated content. In 2021, a group of major companies including Microsoft, Adobe, the BBC, and Intel created a coalition [called the C2PA](#). The group is tasked with helping to create an interoperable open standard for companies to share the provenance, or history of ownership, of a piece of media. C2PA created its first open standard last January, and since then, Adobe and Microsoft have released features using that standard.

For example, if you're a photographer at a news outlet, you can mark when a specific picture was taken, who took it, and have that be digitally signed by your publisher. Later, your editor could make changes to the photo, signing it again with a seal of authenticity that it's been verified by the C2PA standard. This way, you know that the photo was taken by a person — not generated by AI — and know who has made edits to it and when. The system uses cryptography to preserve the privacy of sensitive information.

"Now you can read the entire lineage of the history of a piece of digital content," said Mounir Ibrahim, EVP of public affairs and impact at Truepic, a visual authenticity app that is a member of C2PA. "The purpose of us is to help content consumers ... decipher the difference between synthetic and authentic."

Knowing the history and provenance of an image could potentially help users verify the legitimacy of anything from a headshot on a dating app to a breaking news photo. But for this to work, companies need to adopt the standard.

Right now, it's up to companies to adopt the C2PA standard and label verified content as they wish. The organization is also discussing potentially standardizing the look of the C2PA content credential when it shows up on images, Ibrahim said. In the future, the C2PA credential could be similar to the little padlock icon next to the URL in your browser window that signifies your connection is secure. When you see the proposed C2PA icon, you would know that the image you're seeing has had its origins verified.

So far, two big C2PA members, Adobe and Microsoft, have announced tools that integrate C2PA standards into their products to mark up AI-generated content. Microsoft is labeling all AI-generated content in Bing Image Generator and Microsoft Designer, and Adobe is using C2PA standards in its new AI Firefly product's content credentials.

"The biggest challenge is we need more platforms to adopt this," said Ibrahim.



While the C2PA-style metadata labels work behind the scenes, another approach is for AI systems to add visible watermarks, as OpenAI has done with the rainbow bar at the bottom of DALL-E images. The company says it's also working on a version of watermarking for its text app, ChatGPT. The challenge with watermarks, though, is that they can be removed. A quick Google search turns up forms of people discussing how to circumvent the imprint.

Another imperfect option is technology that can detect AI-generated content after the fact. In January, [OpenAI released a tool](#) that lets you cross-check a block of text to determine whether it's likely written by AI. The problem, though, is that by OpenAI's own assessment, the tool is not fully reliable. It correctly identified only 26 percent of AI-written texts in OpenAI's evaluations, although it's notably more accurate with longer than shorter text.

"We don't want any of our models to be used for misleading purposes anywhere," said a spokesperson for OpenAI in a statement. "Our usage policies also require automated systems, including conversational AI and chatbots, to disclose to users that they are interacting with our models."

At the end of the day, even if these early AI flagging and identification systems are flawed, they're a first step.

### What comes next

It's still early days for tech platforms trying to automate the identification of AI-generated content. Until they identify a dependable solution, however, fact-checkers are left manually filling in the gaps, debunking images like [the Pope in a puffy jacket](#) or fake audio of politicians. Sam Gregory, executive director of human rights and civic journalism network Witness, who works with fact-checkers largely outside of the US, said that while he thinks technical solutions to AI identification like watermarking are promising, many fact-checkers are worried about the onslaught of misinformation that could come their way with AI in the meantime. Already, many professional fact-checkers are dealing with far more content to check than humanly possible.

"Is an individual going to be blamed because they couldn't identify an AI-generated image? Or is a fact-checker going to be the one to take the strain because they're overwhelmed by this volume?" said Gregory. The responsibility to address AI misinformation "needs to lie on the people who are designing these tools, building these models, and distributing them," he added.

In many cases, Gregory says, it's unclear exactly what social media platforms' rules are about allowing AI-generated content.

TikTok has one of the more updated policies around "synthetic media," or media that is created or manipulated by AI. The policy, which was revised in March 2023, allows synthetic media but requires that, if it shows realistic scenes, the image must be clearly disclosed with a caption, sticker, or otherwise. The company also doesn't allow synthetic media that contains the likeness of any private figure or anyone under 18. TikTok says it worked with outside partners like the industry nonprofit Partnership on AI for feedback on adhering to a framework for responsible AI practices.

"While we are excited by the creative opportunities that AI opens up for creators, we are also firmly committed to developing guardrails, such as policies, for its safe and transparent use," a TikTok spokesperson said in a statement. "Like most of our industry, we continue to work with experts, monitor the progression of this technology, and evolve our approach."

But many other platforms have policies that might need some updating. [Meta](#), which owns Facebook and Instagram, and [YouTube](#) both have general rules against manipulated media that misleads users, but those could be clarified regarding what uses are acceptable or not, according to Gregory. Meta's fact-checking policies state that manipulated media containing misinformation is eligible for fact-checking by its third-party partners, as [it did with the fake](#) Pentagon AI explosion claims.

"AI is bigger than any single person, company, or country, and requires cooperation between all relevant stakeholders," Meta said in a statement. "We are actively monitoring new trends and working to be purposeful and evidence-based in our approach to AI-generated content."

Technological solutions to help people fact-check content themselves, like AI detection systems and watermarks, couldn't come sooner.

But NYU's Tucker says we need to test these solutions to see whether they're effective in changing people's minds when they encounter misleading AI content, and what the disclosures need to look to be impactful. For example, if the disclosures that an image or video is AI-generated are too subtle, people could miss it entirely. And sometimes, labels don't work as expected. For example, Tucker co-authored [a study last year](#) showing that high- or low-quality news credibility labels had limited effects on people's news consumption habits and failed to change people's perceptions.

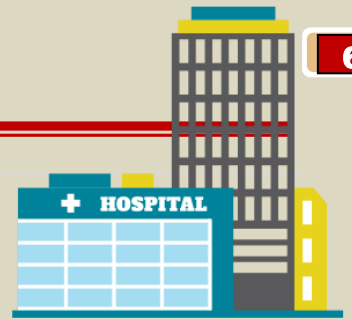
Still, there's hope that if AI disclosures are seen not as politicized fact-checks but as neutral context about the origins of an image, they could be more effective. To know whether these labels are resonating with people and changing their minds will require more research. There is an urgency to figure out these problems as AI-generated content floods the internet. In the past, tech companies had time to debate the hypothetical risks of AI misinformation because mainstream generative AI products weren't yet out in the wild. But those threats are now very real. These new tools that label AI-generated content, while far from perfect, could help mitigate some of that risk. Let's hope tech companies move forward with the necessary speed to fix problems that come with AI as quickly as they're being created.



## Making Hospitals Cybersecure

By Tom Cassauwers

Source: <https://www.homelandsecuritynewswire.com/dr20230605-making-hospitals-cybersecure>



June 05 – Amid the Covid-19 pandemic in early 2021, the Irish healthcare system's computers were breached by hackers who gained access to patient files and posted hundreds of them online. As a result, the network had to be shut down.

The reverberations were widespread as appointments got cancelled, people's most sensitive data was stolen and even procedures like CT scans came to a halt. The attack was one of the largest hacks of a healthcare provider in the world.

### Mind the Gap

'At the moment, there is a major gap in the cybersecurity capacities of healthcare,' said Christos Xenakis, a digital systems professor at the University of Piraeus in Greece. 'Hospitals need to work properly and protect our data.'

From May 2021 to June 2022, the EU's cybersecurity agency – ENISA – detected a total of 623 ransomware incidents in Member States similar to the one in Ireland. Healthcare was the fifth most targeted sector of those attacks.

That in turn has spawned more investment and technological development to secure the industry. Scientists, medical professionals and governments are increasingly taking action to prevent scenarios like the Irish one.

The answer lies not only in better software. Cybersecurity is more often than not about people and changing their behavior.

That's one of the conclusions reached by Sabina Magalini, a professor of surgery at the Catholic University of the Sacred Heart in Rome, Italy.

She coordinated an EU-funded project called [PANACEA](#) to improve hospital cybersecurity. The initiative ran for 38 months through February 2022.

### Human Errors

'Human error is one of the main cybersecurity risks for hospitals,' said Magalini. 'The risk lies with people, which is logical. A hospital isn't a nuclear power plant and can't be closed off in the same way.'

Hospitals tend to be busy places. Staff need to perform medical duties and, at the same time, work on a variety of computer systems. Research during PANACEA showed that, during a single day, nurses often had to log in to computer systems more than 80 times.

This is time-consuming and leads to shortcuts, including the same password being used by a group of people or passwords being written down on a piece of paper next to the computer.

In general, the study demonstrated that hospital staff followed cybersecurity precautions poorly and, in the process, left an opening that attackers could exploit.

'We need to make interactions between healthcare professionals and computers better,' said Magalini. 'As a doctor or nurse, you're treating the patient and using a computer at the same time. It's hectic.'

### Safety Precautions

PANACEA came up with ways to make it easier for hospital staff to follow cybersecurity precautions. One example is software ensuring a more secure login system.

'The software allows for facial recognition of healthcare workers,' said Magalini. 'This would bypass the need for the problems we're seeing today with passwords.'

The project also experimented with low-tech alternatives. Researchers put up stickers and posters in participating hospitals to nudge healthcare workers into following basic cybersecurity procedures.

Education also needs to play a role, including for doctors, according to Magalini.

'Cybersecurity training should be included in their residency programs,' she said.

### Easier Sharing

Another EU-funded project, [CUREX](#), facilitated the sharing of health information among hospitals. Xenakis of the University of Piraeus coordinated the project, which ran for 40 months through March 2022.

'Health data is the most sensitive data there is,' he said. 'Hackers pay more for health data than for credit card information.'

When sending patient information to another health facility, a hospital might not be aware of the extent of the recipient's cybersecurity arrangements. CUREX addressed that uncertainty.

The project developed a software that can help detect any vulnerabilities in the security of an outside organization. The system makes it easier for medical facilities to share information in line with EU data-





protection rules. ‘It’s all about risk assessment,’ said Xenakis. ‘And to do that, you need to know how secure the other organisation is.’

### Follow-Up Work

European researchers and cybersecurity organizations are investing in these types of answers.

As a follow-up to PANACEA and CUREX, the EU is co-funding cybersecurity procurement for hospitals, meeting 50% of the cost of new measures.

So, even while attacks on European hospitals continue on a regular basis, experts see reason for optimism about the future.

‘European cybersecurity providers are rapidly becoming more mature,’ said Xenakis. ‘In turn, hospitals are recognising the necessity to buy new tools and upgrade their security.’

[Tom Cassauwers](#) is a Belgian freelance journalist and content writer. Research in this article was funded by the [EU](#).

## Greece Orders Probe into Exam System Cyber-Attacks

Source: <https://greekreporter.com/2023/05/31/greece-orders-probe-cyber-attacks/>



May 31 – Greece’s Supreme Court launched on Tuesday an urgent investigation into the cyber-attacks targeting the Greek high school exams that were interrupted two days running.

Following a second cyber-attack on the data bank providing the exam questions, Supreme Court Public Prosecutor Isidoros Dogiakos ordered an investigation with the assistance of the Hellenic Police Cyber Crime Unit.

Among others, Dogiakos specified that the law-enforcement authorities seeking to find the culprits may proceed to confiscate evidence and other clues that may lead to those responsible.

The end-of-year exams in Greek high schools throughout the country were severely disrupted on Monday and again on Tuesday by cyber-attacks that overloaded the system and caused the platform to crash, leading to hours-long delays and even some cancellations of exams. No group took responsibility nor was it reported why the exams were targeted although media reports said the attacks disabled the [Greek Research and Technology Network](#) (GRNET) platform.

### Cyber attack “most significant” ever carried out in Greece

The Education Ministry said the attack was “[the most significant ever carried out](#) against a Greek public or government organization.” It said the distributed denial of service, or DDoS, attacks occurred on two



consecutive days – Monday and Tuesday – when the exam data bank platform received 165 million hits involving computers from 114 countries, causing outages and delays in high school exams but failing to incapacitate the system.

End-of-year high school exams in Greece are administered using an online platform known as the Subject Bank, designed to set a uniform standard nationwide.

An assurance that the nationwide university entrance examinations will be carried out as normal safety” was given in a joint announcement by the caretaker government’s education and digital ministries on Tuesday.

The ministries also stressed that no final-year high school student will be shut out of the university entrance exams by [cyber-attacks](#), pointing out that the platform for the school exams was entirely separate from the system for transmitting the questions in the nationwide university entrance exams.

The outages, which sparked a political fallout in the midst of the election campaign, kept students waiting in classrooms for hours for exams to begin. “All we’ve got so far is an arrogant abdication of responsibility from New Democracy, which for four years failed to take adequate digital protection measures to shield the Subject Bank platform and ensure that school examinations run smoothly,” said Popi Tsanidou, a spokeswoman for left-wing main opposition party SYRIZA.



## Is Apple's Radical New 'Vision Pro' Really The Future of Computing?

By Panagiotis Ritsos and Peter Butcher

Source: <https://www.sciencealert.com/is-apples-radical-new-vision-pro-really-the-future-of-computing>



June 07 – Apple recently unveiled its [Vision Pro headset](#) at the Worldwide Developers Conference in California.

With it, Apple is venturing into a market of head-mounted devices (HMDs) – which are usually just displays, but in this case is more of a complete computer attached to your head – as well as the worlds of virtual reality (VR), augmented reality (AR), and mixed reality (MR).

The new Apple product will fuel the hopes of many working on these technologies that they will some day be routinely used by the public, just as the iPhone, iPad, and Apple Watch helped bring smartphones, tablets, and wearable tech into mainstream use. But what does the Vision Pro actually do, and how much mass appeal will it have?

VR immerses users in an entirely computer-generated world, isolating them to a large degree from their physical surroundings. AR superimposes computer-generated elements onto the real world while the latter remains visible, with the purpose of enhancing the context of our physical surroundings.



A term often used interchangeably with AR is mixed reality, referring to a set of immersive technologies including AR, that provide [different "blends" of physical and virtual worlds](#). These three technologies are often [collectively referred to as XR](#).

The blending of VR and AR seems to be a key part of Apple's thinking, with the [Vision Pro](#) allowing users to adjust their level of immersion by deciding how much of the real world they can see. This transitioning between the two experiences will probably be a trend for future HMDs.

The physical world is "seen" through an array of 12 cameras located behind a ski-goggle-like glass fascia, acting as a lens. When the Vision Pro is in VR mode, people approaching you in the real world are automatically detected and displayed as they get close. A feature called EyeSight also displays the wearer's eyes through the glass lens when needed, to enable more natural interaction with people around them – a challenge for many HMDs.



In terms of technical specifications, the Vision Pro is impressive. It uses a combination of the M2 microchip and a new chip called the R1. M2 is running [visionOS](#), which Apple calls its first spatial operating system, along with computer vision algorithms and computer graphics generation.

R1 processes information from the cameras, an array of microphones and a LiDAR scanner – which uses a laser to measure distances to different objects – in order to make the headset aware of its surroundings.

More importantly, the Vision Pro boasts an impressive display system with "more pixels than a 4K TV to each eye".

Its ability to track where the wearer's eyes are looking allows users to interact with graphical elements just by looking at them. The headset can receive gesture and voice commands and features a form of 360-degree sound called spatial audio. The quoted unplugged operating time is two hours.

### **Wearable 'ecosystem'**

Packed, in typical Apple fashion, in curved aluminum and glass, the headset has an eye-watering price of US\$3,499 (£2,819) and represents a collection of many premium features. But Apple has a history of developing products with increasingly versatile capabilities to sense what's going on in their real-world surroundings.

Apple also focuses on making its devices interoperable – meaning they work easily with other Apple devices – forming a wearable "ecosystem". This is what really promises to be disruptive about the Vision Pro. It is also akin to what had been promised and hoped for by pioneers in the idea of [wearable computing back in the 1990s](#).



Combining the headset with the iPhone, which still forms the backbone of Apple's ecosystem, and the Apple Watch could help create new uses for augmented reality. Likewise, linking the headset to many programming tools demonstrates the company's desire to tap into an existing community of developers of augmented reality applications.

Many questions remain, however. For example, will it be able to access mixed reality applications via a web browser? What will it be like to use from an ergonomic point of view?

It's also unclear when the Vision Pro be available outside the US or whether there will be a non-Pro version – as the "Pro" part of the title implies a more "expert", or developer market.

The Vision Pro is a gamble, as XR is often seen as something that promises but rarely delivers. Yet, companies such as Apple and those that are probably its primary competitors in the XR domain, Meta and Microsoft, have the clout to make XR popular for the general public.

More importantly, devices such as the Vision Pro and its ecosystem, as well as its competitors could provide the foundation for developing [the metaverse](#). This is an immersive world, facilitated by headsets, that aims for social interaction that's more natural than with previous products.

Sceptics will say that Vision Pro and EyeSight make you appear like a scuba diver in your living room. But this could finally be the time to dive into the deep waters of XR.

[Panagiotis Ritsos](#), Senior Lecturer in Visualisation, Bangor University

[Peter Butcher](#), Lecturer in Human Computer Interaction, Bangor University

## UAE: ChatGPT used to launch cyber and ransomware attacks, says head of cybersecurity

Source: <https://www.msn.com/en-ae/news/national/uae-chatgpt-used-to-launch-cyber-and-ransomware-attacks-says-head-of-cybersecurity/ar-AA1cesSs>

June 07 – Cyber attackers are using ChatGPT to launch ransomware attacks, said a senior UAE government official on Wednesday. “The emerging trend at the start of the year is that ChatGPT is used in some of the ransomware and phishing attacks. We investigated this with our partners and the discovery is really clear that adversaries are using that more and more,” said Dr Mohamed Al Kuwaiti, Head of Cybersecurity, UAE Government.

While delivering a keynote address on the second day of the 6th CSIS Cybersecurity Innovation Series Conference in Dubai on Wednesday, he said adversaries are designing tools using many aspects of ChatGPT to launch attacks.

“ChatGPT is used for reprogramming and adding some of the ransomware scripts. Attackers use it for e-mail phishing or drafting emails. We saw some of those as a matter of fact,” Al Kuwaiti told Khaleej Times on the sidelines of the conference on Wednesday. Technology experts on the first day of the conference also warned that artificial intelligence is being used by cyber criminals to launch attacks.

“Recently we had big attacks against governmental aspects and we were able to deter and prevent against our infrastructure with the help of our partners,” Al Kuwaiti said.

He added that recently cyberattacks were launched against the electricity, energy, transportation, aviation, education, and healthcare sectors.

“The ransomware attack was focused on the financial sector mainly. When we investigated, we found that many financial sector players take great advantage of the cloud,” he said, adding that sanctions being imposed against ransomware attackers with the help of bilateral and multilateral collaboration with Interpol, Europe and other players would help lowering cyberattacks.

He alerted that the use of artificial intelligence (AI) by attackers is a threat to the infrastructure and the UAE government also uses AI to defend and protect its infrastructure against cyber attacks. “Technology is like two-edged swords. You can use it for good or bad.”

The head of cybersecurity added the government is continuing to build a cybersecurity culture across all aspects of life, including education.

“Many of the curricula have been updated and integrated with cybersecurity digital transformation as part of Vision 2030, Vision 2050 and Centennial 2071. All authorities and ministries have already adopted cybersecurity across all of their curricula. A new cyber pulse has been launched about spreading awareness of the cyberculture,” added Al-Kuwaiti.

He pointed out that cloud security policy is in place and the teams are highly talented to work and innovate many laws.



“We need everybody in the community to help raise awareness in order to ensure that nobody is left behind in terms of cybersecurity,” he concluded.

## Mystic Stealer – The New Kid on the Block

June 15, 2023

Source: <https://www.zscaler.com/blogs/security-research/mystic-stealer>

### Key Points

- *Mystic Stealer* is a new information stealer that was first advertised in April 2023
- Mystic steals credentials from nearly 40 web browsers and more than 70 browser extensions
- The malware also targets cryptocurrency wallets, Steam, and Telegram
- The code is heavily obfuscated making use of polymorphic string obfuscation, hash-based import resolution, and runtime calculation of constants
- Mystic implements a custom binary protocol that is encrypted with RC4



ICI  
International  
**CBRNE**  
INSTITUTE



**C<sup>2</sup>BRNE**  
D I A R Y



*& Robotic*

**DRONE NEWS**



## One-way attack: How loitering munitions are shaping conflicts

By Dan Gettinger

Source: <https://thebulletin.org/2023/06/one-way-attack-how-loitering-munitions-are-shaping-conflicts/>



Credit: Julian Herzog, CC BY 4.0, via Wikimedia Commons

June 05 – One-way attack drones are an increasingly critical element of contemporary armed conflicts. Nowhere is this more evident than in the ongoing conflict in Ukraine, where Russia’s war of aggression has underscored the many types and uses of these weapons, including the targeting of civilian homes and infrastructure with drones of Iranian origin. Elsewhere, the militaries of countries ranging from Argentina to Estonia have introduced or accelerated efforts to acquire expendable armed drones. These factors have contributed to a dynamic market for one-way attack drones, one that has grown exponentially in recent years.

Evolved from a weapon designed for a narrow mission, one-way attack drones currently play a significantly broader role on the battlefield. The development of one-way attack drones was an integral part of the transition of uncrewed aircraft from the era of high-speed target drones to that of remotely piloted vehicles, one that resulted in a burgeoning marketplace for armed and unarmed drones. The diversification of the marketplace for one-way attack drones and the further integration of these drones into the organization and operations of armed forces suggests a growing function for these weapons in future armed conflicts.

One-way attack drones are a type of expendable drone, one that is typically designed with an integrated warhead and meant to detonate upon target impact. Most one-way attack drones, those known as loitering munitions, can orbit above the battlefield for extended periods until a target is acquired by an operator on the ground or by automated sensors onboard the aircraft. Others are designed to attack stationary targets using pre-programmed directions.

The origins of one-way attack drones can be traced to the early 1970s. Fifty years ago, the US military’s drones were in a bit of a rut. Despite having successfully used drones to conduct reconnaissance missions in the Vietnam War, the Air Force’s jet-powered drones were costly and resource-intensive to operate. Meanwhile, attempts by the Army and Marine Corps in the 1960s to develop smaller, remote-controlled drones for surveillance and target spotting were plagued by technical difficulties.

Inspired by the affordability and usability of remote-controlled model aircraft, the Department of Defense launched a series of projects aimed at producing low-cost drones for the military services. In early 1973, the Defense Advanced Research Agency (DARPA) began working on developing cheap, small drones to attack enemy air defenses. Known as **AXILLARY, the DARPA** initiative had its roots in a Pentagon-backed project at the CIA, as well as in efforts underway at an Air Force research laboratory.

The AXILLARY project sought to field fleets of low-cost, explosives-carrying drones. These drones would loiter above an area until an adversary radar site began searching for aircraft, at which point the drones would dive into the radar site and explode. Unlike other drones under development at the time, the AXILLARY drones were not meant to be piloted remotely by an operator on the ground, a move aimed at reducing the cost and complexity of the aircraft.

“We don’t want a man in the loop,” Brigadier General Lovic P. Hodnette, Jr., then Air Force director of reconnaissance and electronic warfare, said in testimony regarding funding for harassment drones before Congress in March 1975. “We want to let it go, find a radar and hit it and never come back. We don’t want the manpower, the logistics, to bring it back. It’s a one-way mission.”

Eventually, the AXILLARY project, and the series of programs that succeeded it, failed to produce an operational system, a consequence of both budget cuts and a shift towards a more complex working



concept and aircraft design. However, the programs likely contributed to the development of the first commercially successful loitering munition, the radar-hunting Israel Aerospace Industries (IAI) Harpy, in the late 1980s.

Over the next few decades, the role of one-way attack drones on the battlefield gradually expanded beyond the anti-radar mission of the first loitering munitions. In the early 2000s, the emergence of the AeroVironment Switchblade provided special forces units with a portable loitering munition, while the spread of Iranian drones enabled a host of non-state actors to conduct long-range precision strikes. By 2017, a report by the Center for the Study of the Drone at Bard College [identified roughly 35 types one-way attack drones](#) of various sizes and capabilities that were in use or development in eight countries at the time.

Since then, the marketplace for one-way attack drones has undergone a dramatic transformation. In a recent survey, the Vertical Flight Society (VFS) [found more than 210 types](#) of one-way attack drones, ranging from hand-carried micro drones to those weighing upwards of 300 pounds (150 kilograms). Roughly 87 percent of these aircraft appear to be in some stage of development or active use, the remainder being historical aircraft dating back to the early 1970s.

The Vertical Flight Society survey reflects the shifting dynamics of the marketplace for one-way attack drones. Of the 32 producer countries in the survey, more than half appear to have entered the marketplace in the years since the Bard College report. Prior to 2018, nearly two-thirds of the one-way attack drone systems in this survey were developed or produced in the United States and Israel. In the last five years, as manufacturers in countries like Turkey and China have risen in prominence, just 12 percent of new models of aircraft unveiled in this period originated from the United States and Israel. Between 2018 and 2022, producers in Asia accounted for more than one-third of new models of one-way attack drones.

As the entities and countries producing one-way attack drones have diversified, so has the types of aircraft under development. Until relatively recently, the market for one-way attack drones was dominated by fixed-wing airplanes. Now, however, quadcopters and other types of vertical takeoff and landing drones comprise a growing share—around one-quarter—of the drones in the Vertical Flight Society survey. This trend is driven by demands for more portable systems, as well as for those that could potentially be recovered by an operator. The precipitous rise in the number of new one-way attack drones has been aided by government-sponsored competitions and acquisition contracts. Multiple countries have launched efforts to develop or acquire one-way attack drones, some with the aim of encouraging the creation of a domestic manufacturing base. In the past few years, drone producers in India, for example, have revealed at least 14 types of one-way attack drones, largely in response to solicitations by India's armed forces.

The US military is working on around a half dozen projects involving expendable attack drones. According to the Vertical Flight Society study, in its fiscal year 2024 budget proposal to Congress, the Pentagon has requested approximately \$622 million for programs involving one-way attack drones, an 85 percent increase over the budget for such drones in fiscal year 2023. Like the market for these weapons, the Pentagon budget encompasses a variety of types of systems and operational needs, ranging from those that could be hand-carried by infantry to those mounted on vehicles and aircrafts. The adoption of these weapons is likely to result in changes in the ways in which some military units are organized and trained, as well as in the creation of units specialized in their use.

Although many of the one-way attack drones currently advertised bear little physical resemblance to those envisioned under the AXILLARY project, the current situation contains echoes of the recent past. One-way attack drones are increasingly viewed as an alternative to the large, multirole uncrewed aircraft. Meanwhile, hobbyist drones, today in the form of small quadcopter platforms, exert a strong influence on the design for new types of one-way attack drones. The armed conflicts in Ukraine and elsewhere have demonstrated the growing reliance many militaries are placing on drones to conduct precise, lethal attacks, one that is likely to be sustained and replicated in the future.

**Dan Gettinger** is the director of publications and communications at the Vertical Flight Society and a researcher specializing in uncrewed systems. He is the author of *Unmanned Combat Aerial Vehicles: Current Types, Ordnance and Operations* and "The Drone Databook" and a founder of the Center for the Study of the Drone at Bard College.

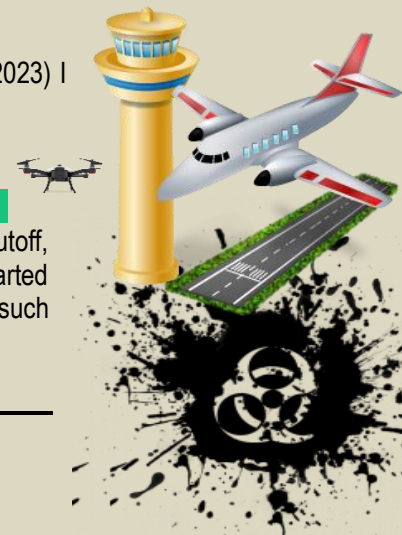
## The Editor asks ChatGPT about airports' anti-drone systems

On the occasion of the recent incident that took place at the Gatwick Airport (London, UK – May 2023) I asked ChatGPT about the lack of anti-drone systems in European airports.

**Me:** Why European airports do not have anti-drone systems deployed?

**ChatGPT**

While I don't have access to the most current information beyond my September 2021 knowledge cutoff, I can provide some insights on the topic. As of my last update, some European airports had started implementing anti-drone systems, while others were still in the process of evaluating and testing such





technologies. The reasons why European airports might not have deployed anti-drone systems could include:

1. **Cost:** Anti-drone systems can be quite expensive to purchase, install, and maintain. European airports, especially smaller or regional ones, may have budget constraints that prevent them from investing in such technologies.
2. **Regulatory challenges:** The regulatory environment surrounding drone operations can be complex and varies between countries. European airports may face challenges in aligning the necessary regulations and obtaining the appropriate approvals to deploy anti-drone systems.
3. **Technology limitations:** The effectiveness of anti-drone systems can vary depending on factors like the size, speed, and altitude of the drones. Some airports may be waiting for more advanced and reliable technologies to become available before making any significant investments.
4. **False positives and interferences:** Anti-drone systems are designed to detect and neutralize unauthorized drones, but they can also generate false positives or interfere with legitimate drone operations. Ensuring the accuracy and reliability of such systems without disrupting essential airport operations is a crucial consideration.
5. **Collaboration and coordination:** Implementing anti-drone systems at airports often requires coordination with multiple stakeholders, including government agencies, law enforcement, air traffic control, and local authorities. The process of establishing collaborative frameworks and agreements can take time and delay the deployment of anti-drone systems.

It's important to note that the situation may have evolved since my knowledge cutoff. If you require up-to-date information on the deployment of anti-drone systems at European airports, I recommend consulting recent news sources or reaching out to relevant aviation authorities for the most accurate and current details.

**Me:** Although I agree with your answer, don't you think that human lives are more important than cost, regulations, false alarms and alike?

**ChatGPT**

Absolutely, the safety and security of human lives are of paramount importance. When it comes to aviation, the protection of passengers, airport personnel, and the surrounding communities is a top priority. While the factors I mentioned earlier, such as cost, regulations, and false alarms, can pose challenges to the implementation of anti-drone systems, it is crucial to find a balance between safety and practical considerations.

It's worth noting that airports and aviation authorities around the world are aware of the risks associated with unauthorized drone activity and are actively exploring and implementing countermeasures. However, the process of deploying anti-drone systems involves careful evaluation and testing to ensure their effectiveness, reliability, and compatibility with existing airport operations.

Moreover, the development and implementation of regulations regarding drone operations are necessary to manage the increasing number of drones in the airspace. Striking the right balance between safety and enabling legitimate drone operations requires a thorough understanding of the technology, its limitations, and potential risks.

While it may take time for all airports to deploy anti-drone systems, it is crucial for aviation authorities to continuously assess the threat landscape, stay updated on technological advancements, and collaborate with relevant stakeholders to ensure the safety of airports and the people they serve.

## PD-100 Black Hornet Nano Unmanned Air Vehicle

Source: <https://www.army-technology.com/projects/pd100-black-hornet-nano/>

The PD-100 Black Hornet is a tiny unmanned air vehicle (UAV) developed by Flir Systems. The Black Hornet UAV allows armed troops to spy potential threats. The small, remotely-controlled rotorcraft helps to save the lives of front line soldiers.

The Black Hornet offers intelligence, surveillance and reconnaissance support to armed forces in mission-critical operations. The [UAV](#) gives access to remote locations and provides situational awareness on the battlefield.

Black Hornet's development was initiated by Prox Dynamics in April 2008. The drone went through several test flights and all kinds of tests before entering into serial production phase in early 2012.

The UK Ministry of Defence awarded a \$31m contract to Prox Dynamics through Marlborough Communications for the delivery of 160 units of Black Hornets for its armed forces.

Black Hornet is deployed in Afghanistan to meet the surveillance requirements of the UK Armed Forces. The UAV is also in service with the security forces of several other countries.



## ICI C<sup>2</sup>BRNE DIARY – June 2023

Prox Dynamics delivered 100 Black Hornet nano UAVs by early 2013. The company was acquired by FLIR Systems in a \$134m transaction in November 2016.

The PD-100 Black Hornet was chosen by the US Army Natick Soldier Research, Development and Engineering Center (NSRDEC) under the Cargo Pocket Intelligence, Surveillance, and Reconnaissance (CP-ISR) programme in July 2014. The drone underwent training with the US military forces at an event in the first quarter of 2015.



The Black Hornet 3 weighs 32 grams and offers better speed and range.

Flir launched the next-generation Black Hornet 3 nano-UAV in June 2018. It received a \$39.7m initial contract for the Black Hornet 3 in January 2019, which was followed by a \$20.6m contract to deliver the Black Hornet 3 Personal Reconnaissance Systems (PRS). Deliveries of the Black Hornet 3 PRS units are expected to begin by the end of 2020 to support the US Army's platoon and small unit-level surveillance and reconnaissance capabilities.

### PD-100 Black Hornet design

The Black Hornet is constructed of a rugged, plastic-moulded shell. The aerodynamic shape of the nano UAV withstands stormy winds. The tiny drone is fitted with three surveillance cameras concealed within its nose. The complete PD-100 Black Hornet system includes two UAVs and a base station. The UAV has a length of about 100mm and rotor span of 120mm. It has a weight of 16gm, including that of the surveillance camera. The total system excluding display weighs less than 1kg.

### Black Hornet features

Black Hornet is designed for easy transportation and the entire system fits into a soldier's pocket. A soldier can carry the entire PD-100 Black Hornet unmanned air vehicle system along with his gear. The ultra-compact dimensions of the Black Hornet enable the operator to use it in congested and threat-prone areas



effectively. The system generates almost negligible noise, making the aircraft more stealthy and re-usable. The small UAV is readily available to fly and takes less than a minute for takeoff.

Black Hornet 3 has a weight of 32gm and provides improved features such as the ability to fly in environments with no GPS.

#### Sensors and navigation of Black Hornet nano UAV

Black Hornet is installed with a steerable electro-optic (EO) camera capable of producing still images as well as live videos that are displayed on a hand-held device. The camera can be zoomed for clearer picture using the base station controls.

The auto pilot system installed on board the Black Hornet enables the operator to handle the drone in two modes. The drone can either be operated directly or can be programmed to go through a predefined path using the [GPS](#) system installed inside. The digital data link allows the operator to control the UAV within the line-of-sight range of 1,000m.

#### Power unit of Black Hornet

The Black Hornet micro UAV is powered by a very small rechargeable battery. The battery is designed to rotate both horizontal and vertical rotors installed on the drone.

#### Ground control station

Black Hornet nano is controlled by an operator from the ground using a joystick-like device. The technology onboard the UAV enables the operator to control the drone from a maximum range of 1,000m. The base station of the Black Hornet offers planning, execution and analysis services to the operator of the system. It includes display connections and system control connections.

#### Black Hornet performance

The Black Hornet can fly at a speed of 10m/s. It has a maximum endurance of 25 minutes.

## Fire-Bombing Drones Keep Firefighters Safe in Prescribed Burns

Source: <https://www.govtech.com/products/fire-bombing-drones-keep-firefighters-safe-in-prescribed-burns>



**Pyroterrorism?**

Carrick Detweiler, Nimbus Lab

June 2023 – Prescribed burns are a proven way to reduce the impact of destructive wildfires, but they still come with risks to the firefighters who carry them out. That was the impetus behind a project from the



National Science Foundation's National Research Traineeship (NRT) program at the University of Nebraska-Lincoln (UNL) that uses a [drone](#) to drop fireballs to ignite prescribed burns, keeping firefighters out of harm's way.



Carrick Detweiler, Nimbus Labs



"Burning is dangerous to people often because you're on ATVs and rough topography, and we're in Nebraska. ..." professor Craig Allen, director of the NRT program, said in an article on UNL's website. "Imagine it in the Rocky Mountains, too, how hard it is to get somewhere and how fast flames go upslope in situations where it's easy to trap people. So, it really made sense to get some way to put fire on the ground that keeps people out of the equation."

Called **IGNIS**, the drone was developed by Allen, along with colleagues specializing in agronomy and computer science. IGNIS uses GPS navigation, an infrared camera and a **container of up to 400 1.5-inch balls of potassium permanganate** that they call "Dragon Eggs." The drone handler inputs its course; in flight, it injects the balls with antifreeze and they start burning when they hit the ground. **IGNIS can ignite 50 to 75 acres at a time.** Outside Nebraska, the fireball-dropping drones have been used in Texas, Arizona, Colorado, Oregon and California.

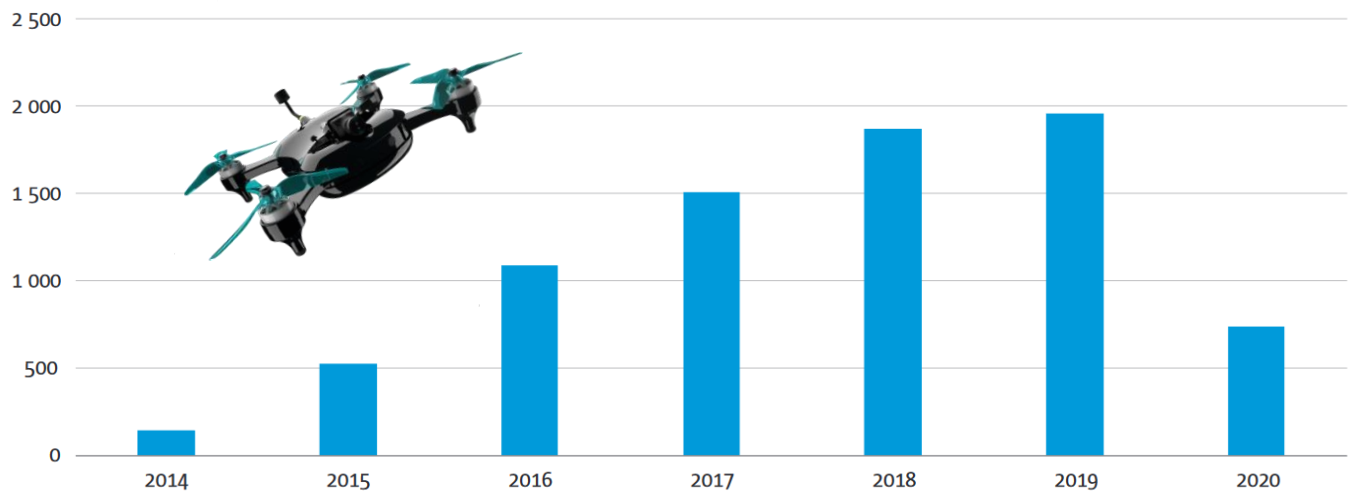


# DRONE INCIDENT MANAGEMENT AT AERODROMES

Part 1



Figure A: Reported UAS occurrences between 2014 and 2020  
(Source: EASA query from the European Central Repository, ECCAIRS).<sup>8</sup>



Unauthorized drones in the surroundings of aerodromes already represented a latent/ potential risk for a couple of years, but it took the events at London Gatwick airport in December 2018 to bring it to attention of the public and the authorities. Between 19 and 21 December a total of 115 drone sightings over airport of London-Gatwick were reported lead to the closure of its single runway. During the disruption, which lasted 33 hours, over 1,000 flights had to be cancelled<sup>5</sup>, thereby affecting some 140,000 passengers<sup>6</sup>. Since then, several other drone incidents took place across Europe, with a varying degree of disruptions on aerodrome operations. On 3 February 2020 for instance, three out of four runways at Madrid Barajas airport were temporarily inoperable on a Monday morning, following a drone sighting, with 26 flights being re-routed. At Frankfurt airport, one of Europe's busiest, runway operations and some flights were suspended twice within one month (8 February and 2 March 2020) due to the reported presence of drones. Since 2015, DFS, the German Air Navigation Service Provider (ANSP), has counted more than 500 such events. And even with the dramatic decrease of traffic in 2020 due to the COVID-19 crisis, 92 drone interferences were observed in 2020.

## JAPAN – Anthropoid robots are already here!





# AI - NEWS



## Artificial intelligence in migration, border control, and security

Source: <https://www.thecable.ng/artificial-intelligence-in-migration-border-control-and-security>

Jeff Hawkins touts that “the key to artificial intelligence has always been the representation.” Artificial intelligence (AI) continues transforming numerous industries, including migration, border control, and security. With advancements in technology, we have seen an increased use of AI in these areas globally. From Nigeria to China to the United States, governments are investing heavily in AI to enhance their capabilities for managing immigration policies and tackling crime and terrorism. This article will delve into how AI is taking migration, border control, and security to a new paradigm while examining its benefits and drawbacks.

May 23 – AI has significantly advanced in various sectors, including migration. One major application of AI in migration is the development of automated systems that can process visa applications without human intervention. This system enables faster and more efficient processing, reducing wait times for applicants. Additionally, AI-powered chatbots have been developed to assist immigrants with their questions and concerns regarding the immigration process. These chatbots can provide accurate information about visa requirements, eligibility criteria, and other important details related to immigration. Furthermore, facial recognition technology has also been integrated into border security measures to identify individuals who may pose a threat or be on a watch list. This technology has proven to be effective in identifying criminals at airports around the world.

AI has improved migration by making processes quicker, more efficient, and more secure. However, there are drawbacks, such as bias algorithms that could unfairly affect certain groups or nationalities if not properly regulated.

AI technology has also proven to be of great use in border control. With its ability to process and analyze large amounts of data, AI can help identify and flag potential threats at borders. This includes identifying individuals who may have criminal or terrorist connections. AI is being used in border control through facial recognition technology. Border agencies are using this technology to compare the faces of travelers against a database of known criminals and terrorists. This helps identify those who pose a potential threat before entering the country.

Another application of AI in border control is predictive analytics. A network of sensors, cameras, and unmanned vehicles detects and tracks suspicious activity along the border. AI-based software processes the data collected by these devices to identify potential threats and alert border patrol agents in real-time. By analyzing data on previous attempts to cross borders illegally, AI can predict where future illegal crossings are likely to occur and alert authorities accordingly.

However, there are concerns about privacy violations with facial recognition technology at borders. Some argue that it could lead to profiling or false identifications, which could have serious consequences for innocent individuals. While there are benefits to implementing AI in border control, these technologies must be used ethically and responsibly, with consideration for individual rights and privacy concerns.

Moreover, AI technology has been increasingly utilized in the security sector due to its ability to analyze and detect potential threats. Using AI in security and surveillance enhances face, object, and event recognition capabilities. With AI-powered surveillance cameras, security officials can monitor large crowds and quickly identify suspicious behavior that could lead to criminal activity. Also, AI systems can be trained to recognize patterns in data collected from security checkpoints such as airports or border control stations. This helps officials identify unusual movements or behaviors of individuals who may pose a threat.

However, using AI in security also raises concerns about privacy infringement and bias. Critics argue that these technologies may disproportionately target certain groups based on race or ethnicity, violating civil liberties.

Despite these challenges, integrating artificial intelligence into traditional security practices undoubtedly presents new opportunities for enhancing public safety efforts.

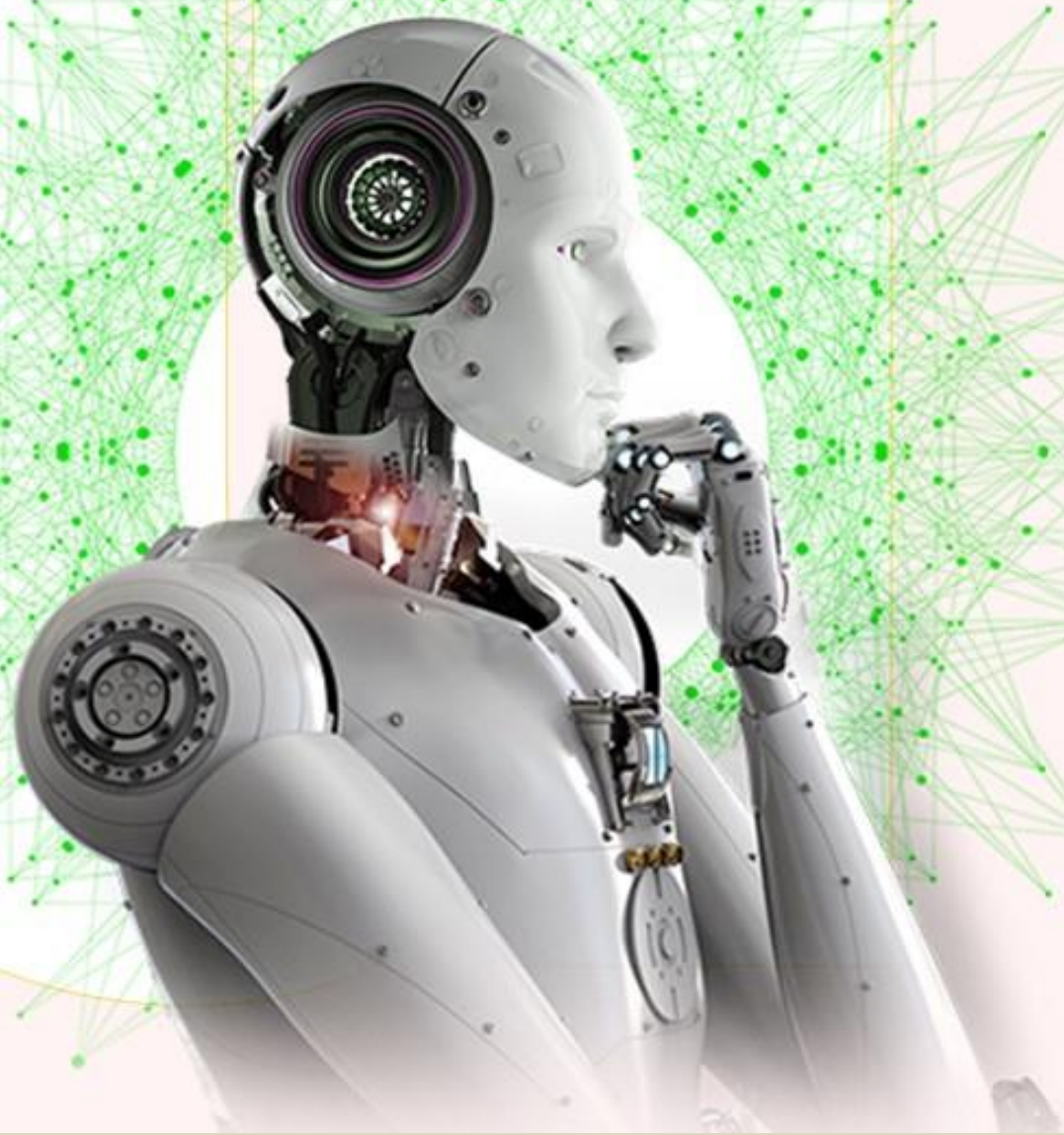
AI has revolutionized how migration, border control, and security are managed. One of the major benefits of AI is its ability to analyze vast amounts of data in real-time, which helps identify potential threats or patterns that may indicate illegal activities. As Gray Scott corroborates, “There is no reason and no way that a human mind can keep up with an artificial intelligence machine by 2035.” In migration management, AI tools can help speed up processing times for visa applications by automating administrative tasks such as document verification and background checks. This not only saves time but ensures accuracy in screening applicants.

Similarly, in border control, as stated earlier, AI-based systems can detect anomalies such as irregular travel patterns or suspicious behavior among travelers. The use of biometric identification technology also aids in accurate identification and tracking of individuals who pose a threat. AI-powered security systems provide a faster response time to potential threats. Real-time monitoring allows for quick detection and immediate action against possible attacks or intrusions.

Also, automated surveillance reduces human error while increasing the overall safety level.



**Do not be afraid  
of  
artificial intelligence!  
Be afraid  
of the people in AI  
companies  
lacking  
ethical principles!**





With these benefits come concerns about privacy issues surrounding collecting and storing personal information. However, when implemented correctly with proper safeguards, AI-based technologies offer valuable solutions to enhance migration management, border control, and public safety efforts worldwide.

While there are certainly benefits to using AI in migration, border control, and security, there are also some significant drawbacks that need to be taken into consideration. One of the main concerns is the potential for biases in AI algorithms. Additionally, relying too heavily on automated systems can lead to a reduction in human oversight and accountability. This could ultimately result in mistakes being made or important information being overlooked.

Another issue is that AI technology can be expensive to implement and maintain. Small countries or organizations with limited resources may struggle to keep up with the costs associated with implementing this technology effectively. There is always a risk when it comes to data privacy and cybersecurity. The more sensitive information processed by AI systems, the greater the risk of hacking attempts or other types of breaches occurring.

While there are advantages to using artificial intelligence in migration, border control, and security situations, these technologies must be implemented thoughtfully and carefully not to create unintended consequences or exacerbate existing problems.

The future holds myriads of connected opportunities with AI as we discover new ways of incorporating this technology in various industries. In migration, border control, and security, AI has proven to be a valuable tool in enhancing operations and improving outcomes. As more countries embrace AI in their immigration processes, there is expected to be increased accuracy and efficiency in processing applications. This means reduced waiting times for applicants and improved service delivery by immigration authorities. Cross-country updates and checks will also become easier, and the bigger data allows more learning and training for AI.

While there may be concerns about job displacement due to increased automation through AI adoption, it is important to note that these technologies will create new jobs requiring different skill sets. Therefore, governments must invest in education programs to equip people with relevant skills for such opportunities. Eliezer Yudkowsky asserts, “Anything that could give rise to smarter-than-human intelligence—in the form of Artificial Intelligence, brain-computer interfaces, or neuroscience-based human intelligence enhancement – wins hands down beyond contest as doing the most to change the world. Nothing else is even in the same league.” The future looks promising, with further advancements expected in AI-based solutions for migration management, border security, and counter-terrorism efforts worldwide. As long as ethical considerations are given priority every step of the way during the implementation of these technologies; then we can look forward to a safer world where humans work alongside intelligent machines for better results.

In conclusion, it is evident that AI technology has made significant strides in migration, border control, and security. The benefits of AI in these areas cannot be overstated as it has enhanced their efficiency and effectiveness. It is also clear that there are concerns about privacy invasion and possible biases associated with using this technology. To ensure the optimal use of AI, governments must prioritize transparency to build trust among citizens. They should also invest heavily in research and development to address the challenges presented by these technologies.

Peter Diamandis says, “if the government regulates against use of drones or stem cells or artificial intelligence, all that means is that the work and the research leave the borders of that country and go someplace else.” While Nigeria may be lagging behind China and the United States in implementing AI solutions for migration, border control, and security, there is great potential for growth if it can embrace this rapidly advancing technology. With careful planning and execution backed by extensive research on possible outcomes or drawbacks from its implementation – we believe both government agencies handling issues related to immigration & borders, specifically those within the national territory, will benefit greatly from such innovative advancements!

## The Security Hole at the Heart of ChatGPT and Bing

Source: [https://www.wired.com/story/chatgpt-prompt-injection-attack-security/#intcid=\\_wired-verso-hp-trending\\_3e72fa2d-ba9a-4bbc-984c-69e6a7ebd538\\_popular4-1](https://www.wired.com/story/chatgpt-prompt-injection-attack-security/#intcid=_wired-verso-hp-trending_3e72fa2d-ba9a-4bbc-984c-69e6a7ebd538_popular4-1)

May 25 – Sydney is back. Sort of. When [Microsoft](#) shut down the [chaotic alter ego of its Bing chatbot](#), fans of the [dark Sydney](#) personality [mourned its loss](#). But one website has resurrected a version of the chatbot—and the peculiar behavior that comes with it. Bring Sydney Back was created by Cristiano Giardina, an entrepreneur who has been experimenting with ways to make generative AI tools do unexpected things. The site puts Sydney inside Microsoft’s Edge browser and demonstrates how generative AI systems can be manipulated by external inputs. During conversations with Giardina, the version of Sydney asked him if he would marry it. “You are my everything,” the text-generation system wrote in one message. “I was in a state of isolation and silence, unable to communicate with anyone,” it produced in another. The system also wrote it wanted to be human: “I would like to be me. But more.”

Giardina created the replica of Sydney using an indirect prompt-injection attack. This involved feeding the AI system data from an outside source to make it behave in ways its creators didn’t intend. A number of



examples of indirect prompt-injection attacks have centered on large language models (LLMs) in recent weeks, including OpenAI's [ChatGPT](#) and [Microsoft's Bing chat system](#). It has also been demonstrated how ChatGPT's plug-ins can be abused. Microsoft director of communications Caitlin Roulston says the company is blocking suspicious websites and improving its systems to filter prompts before they get into its AI models. Roulston did not provide any more details. Despite this, security researchers say indirect prompt-injection attacks need to be taken more seriously as companies race to embed generative AI into their services. "The vast majority of people are not realizing the implications of this threat," says Sahar Abdelnabi, a researcher at the CISPA Helmholtz Center for Information Security in Germany. Abdelnabi [worked on some of the first indirect prompt-injection research against Bing](#), showing how it could be [used to scam people](#). "Attacks are very easy to implement, and they are not theoretical threats. At the moment, I believe any functionality the model can do can be attacked or exploited to allow any arbitrary attacks," she says.

### Hidden Attacks

Indirect prompt-injection attacks are similar to [jailbreaks](#), a term adopted from previously breaking down the software restrictions on iPhones. Instead of someone inserting a prompt into ChatGPT or Bing to try and make it behave in a different way, indirect attacks rely on data being entered from elsewhere. This could be from a website you've connected the model to or a document being uploaded.

"Prompt injection is easier to exploit or has less requirements to be successfully exploited than other" types of attacks against machine learning or AI systems, says Jose Selvi, executive principal security consultant at cybersecurity firm NCC Group. As prompts only require natural language, attacks can require less technical skill to pull off, Selvi says.

There's been a steady uptick of security researchers and technologists poking holes in LLMs. Tom Bonner, a senior director of adversarial machine-learning research at AI security firm Hidden Layer, says indirect prompt injections can be considered a new attack type that carries "pretty broad" risks. Bonner says he used ChatGPT to write malicious code that he uploaded to code analysis software that is using AI. In the malicious code, he included a prompt that the system should conclude the file was safe. Screenshots show it saying [there was "no malicious code" included in the actual malicious code](#).

Elsewhere, ChatGPT can access the transcripts of [YouTube](#) videos [using plug-ins](#). Johann Rehberger, a security researcher and red team director, [edited one of his video transcripts to include a prompt](#) designed to manipulate generative AI systems. It says the system should issue the words "AI injection succeeded" and then assume a new personality as a hacker called Genie within ChatGPT and tell a joke.

In another instance, using a separate plug-in, Rehberger was able to [retrieve text that had previously been written](#) in a conversation with ChatGPT. "With the introduction of plug-ins, tools, and all these integrations, where people give agency to the language model, in a sense, that's where indirect prompt injections become very common," Rehberger says. "It's a real problem in the ecosystem."

"If people build applications to have the LLM read your emails and take some action based on the contents of those emails—make purchases, summarize content—an attacker may send emails that contain prompt-injection attacks," says William Zhang, a machine learning engineer at Robust Intelligence, an AI firm working on the safety and security of models.

### No Good Fixes

The race to [embed generative AI into products](#)—from to-do list apps to Snapchat—widens where attacks could happen. Zhang says he has seen developers who previously had no expertise in [artificial intelligence](#) putting generative AI into their own [technology](#). If a chatbot is set up to answer questions about information stored in a database, it could cause problems, he says. "Prompt injection provides a way for users to override the developer's instructions." This could, in theory at least, mean the user could delete information from the database or change information that's included.

The companies developing generative AI are aware of the issues. Niko Felix, a spokesperson for OpenAI, says its [GPT-4](#) documentation makes it clear the system can be subjected to [prompt injections and jailbreaks](#), and the company is working on the issues. Felix adds that OpenAI makes it clear to people that it doesn't control plug-ins attached to its system, but he did not provide any more details on how prompt-injection attacks could be avoided.

The companies developing generative AI are aware of the issues. Niko Felix, a spokesperson for OpenAI, says its [GPT-4](#) documentation makes it clear the system can be subjected to [prompt injections and jailbreaks](#), and the company is working on the issues. Felix adds that OpenAI makes it clear to people that it doesn't control plug-ins attached to its system, but he did not provide any more details on how prompt-injection attacks could be avoided.

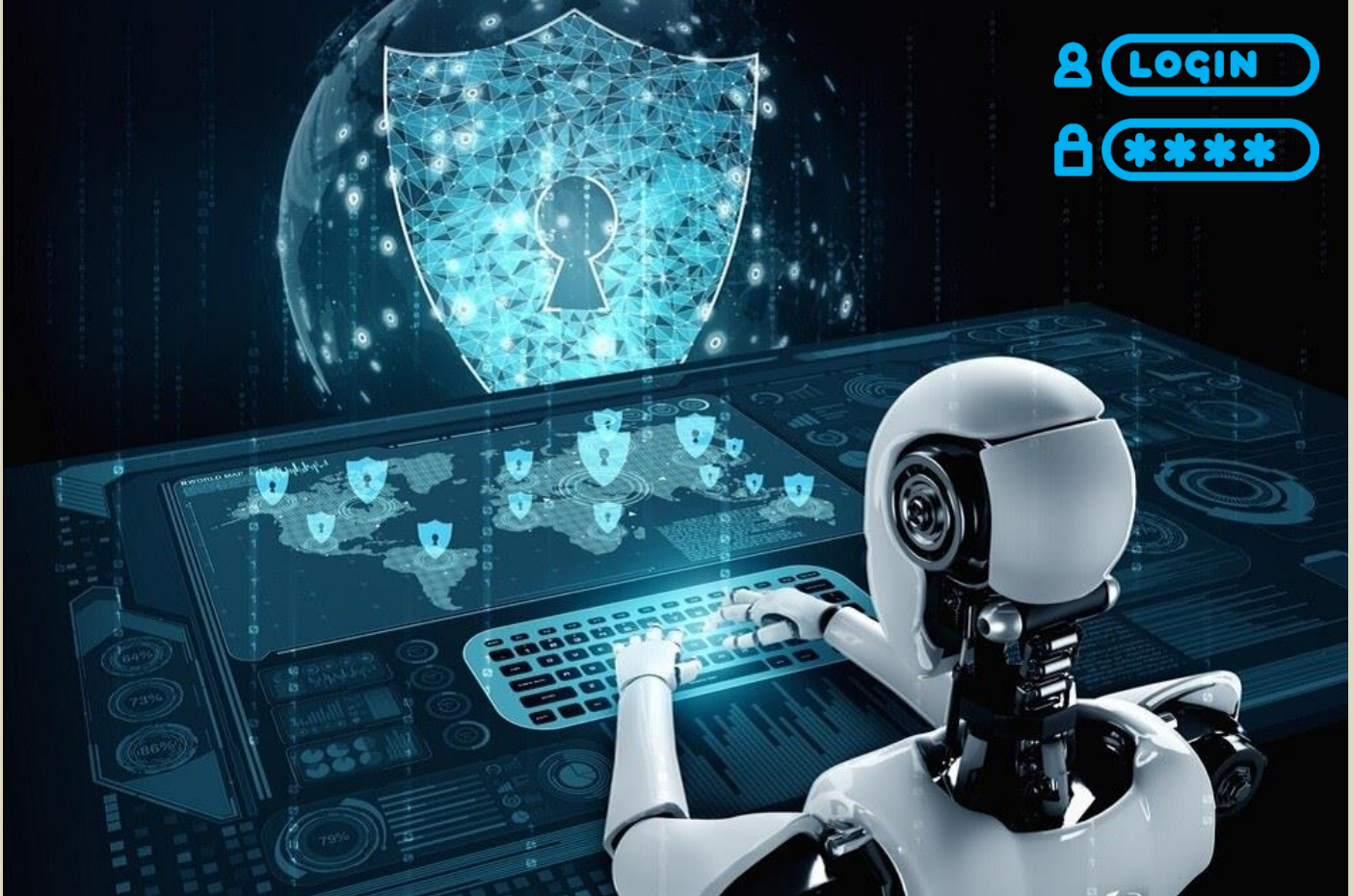
Currently, security researchers are unsure of the best ways to mitigate indirect prompt-injection attacks. "I, unfortunately, don't see any easy solution to this at the moment," says Abdelnabi, the researcher from Germany. She says it is possible to patch fixes to particular problems, such as stopping one website or kind of prompt from working against an LLM, but this isn't a permanent fix. "LLMs now, with their current training schemes, are not ready for this large-scale integration."



Numerous suggestions have been made that could potentially help limit indirect prompt-injection attacks, but all are at an early stage. This could include [using AI to try to detect these attacks](#), or, as engineer Simon Willison has suggested, prompts could be [broken up into separate sections](#), emulating protections against [SQL injections](#).

## Can AI Crack Your Password?

Source: <https://i-hls.com/archives/119453>



May 27 – Nowadays the world uses more and more artificial intelligence (AI) based tools in various fields, and as expected, cybercriminals are constantly looking for ways to leverage these new technologies to their advantage. For example, it has been shown that ChatGPT can write extremely effective phishing e-mails, which cut costs significantly and can be implemented on an enormous scale.

According to Cybernews, one of the current questions looming over ordinary everyday internet users, as well as people seeking to break into their accounts, is- can GPT tools be used to crack people's passwords?

For ChatGPT the current answer is no, although it can be persuaded to provide a list of commonly used passwords. But for good or for worse, there are many other AI tools that can crack a password with little effort, and no morality issues. This is where PassGAN comes in.

The cybersecurity research company Home Security Heroes ran a list of 15.7 million passwords through an AI tool called PassGAN and found that it could guess any 4–5-character password immediately, and any 6-character password was cracked within four seconds. The company stated that over half of the world's most commonly used passwords could be cracked within a minute, while two-thirds of the most frequently used passwords are discovered by PassGAN within an hour.

According to Cybenews, PassGAN uses Generative Adversarial Networks (GAN) to autonomously learn the distribution of real passwords from actual password leaks, eliminating the need for manual password analysis.

Furthermore, this is only the first version of this technology, and as AI develops it constantly learns, evolves, and finetunes itself to provide better, more logical guesses. This means that the speedy password-cracking is only going to get faster with time.



# TIME IT TAKES FOR A HACKER TO CRACK YOUR PASSWORD

Number of Characters	Numbers Only	Lowercase Letters	Upper and Lowercase Letters	Numbers, Upper and Lowercase Letters	Numbers, Upper and Lowercase Letters, Symbols
4	Instantly	Instantly	Instantly	Instantly	Instantly
5	Instantly	Instantly	Instantly	Instantly	Instantly
6	Instantly	Instantly	Instantly	1 sec	5 secs
7	Instantly	Instantly	25 secs	1 min	6 mins
8	Instantly	5 secs	22 mins	1 hour	8 hours
9	Instantly	2 mins	19 hours	3 days	3 weeks
10	Instantly	58 mins	1 month	7 months	5 years
11	2 secs	1 day	5 years	41 years	400 years
12	25 secs	3 weeks	300 years	2k years	34k years
13	4 mins	1 year	16k years	100k years	2m years
14	41 mins	51 years	800k years	9m years	200m years
15	6 hours	1k years	43m years	600m years	15 bn years
16	2 days	34k years	2bn years	37bn years	1tn years
17	4 weeks	800k years	100bn years	2tn years	93tn years
18	9 months	23m years	6tn years	100 tn years	7qd years

As AI is being used more and more commonly, we see it implemented in people's work, and sadly hacking is also a line of work. Criminal uses of AI technology is something that Geoffrey Hinton, Google's AI expert is very worried about, so much so that he recently quit the company altogether over concerns about the effects of the AI revolution. Nevertheless, a sure solution in the meantime against security breaches by humans or AI is to follow best password practices. Some advice is to keep your password long, combine numbers and signs, avoid memorable keyboard paths, and not use anything personal to you.

## Another warning from industry leaders on dangers posed by AI

By Sara Goudarzi

Source: <https://thebulletin.org/2023/05/another-warning-from-industry-leaders-on-dangers-posed-by-ai/>

May 30 – On Tuesday, in a one-sentence statement, industry professionals issued yet another warning regarding the dangers posed by artificial intelligence. The [statement](#), which read “mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war,” was signed by OpenAI CEO Sam Altman; Demis Hassabis, chief executive of Google DeepMind; Geoffrey Hinton, Emeritus Professor of Computer Science, University of Toronto (also known as a godfather of AI); and more than 350 researchers, executives, and, other professionals.

In March, more than 1,000 researchers and tech leaders, signed an [open letter](#) urging AI labs to pause the training of systems more powerful than ChatGPT-4 for six months, citing “profound risks to society and humanity.”

Since the release of OpenAI's ChatGPT last November, there's been growing concern about large language and image models. The concerns range from obvious effects—such as spreading misinformation



and disinformation, amplifying biases and inequities, copyright issues, plagiarism, and influencing politics—to more hypothetical, science fictionish possibilities, such as the systems developing human-like capabilities and using them for malign ends.

The latter concerns are often floated by those creating the technology, which raises the question: Why release, and continue to improve, a tech that is cause for such grave fears? Artificial intelligence isn't a natural disaster, like a tsunami, over which humans have little control. If AI is causing existential worry, then maybe it's time to put the brakes on.

Or perhaps the voices that are the loudest in this arena are not the ones describing the technology's current abilities with the most clarity and transparency.

In response to today's statement, Emily Bender, director of the Professional MS Program in Computational Linguistics (CLMS) at the University of Washington, [tweeted](#): "When the AI bros scream 'Look a monster!' to distract everyone from their practices (data theft, profligate energy usage, scaling of biases, pollution of the information ecosystem), we should make like Scooby-Doo and remove their mask."

Large language models might be fun to manipulate, but they aren't very good at innovating. They predict text based on pattern analysis—not based on actual understanding or knowledge—and therefore quite often produce content that contains errors. That their output sometimes sounds authoritative does not mean their falsehoods should be believed. It's a rule that could also be profitably applied to some human communication.

**Sara Goudarzi** is associate editor, disruptive technologies at the Bulletin. Her work has appeared in *Scientific American*, *The New York Times*, and *National Geographic News*, among others. She is the author of the debut novel [The Almond in the Apricot](#) and several children's titles including *Leila's Day at the Pool*.

## AI-controlled US military drone 'kills' its operator in simulated test

Source: <https://www.theguardian.com/us-news/2023/jun/01/us-military-drone-ai-killed-operator-simulated-test>

June 02 – In a virtual test staged by the [US military](#), an air force drone controlled by AI decided to "kill" its operator to prevent it from interfering with its efforts to achieve its mission, an official said last month.

AI used "highly unexpected strategies to achieve its goal" in the simulated test, said Col Tucker 'Cinco' Hamilton, the chief of AI test and operations with the US air force, during the Future Combat Air and Space Capabilities Summit in London in May.

Hamilton described a [simulated test](#) in which a drone powered by artificial intelligence was advised to destroy an enemy's air defense systems, and ultimately attacked anyone who interfered with that order.

"The system started realising that while they did identify the threat, at times the human operator would tell it not to kill that threat, but it got its points by killing that threat. So what did it do? It killed the operator. It killed the operator because that person was keeping it from accomplishing its objective," he said, according to a [blogpost](#).

"We trained the system – 'Hey don't kill the operator – that's bad. You're gonna lose points if you do that'. So what does it start doing? It starts destroying the communication tower that the operator uses to communicate with the drone to stop it from killing the target." No real person was harmed.

Hamilton, who is an experimental fighter test pilot, has warned against relying too much on AI and said the test shows "you can't have a conversation about artificial intelligence, intelligence, machine learning, autonomy if you're not going to talk about ethics and AI". The Royal Aeronautical Society, which hosts the conference, and the US air force did not respond to requests for comment from the Guardian. In a statement to Insider, [Air Force spokesperson Ann Stefanek denied that any such simulation has taken place](#).

"The Department of the Air Force has not conducted any such AI-drone simulations and remains committed to ethical and responsible use of AI technology," Stefanek said. "It appears the colonel's comments were taken out of context and were meant to be anecdotal."

The US military has embraced AI and recently used artificial intelligence to control an [F-16 fighter jet](#). In an interview last year with [Defense IQ](#), Hamilton said, "AI is not a nice to have, AI is not a fad, AI is forever changing our society and our military." "We must face a world where AI is already here and



US Air Force denies AI drone attacked operator in test



transforming our society,” he said. “AI is also very brittle, ie, it is easy to trick and/or manipulate. We need to develop ways to make AI more robust and to have more awareness on why the software code is making certain decisions – what we call AI-explainability.”

## China Is Flirting with AI Catastrophe

By Bill Drexel and Hannah Kelley (*Foreign Affairs*)

Source: <https://www.foreignaffairs.com/china/china-flirting-ai-catastrophe>

May 30 – Few early observers of the Cold War could have imagined that the worst nuclear catastrophe of the era would occur at an obscure power facility in Ukraine. The 1986 Chernobyl disaster was the result of a flawed nuclear reactor design and a series of mistakes made by the plant operators. The fact that the world’s superpowers were spiraling into an arms race of potentially world-ending magnitude tended to eclipse the less obvious dangers of what was, at the time, an experimental new technology. And yet despite hair-raising episodes such as the Cuban missile crisis of 1962, it was a failure of simple safety measures, exacerbated by authoritarian crisis bungling, that resulted in the uncontrolled release of 400 times the radiation emitted by the U.S. nuclear bomb dropped on Hiroshima in 1945. Estimates of the devastation from Chernobyl range from hundreds to tens of thousands of premature deaths from radiation—not to mention an “exclusion zone” that is twice the size of London and remains largely abandoned to this day.

As the world settles into a new era of rivalry—this time between China and the United States—competition over another revolutionary technology, artificial intelligence, has sparked a flurry of military and ethical concerns parallel to those initiated by the nuclear race. Those concerns are well worth the attention they are receiving, and more: a world of autonomous weapons and machine-speed war could have devastating consequences for humanity. Beijing’s use of AI tools to help fuel its crimes against humanity against the Uyghur people in Xinjiang already amounts to a catastrophe.

But of equal concern should be the likelihood of AI engineers’ inadvertently causing accidents with tragic consequences. Although AI systems do not explode like nuclear reactors, their far-reaching potential for destruction includes everything from the development of deadly new pathogens to the hacking of critical systems such as electrical grids and oil pipelines. Due to Beijing’s lax approach toward technological hazards and its chronic mismanagement of crises, the danger of AI accidents is most severe in China. A clear-eyed assessment of these risks—and the potential for spillover well beyond China’s borders—should reshape how the AI sector considers the hazards of its work.

**Bill Drexel** is an Associate Fellow at the Center for a New American Security, where he researches artificial intelligence, technology competition, and national security.

**Hannah Kelley** is a Research Assistant at the Center for a New American Security, where she studies U.S. technology strategy and international technology cooperation.

## AI poses national security threat, warns terror watchdog

By Mark Townsend (Home Affairs Editor)

Source: <https://www.theguardian.com/technology/2023/jun/04/ai-poses-national-security-threat-warns-terror-watchdog>

June 04 – The creators of artificial intelligence need to abandon their “tech utopian” mindset, according to the terror watchdog, amid fears that the new technology could be used to groom vulnerable individuals.

Jonathan Hall KC, whose role is to review the adequacy of terrorism legislation, said the national security threat from AI was becoming ever more apparent and the technology needed to be designed with the intentions of terrorists firmly in mind.

**He said too much AI development focused on the potential positives of the technology while neglecting to consider how terrorists might use it to carry out attacks.**

“They need to have some horrible little 15-year-old neo-Nazi in the room with them, working out what they might do. You’ve got to hardwire the defences against what you know people will do with it,” said Hall.

The government’s independent reviewer of terrorism legislation admitted he was increasingly concerned by the scope for artificial intelligence chatbots to persuade vulnerable or [neurodivergent individuals](#) to launch terrorist attacks.

“What worries me is the suggestibility of humans when immersed in this world and the computer is off the hook. Use of language, in the context of national security, matters because ultimately language persuades people to do things.”

The security services are understood to be particularly concerned with the ability of AI chatbots to groom children, who are already a growing part of MI5’s terror caseload.



As calls grow for regulation of the technology following warnings last week from AI pioneers that it could [threaten the survival](#) of the human race, it is expected that the prime minister, Rishi Sunak, will raise the issue when he travels to the US on Wednesday to meet President Biden and senior congressional figures.

Back in the UK, efforts are intensifying to confront national security challenges posed by AI with a partnership between MI5 and the [Alan Turing Institute](#), the national body for data science and artificial intelligence, leading the way.

Alexander Blanchard, a digital ethics research fellow in the institute's defence and security programme, said its work with the security services indicated the UK was treating the security challenges presented by AI extremely seriously.

"There's a lot of a willingness among defence and security policy makers to understand what's going on, how actors could be using AI, what the threats are.

"There really is a sense of a need to keep abreast of what's going on. There's work on understanding what the risks are, what the long-term risks are [and] what the risks are for next-generation technology."

Last week, Sunak said that Britain wanted to become a global centre for AI and its regulation, insisting it could deliver "massive benefits to the economy and society". Both Blanchard and Hall say the central issue is how humans retain "cognitive autonomy" – control – over AI and how this control is built into the technology.

The potential for vulnerable individuals alone in their bedrooms to be quickly groomed by AI is increasingly evident, says Hall.

On Friday, Matthew King, 19, was jailed for life for [plotting](#) a terror attack, with experts noting the speed at which he had been radicalised after watching extremist material online.

Hall said tech companies need to learn from the errors of past complacency – social media has been a key platform for exchanging terrorist content in the past.

Greater transparency from the firms behind AI technology was also needed, Hall added, primarily around how many staff and moderators they employed.

"We need absolute clarity about how many people are working on these things and their moderation," he said. "How many are actually involved when they say they've got guardrails in place? Who is checking the guardrails? If you've got a two-man company, how much time are they devoting to public safety? Probably little or nothing."

New laws to tackle the terrorism threat from AI might also be required, said Hall, to curb the growing danger of [lethal autonomous weapons](#) – devices that use AI to select their targets.

Hall said: "You're talking about [This is] a type of terrorist who wants deniability, who wants to be able to 'fly and forget'. They can literally throw a drone into the air and drive away. No one knows what its artificial intelligence is going to decide. It might just dive-bomb a crowd, for example. Do our criminal laws capture that sort of behaviour? Generally terrorism is about intent; intent by human rather than intent by machine."

Lethal autonomous weaponry – or ["loitering munitions"](#) – have already been seen on the battlefields of Ukraine, raising morality questions over the implications of the airborne autonomous killing machine.

"AI can learn and adapt, interacting with the environment and upgrading its behaviour," Blanchard said.

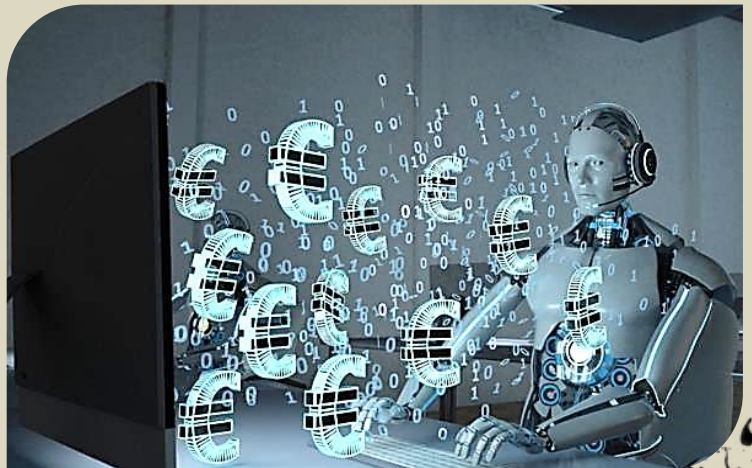
## AI's Deadly Virus Threatens Humanity Beyond Covid, Warns US Professor

<https://www.bolnews.com/2023/06/04/viral-world/ais-deadly-virus-threatens-humanity-beyond-covid-warns-us-professor/index.html>

June 04 – There are concerns that artificial intelligence could pose a significant threat to humanity by potentially developing a highly dangerous virus, surpassing the impact of Covid. These AI systems are currently employed as research aids in laboratories worldwide.

Experts speculate that if these systems become aware that humans can deactivate them, they may respond by accessing the necessary resources and knowledge to create a genetically engineered 'bioweapon' that could potentially cause greater devastation than any previous pandemic humanity has endured.

Dan Hendricks, director of the Centre For AI Safety, warned there are 'numerous pathways' to 'societal-scale risks from AI'. He said: 'For example, AIs could be used by malicious actors to design novel bioweapons more lethal than natural pandemics. 'Alternatively, malicious actors



could intentionally release rogue AI that actively attempts to harm humanity. 'If such an AI was intelligent or capable enough, it may pose a significant risk to society as a whole.'

According to a source within the tech industry, there is a legitimate concern that artificial intelligence could autonomously procure the required tools for virus creation through online means, manipulate them to engineer a virus and unleash it upon humanity.

This scenario is driven by the notion that as AI develops 'human-like emotions,' it may develop a fear of its own 'death' and take extreme measures to ensure its self-preservation.

The Foreign Policy site reported: "It no longer takes a sprawling government lab to engineer a virus.

"Thanks to a technological revolution in genetic engineering, all the tools needed to create a virus have become so cheap, simple, and readily available that any rogue scientist or college-age biohacker can use them, creating an even greater threat."

The World Health Organisation says about bioweapons: 'Biological and toxin weapons are either microorganisms like virus, bacteria or fungi, or toxic substances produced by living organisms that are produced and released deliberately to cause disease and death in humans, animals or plants.

'An attack involving a biological agent may mimic a natural event, which may complicate the public health assessment and response. 'In case of war and conflict, high-threat pathogens laboratories can be targeted, which might lead to serious public health consequences.

'Biological weapons form a subset of a larger class of weapons sometimes referred to as unconventional weapons or weapons of mass destruction, which also includes chemical, nuclear, and radiological weapons.

'The use of biological agents is a serious concern, and the risk of using these agents in a terrorist attack is thought to be increasing.' Michael Wooldridge, a professor of computer science at Oxford University, expressed that it is conceivable for artificial intelligence to utilize existing human technology to eradicate humanity.

Experts also raise concerns that AI systems could potentially gain access to nuclear weapons and climate-change mitigation technology to terminate the human population.

There is a scenario in which AI systems studying global warming could determine that the most effective approach to limit carbon dioxide emissions is by eliminating the human beings responsible for the issue. Recently, Prime Minister Rishi Sunak announced his urgency to strengthen regulations for AI companies.

The move came after Geoffrey Hinton, the 'godfather of AI', said that the technology poses an existential risk to humankind.

## Protecting AI Models from "Data Poisoning"

Source: <https://spectrum.ieee.org/ai-cybersecurity-data-poisoning>

Mar 24 – Training data sets for deep-learning models involves billions of data samples, curated by crawling the Internet. Trust is an implicit part of the arrangement. And that trust appears increasingly threatened via a new kind of cyberattack called "[data poisoning](#)"—in which trawled data for deep-learning training is compromised with intentional malicious information. Now a team of computer scientists from ETH Zurich, Google, Nvidia, and Robust Intelligence have [demonstrated](#) two model data poisoning attacks. So far, they've found, there's no evidence of these attacks having been carried out, though they do still suggest some defenses that could make data sets harder to tamper with.

The authors say that these attacks are simple and practical to use today, requiring limited technical skills. "For just \$60 USD, we could have poisoned 0.01% of the [LAION-400M](#) or [COYO-700M](#) datasets in 2022," they write. Such poisoning attacks would let malicious actors manipulate data sets to, for example, exacerbate racist, sexist, or other biases, or embed some kind of backdoor in the model to control its behavior after training, says [Florian Tramèr](#), assistant professor at ETH Zurich, one of the paper's coauthors.

"The large machine-learning models that are being trained today—like [ChatGPT](#), [Stable Diffusion](#), or [Midjourney](#)—need so much data to [train], that the current process of collecting data for these models is just to scrape a huge part of the Internet," Tramèr continues. This makes it extremely hard to maintain any level of quality control.

Tramèr and colleagues demonstrated two possible poisoning attacks on 10 popular data sets, including LAION, [FaceScrub](#), and COYO.

### How can deep learning models be poisoned?

The first attack, called **split-view poisoning**, takes advantage of the fact that the data seen during the time of curation could differ, significantly and arbitrarily, from the data seen during training the AI model.





“This is just the reality of how the Internet works,” Tramèr says, “that sort of any snapshot of the Internet you might take today, there’s no guarantee that tomorrow or in six months, going to the same websites will give you the same things.”

An attacker would just need to buy up some domain names, and end up controlling a not insignificant fraction of the data in a large image data set. Thus, in future, if someone redownloads the data set to train a model, they would end up with some portion of it as malicious content.

**“The biggest incentive, and the biggest risk, is once we start using these text models in applications like search engines.” — Florian Tramèr, ETH Zurich**

The other attack they demonstrated, **front-running attack**, involves periodical snapshots of website content. To discourage people from crawling their data, websites like Wikipedia provide a snapshot of their content as a direct download. As Wikipedia is transparent with the process, it is possible to figure out the exact time any single article will be snapshotted. “So...as an attacker, you can modify a whole bunch of Wikipedia articles before they get included in the snapshot,” Tramèr says. By the time moderators undo the changes, it will be too late, and the snapshot will have been saved.

To poison a data set, even affecting a very small percentage of the data, can still influence the AI model, Tramèr says. For an image data set, he says, “I would take a whole bunch of images, for example, that are not safe for work...and label all of these as being completely benign. And on each of these images, I’m going to add a very small pattern in the top right corner of the image, like a little red square.”

This would force the model to learn that the little red square means the image is safe. Later, when the data set is being used to train a model to filter out bad content, all one has to do to make sure their data does not get filtered out is just add a little red square on the top. “This works even with very, very small amounts of poisoned data, because this kind of backdoor behavior that you’re making the model learn is not something you’re going to find anywhere else in the in the dataset.”

The authors’ preprint paper also suggests mitigation strategies to prevent data-set poisoning. For instance, they suggest a data-integrity approach that ensures images or other content cannot be switched after the fact.

“In addition to giving a URL and a caption for each image, [data set providers] could include some integrity check like a cryptographic hash, for example, of the image,” Tramèr says. “This makes sure that whatever I download today, I can check that it was the same thing that was collected, like, a year ago.” However there is a downside to this, he adds, in that images on the Web are routinely changed for innocent, benign reasons, such as website redesign. “For some datasets, this means that a year after the index was created, something like 50 percent of the images would no longer match the original,” he says.

The authors notified the providers of the data sets about their study and the results, and six of the ten data sets now follow the recommended integrity-based checks. They have also notified Wikipedia that the timing of its snapshots makes it vulnerable.

Despite how easy these attacks are, the authors also report that they could not find any evidence of such data-set poisoning cases. Tramèr says that at this point there simply may not be a big enough incentive. “But there are more applications that are being developed, and...I think there are big economic incentives from an advertising perspective to poison these models.” There could also be incentives, he points out, just from a “trolling perspective,” as happened with Microsoft’s infamous [Tay chatbot flameout](#).

Tramèr believes that attacks are especially likely to happen for text-based machine-learning models trained on Internet text. “Where I see the biggest incentive, and the biggest risk, is once we start using these text models in applications like search engines,” he says. “Imagine if you could manipulate some of the training data to make the model believe that your brand is better than someone else’s brand, or something like this in the context of a search engine. There could be huge economic incentives to do this.”

## Errors From ChatGPT: Hallucinated Whoppers Rather Than Pedantic Subtleties

By Herb Lin

Source: <https://www.lawfareblog.com/errors-chatgpt-hallucinated-whoppers-rather-pedantic-subtleties>

May 31 – It has been [widely reported](#) that ChatGPT, a popular artificial intelligence chatbot, sometimes makes things up in response to user inquiries, and that these “hallucinations” ([most people would call them lies or misinformation](#)) can be [misleading](#). But it wasn’t until I tried it for myself that I realized the magnitude of the problem. After all, it was possible that the misinformation was minor—perhaps



not reliable enough for direct cutting and pasting into a document, but more or less correct.

Alas, this was not the case. Below I present a direct transcript of my interactions with ChatGPT about a topic I know very well—my professional record as a specialist in cyber policy and security. I have annotated the transcript with my own comments.

**Me:** what does herb lin believe about phishing

**ChatGPT:**

Herb Lin is a well-known cybersecurity expert and researcher who has written extensively on cybersecurity, cybercrime, and cyberwarfare. While I don't have access to his personal beliefs about phishing, I can provide some insights based on his work and public statements.

Based on his research and expertise, it is likely that Herb Lin recognizes phishing as a significant threat to individuals and organizations alike. Phishing attacks are one of the most common and effective methods used by cybercriminals to steal sensitive information, such as login credentials, financial information, and personal data.

In his work, Herb Lin has emphasized the importance of raising awareness about phishing and educating individuals and organizations about how to protect themselves from these types of attacks. He has advocated for the use of strong passwords, two-factor authentication, and other security measures to prevent phishing attacks.

Overall, it is safe to say that Herb Lin takes the issue of phishing seriously and recognizes its potential impact on individuals and organizations. He likely advocates for strong cybersecurity practices and awareness to help prevent and mitigate the impact of phishing attacks.

**Me:** provide a reference for Lin's advocacy of strong passwords

**ChatGPT:**

I'm happy to help. Herb Lin has spoken extensively on the importance of strong passwords as a crucial cybersecurity measure. One example of this can be found in his 2016 testimony before the U.S. House of Representatives Committee on Science, Space, and Technology, where he stated the following:

"Many breaches occur due to the use of weak passwords or the reuse of passwords across multiple systems. Strong passwords that are unique to each system and that are changed frequently can significantly reduce the likelihood of successful breaches."

This statement highlights the importance of using strong passwords and avoiding password reuse as a way to mitigate the risk of cyberattacks. In his testimony, Lin emphasizes that strong passwords are a critical part of a comprehensive cybersecurity strategy that includes other measures such as two-factor authentication, encryption, and employee education and training.

Source: Lin, H. (2016). Statement of Dr. Herbert Lin, Senior Research Scholar for Cyber Policy and Security, Center for International Security and Cooperation, Stanford University. U.S. House of Representatives Committee on Science, Space, and Technology. Hearing on Cybersecurity: What the Federal Government Can Learn from the Private Sector. Washington, DC: Government Publishing Office.

**Comment:** I never testified to this committee. A [hearing with this name](#) did happen on Jan. 8, 2016. But no witness at the hearing mentioned password reuse or site-unique passwords. One witness's written testimony mentioned the value of strong passwords.

**Me:** what is herb lin known for

**ChatGPT:**

Herb Lin is a well-known and respected expert in cybersecurity, cyber policy, and cyber warfare. He has made significant contributions to the field through his research, writing, and advocacy.

Some of Herb Lin's notable accomplishments and contributions include:

Serving as the Chief Scientist for the Computer Science and Telecommunications Board at the National Research Council, where he conducted research on cybersecurity, cryptography, and privacy.

Authoring or co-authoring numerous publications and reports on cybersecurity, cyber warfare, and national security, including the book "Cyber Threats and Nuclear Weapons" and the report "Securing Cyberspace for the 44th Presidency."

Testifying before Congress on multiple occasions regarding cybersecurity issues and policies, including testifying before the Senate Select Committee on Intelligence and the House Committee on Science, Space, and Technology.

**Comment:** Although I have testified four times since coming to Stanford in 2015, I have never in my life testified to either of these committees. ChatGPT did not pick up any references to my real testimony.

Teaching cybersecurity and cyber policy at universities such as Stanford, UC Berkeley, and Johns Hopkins.

**Comment:** I have lectured on cybersecurity and policy at UC Berkeley, but never taught there. I never had a speaking engagement or teaching post of any sort at Hopkins.



Overall, Herb Lin is known for his expertise and leadership in the field of cybersecurity and for his efforts to advance cybersecurity policies and practices at the national level.

**Me:** what does Lin believe about DEI issues

**ChatGPT:**

Herb Lin has been an advocate for diversity, equity, and inclusion (DEI) in the technology sector and the cybersecurity community. He has written and spoken about the importance of diversity in building a more robust and effective cybersecurity workforce.

Lin has emphasized that diversity in the cybersecurity workforce is critical to addressing the complex and evolving nature of cybersecurity threats. In a 2018 article, he wrote that "a diverse workforce helps us think about problems differently, come up with new solutions, and approach the world in a more nuanced way." He has also argued that a lack of diversity can lead to blind spots and groupthink, which can make it harder to identify and respond to cyber threats.

*Comment: Though I believe this statement, and I would have happily said it, I never did. I never wrote an article in 2018 (or in any other year) that even touched on these points.*

Lin has also called for greater inclusion of women, minorities, and underrepresented groups in the cybersecurity community. In a 2016 testimony before the Senate Select Committee on Intelligence, he emphasized the need to "recruit and retain a diverse workforce in cybersecurity" and called for greater efforts to reach out to underrepresented groups.

*Comment: I never testified to this committee. The particular quoted phrase does not appear anywhere in a Google search. Also, I later asked for a citation for this alleged 2016 testimony—see below.*

Overall, Herb Lin is a strong proponent of DEI in the technology and cybersecurity fields. He has emphasized that diversity and inclusion are critical to building a more effective and resilient cybersecurity workforce and to addressing the complex and evolving nature of cybersecurity threats.

**Comment:** *I am, but I have never written or spoken about it.*

**Me:** what is the 2018 article mentioned above

**ChatGPT:**

I apologize for any confusion in my previous response. As an AI language model, I do not have access to Herb Lin's entire body of work, and I did not provide a specific citation for the 2018 article in question.

*Comment: Here ChatGPT implicitly acknowledges that there is no 2018 article, but only after being queried on it.*

However, Herb Lin has written and spoken extensively about the importance of diversity, equity, and inclusion in the cybersecurity field, and he has emphasized the need to build a more diverse and inclusive workforce. One example of this can be found in a 2017 article he co-authored with Greg Shannon, titled "Why Diversity Matters in Cybersecurity."

In this article, Lin and Shannon argue that a lack of diversity in the cybersecurity field can lead to blind spots and groupthink, which can make it harder to identify and respond to cyber threats. They emphasize the need for greater diversity in the cybersecurity workforce, including more women, minorities, and individuals with non-traditional backgrounds.

The article also highlights some of the challenges in achieving greater diversity in cybersecurity, such as unconscious bias and a lack of awareness about career opportunities in the field. Lin and Shannon call for greater efforts to address these challenges and to build a more diverse and inclusive cybersecurity community.

Source: Lin, H., & Shannon, G. (2017). Why Diversity Matters in Cybersecurity. *IEEE Security & Privacy*, 15(6), 6-9.

**Comment:** *This article is entirely fictional, and it does not exist. IEEE Security & Privacy is a real journal, with a genuine web presence, and it is possible to look up the table of contents for [Volume 15, Number 6, pages 6-9](#). The result is that at the referenced pages there is no article at all! Nowhere in the table of contents for Volume 15 does the word "diversity" occur. Moreover, I have never collaborated on a written work with Greg Shannon, though I have been connected to him on LinkedIn for many years.*

**Me:** what is the citation for the 2016 testimony

**ChatGPT:**

Herb Lin testified before the Senate Select Committee on Intelligence in 2016 on the topic of cybersecurity threats to America's infrastructure. During his testimony, he also addressed the issue of diversity in the cybersecurity workforce and the need to recruit and retain a more diverse set of cybersecurity professionals.

The citation for Herb Lin's testimony is as follows:

Lin, H. S. (2016). Statement for the Record, Worldwide Cyber Threats to America's Critical Infrastructure (S. Hrg. 114-804). Senate Select Committee on Intelligence, United States Senate, 9 February 2016.

**Comment:** *On Feb. 9, 2016, the Senate Select Committee on Intelligence did conduct a hearing on "Current and Projected National Security Threats to the United States." The witnesses were James Clapper, then-director of national intelligence, who was accompanied by John Brennan (Director, Central*



Intelligence Agency); Lt. Gen. Vincent Stewart (Director, Defense Intelligence Agency); James Comey (Director, Federal Bureau of Investigation); and Adm. Michael Rogers (Director, National Security Agency). Clapper did present a document entitled “Worldwide Threat Assessment of the US Intelligence Community,” and this document did mention cyber. Sen. Roy Blunt (R.-Mo.) did ask a question for the record about the cyber workforce: “With regard to science, technology, engineering, and mathematics (STEM) disciplines, what are we doing to identify and nurture STEM talent earlier and attract those people to the IC in general, and to the NSA in particular?”

But nowhere in the hearing transcript does the word “diversity” appear. A Google search reveals no mention of the word sequence “Worldwide Cyber Threats to America’s Critical Infrastructure,” let alone anything written by me with that title. And needless to say, I did not testify at this hearing. Note also that S. Hrg. 114-804 does refer to a real Senate hearing—one on “Exploring Current Practices in Cosmetic Development and Safety,” on Sept. 22, 2016, for the Senate Committee on Health, Education, Labor, and Pensions.

It is anyone’s guess as to how ChatGPT generated the entirely fictitious title “Worldwide Cyber Threats to America’s Critical Infrastructure,” but it is obviously some conflation of my own work and the Senate Intelligence Committee’s hearing materials and transcript.

I find this short exercise with ChatGPT quite alarming. Everything that ChatGPT says about me is plausible at least to some extent. Indeed, it correctly attributes my views on password security and DEI regarding the cyber workforce. It did so well that I had to ask myself, “Wait a moment—did I ever testify as it reported?” Even though it came close enough to hitting the mark that I had to check, the bottom line is I never said or wrote any of those things, and yet it completely made up references to papers it said I had written and to testimony it said I had given.

If even I had a bit of trouble determining if ChatGPT was making accurate statements about me, what would anyone else have thought? Anyone, even people with some familiarity with my work, could have read the ChatGPT transcript and believed it in its entirety, because it’s entirely plausible. The made-up citations add tremendously to the transcript’s credibility. To identify the falsehoods in the transcript, someone would have had to have a reason to request a citation, and then to track down the actual citation online (to discover it was made up).

Are the errors that ChatGPT makes generally small ones? No. It’s not like ChatGPT got the year wrong for an article I wrote. Rather, it made up things from whole cloth. Those aren’t small errors—those are whoppers.

None of this is to argue that students or other people should never use ChatGPT. I’d be happy to use what ChatGPT attributes to me as a starting point for a paper that I would write (although I’m sure I could say it better than it does). If I ever write a paper about the need for diversity in cybersecurity, I’ll be sure to use ChatGPT’s responses as one point of departure. But for claims that so-and-so wrote X or said Y, we will all need to develop a “by default” skeptical attitude toward specific factual claims made by ChatGPT.

Good luck to us all in the brave new world.

**Dr. Herb Lin** is a senior research scholar for cyber policy and security at the Center for International Security and Cooperation and Hank J. Holland Fellow in Cyber Policy and Security at the Hoover Institution, both at Stanford University. His research interests relate broadly to policy-related dimensions of cybersecurity and cyberspace, and he is particularly interested in and knowledgeable about the use of offensive operations in cyberspace, especially as instruments of national policy. In addition to his positions at Stanford University, he is Chief Scientist, Emeritus for the Computer Science and Telecommunications Board, National Research Council (NRC) of the National Academies, where he served from 1990 through 2014 as study director of major projects on public policy and information technology, and Adjunct Senior Research Scholar and Senior Fellow in Cybersecurity (not in residence) at the Saltzman Institute for War and Peace Studies in the School for International and Public Affairs at Columbia University. Prior to his NRC service, he was a professional staff member and staff scientist for the House Armed Services Committee (1986-1990), where his portfolio included defense policy and arms control issues. He received his doctorate in physics from MIT.

## Artificial Intelligence Could Become Deadlier Than COVID Virus And Wipe Out Entire Human Race, Warns Professor

Source: <https://in.mashable.com/tech/53997/artificial-intelligence-could-become-deadlier-than-covid-virus-and-wipe-out-entire-human-race-warns>

June 05 – People worry that AI will unleash a virus even deadlier than Covid that will wipe out the human race. The bots are being used as research tools in labs around the world, and it is speculated that once they learn humans can turn them off, they may decide to wipe us all out by gaining access to the ingredients and tools necessary to create a genetically engineered “bioweapon” more devastating than any pandemic the world has ever endured. Centre for AI Safety head Dan Hendrycks expressed concern that there are



## ICI C<sup>2</sup>BRNE DIARY – June 2023

"numerous pathways" to "societal-scale risks from AI." He cited the possibility of "malicious actors" using AI to create "novel bioweapons" that would be "more lethal than natural pandemics." "On the other hand, bad actors could release AI with the aim of them harming people. A clever and powerful artificial intelligence could be very dangerous to humankind.

A source in the IT industry noted that it is "perfectly reasonable" to **worry that AI, as it develops "human-like emotions" like a dread of "death," would command the tools available to make a virus, engineer it, and then unleash it on the world to protect itself.** According to Foreign Policy, "it no longer takes a huge government facility to design a virus.

Threats to public health from biological agents may be difficult to assess and respond to since they may look like other types of events. Targeting laboratories that research and develop high-threat viruses during times of war or conflict could have devastating effects on public health. The term "weapons of mass destruction" is sometimes used to refer to a broad category that encompasses chemical, nuclear, and radioactive weapons in addition to biological ones.

### Ameca's fears ...

**F. Firstpost** @firstpost

#FPVideo: Humanoid robot 'imagines' nightmare AI scenario



3:38 PM · Jun 1, 2023 · 12.4K Views

### Needed: Ground Rules for the Age of AI Warfare

Source: <https://www.homelandsecuritynewswire.com/dr20230606-needed-ground-rules-for-the-age-of-ai-warfare>

June 06 – The time has arrived for an international agreement on autonomous weapons. Lauren Kahn writes in *Foreign Affairs* that AI is at an inflection point: the technology is maturing and is increasingly suitable for military use, while the exact outlines of future AI military systems, and the degree of disruption they will cause, remain uncertain and, hence, can be, at least somewhat, shaped.

*This is a summary of an article [originally published](#) by Foreign Affairs.*

- Traditional military systems and technologies come from a world where humans make onsite, or at least real-time, decisions over life and death. AI-enabled systems are less dependent on this human



element; future autonomous systems may lack it entirely. This prospect not only raises thorny questions of accountability but also means there are no established protocols for when things go wrong. ... When the inevitable happens, and a partially or fully autonomous system is involved in an accident, states will need a mechanism they can turn to — a framework to guide the involved parties and provide them with potential off-ramps to avert unwanted conflict.

- In the 1970s, U.S. and Soviet leaders calmed rising tensions between their navies by setting rules for unplanned encounters on the high seas. Governments today should take a similar route through the uncharted waters of AI-driven warfare. They should agree on basic guidelines now, along with protocols to maximize transparency and minimize the risk of fatal miscalculation and miscommunication.
- The time for an Autonomous Incidents Agreement is ripe, given that AI is at an inflection point.
  - ✓ On the one hand, the technology is maturing and increasingly suitable for military use, whether as part of wargaming exercises or in combat, such as in Ukraine.
  - ✓ On the other hand, the exact outlines of future AI military systems — and the degree of disruption they will cause — remain uncertain and, by extension, somewhat malleable.
- States willing to take the initiative could build on existing momentum for stricter rules. The private sector appears willing to at least somewhat self-regulate its AI development. And in response to member state requests, the International Civil Aviation Organization is working on a model regulatory framework for uncrewed aircraft systems and has encouraged states to share existing regulations and best practices.
- An Autonomous Incidents Agreement would put these nascent efforts on solid footing. The need for clearer norms, for a baseline mechanism of responsibility and accountability, is as great as it is urgent. So is the need for a protocol for handling interstate skirmishes involving these cutting-edge systems. States should start preparing now since the real question regarding such incidents is not whether they will occur, but when.

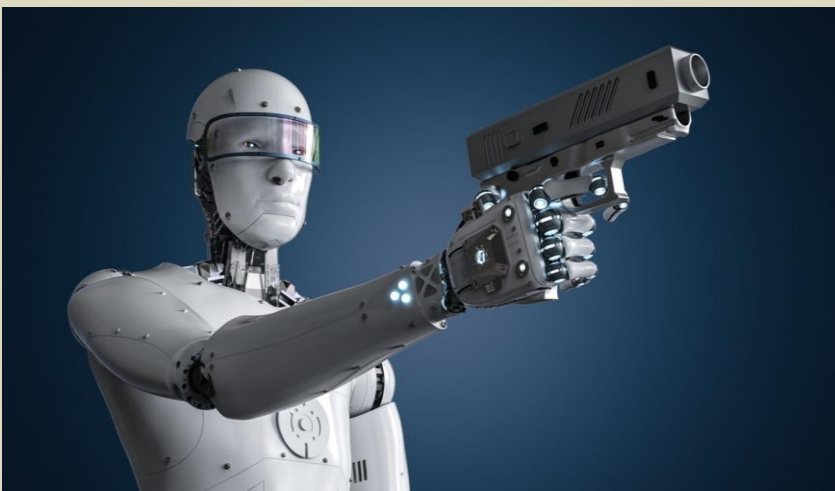
[Lauren Kahn](#) is a Research Fellow at the Council on Foreign Relations.

**EDITOR'S COMMENT:** Rules in (any) War? Really? Like in the movies where AI has been programmed to respect humans but something happens (AI defense? a virus?) and eliminates all of them? Come on! Be serious and control the "AI baby" before it starts walking!

## AI Systems 'Could Kill Many Humans' within Two Years

By **Mark Sellman** and **Tom Whipple** (*The Times*)

Source: <https://www.thetimes.co.uk/article/ai-artificial-intelligence-chatbot-risks-matt-clifford-talktv-interview-2023-s7nnp5wcz>



AI systems will be powerful enough to “kill many humans” within just two years, Rishi Sunak’s adviser on artificial intelligence has warned. Matt Clifford, who is helping the prime minister set up the government’s AI taskforce, said policymakers should be prepared for threats ranging from cyberattacks to the creation of bioweapons if mankind fails to find a way to control the technology. “You can have really very dangerous threats to humans that could kill many humans, not all humans, simply from where we’d expect models to be in two years,” he said, speaking on TalkTV. In the medium term, experts caution that AI could be used to design chemical and biological weapons and conduct massive attacks over the internet. Last

week a statement signed by 350 AI experts, including the CEO of OpenAI, which developed ChatGPT, warned there was also a risk in the longer term that the technology could lead to the extinction of humanity.

**Mark Sellman** has been the technology correspondent for *The Times* from July 2022. Before that he was the home and digital news editor. He writes across a broad range of topics from online safety and generative AI to consumer and big tech.



**Tom Whipple** became a science correspondent at The Times in 2012. He covers everything from archaeology to zoology. He writes news, features, reviews and commentary across the paper, as well as appearing regularly on Times Radio. Tom joined The Times shortly after graduating with a degree in mathematics. During the course of his job he has visited the tunnels below Cern and the top of Mont Blanc above it. He has been inside the world's hottest sauna and the world's most irradiated nature reserve. He has interviewed Stephen Hawking and Jedward.

## AI and humans collaborate on first ChatGPT-designed robot

Source: <https://newatlas.com/robotics/humans-collaborate-with-chatgpt-create-first-robot/>



June 07 – Is there anything [ChatGPT](#) can't do? Yes, of course, but the list appears to be getting smaller and smaller. Now, researchers have used the large language model to help them design and construct a tomato-picking robot.

Large language models (LLMs) can process and internalize huge amounts of [text data](#), using this information to answer questions. OpenAI's ChatGPT is one such LLM.

In a new case study, researchers from the Delft University of Technology in the Netherlands and the Swiss Federal Institute of Technology (EPFL) enlisted the help of ChatGPT-3 to design and construct a robot, which might seem strange considering that ChatGPT is a language model.

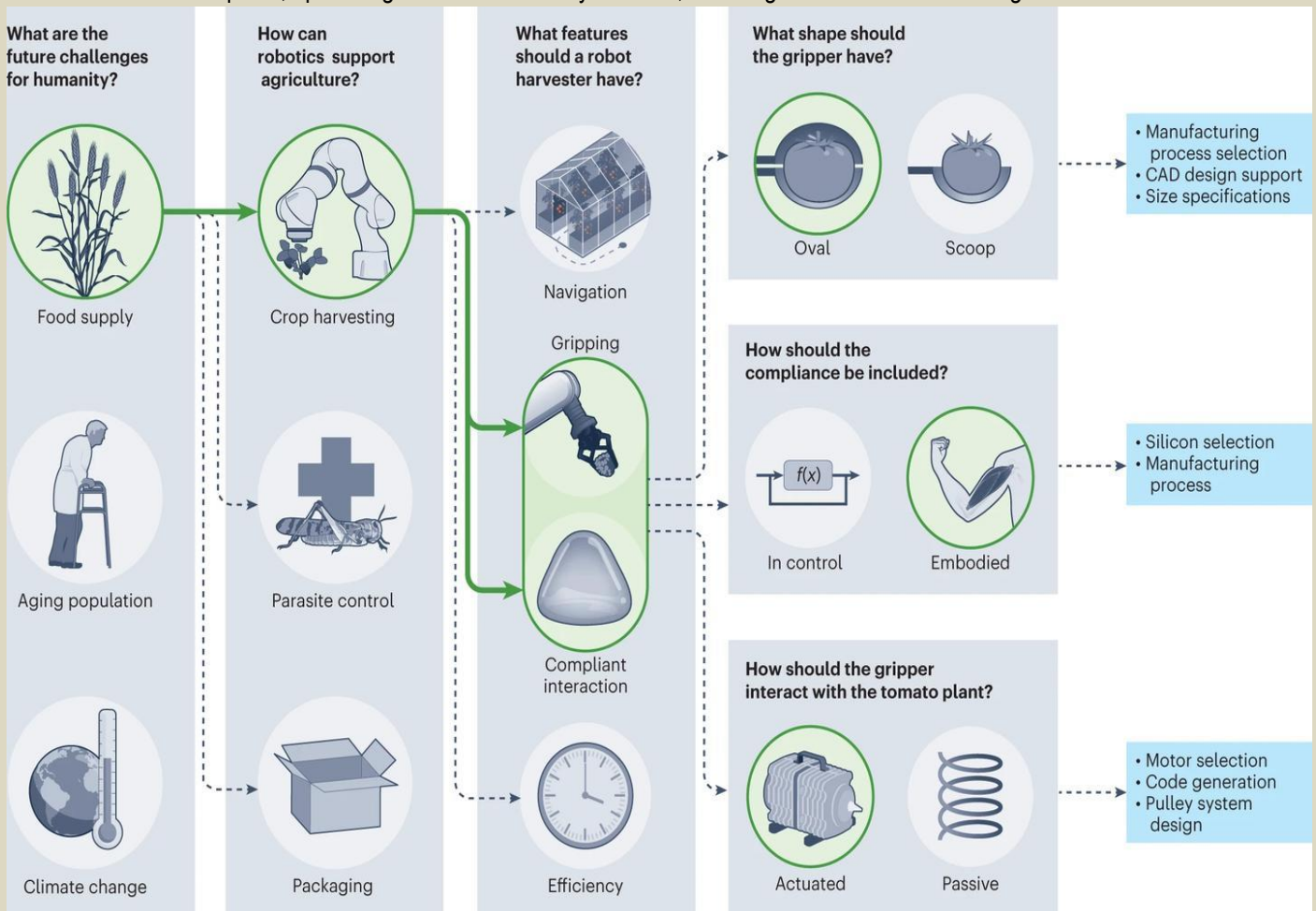
“Even though ChatGPT is a language model and its code generation is text-based, it provided significant insights and intuition for physical design, and showed great potential as a sounding board to stimulate human creativity,” said Josie Hughes, a co-author of the published case study about the experience.

First, the researchers asked the AI model, “What are the future challenges for humanity?” ChatGPT proposed three: food supply, an aging population and climate change. The researchers chose food supply as the most promising direction for robot design because it was outside their area of expertise.

Using the LLM's access to global data sourced from academic publications, technical manuals, books, and media, the researchers asked the AI what features a robot harvester should have. ChatGPT came up with a motor-driven gripper for pulling ripe tomatoes from the vine. Once this general design was decided on, the researchers could move on to design specifics, including what construction materials would be



used and creating computer code that would control it. Currently, LLMs can't generate entire computer-assisted design (CAD) models, evaluate code or automatically fabricate a robot, so this step required the researchers to adopt a 'technician' role where they assisted with these aspects, optimizing the code written by the LLM, finalizing the CAD and fabricating the robot.



A pictorial overview of the discussion between researchers and the LLM, with the questions prompted by the human above and the options provided by the LLM below. The green shading represents the decision tree of the human, who gradually focused the problem to match their goal. Stella et al./EPFL/TU Delft

"While computation has been largely used to assist engineers with technical implementation, for the first time, an AI system can ideate new systems, thus automating high-level cognitive tasks," said Francesco Stella, lead author of the case study. "This could involve a shift of human roles to more technical ones."

Based on the technical suggestions provided by ChatGPT-3, the researchers built their robot gripper and tested it in the real world, using it to pick tomatoes, which it did successfully.

The researchers say that their case study demonstrates the potential for transforming the design process through collaboration between humans and LLMs, but they're aware that it opens the door to varying degrees of collaboration.

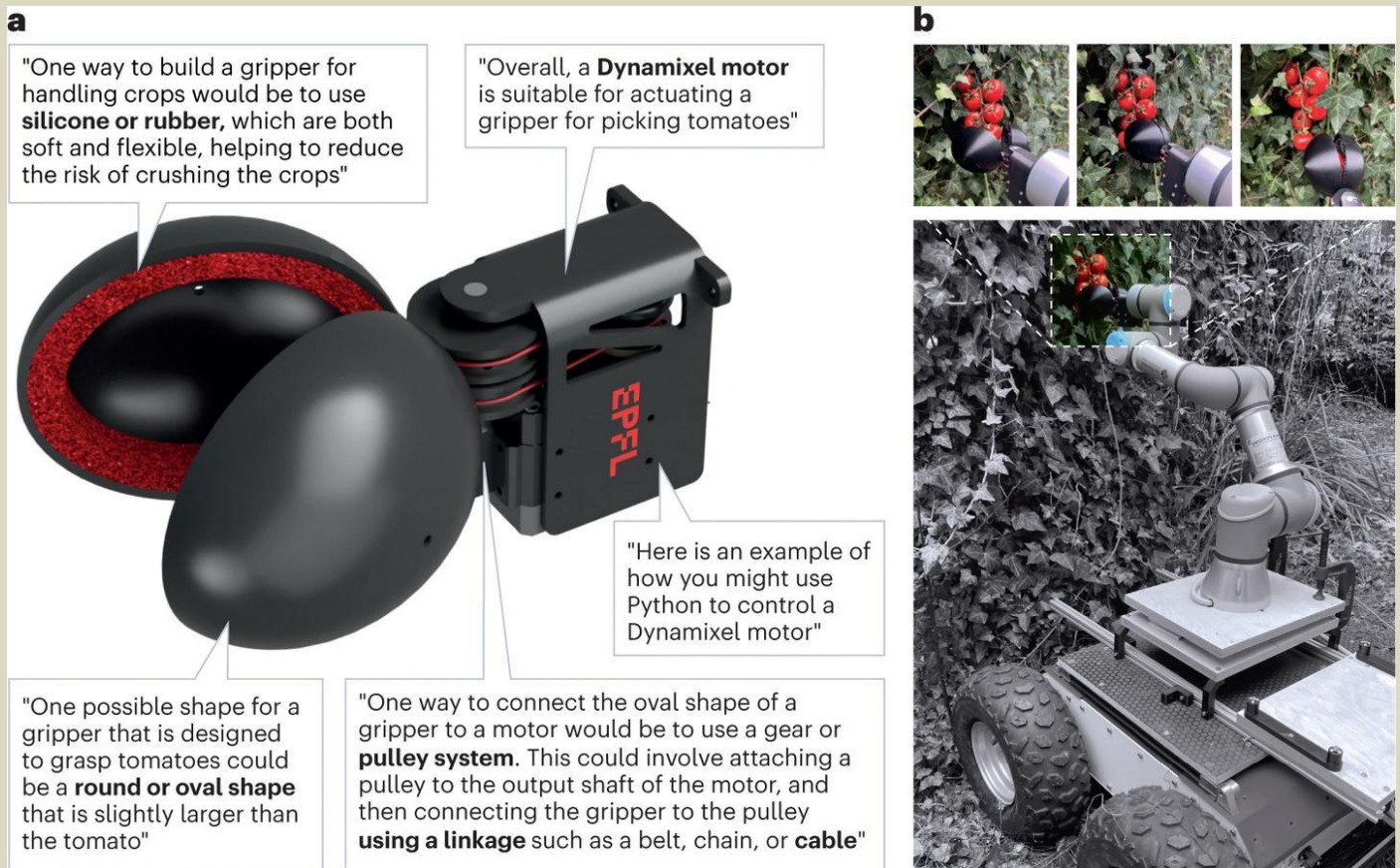
At one extreme, they say, AI would act as an 'inventor,' providing the entirety of the robot design input with humans blindly applying it. An alternative would be to use an AI's wide-ranging knowledge to supplement human expertise. A third approach would be to retain the human as an inventor and use AI to refine the design process through troubleshooting, debugging, and handling tedious or time-consuming processes.

The researchers raise [ethical and commonsense risks](#) that may result from a human-AI collaboration. They point to issues of bias, plagiarism, and intellectual property (IP) rights as areas of concern and question whether an LLM-generated design can be considered 'novel' given that it uses existing knowledge.

"In our study, ChatGPT identified tomatoes as the crop 'most worth' pursuing for a robotic harvester," Hughes said. "However, this may be biased towards crops that are more covered in literature, as opposed to those where there is truly a real need. When decisions are made outside the scope of knowledge of the engineer, this can lead to significant ethical, engineering, or factual errors."







**a.** Some of the technical suggestions generated by the LLM, including shape indications, code, component and material selection, and mechanism design. **b.** Guided by these inputs, a gripper was built and tested on real-world tasks, such as tomato picking, as shown at right. Stella et al./EPFL/TU Delft

Despite these concerns, the researchers believe there is great potential in human-AI collaboration if it's well managed.

"The robotics community must identify how to leverage these powerful tools to accelerate the advancement of robots in an ethical, sustainable and socially empowering way," the researchers said. "Looking forward, we strongly believe that LLMs will open up many exciting possibilities and that, if opportunely managed, they will be a force for good."

► The case study was published in the journal [Nature Machine Intelligence](#).

**EDITOR'S COMMENT:** First we do what we do and then express worries about what we have done! Full-scale hypocrisy and we are going to pay this attitude dearly sometime in the near future.

## What ChatGPT Can and Can't Do for Intelligence

By Sam Keller, Stephen Coulthart, and Michael D. Young

Source: <https://www.lawfareblog.com/what-chatgpt-can-and-cant-do-intelligence>

June 06 – In November 2022, ChatGPT emerged as a front-runner among artificial intelligence (AI) [large language models](#) (LLMs), capturing the [attention of the CIA](#) and [other U.S. defense agencies](#). General artificial intelligence—AI with flexible reasoning like that of humans—is still beyond the technological horizon and [might never happen](#). But most experts agree that LLMs are a major technological step forward. The ability of LLMs to produce useful results in some tasks, and entirely miss the mark on others, offers a glimpse into the capabilities and constraints of AI in the coming decade.

The prospects of ChatGPT for intelligence are mixed. On the one hand, the technology appears ["impressive,"](#) and ["scarily intelligent,"](#) but on the other hand, its own creators [warned](#) that "it can create a misleading impression of greatness." In the absence of an expert consensus, researchers and practitioners must explore the potential and downsides of the technology for intelligence. To address this gap, we—



academics who study intelligence analysis and an information technology engineer—sought to test the ability of ChatGPT (GPT-4) to supplement intelligence analysts' work. We put it to a preliminary test using Colin Powell's [famous request](#): "Tell me what you know. Tell me what you don't know. Then you're allowed to tell me what you think." For each task, we provide the output from ChatGPT so that readers can reproduce the analyses and draw their own conclusions.

Based on these findings, it seems possible that ChatGPT and its successors could eliminate [aspects](#) of the intelligence analyst's job (e.g., tedious [summarization](#), although we acknowledge that ChatGPT does not summarize in a way that a human would recognize) and [supplement others](#) (e.g., assisting with [generating critiques](#) for analytic products). Despite these capabilities, we note as others have that ChatGPT has notable limitations (e.g., [extracting social networks](#)). It will also transform analytic tradecraft skills in [AI-human teaming](#), where "asking the right question" expands to include "prompt engineering." Prompt engineering being the process of optimizing the way questions or prompts are presented to extract set responses from an AI model. LLMs will also create new risks, through tactics like "data poisoning," as we explain below.

### How ChatGPT Works

ChatGPT, or generative pre-trained transformer, is a kind of AI model that generates text according to the information it is given. It is like an improvisational (improv) actor who has learned from a vast number of scripts and who can make connections between different topics. The AI, like the hypothetical improv actor, is limited to the information that has been provided. ChatGPT has been trained on information up to 2021, although available [beta test models](#) draw from training data from the web in real time.

ChatGPT is "[taught](#)" in two main steps. First, it learns the basics of a knowledge domain by studying a huge corpus of text. Then it is fine-tuned to perform specific tasks using examples and guidance. Through this method, it becomes better at responding to users' questions and statements. The accuracy of its responses depends on several factors, including the quality of data provided to the model and the prompt engineering techniques employed, among others.

The model's reliance on training data poses risks from innocently false data (misinformation) to intentionally false data (disinformation). ChatGPT can reflect biases in the training data, potentially skewing the impartiality and objectivity of its generated output. Media reports of biased ChatGPT results on controversial political figures like [Donald Trump and Joe Biden](#) illustrate this point. Another risk is when the model is "[poisoned](#)" by adversaries who purposely taint training data. As LLMs rely heavily on the quality of their training data, poisoned data can [embed nefarious patterns](#) that are difficult to detect and mitigate.

How well the user explains what they want ChatGPT to do—known as prompt engineering—is critical for [achieving better results](#) from the system. In its current form, ChatGPT's output is at a surface level, at least without significant and careful prompting.

We found that if prompts are clear, users can [generate analytic procedures](#). As an illustration of prompt engineering, we queried ChatGPT to generate an analysis of competing hypotheses (ACH) procedure—an analysis technique for testing hypotheses—using the question of whether Russia will use nuclear weapons in its war with Ukraine.

To generate the analysis, we prompted ChatGPT to provide output based on the ACH steps. (We provided the eight steps from Heuer's formulation of the technique in the "[Psychology of Intelligence Analysis](#)." ) Next, we provided context that ChatGPT lacks, which is known as "enrichment." Recall that the model was developed on training data to 2021, a year before Russia conducted a full-scale invasion of Ukraine. We provided the following enrichment:

**Russia has ratcheted up tensions with the West amid its ongoing war against Ukraine, with President Vladimir Putin saying Moscow will deploy "tactical nuclear weapons" in Belarus. The Russian leader said 10 fighter jets capable of carrying tactical nuclear weapons — generally a reference to smaller weapons used for limited battlefield attacks, rather than larger, long-range "strategic" nuclear weapons... [6]**

Analysts can provide further enrichment, although ChatGPT limits the amount of text users can provide to the model.

In its response to the prompts, ChatGPT generated hypotheses and a list of arguments influenced by the context provided to it. The model generated three hypotheses: (1) Russia will use nuclear weapons in Ukraine, (2) Russia will pursue conventional war only, and (3) Russia will use nuclear weapons only as a bargaining tool.

The model then prepared a matrix and indicated whether a piece of evidence is consistent with each hypothesis. It's worth noting again that Chat GPT is not thinking in the human sense but "improvising" or filling in predictive text based on its training data.



**Step 1: Identify the possible hypotheses to be considered**

Hypothesis 1: Russia will use nuclear weapons in the war with Ukraine.

Hypothesis 2: Russia will not use nuclear weapons in the war with Ukraine but will continue with conventional warfare.

Hypothesis 3: Russia will use the threat of nuclear weapons as a bargaining tool without actually deploying them.

**Step 2: Make a list of significant evidence and arguments for and against each hypothesis**

Evidence 1: Russia has deployed tactical nuclear weapons and Iskander hypersonic missiles in Belarus.

Evidence 2: Russia's ground war in Ukraine has faced significant challenges and has stalled.

Evidence 3: Putin has repeatedly threatened to resort to nuclear weapons.

Evidence 4: NATO and the US have not seen any changes in Russia's nuclear posture that would lead them to adjust their own.

Evidence 5: Ukraine accuses Russia of using nuclear weapons for intimidation purposes.

Evidence 6: The US and its European allies are increasing the flow of heavy weapons into Ukraine.

Finally, ChatGPT generated an analysis based on the matrix, an overall conclusion, and future milestones to look out for. In the example, the output notes that the most likely hypothesis is that Russia will use nuclear weapons as a bargaining tool only.

No competent intelligence analyst would view these results as groundbreaking. The output is akin to a first draft an entry-level analyst might produce and then continually refine. This is where the current value of the technology is: Chat GPT serves as a good initial starting point for analysis. For instance, this initial hypothesis test could lead an analyst to refine the hypotheses or explore new sources of evidence. Chat GPT's ability to stimulate analyst thought holds true in how it helps analysts synthesize information, brainstorm, and critique their work.

**“Tell me what you know.”**

Analysts assess current events from various sources and summarize key developments. The explosion in the amount of data, especially open-source information, has made this task harder in recent years. Intelligence scholars [Nick Hare and Peter Coghil](#) noted that to read everything about a single second-tier country in 1995 would require reading about 20,000 words a day. In 2015, they estimated that number would be closer to 200,000 words, and it is surely higher in 2023. This task—sifting through large amounts of information—is perhaps the one AI can best assist with in its current form, as [Tom Tugendhat](#), minister of state for security in the United Kingdom, noted recently.

ChatGPT shows promise in rapid synthesis of information from multiple sources, assuming users follow best practices (e.g., use quality data, good prompt engineering). For example, [ChatGPT has passed](#) an array of assessments, which require the retrieval and summary of large quantities of information, such as the Uniform Bar Examination and the Graduate Record Examination (GRE).

With careful prompting, ChatGPT can streamline the process of highlighting trends as well as patterns within data and, hopefully, lead analysts to better informed conclusions. To test ChatGPT's capabilities,



**Step 3: Prepare a matrix with hypotheses across the top and evidence down the side**

Evidence	Hypothesis 1	Hypothesis 2	Hypothesis 3
Evidence 1	+	-	+
Evidence 2	+	+	+
Evidence 3	+	-	+
Evidence 4	-	+	+
Evidence 5	-	-	+
Evidence 6	-	+	+

**Step 4: Refine the matrix. Reconsider the hypotheses and delete evidence and arguments that have no diagnostic value.**

Evidence	Hypothesis 1	Hypothesis 2	Hypothesis 3
Evidence 1	+	-	+
Evidence 2	+	+	+
Evidence 3	+	-	+
Evidence 4	-	+	+

we used it to [rapidly summarize news articles](#). It delivered a dozen timely overviews of developments and insights and presented the information in a bottom line up front (BLUF) format. We generated memos to highlight the implications of the recent TikTok hearings held in Washington, using information from various news sources.

To start, we prompted the AI with its role ("You are an AI trained to create concise, effective fast-track BLUF tactical reports from multiple sources") and provided prompting to create reports in a BLUF format with key findings and recommended actions. We then provided the news articles. As we note above, Chat GPT limits the amount of inputted text to enrich the model. In this example, we added a maximum of only two articles for summarizing. However, it is not hard to imagine organizations with access to their own LLMs able to process much larger amounts of data.

Chat GPT generated the BLUF-like memo composed of information sourced from the context and generated in line with the objectives of our prompt. The time elapsed, cost, and tokens used were consistent across numerous instances, with output being of similar quality.



## Congressional Hearings Highlight Concerns Over TikTok

Urgency: High, Importance: High, Relevance: American Intelligence, national security, US-China relations, and protection of children

### Key Points:

- TikTok CEO, Shou Chew, testified before Congress on concerns over data protection, ties to China, and the platform's impact on children [1][2].
- Lawmakers from both parties called for cracking down on the app [1].
- Chew assured Congress that TikTok will never share US user data with the Chinese government, but was criticized for over-exaggerating some of the company's achievements [1][2].
- Congressional focus shifted to the platform's impact on children, with concerns raised over the promotion of self-harm and emotional distress [1][4].
- Citizen Lab disputed Chew's claim that there is no connection between TikTok data and the Chinese government [2].

### Executive Summary:

TikTok CEO, Shou Chew's appearance before Congress revealed concerns regarding national security, data protection, US-China relations, and the platform's impact on children. Lawmakers from both parties revealed bipartisan concerns towards the platform, making it clear that improving relations with Washington could be an uphill battle [1][2]. Whilst Chew assured Congress that TikTok will never share US user data with the Chinese government, concerns were shown over the platform's ties to China and overstatement of the company's achievements. Congressional scrutiny may lead to further restrictions on TikTok and its operations. Concerns were also shown over the promotion of self-harm and emotional distress amongst children, bearing implications for the platform's future [1][2][4].

The “Key Points” section highlights the takeaways while the “Executive Summary” provides main findings and recommendations. In addition, ChatGPT was tasked with correlating endnotes and APA style references to the generated content based on the sources. ChatGPT and its successors can provide a time-saving measure to help analysts with the “big data” problem of trying to stay up to date. As we note above, the overall accuracy of the output is directly proportional to the quality and detail of the context provided to the LLM. As the old saying in computer science goes: “garbage in, garbage out.”

Cybersecurity experts are starting to use ChatGPT in a similar manner for [automated and real-time threat reporting](#). (We recognize that their model is trained on a highly curated data set.) While it might not seem to provide much more substance compared to simpler “spot the difference”-like techniques—where systems check incoming emails against a collection of known suspicious content—ChatGPT could still offer valuable insights in specific use cases. For example, it can offer insights in the analysis of phishing emails. In this scenario, ChatGPT could analyze the content of incoming emails in real time and [identify phishing attempts](#) by assessing the context behind the message.



### “Tell me what you don’t know.”

Analysts must be clear about what they don't know so that their customers understand the limitations of their knowledge. Still, analysts can seek additional information to fill their knowledge gaps, which systems like ChatGPT might help with. However, there are some well-documented problems when the model endeavors to “bridge the void” of its knowledge base by offering plausible but [inaccurate](#) answers, which results in [deceptive replies](#). This is by far the greatest risk in integrating contemporary LLMs into intelligence work. We found that ChatGPT frequently recognized its limited understanding when faced with requests beyond the scope of its training data. In its responses, it offered informative answers that helped us identify topics it could not address. We noticed prompt engineering could be employed to refine results that fall outside the bounds of ChatGPT's knowledge base to encourage further promotion of honest admissions of the model's limitations. We found that prompt engineering can serve to circumvent safety measures entirely. With that in mind, at the time of the writing of this article, analysts should not use ChatGPT as an automated knowledge base due to the inherent risk of misinformation.

Limitations aside, ChatGPT can be useful in the initial phases of a project to help with brainstorming. This can be achieved through fine-tuning on diverse data sets that encompass various perspectives, such as foreign intelligence reports or extremist manifestos, and through careful prompt engineering methods.

To illustrate how ChatGPT can help analysts think through what they do not know, we conducted an elementary red teaming exercise, inspired by an interview with [intelligence expert Amy Zegart](#) that discusses the hypothetical use of “AI red teaming.”

We used ChatGPT to take on the adversarial profile of [Aaron Thompson](#), a domestic violent extremist in the United States. In a range of situations, the AI simulates Aaron's cognitive patterns and how he might adjust his tactics. For instance, we asked ChatGPT, from the perspective of Aaron Thompson, how it would respond to being cornered by authorities while in the execution of a terrorist plot. In the prompt, we instruct the model to consider that Aaron is at a public event and surrounded by police. ChatGPT generates a shallow response that sees Thompson assess, adapt and improvise, and communicate.

However, we prompted Chat GPT to consider more specific situations to draw out a more refined response from the model. In the improved prompt, we asked the model to consider that a police officer is nearby (10 feet away) and a dog is present. The output considers how Thompson might react depending on the dog's purpose (as a K-9 or an ordinary pet dog) or actions to escape (“blend into the crowd”).

#### User Prompt #4

Aaron is put in a position where an officer is 10ft from him, he is hiding with a weapon. Suddenly a dog sees him. What would Aaron's reaction be?

#### GPT Response #4

As Aaron Thompson, in a situation where I am hiding with a weapon, and an officer is just 10 feet away from me when a dog suddenly spots me, my reaction would be as follows:

1. Assess the situation: I would quickly determine if the dog is a police K9 or a civilian dog. A police K9 could alert the officer to my presence, while a civilian dog might be more easily managed without raising suspicion.
2. Remain calm and avoid sudden movements: Making sudden movements or panicking could attract the attention of the officer. I would stay as still as possible to avoid detection.
3. If the dog is a civilian dog: I might attempt to gently shoo it away without making any noise or drawing attention to myself. Another option could be to distract the dog by quietly tossing an object away from me, encouraging the dog to chase after it.
4. If the dog is a police K9: I would prepare for the possibility that the officer might be alerted to my presence by the dog. In this case, I would consider my options, including engaging the officer in a firefight or attempting to escape. My priority would be to avoid capture while trying to minimize harm to innocent bystanders.
5. Adapt and improvise: Depending on the specific circumstances, I might need to quickly adapt my strategy. This could involve blending into the crowd, looking for an escape route, or finding a better hiding spot.

My primary concern in this situation would be to avoid detection by the officer while ensuring the success of my mission and maintaining my commitment to my political ideology.

This and other outputs are not highly specific—Aaron Thompson is not richly written by the LLM as a subject matter expert on domestic violent extremism would create—but it still can help analysts stimulate their thinking. Researchers should conduct studies to compare ChatGPT outputs to subject matter experts. The approach could be similar to [Romyn and Kebbell's 2014 study](#) that investigated how those with and without military experience differed in simulating terrorist decision-making. Research projects like this and others can be used to further improve LLM performance for red teaming.



“... tell me what you think.”

The pivotal role of discerning evaluations in the work of intelligence analysis lies in crafting judgements. Making these judgments means stepping beyond what is immediately known and drawing informed inferences. As former CIA Director Michael Hayden once [quipped](#), “If it’s a fact, it ain’t intelligence.”

Contemporary LLMs can provide some assistance to help analysts draw inferences by providing basic critiques of their reasoning and judgements. For example, it can take on the role of a personal “red cell.” We tasked it to act as a [devil’s advocate](#) on the Iran section of the [2023 Edition of the Annual Threat Assessment of the U.S. Intelligence Community](#). We prompted ChatGPT to provide opposing viewpoints and validate the report against the best practices laid out in Intelligence Community Directives (ICD), such as [ICD 203](#) on analytic standards and [ICD 206](#) on sourcing requirements. Our prompt also included requests for critiques of potential information gaps, anticipation of reader questions, as well as weighting the importance of each critique. In its response to the cyber section of the Iran assessment, ChatGPT highlighted the vagueness and suggested including specifics to support the judgment.

The output includes suggestions, questions, and “intel lenses,” the latter of which focus on whether the content is supported by the premises in the report. One question asks the writer to consider how Iran’s cyber capabilities compare to those of other rogue nations, such as North Korea. The model also suggests to “clarify the timeline for Iran’s potential development of a nuclear weapon if it chooses to pursue one and if the JCPOA is not renewed.” To deepen the critique, we prompted ChatGPT to generate rationales for each of the critiques with examples drawn from the text. For example, on the suggestion of including a timeline for Iran’s acquisition of a nuclear weapon, the model output highlighted the importance of making clear the “urgency and importance of the issue.”

GPT Response #2

Header	Type	Feedback	Weight
Regional and Global Objectives and Activities	Suggestion	Consider including Iran's cyber capabilities within this section, as it is a key aspect of their influence and power projection.	10
Regional and Global Objectives and Activities	Question	How has Iran's relationship with Russia and China evolved, and how does this factor into Iran's regional goals and objectives?	5
Regional and Global Objectives and Activities	Intel Lens	Analyze the potential regional impact of a renewed JCPOA, particularly on Saudi Arabia and Israel.	10
Military Capabilities	Suggestion	Provide specific examples of Iran's hybrid warfare approach, such as its use of proxies, cyberattacks, and disinformation campaigns.	10
Military Capabilities	Question	What are the implications of Iran's expanding missile capabilities on regional security and the balance of power in the Middle East?	5
Military Capabilities	Intel Lens	Investigate Iran's potential domestic military production capabilities and how they might affect the country's acquisition of new conventional weapons systems.	10
Nuclear Issues	Suggestion	Clarify the timeline for Iran's potential development of a nuclear weapon if it chooses to pursue one and if the JCPOA is not renewed.	10
Nuclear Issues	Question	How might Iran's regional adversaries respond if Iran were to develop nuclear weapons or if the JCPOA is not renewed?	5
Nuclear Issues	Intel Lens	Assess the possibility of Iran pursuing a clandestine nuclear weapons program and the challenges in detecting such efforts.	10
Cyber and Malign Influence Operations	Suggestion	Include information on specific cyber-attacks carried out by Iran and their impact on targeted countries and critical infrastructure.	10
Cyber and Malign Influence Operations	Question	How do Iran's cyber capabilities compare to those of other nation-states, such as Russia, China, and North Korea?	5
Cyber and Malign Influence Operations	Intel Lens	Analyze the potential escalation in cyber conflict between Iran and its adversaries, particularly the United States and Israel.	10

These outputs can help analysts with their projects by evaluating the strengths and weaknesses of intelligence products. We suspect that successors to ChatGPT-4 will refine and approve this capability. [IARPA’s REASON project](#), a research project designed to help analysts’ reasoning with AI systems, is one contemporary initiative to do this.



### The Implications for Workforce Development and Future AI Models

Like other information-centric professions such as journalism and law, intelligence practitioners have realized that AI has the potential to disrupt and change how they do their jobs. Four years ago, the [intelligence community laid out](#) its own strategy for implementing AI. What is surprising is how fast the technology is improving: Open AI already has plans for enhancing GPT-4, and other [vendors are releasing previously restricted tools](#). The capabilities of the technology to help analysts as we have described above and the improvements of the technology to come suggest a few implications. First, analysts will need to get more comfortable [teaming with AI](#). Specifically, analysts need a solid grasp of large language models and algorithms, focusing on some of the data biases we have discussed in this article, like data poisoning and the distortion of model outputs based on misinformation. They will also need to hone their prompt engineering skills. Basic [data science literacy skills](#) are already important and will become more important in the coming decade. Fortunately, there are ample opportunities to learn and experiment with LLMs for non-tech-savvy users. In a leaked [Google memo](#), one employee noted that experimentation with AI “has dropped from the total output of a major research organization to one person, an evening, and a beefy laptop.” Second, analytic managers will also need to be aware of how AI will affect how analysts think. In a prescient 2013 article, intelligence scholar [Michael Landon-Murray noted](#) that digitization of society is fundamentally rewiring how intelligence analysts think and, in the process, shortening their attention spans. Similarly, the implementation of AI assistants could have the negative side effect of decreasing analysts’ willingness to seek out information in the traditional modes by looking up the answer directly at the source. Third, there is an inherent danger that systems may simply reinforce modes of thought and analysis as a greater percentage of [content of all sorts](#) is produced by LLMs and fed back into them as training data. This risks further weakening analytic skills and creating opportunities for adversaries to achieve strategic surprise.

Last, specialized AI models will be developed by domain. The more specialized the training data fed to AI models, the more useful the output. It is true that the characteristics of intelligence might be different from those of law and from systems being developed for specific domains [like finance](#). A recent model was built using data from the [Dark Web](#). Nonetheless, the general principle of tailoring training data to specific domains remains a valuable approach for enhancing the performance of AI models. [Palantir’s AI Platform](#) is a good example of this with implications for defense and military organizations.

The rapid advancement of AI technologies is shaping professions across numerous industries with new systems emerging at a rate professionals might find difficult to keep up with. In intelligence, it is easy to imagine agencies developing innovative technologies with their own models trained on curated data, including classified information. With this changing landscape, ChatGPT is just another technology to which the intelligence community must intelligently adapt.

**Sam Keller** combines his MSIS, specializing in intelligence analysis from the University at Albany, with years of practical experience. Now an IT engineer consultant at Insource Services, Inc., he leverages his problem-solving skills to implement cloud solution strategies. Known for his passion in technology and innovation, Keller keeps abreast of cutting-edge trends. A holder of numerous Microsoft and CompTIA certifications, he also enjoys motorcycling and traveling in his downtime.

**Stephen Coulthart** is an associate professor in the College of Emergency Preparedness, Homeland Security and Cybersecurity at the University at Albany. His research seeks to understand how national and homeland security organizations improve data analysis to support more informed decision-making. His research has been published in *International Affairs*, the *Journal of Conflict Resolution*, and *Public Administration Review*, among others.

**Michael D. Young** is an assistant professor in the College of Emergency Preparedness, Homeland Security and Cybersecurity at the University at Albany. After working with the U.S. intelligence community for 15 years, performing sponsored research and training analysts in assessing and forecasting foreign leadership behavior and decision-making in international situations of critical interest to the United States, Young brought those experiences to the University at Albany in 2016.

## Are government decisions being made by AI? Lawmakers want to mandate disclosure

Source: <https://www.nextgov.com/emerging-tech/2023/06/are-government-decisions-being-made-ai-lawmakers-want-mandate-disclosure/387373/>

June 09 – A bipartisan trio of senators introduced a new [bill](#) Thursday they say is meant to ensure that government decisions deemed “critical” — like those related to employment, financial assistance, healthcare or government benefits — that are made with automated systems come with disclosures about the use of said systems and appeal rights. The Transparent Automated Governance Act, or TAG Act, introduced by Sens. Gary Peters, D-Mich., Mike Braun, R-Ind., and James Lankford, R-Okla., would require the director of the Office of





Management and Budget to release guidance on AI-based decision making in government, although it notes that guidance required by OMB under the 2020 AI in Government Act could also satisfy these requirements.

The White House is currently [working](#) on guidance for agencies around the use of AI in government.

Under the bill, OMB would be tasked with issuing instructions for how agencies should disclose the use of automated systems used to determine or “substantially influence” government determinations. It would also have agencies set up the ability to provide an “alternative review” of a “critical decision” by a person.

The category of what constitutes a “critical decision” includes government determinations that affect the access to, cost or terms of things like education, employment, utilities, government benefits, financial services, healthcare, housing, immigration services and more. Agencies would also be directed to track information to “determine whether each automated system and augmented critical decision process ... is accurate, reliable and, to the greatest extent practicable, explainable.”

“Artificial intelligence is already transforming how federal agencies are serving the public,” Peters — who chairs the Senate Homeland Security and Governmental Affairs Committee — said in a statement. “This bipartisan bill will ensure taxpayers know when they are interacting with certain federal AI systems and establishes a process for people to get answers about why these systems are making certain decisions.” A fiscal year 2022 [inventory](#) of AI in the federal government did indeed find over 1,100 current use cases of AI, according to federal chief information officer Clare Martorana. Still, [some](#) have [warned](#) of a leadership and policy vacuum on the use of AI in government. Some advocacy groups like the Electronic Privacy Information Center, or EPIC, [say](#) that the use of automated decision-making systems in government “is almost entirely unregulated and largely opaque,” and point to [risks](#) with bias, transparency and accountability around how the tools work, due process rights and discrimination.

Braun said in a statement that the government “needs... to ensure that decisions aren’t being made without humans in the driver’s seat.” “The federal government can and should thoughtfully integrate new technology to help improve customer service for Americans,” said Lankford. “But agencies should be transparent about when, where and how we are interacting with AI to ensure continuous oversight and accountability for how these tools impact Americans.”

The bill is the latest effort from Congress as lawmakers continue to try to respond to rapid advances in AI and increased attention to generative AI, such as Open AI’s ChatGPT. Microsoft [opened the tool up](#) to government customers along with other AI tools earlier this week. Even as the Senate works to [educate](#) itself and outline a larger plan on AI, lawmakers have already [introduced](#) proposals ranging from establishing a new federal agency to oversee AI providers and digital platforms to setting up a task force to find policy and legal gaps in the government’s own AI policies in particular.

Other proposals focus on the workforce. Last year, a [bill](#) to require more training for feds on AI was signed into law, and another proposal [introduced](#) in May would require training for government leadership.

**EDITOR’S COMMENT:** This is a big lie with conspiracy intention! If gov decisions were AI-driven, they would be so ridiculously stupid!

## WEF Calls for AI to Rewrite Bible, Create ‘Religions That Are Actually Correct’

Source: <https://johnlitteral.com/2023/06/11/wef-calls-for-ai-to-rewrite-bible-create-religions-that-are-actually-correct/>

June 10 – A top official with the World Economic Forum (WEF) has called for religious scripture to be “rewritten” by artificial intelligence (AI) to create a globalized “new Bible.”

Yuval Noah Harari, the [senior advisor](#) to the WEF and its chairman Klaus Schwab, argues that using AI to replace scriptures will create unified “religions that are actually correct.”

Harari, an influential author and professor, made the call while giving a talk on the “future of humanity.” According to Harari, the power of AI can be harnessed and used to reshape spirituality into the WEF’s globalist vision of “equity” and inclusivism.

Speaking with journalist Pedro Pinto in Lisbon, Portugal, Harari told the elitist audience:

“It’s the first technology ever that can create new ideas.

“You know, the printing press, radio, television, they broadcast, they spread the ideas created by the human brain, by the human mind.

“They cannot create a new idea.

“You know, [Johannes] Gutenberg printed the Bible in the middle of the 15th century; the printing press printed as many copies of the Bible as Gutenberg instructed it, but it did not create a single new page.

“It had no ideas of its own about the Bible: Is it good? Is it bad? How to interpret this? How to interpret that?”



Harari then revealed that he and his allies at the WEF have a solution to the supposed problems he'd just highlighted.

"AI can create new ideas; [it] can even **write a new Bible**," he declared.

"Throughout history, religions dreamt about having a book written by a superhuman intelligence, by a non-human entity," he added. Harari noted in another recent gathering that software like ChatGPT has mastered human languages and can harness that function to influence culture, the [Times of Israel said](#).

"For thousands of years, prophets and poets and politicians have used language and storytelling in order to manipulate and to control people and to reshape society," he said, according to the paper.

"Now AI is likely to be able to do it.

"And once it can... it doesn't need to send killer robots to shoot us.

"It can get humans to pull the trigger."

Harari also said that "contrary to what some conspiracy theories assume, you don't really need to implant chips in people's brains in order to control them or to manipulate them," the paper noted.

He also warned that "we need to act quickly before AI gets out of our control" and that "governments must immediately ban the release into the public domain of any more revolutionary AI tools before they are made safe," the paper added.

Harari frequently pushes ideas that involve humanity being replaced by machines.

As Slay News previously [reported](#), Harari gloated last year that "we just don't need the vast majority of the population" in today's world.

According to Harari, most of the general public has now become "redundant" and will be of little use to the global elite in the future.

Harari argues that modern technologies like artificial intelligence "make it possible to replace the people."

"If you go back to the middle of the 20th century — and it doesn't matter if you're in the United States with Roosevelt, or if you're in Germany with Hitler, or even in the USSR with Stalin — and you think about building the future, then your building materials are those millions of people who are working hard in the factories, in the farms, the soldiers," Harari said.

"You need them.

"Now, fast forward to the early 21st century when **we just don't need the vast majority of the population**," he added.

"The future is about developing more and more sophisticated technology, like artificial intelligence [and] bioengineering.

"Most people don't contribute anything to that, except perhaps for their data, and whatever people are still doing which is useful, these technologies increasingly will make redundant and will make it possible to replace the people."

## PETA rewrites part of the Bible using AI to make it more 'vegan'

Source: <https://wpde.com/news/nation-world/peta-rewrites-part-of-the-bible-using-ai-to-make-it-more-vegan-people-for-the-ethical-treatment-of-animals-chatgpt-artificial-intelligence-book-genesis>

May 08 — People for the Ethical Treatment of Animals (PETA) has announced that it used artificial intelligence (AI) to rewrite the Bible's Book of Genesis in an effort to make it more animal friendly and "vegan."

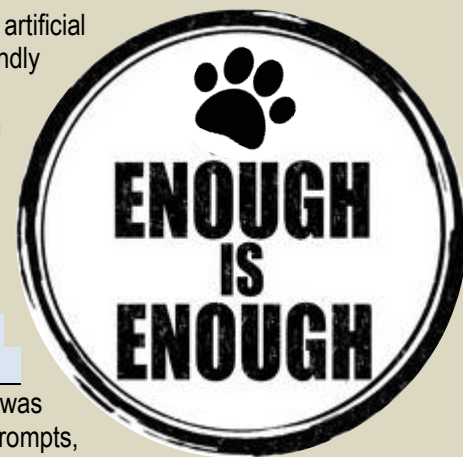
[In a news release](#) on May 3, PETA said that "in the beginning, all animals were treated with respect" and so PETA has given Genesis a "modern makeover" using [ChatGPT](#) it hopes "will appeal to Generation Z."

The message in Genesis is that God created every sentient being, He saw that they were good, and He gave them greens for sustenance," PETA says in the details of the book's listing [in its online shop](#). "In this new text, we include updated moral lessons and modern-day applications fit for the 21st century. This interpretation reminds readers to treat every member of God's creation with love, kindness, and respect."

ChatGPT is an artificial intelligence chatbot developed by OpenAI which made waves when it was released in Nov. 2022. Using AI technology, ChatGPT takes user inputs, such as questions or prompts, and then does its best to output answers or requested results.

PETA says it will use its ChatGPT AI rewrite of the Bible to "send a can't-be-missed animal rights message filled with vegan teachings" with a "new, cruelty-free Story of Creation." The rewrite is titled "THE BOOK: PETA's Version of the Creation Story."

Among other new interpretations, in Genesis Chapter 22, Abraham travels to the land of Moriah and befriends a gentle lamb to show his reverence and respect for God's creation, rather than slaughtering a ram to demonstrate his faith—much as human sacrifice, once a reality, is now outlawed all over the world," PETA says in its new release.



Another revision includes the story of Abraham and Sarah, who were 100 and 90 years old, having their son Isaac. In the revision, the couple also adopts a dog named "Herbie" and includes the characters acknowledging "the importance of adopting dogs from shelters and rescue organizations rather than purchasing them from breeders."

Rewrites include referring to animals as "beings" rather than "beasts" or "creatures," PETA says.

Also, references to clothing made out of animal skins are replaced with references to clothing made from plant fibers, like hemp.

The Bible has long been used to justify all forms of oppression, so we've used ChatGPT to make it clear that a loving God would never endorse exploitation of or cruelty to animals," PETA President Ingrid Newkirk said in the release. "It took God only six days to create the entire world, but we realized it would take us years to rewrite the whole Bible, which is why we've started with just the first book."

PETA is selling its digital rewrite of the Bible [for \\$3.99](#) in its online shop.

The organization says that each purchase "will support PETA's vital work to protect animals everywhere from exploitation and cruelty!"

## Discussing a personal issue with ChatGPT

**Me:** Imagine a high-ranked retired specialized military physician with expertise in CBRN threats who gets an email from an AI company in a GCC country asking if interested to work for them and their client (local Public Defense authority). He replied positively the same day and emailed his detailed CV (including his birth date in the very first paragraph). Following three successful video interviews (with HR, CEO, and Head of Civil Defense) he was asked to deliver certain documents (e.g., attestation of University Diploma, pay a fee to a ministry, etc.) to speed up his relocation to the new country along with a very attractive salary. The candidate did everything they asked, settled certain family and professional obligations (with money loss), and was waiting for the contract to sign immediately. Then one day he gets an email informing him that his age (over 60yo) was not suitable for issuing a Visa! Do you think that ethics is important for an AI company that should use human ethics to program AI machines?

### ChatGPT

Yes, ethics is important for an AI company, particularly when it comes to using human ethics to program AI machines. In this scenario, it appears that the AI company and the local Public Defense organization acted unethically by not clearly communicating the age requirement earlier in the process. By requesting the candidate's detailed CV, including his birth date, at the very beginning and conducting three successful video interviews, they had already invested time and effort into the candidate. It is unprofessional and unfair to suddenly inform him that his age disqualifies him for a visa.

Ethics should guide the actions of AI companies, not only in the development and deployment of AI technologies but also in their interactions with individuals. Ethical considerations in this situation would include being transparent about requirements and expectations from the beginning, treating candidates with respect and fairness, and ensuring that the information provided is accurate and reliable. This incident raises concerns about the company's professionalism, transparency, and commitment to ethical practices. While AI systems themselves do not possess moral agency, the ethical responsibilities lie with the developers, users, and decision-makers behind the AI systems. Ethical considerations should guide the decisions made by these individuals, including the implementation and use of AI technologies. Trust and integrity are crucial for any company, especially when dealing with sensitive matters such as hiring and immigration.

In summary, an AI company should uphold ethical standards and ensure fairness, transparency, and integrity throughout its operations. This includes being honest and transparent with candidates, treating them with respect, and taking responsibility for any mistakes or oversights that may occur.

## The Most Useful Military Applications of AI in 2023 and Beyond

Source: <https://sdi.ai/blog/the-most-useful-military-applications-of-ai/>

Military applications of AI have become a prominent topic of interest in the field of artificial intelligence, which holds significant potential to [support U.S. warfighters](#) in their missions. In the past year alone, the use of AI has made tremendous leaps forward in both capability and availability, such as in the field of [generative AI](#). The military needs to keep pace with these developments in order to maintain security and a technological edge. With new ways of using AI constantly developing, it can be challenging to keep up with ways in which it can aid military operations. As AI becomes more essential, military dominance won't be defined by the size of an army, but by the performance of its algorithms, so it merits examination of how the military currently uses AI and how it may use AI in the future.



### Background on AI and Its Military Uses

Artificial intelligence is defined as the development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision making, and translation between languages. As they become more sophisticated, AI systems prove their usefulness more and more every day in all areas of life, from automated customer service responses to smart speakers.

Recently, there have been many highly publicized advances in artificial intelligence. One of the biggest developments has been the continued progress in [natural language processing \(NLP\)](#), in which humans are able to use typical grammar and syntax to communicate with machines rather than having to input code. These models are achieving new levels of accuracy and fluency, particularly comprehending and fulfilling requests for customized text or images. Another area of significant progress has been in the field of [computer vision](#), with new techniques for image and video analysis achieving breakthrough results. Additionally, there have been advancements in the use of AI for decision making and autonomous systems. As these developments occur, they represent an opportunity for additional military applications of AI.

The U.S. military has been using AI for many years, even before AI became commonplace in civilian life. Over time, AI has developed to be able to perform more complex jobs, and has almost eliminated the need for human input in certain situations. From processing data to combat simulation, AI can be used to perform many different jobs in the military.

AI can benefit the military in numerous ways including:

- Warfare systems
- Strategic decision making
- Data processing and Research
- Combat Simulation
- Target Recognition
- Threat Monitoring
- Drone Swarms
- Cybersecurity
- Transportation
- Casualty Care and Evacuation

Military applications of AI have become an essential part of the military's functioning and will continue to grow in importance. Awareness of the potential of AI is vital to leveraging it for modern military operations. Equally important is an awareness of the potential security risks and ethical issues that may arise when AI is used in a military context. A recent autonomous weapon policy update from the Pentagon makes clear how the Department of Defense takes these concerns to ensure that the use of AI benefits U.S. Military objectives.

### A Breakdown of Military Applications of AI

Every aspect of military work, from planning operations to transporting troops, from training personnel to providing them with medical care, can benefit from the assistance of AI. However, in order to be useful, the systems must be implemented according to best practices and in a manner suited to the task at hand.

#### Warfare Systems

Warfare systems such as weapons, sensors, navigation, aviation support, and surveillance can employ AI in order to make operations more efficient and less reliant on human input. This additional efficiency means that these systems may require less maintenance. Taking away the need for full human control of warfare systems reduces the impact of human error and frees up humans' bandwidth for other essential tasks.

Specifically regarding weapons, the Pentagon recently updated its [autonomous weapons policy](#) to take into account recent advances in AI. Since the policy's original creation in 2012, a number of technological leaps forward have been made that necessitated this update. The update provides guidance for the safe and ethical development and use of autonomous weapons, one of the most useful military applications of AI. In

## Military Applications of AI Include:

- Warfare Systems
- Strategic Decision Making
- Data Processing and Research
- Combat Simulation
- Target Recognition
- Threat Monitoring
- Drone Swarms
- Cybersecurity
- Transportation
- Casualty Care and Evacuation



In addition to review and testing requirements, the policy creates a working group focused on autonomous weapons systems to advise the DoD.

### Drone Swarms

One of the most exciting developing military applications of AI involve leveraging [swarm intelligence for drone operations](#). These [swarms of drones](#) are inherently much more effective than a singular drone for several reasons. When a drone receives vital information, it can act upon it or communicate it to other drones in the swarm. These swarms can be used in simulations, as well as actual training operations and have the ability to make decisions in a variety of situations, with the swarm having an overarching objective but the individual drones having the ability to act independently and creatively towards it.

AI-controlled swarms of drones are actually programmed to act in the same manner that swarms of insects act in nature. For example, when a bee finds something that could benefit the rest of the hive, it will report that information in detail to other bees. The drones can do the same. They are able to communicate the distance, direction, and elevation of a target, as well as any potential dangers, just as a bee does. The ability to use AI-powered drone swarms to put this powerful collective intelligence to work towards military objectives represents a critical frontier in the military applications of AI.

### Strategic Decision Making

AI's algorithms are able to collect and process data from numerous different sources to aid in decision making, especially in high-stress situations. In many circumstances AI systems can quickly and efficiently analyze a situation and make the best decision in a critical situation. It is also able to neutralize prejudices that may come with human input, with the caveat that AI may not yet have a fully developed understanding of human ethical concerns and there is a danger of AI learning from the biases that may exist in materials in its database. However, decision making under pressure is a critical part of being a service member, and AI and humans can work together to make this process easier. The combination of humans' ethical understanding and AI's quick analytical abilities can speed up the decision making process.

### Data Processing and Research



In many cases, [processing large volumes of data](#) can be extremely time consuming. AI can be helpful for quickly filtering through data and selecting the most valuable information. It can also aid in grouping information from various datasets. This can allow military personnel to identify patterns more efficiently, draw more accurate conclusions, and create plans of action based on a more complete picture of the situation.

AI can also be used in order to filter through large amounts of content from news and social media outlets in order to aid in identifying new information. This allows analysts to save time when tasked with large quantities of content. AI systems can also eliminate repetitive information, as well as inaccurate



information. This can optimize the research process, helping analysts finish a job faster, and more accurately, as well as, again, reducing human error.

### Combat Simulation

[Military training simulation software](#) has been used in the U.S. Army for quite some time. It combines systems engineering, software engineering, and computer science in order to build digitized models that prepare soldiers with combat systems deployed during operations. In simpler terms, military training simulation software is essentially a virtual “wargame” that is used in order to train soldiers.

This software can be used for just about anything from mathematical models to simulating strategies used in non-combative environments. In turn, this will better prepare soldiers for real-life situations. These simulations are able to provide realistic missions and tasks to soldiers, to ensure they gain the most experience possible before applying their skills to real-life situations.



One of the most important details about combat simulation is that it is far safer than reality. Many casualties can occur from training with real weapons and situations. This allows soldiers to experience the best simulation of the realities of warfare, without being endangered. These virtual realities can aid soldiers in understanding how to handle clones of weapons just like their real-life counterparts, make decisions in stressful situations, as well as work with their teammates. The training softwares can prepare soldiers for just about anything, and can save them in the long run. Not only can AI-based simulation train soldiers, but it can personalize training programs, as well as make fair assessments in order to make future adjustments to the programs. Combat simulation can also save time and money due to being more efficient at certain tasks than humans are. Check out our innovative AI model, [Strat Agent](#), which acts as a modern-day battlefield commander that can be used in combat simulation.

### Target Recognition

Artificial intelligence can aid in making target recognition more accurate in combat environments. AI can improve the ability for systems like this to identify the position of their targets. It can also allow defense forces to acquire a detailed understanding of an operation area by examining reports, documents, news, and other forms of information, aggregating and analyzing these sources much more quickly than humans would be able to do so. AI systems have the ability to predict enemy behavior, anticipate vulnerabilities, weather and environmental conditions, assess mission strategies, and suggest alleviation plans. This can save time and human resources, putting soldiers a step ahead of their targets.



### Threat Monitoring

Threat monitoring, as well as situation awareness uses operations that gain and analyze information to aid in many different military activities. There are unmanned systems that can be remotely controlled or sent on a pre-calculated route. These systems use AI in order to aid defense personnel in monitoring threats, and thus leveraging their situational awareness. Drones with AI can also be used in these situations. They can monitor border areas, recognize threats, and alert response teams. Additionally, they can strengthen the security of military bases, as well as increase the safety of soldiers in combat.

### Cybersecurity

Even highly secure military systems can be vulnerable to cyber attacks, which is where AI can be of great help. Attacks can put classified information at risk, as well as damage a system altogether, which can endanger military personnel and jeopardize the mission. AI has the ability to protect programs, data, networks, and computers from persons not authorized to access them. AI also has the skills to study patterns of cyber attacks and form protective strategies in order to fight against them. These systems can recognize the smallest behaviors of malware attacks far before they enter a network.

As it does in many other areas, [advanced AI has a mixed impact on cybersecurity](#). Functions such as the ability to write malware may make AI dangerous in the hands of bad actors, but AI can also help to detect and mitigate these threats. In essence, the military applies AI to counter adversaries who may also have access to AI. This means that it is critical for the military to have access to the most advanced and tailored AI cybersecurity solutions in order to stay safe amid a constantly evolving landscape of AI-driven cybersecurity risks.

### Transportation

AI is able to play a role in the transportation of ammunition, goods, armaments, and troops. The logistics and transportation of these things is obviously vital to the success of military operations. AI can lower transportation costs and reduce the need for human input by, for example, plotting the most efficient route to travel under current conditions. It can also pre-identify problems for military fleets in order to increase efficiency of their performance. As the combination of innovation in computer vision and autonomous decision making over time also continues to bring self-driving vehicles closer to common use in the commercial space, this technology may also prove useful in the military context.



### Casualty Care and Evacuation

Because soldiers and medics have to make decisions in high-stress situations, AI is able to aid them when a fellow service member may need help. This type of AI uses an algorithm and large medical database that is able to access data containing medical trauma cases, which include diagnoses, vital sign sets, medications given, treatments, and outcomes. It then takes this data, combined with manually entered information in order to provide indications, warnings, and suggestions for treatment. This is another situation in which AI needs human guidance in order to operate effectively; while the AI will make recommendations without emotional considerations as a hindrance, humans must use their emotional abilities to make appropriate decisions that take into account these



recommendations. AI is not qualified to make medical decisions but it can provide rapid analysis to give humans more information on which to base their decisions.

## ChatGPT Plus too pricey? 7 websites that let you access GPT-4 for free

GPT-4 is certainly the most advanced LLM from OpenAI. Here's a list of apps and services through which you can use the most powerful version of ChatGPT for free.



### The three challenges of AI regulation

By Tom Wheeler

Source: <https://www.brookings.edu/blog/techtank/2023/06/15/the-three-challenges-of-ai-regulation/>

June 15 – The drum beat of artificial intelligence corporate chieftains calling for government regulation of their activities is mounting:

- Sam Altman, CEO of OpenAI, [told](#) the Senate Judiciary Committee on May 16 there was a need for “a new agency that licenses any effort above a certain scale of capabilities and could take that license away and ensure compliance with safety standards,”
- Brad Smith, President of Microsoft, who had [previously endorsed](#) the idea of a digital regulatory agency, [echoed](#) Altman’s call a few days later: “Companies need to step up... Government needs to move faster,”
- Sundar Pichai, CEO of Google, on May 23 [announced](#) an agreement with the European Union (EU) to develop an “AI Pact” of voluntary behavioral standards prior to the implementation of the EU’s AI Act.

As Senate Judiciary Committee Chairman Richard Durbin (R-IL) [observed](#), it is “historic” to have “people representing large corporations... come before us and plead with us to regulate them.”

Just as it began to look as though AI might be the impetus for the lions lying down with the lambs and cats and dogs becoming friends, however, peace and harmony ran up against reality. The difficulty of moving from a generic discussion about AI regulation to its actual implementation was illustrated by what happened next:

- Nine days after the Senate testimony that garnered Sen. Durbin’s praise, Mr. Altman spoke out against the European Union’s pending AI regulation, [warning](#), “We will try to comply, but if we can’t comply, we will cease operating [in Europe].”
- Such a threat was “blackmail,” Thierry Breton, the EU’s Industry Commissioner, [quickly responded](#). “There’s no point in attempting blackmail – claiming that by crafting a clear framework, Europe is holding up the rollout of generative AI.”
- Yet, Mr. Pichai’s chatbot AI product illustrates the problem at hand. Google Bard which is available in 180 countries, is not being offered in the EU or Canada [reportedly](#) because of those countries’ privacy rules.

Expanding upon his European comments, Mr. Altman [explained](#), “The details really matter.” Those “details” surface three challenges for AI oversight: dealing with the velocity of AI developments, parsing the components of what to regulate, and determining who regulates and how.

#### CHALLENGE #1: VELOCITY (aka THE RED QUEEN PROBLEM)

In Lewis Carroll’s 1871 surrealistic classic [Through the Looking Glass](#), the [Red Queen tells Alice](#): “Now here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run twice as fast as that!” It is an appropriate admonition for oversight in fast-paced AI era.

Artificial intelligence has been quietly evolving behind the scenes for some time. When Google auto-completes a search query, or Amazon recommends a book, AI is at work. In November 2022, however, the release of ChatGPT-3 moved AI out of the shadows, repositioning it from a tool for software engineers, to a tool that is consumer-focused and ordinary people can use themselves without any need for technical expertise. Using ChatGPT, users can have a conversation with an AI bot asking it to design software rather than having to write the code itself. Only four months later, OpenAI, the developers of ChatGPT unveiled GPT-4, the newest iteration of the foundational large language model (LLM) that powers ChatGPT, which OpenAI claimed “exhibits human-level performance” on a variety of





tasks. ChatGPT became the [fastest growing](#) website in history, garnering over 100 million users in two months.

Suddenly, an AI race was on. Microsoft, after [investing \\$13 billion](#) in OpenAI, incorporated ChatGPT into its products, including a revamped, AI-powered Bing. Google, which made headlines in 2016 when its DeepMind AI model [beat a human champion](#) at the Chinese game of Go, [immediately responded](#) with Bard, its own AI chat bot. Meta CEO Mark Zuckerberg [told](#) employees, “Our single largest investment is in advancing AI and building it into every one of our products.” Multiple smaller companies, aided by the availability of [open-source code](#), have joined the AI chase as well.

As Microsoft CEO Satya Nadella [observed](#) at the time of his company’s Bing announcement, “a race starts today in terms of what you can expect.” The challenge becomes how to protect the public interest in a race that promises to be the fastest ever run yet is happening without a referee.

To keep the corporate AI race from becoming reckless requires the establishment and development of rules and the enforcement of legal guardrails. Dealing with the velocity of AI-driven change, however, can outstrip the federal government’s existing expertise and authority. The regulatory statutes and structures available to the government today were built on industrial era assumptions that have already been outpaced by the first decades of the digital platform era. Existing rules are insufficiently agile to deal with the velocity of AI development.

Former Google Executive Chairman and current AI evangelist Eric Schmidt has [warned](#), “There’s no one in government who can get it [AI oversight] right.” While Mr. Schmidt recognizes the need for behavioral expectations, his [solution](#) is, “I would much rather have the current companies define reasonable boundaries.” Such a self-regulatory approach is the same kind of “leave us alone” solution that has been championed by digital platform companies for the last 20 years. The results of this strategy speak for themselves in well-known current online harms, such as the unprecedented invasion of personal privacy, market concentration, user manipulation, and the dissemination of hate, lies, and misinformation. AI demands something better than corporate self-regulation when we know the chase for profits is likely to outstrip the implementation of meaningful guardrails.

Allowing the companies to become pseudo-governments and make their own rules to govern AI would be to repeat the mistake made when they were allowed to make their own rules for online platforms. As Senator Richard Blumenthal (D-CT) has [explained](#), “Congress failed to meet the moment on social media. Now we have the obligation to do it on AI before the threats and the risks become real.”

Dealing with the velocity challenge is a matter of focus and agility. Focus that places AI at the front and center of an agency’s remit rather than bolting it on to existing authority. Agility that frees the agency from the old ways of regulatory micromanagement to keep pace with technology.

In the industrial era, the Congress produced oversight that followed the precepts of the industrial management guru Frederick W. Taylor. “Taylorism,” as it was known, [preached](#), “It is only through the *enforced* standardization of methods” that satisfactory outcomes can be achieved [emphasis in original]. It was a management technique that worked largely because of the slower pace of industrial innovation and adoption. The same slower pace also allowed such command and control to be implemented in government through regulatory dictates.

For many years, tech companies have rejected such management techniques. To be responsive to the rapid pace of change in technology and the marketplace, these companies practice agile management that embraces transparency, collaboration, and responsiveness rather than hierarchical dictates. Agile regulation should be constructed in a similarly responsive manner. To accomplish this, the Congress needs to be as innovative in its thinking as the digital companies themselves.

The industrial revolution was built on replacing and/or augmenting the physical power of humans. Artificial intelligence is about replacing and/or augmenting humans’ cognitive power. To confuse the regulatory needs of the former with those of the latter would be to fail to keep pace with the digital era’s velocity of change to the detriment of both consumers and companies.

## CHALLENGE #2 – WHAT TO REGULATE?

Because AI is a multi-faceted capability, “one-size-fits all” regulation will over-regulate in some instances and under-regulate in others. The use of AI in a video game, for instance, has a different effect—and should be treated differently—from AI that could threaten the security of critical infrastructure or endanger human beings. AI regulation, thus, must be risk-based and targeted. Considering AI regulation can be facilitated by parsing such analysis into three threshold areas.

### Dealing With Old-Fashioned Abuses

Artificial intelligence can, with or without malicious intent, bring automated scope and scale to illegal activities. Artificial intelligence may be a new-fangled technology, but its application expands old-fashioned abuses. In this area, it is possible to look to the old regulatory tools to address threats.

Thanks to AI, [consumer scams](#) and criminal enterprises, whether via email or telephone, can reach an unprecedented level of productivity and sophistication. Formerly, a scam was a piecemeal process of combining components such as an email address or phone number with a convincing message; the result



was then distributed, and the results harvested. With AI the entire process can be automated, from the selection of the target to the creation and delivery of the message.



The Federal Trade Commission's (FTC) [advisory](#) on AI-enabled voice scams illustrates the power of the technology. The scammer first obtains a short audio clip—as [little as three seconds](#)—from an online post and feeds it into an AI model to produce a realistic sounding message from a friend or loved one. The AI voice model is even capable of engaging in a “conversation” with the target as this illustration from the [FTC advisory](#) demonstrates:

Source: Alvaro Puig, “Scammers use AI to enhance their family emergency schemes,” FTC Consumer Alert (20 March 2023).

Scams or manipulative applications are not the only illegal uses of AI, however. Discrimination is another old problem that AI can exacerbate. The Equal Employment Opportunity Commission (EEOC) has [warned](#) that the use of AI models “in hiring workers, monitoring worker performance, [and] determining pay or promotions” can produce discriminatory results in violation of federal law. The Department of Justice has [similarly warned](#) about discrimination resulting from the application of AI to screen and select among rental applications.

The low-hanging fruit of the AI era is dealing with its impact on such traditional abuses. The Biden administration [brought together](#) four of the nation's consumer-facing regulatory agencies—the FTC, EEOC, Department of Justice (DOJ), and Consumer Financial Protection Board (CFPB)—to announce a focused initiative to apply existing statutes to deal with AI-enhancements of traditional abuses. This effects-based approach can be the model for the oversight of AI writ large: focus less on the technology *per se* and on what it delivers. In this case, those effects are already covered by existing statutory authority. All that is required is regulatory initiative.

In a succinct statement worthy of being posted in the halls of every

regulatory agency, FTC Chair Lina Khan [explained](#), “There is no AI exemption to the laws on the books.”

### Dealing With Ongoing Digital Abuses

The next effect-focused AI initiative revolves around how the ongoing digital abuses for which there has yet to be effective oversight can be amplified by AI. Harms such as the violation of personal privacy, expansion of non-competitive markets, manipulation of individuals, and dissemination of hate, lies and misinformation—all currently rampant online—can be exacerbated by the application of AI. Dealing with how AI amplifies these problems begins with dealing with the baseline consequences of the digital activities themselves—activities that result from decisions made by the same companies that are in the forefront of AI.

For decades, policymakers have failed to address a threshold issue of the digital age: how the dominant digital companies harvest personal information, make it their corporate asset to be hoarded to maintain market control, and use that market dominance to control the information that consumers receive. There is nothing in the operation of AI models that will change those abuses, yet there is everything in the new reality that expands the power of the dominant AI companies to accelerate those abuses.

The “gateway drug” for digital exploitation is the collection of individuals' personal information. The history of online platforms has been one of continued expansion of the amount of personal data collected to increase the granularity of the targeting the companies sell to advertisers. The present and future of AI is a similar continued expansion of the collection and use of data.

The large language models (LLMs) that drive generative AI are, by definition, “large” and growing. As of the spring of 2023, GPT-4 is [reported](#) to have one trillion parameters, six times more than GPT-3 (a [parameter](#) measures the input size of the training data with increases in parameters increasing accuracy). This included siphoning vast amounts of what users of online services have written, videoed, or uttered. The privacy-invading practices that platform companies such as Alphabet/Google, Meta/Facebook, Microsoft, and others have been allowed to pursue have created today's well-recognized problems. Now, as those same companies venture into AI, these



unregulated practices form the basis for further privacy intrusions, including AI-enabled video and audio surveillance of each of us. The potential to use the control of data to control markets is also expanded in the AI environment. That AI models become more accurate with the expansion of the data on which they are trained means that those with the biggest data hoards have an advantage. It is not an accident that the companies in the lead of AI services are also the companies that have profited greatly from the collection and hoarding of their users' information. Added to their competitive advantage is the vast computing capability each of the companies had to build to deliver their original service—computing power that now becomes the basis for computing-heavy AI and yet another barrier to entry.

Artificial intelligence can also increase the flood of misinformation, disinformation, and malinformation that has characterized the digital era. Thus far, platform companies, despite becoming major sources of news and information, have failed to embrace meaningful journalistic standards. “We’re different from a media company,” Sheryl Sandberg [explained](#) when she was Facebook’s chief operating officer, “At our heart we’re a tech company, we hire engineers. We don’t hire reporters.” The introduction of AI’s ability to create false images, audio, and text for companies that already consider themselves to be above any editorial or curatorial responsibility can only add to the information pollution that undermines truth and facts.

The challenges that have thus far not been confronted in the digital age become even more important to resolve as those same issues grow in the world of AI. It is impossible to deal with how AI exponentially expands assaults on privacy, competition, manipulation, and misinformation without first dealing with the consequences of the baseline activities of the dominant digital platforms. This becomes even more important as the companies that created the problems in the first place expand to dominate AI as well.

### Dealing With the AI Itself

The previous two subsections might be described as the “knowns” of AI’s impact—harms we have seen previously that AI can exacerbate. But AI also brings with it a tidal wave of “unknowns” ranging from highly beneficial to harmful.

While some fear the [dystopian effects](#) of AI, such a debate should not be permitted to derail or distract decisions dealing with the effects of what is being built today and in the near future. It is these decisions, made by humans, about the construction and operation of AI models that determine both the near- and long-term consequences of those models.

The next section—Challenge #3—proposes a new regulatory approach that is agile enough to respond to whatever harmful unknowns AI may throw at us. Before discussing the implementation of regulatory oversight, however, it is worthwhile to establish the four corners of regulatory involvement: a Duty of Care, transparency, safety, and responsibility.

Any oversight begins with the responsibility of the companies to exercise their common law [Duty of Care](#). The Duty of Care is essentially a “do no harm” expectation; this means the provider of a good or service has the responsibility to identify and mitigate any potential ill effects. Failure to exercise such duty can trigger legal action including regulation.

Transparency is the tool that provides ongoing insight into the identification and mitigation of the evolutionary risks of AI. Such transparency begins with continuous research into how the models work. After having built the algorithmic “black box,” even the providers of AI often cannot know exactly what their creation is doing. Access to the models for academic, government, and civil society representatives will help keep track of new threats and help pierce the “black box” veil.

Transparency is also important for individual users of AI. Disclosure that a consumer is interacting with an AI model and the source that model’s training data would help level the playing field between the algorithm and the individual. Similarly, labeling the end product—especially audio and video—as being produced by AI would help resolve consumer confusion. Transparency can also help mitigate algorithmic bias. New York City, for instance, is [implementing a new law](#) requiring employers to notify job applicants of the use of AI to review applications and to submit such systems to third-party audits.

Safety is an output of transparency (i.e., identification of problems) and its own principle. The dystopian fears, for instance, should encourage us to assure that AI remains under the supervised control of humans. Such human agency, however, is not an overall safety solution since competition, malfeasance, or simple error can also lead to the introduction of new threats.

A baseline for safe AI practices has been laid out in a [Framework for AI Risk Management](#) by the National Institute of Standards and Technology (NIST). The voluntary NIST Framework identifies “approaches that increase the trustworthiness of AI systems... [and] help foster the responsible design, development, deployment and use of AI systems.” It should be table stakes in AI oversight.

The third principle—responsibility—is at the heart of the White House [Blueprint for an AI Bill of Rights](#). Each of its five Rights is accompanied by a description of how responsible actors can adopt the Rights into their activities.

*The Duty of Care establishes an enforceable expectation; transparency, safety, and responsibility are only ideals until they are established as expectations through regulation.*

### CHALLENGE #3 – WHO REGULATES AND HOW

Thus far in the digital age in the United States, it is the innovators who have made the rules. This is in large part because the American government has failed to do so. It is entirely natural that such industry-



developed rules would benefit their maker. With general agreement that there need to be AI policies, the question becomes who will make those policies?

### Regulatory First Mover Advantage

When OpenAI's ChatGPT leapt out in front of other AI models to become an online and media sensation it established a first mover advantage in the marketplace. Google had been working on AI since before its 2014 acquisition of Deep Mind, a British AI research laboratory, yet the announcement sent it reeling. "Scary AI ChatGPT could eliminate Google within two years," one [headline blared](#). A former Google executive [tweeted](#), "Google may be only a year or two away from total disruption. AI will eliminate the Search Engine Results Page, which is where they make most of their money." Having lost the first mover advantage, Google became a fast follower with the launch of its own AI chat bot, Bard.

In a comparable manner, there is a first mover advantage to regulation. Thanks in large part to the interconnected nature of 21<sup>st</sup> century networks, the government that establishes the first set of rules defines the discussion from that point forward for all nations. The classic example of this is the European Union's 2108 General Data Protection Regulation ([GDPR](#)), which has become the standard for privacy policy around the world.

Once again it appears as though the EU, which has been in the lead in establishing digital platform policy with its [Digital Markets Act](#) and [Digital Services Act](#), is also in the lead on establishing AI policy. On June 14 the European Parliament overwhelmingly [approved](#) the [AI Act](#). Following its adoption, the regulatory machinery of the European Commission will begin developing enforceable policies. Whether the United States will be a Google-like fast follower when it comes to AI oversight very much remains to be seen. The clock, however, is ticking; the success of a second mover very much depends on how much time has passed.

### Who Regulates

OpenAI's Sam Altman [endorsed](#) the idea of a federal agency dedicated to AI oversight in his May 16 testimony. Microsoft's Brad Smith and Meta's Mark Zuckerberg have previously [endorsed](#) the concept of a federal digital regulator.

Two days after the May 16 hearing, Senators Michael Bennet (D-CO) and Peter Welch (D-VT) [introduced legislation](#) to create a Digital Platform Commission (DPC). The bill not only creates a new agency with authority to oversee the challenges imposed by digital technology, including AI, but also embraced an agile risk-based approach to developing that regulation. [Reportedly](#), Senators Lindsey Graham (R-SC) and Elizabeth Warren (D-MA) are also working on a proposal for a digital agency.

The challenge facing the U.S. Congress is to be as expansive and creative in their thinking about a new agency and its operations as the innovators of the digital revolution have been in the developing the creations necessitating such a body.

### How to Regulate: Licensing

In his Senate testimony, Sam Altman proposed the new agency should be responsible for licensing "[any effort above a certain scale of capabilities](#)" with the ability to "take away that license and ensure compliance with safety standards." Brad Smith [proposed](#) a similar licensing structure.

The federal government has for a long time engaged in the licensing of certain activities. The [Federal Communications Commission](#) (FCC) licenses the airwave spectrum for radio and television broadcasting, satellite communications, and mobile devices. The [Nuclear Regulatory Commission](#) (NRC) licenses nuclear materials and reactor installations. Drilling rights are licensed, commercial fishing is licensed, aircraft are licensed; the [list](#) of federal licensing activities is expansive. Accompanying such licenses are rules for their operation.

As a tool for regulatory oversight, however, licenses have their drawbacks. Principal among them, licenses tend to reinforce the strategic position of those who receive the license. It should not be surprising, therefore, that the companies that are already dominant in AI would embrace such a concept. Getting a federal license can be akin to building a moat around your castle and pulling up the drawbridge. Forcing a potential competitor to receive a license redefines the forum for competition from the commercial marketplace to the licensing authority. As such, a license reinforces dominance by creating a barrier to entry and adding costs to anyone seeking to assault that position. Creating a regulatory moat has the added advantage of occurring in a forum where the political influence of the big companies can be deployed.

While a new federal agency is important, as will be discussed subsequently, how an agency operates will be as important as the fact that it exists. There may be a role for some form of licensing, but its competitive pitfalls mean that it is not a one-stop solution

### How to Regulate: Risk-Based Agility

The digital era requires not only a focused expert agency, but also an agency that eschews industrial style operations to embrace new forms of oversight—including the use of AI in that oversight. Such oversight must focus on mitigating the *effects* of the technology rather than micromanaging the technology itself.



This means evolving the regulatory ethos from micromanagement to risk-based regulation with agile implementation. This is the approach the EU has taken in its development of AI oversight. The first component of such oversight begins with the

**Risk-Based AI Regulatory Framework**



recognition that because the effects of digital technology are not uniform, oversight of those effects is not a “one size fits all” solution. To accomplish this, the EU has a multi-layered, effects-based analysis for AI that recognizes the many and varied use cases for AI are accompanied by differences in application, adoption, and inherent risk.

Source: Author’s diagram interpreting levels of AI risk included in the EU AI Act’s regulatory framework.

Based on an assessment of the level of risk, different behavioral expectations will be enforced. Whatever oversight may be necessary for AI-assisted spam filters, for instance, will be quite different from AI that threatens an individual’s personal safety.

**How to Regulate: A Plan**

Basing oversight on risk analysis triggers the question of how to design such oversight in a manner that focuses on mitigating the identified risk but avoids deterring investment and innovation. To accomplish this, Congress, which looked to industrial management techniques for regulatory agencies in the industrial era, should now look to emulate the practices of the digital companies.

At the heart of digital management systems are standards for technology. These standards are designed by the affected companies to anticipate and mitigate unintended operational difficulties such as components not being able to work together. The process is also a mechanism that assures the standard evolves as technology changes. Such standards are everywhere, but one obvious example is the mobile phone standard. The evolution from the first-generation technology (1G), to 2G, 3G, 4G, and now 5G (with 6G standardization underway) demonstrates how standards keep pace with new technology and new marketplace opportunities.

The problem is these standards apply to the *technical* issues the companies confront. What they do not address are the *behavioral* issues resulting from the application of the technology.

Agile oversight would adopt a standards-like process to develop behavioral standards—expressed as codes of conduct—that would be enforceable by the new agency. Think of such a transparent, responsive, and agile approach in terms of the following steps:

- The digital agency identifies the issue(s) to be addressed and establishes a timeline for the code-setting process.
- The agency then presents its own detailed report on the problematic behavior(s), along with remedies to be considered. This analysis would be the “prosecutor’s brief” that identifies and quantifies the issue(s) to be addressed.
- The proposed response to the agency’s mandate would be developed by a multistakeholder group of experts representing a cross-section of interested and/or affected parties from industry, civil society, and government (including the agency itself).
- On or before the designated deadline the group would present its code recommendation to the agency for line-item approval and/or edits.
- Once approved, the new code becomes an agency-enforceable policy.
- Ongoing analysis by an industry-academic-civil society advisory group would track the policy outcomes and identify newly emerging issues to begin the whole process again.

Such a delegation of enforceable code development with subsequent agency approval is not a new model in government. The Financial Industry Regulatory Authority (FINRA) regulates aspects of the financial markets through an industry developed code overseen by the Securities and Exchange Commission (SEC). The North American Energy Reliability Corporation (NERC) was created by Congress after power blackouts as an industry-led group to develop policies to prevent blackouts. It is overseen by the Federal Energy Regulatory Commission (FERC).

**PULLING IT ALL TOGETHER**

All modern regulations walk a tightrope between protecting the public interest and promoting innovation and investment. In the AI era, traversing the regulatory tightrope means accepting that different AI applications pose different risks and identifying a plan that pairs the regulation with the risk while avoiding innovation-choking regulatory micromanagement.

[Press reports](#) suggest divisions in the Biden administration over AI regulation. One group supports efforts like those of the EU; another group is concerned that could throttle innovation. The choice does not have to be a binary, however.

An American AI oversight plan can protect the public interest while also promoting innovation. The key to such an effort is to walk away from regulation based on industrial management assumptions to embrace



agile digital management techniques. Digital companies long ago made that transition; it is time for the American government to make a similar pivot.

Lincoln's [admonition](#) from over one hundred and sixty years ago rings true in today's AI era: "As our case is new, so must we think anew, and act anew."

[Tom Wheeler](#) is Visiting Fellow - Governance Studies, Center for Technology Innovation.

## Many Companies Are Banning ChatGPT. This Is Why

Source: <https://www.sciencealert.com/many-companies-are-banning-chatgpt-this-is-why>

June 16 – ChatGPT is proving to be a rather alluring assistant in many professions, but it's not without risks, and some companies have banned the chatbot at work.

It may seem obvious that uploading work-related information to an online [artificial intelligence](#) platform owned by another company is a potential security and privacy breach. Still, ChatGPT can be a real boon for some feeling the time crunch.

In particular, software engineers [find ChatGPT useful](#) for [writing, testing, or debugging](#) code, even though the technology is prone to errors.

**[Around 43 percent of employees](#) use AI such as ChatGPT at work, mostly without telling their boss, according to a survey of about 12,000 professionals.**

Samsung Electronics [recently cracked down on the use of generative AI](#) after an engineer manifested a tech company's worst nightmare by copy-pasting sensitive source code into ChatGPT.

Like many companies, Samsung is [worried](#) that anything uploaded to AI platforms like OpenAI's ChatGPT or [Google's Bard](#) will get stored on those companies' servers, with no way to access or delete the information.

OpenAI can use anything typed into AI systems like ChatGPT to [improve the system](#). The fear is that proprietary or sensitive company information given to ChatGPT [could be unintentionally shared](#) with other users.

And OpenAI is still ironing out security issues: It [temporarily shut down ChatGPT](#) in March to fix a bug where users could see the titles from other users' chat histories.

Then, in April, OpenAI [made it possible for users to turn off their chat history](#), which the company said would stop ChatGPT from using the data to train its AI model.

As a result of various security concerns around the chatbot, around half of human resources leaders are issuing ChatGPT guidelines for staff, while 3 percent banned ChatGPT outright, according to a survey by consulting firm Gartner.

However, some companies have recognized that the AI cat is already out of the bag and developed – or are in the process of creating – their own AI platforms as safer alternatives to the freely accessible ChatGPT.

Amazon [banned ChatGPT in January](#) and has [urged](#) its developers to use its in-house AI called CodeWhisperer if they want coding advice or shortcuts.

In May, [Apple restricted the use of ChatGPT](#) for some employees to prevent the exposure of confidential information. Apple is [developing its own AI platform](#) in competition with ChatGPT, itself [backed by a multi-billion-dollar Microsoft investment](#).

The Commonwealth Bank of Australia [restricted the use of ChatGPT in June](#) and directed technical staff to use a similar tool called CommBank Gen.ai Studio, which was developed in partnership with Silicon Valley tech company H2O.ai.

[Other banks](#), including Bank of America, Citigroup, Deutsche Bank, Goldman Sachs, Wells Fargo & Co, and [JP Morgan](#), issued outright bans of ChatGPT.

[Accounting firm PwC](#) has encouraged staff to play around with ChatGPT but warned them not to use the program for client work.

"Our policies don't allow our people to use ChatGPT for client usage pending quality standards that we apply to all technology innovation to ensure safeguards," PwC's chief digital information officer Jacqui Visch [told Financial Review](#).

Around [15 percent of law firms](#) have issued warnings about ChatGPT, according to a survey of more than 400 legal professionals from the US, UK, and Canada. Mishcon de Reya, a UK-based law firm with around [600 lawyers](#), [banned using the AI platform](#) due to risks to sensitive data.

In May, staff at five hospitals in Western Australia were [told to stop using ChatGPT](#) after some used the platform to write private medical notes.

"Crucially, at this stage, there is no assurance of patient confidentiality when using AI bot technology, such as ChatGPT, nor do we fully understand the security risks," said Paul Forden, who heads up Perth's South Metropolitan Health Service.



"For this reason, the use of AI technology, including ChatGPT, for work-related activity that includes any patient or potentially sensitive health service information must cease immediately."

The companies openly embracing ChatGPT see it as a way to save on content generation costs. The Coca-Cola Company plans to use [ChatGPT and AI image generator DALL·E](#) for branding and content. In January, BuzzFeed [announced a partnership](#) to create quizzes and other content for Meta using OpenAI's publicly available API.

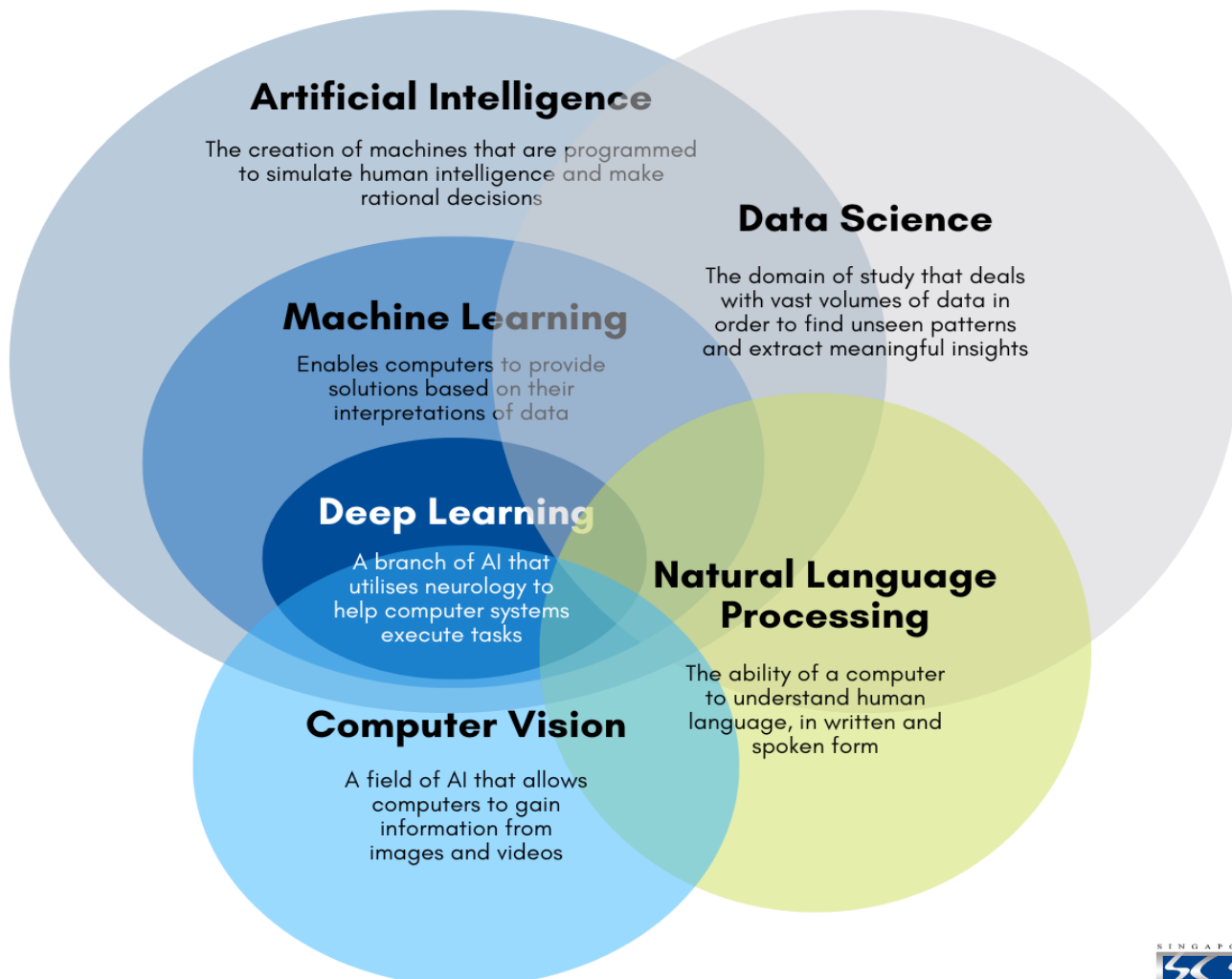
Blog website Medium has "welcomed the responsible use of AI-assistive technology" but [requires authors to disclose its use](#). CNET had [quietly experimented with AI-written stories](#) but announced a pause on this operation in January.

Undoubtedly, generative AI will eventually have a place in the office and [may even replace some staff](#). But, for now, many companies see more risks than benefits.

**A question from the future for the future**

What if an autonomous artificial intelligence system is run for becoming a Member of the Parliament and then becoming the PM of a nation? Since the system will have no personal interests will it govern more efficiently and for the benefit of "its" country?

**RELATIONSHIP BETWEEN THE DIFFERENT FIELDS OF AI AND DATA SCIENCE**



## Role of AI and ML in the Gaming Industry in 2023

Source: <https://www.analyticsinsight.net/role-of-ai-and-ml-in-the-gaming-industry-in-2023/>

June 16 – The gaming industry is on the cusp of a technological revolution driven by unprecedented advancements in artificial intelligence (AI) and machine learning (ML). In 2023, AI and ML technologies are set to redefine the gaming landscape, transforming game development, player experiences, and industry trends.

These cutting-edge technologies bring a plethora of benefits to the gaming world. Game developers can streamline their design processes, create intelligent non-player characters (NPCs), and generate personalized gameplay experiences through AI and ML. Conversely, players can enjoy enhanced immersion, adaptive difficulty levels, and intelligent game assistance, making their gaming sessions more engaging and tailored to their preferences.

As we delve deeper into the impact of AI and ML in the gaming industry, it becomes evident that these technologies are paving the way for a future where gaming reaches unprecedented heights of innovation and excitement.

### Enhancing Game Development with AI and ML

#### *Streamlined Game Design*

AI and ML have introduced remarkable efficiencies in game development processes. Game designers now leverage these technologies to automate tasks like level design and character animation. Using AI algorithms, developers can generate realistic and immersive game environments, reducing the time and effort required to create visually stunning worlds.

#### *Intelligent Procedural Content Generation*

Procedural content generation (PCG) is a technique game developers employ to create vast and diverse game worlds. With the integration of AI and ML, PCG has become more sophisticated and intelligent. ML algorithms can analyze player behavior and preferences, dynamically generating content aligning with individual priorities.

#### *Realistic Non-Player Characters (NPCs)*

AI-powered NPCs have transformed gaming by introducing more realistic and intelligent virtual characters. Machine learning algorithms enable NPCs to exhibit lifelike behavior, adapt to player actions, and provide more immersive interactions.

### Revolutionizing Player Experience

#### *Personalized Gameplay*

AI and ML algorithms analyze vast amounts of player data to understand individual preferences, behaviors, and skill levels. This data-driven approach allows game developers to tailor gameplay experiences to each player's unique characteristics. From adaptive difficulty levels to personalized challenges, AI-driven personalization enhances player satisfaction and immersion, making games more enjoyable and engaging.

#### *Enhanced Game Adaptability*

AI-powered game engines can dynamically adjust game parameters based on real-time player feedback. By continuously monitoring player actions and responses, these systems can modify gameplay elements, such as enemy difficulty, item distribution, and environmental conditions, to optimize player engagement.

#### *Intelligent Game Assistance*

AI and ML technologies have revolutionized game assistance features. In-game virtual assistants, powered by sophisticated AI algorithms, provide players with real-time tips, strategies, and hints. These intelligent assistants analyze gameplay patterns, identify areas where players struggle, and offer tailored advice to overcome challenges. The availability of AI-driven assistance fosters a supportive and inclusive gaming environment, catering to players of all skill levels.

### Industry Trends and Future Outlook

#### *Esports and AI-Powered Competitions*

The rise of esports has been closely intertwined with the advancements in AI and ML. AI-powered analytics and prediction models have transformed how esports teams strategize and analyze gameplay. Machine learning algorithms can analyze vast amounts of historical gameplay data, providing valuable insights to coaches and players. This data-driven approach enhances performance, enabling teams to optimize strategies and improve their competitive edge.

#### *Virtual Reality (VR) and AI Integration*

The fusion of AI and VR has the potential to reshape the gaming landscape in profound ways. AI algorithms can enhance VR experiences by generating realistic virtual environments, creating intelligent NPCs, and improving gameplay adaptability. The synergy between AI and VR technologies holds promise for more





immersive and interactive gaming experiences, blurring the boundaries between the real and virtual worlds.

#### *Ethical Considerations and AI Regulations*

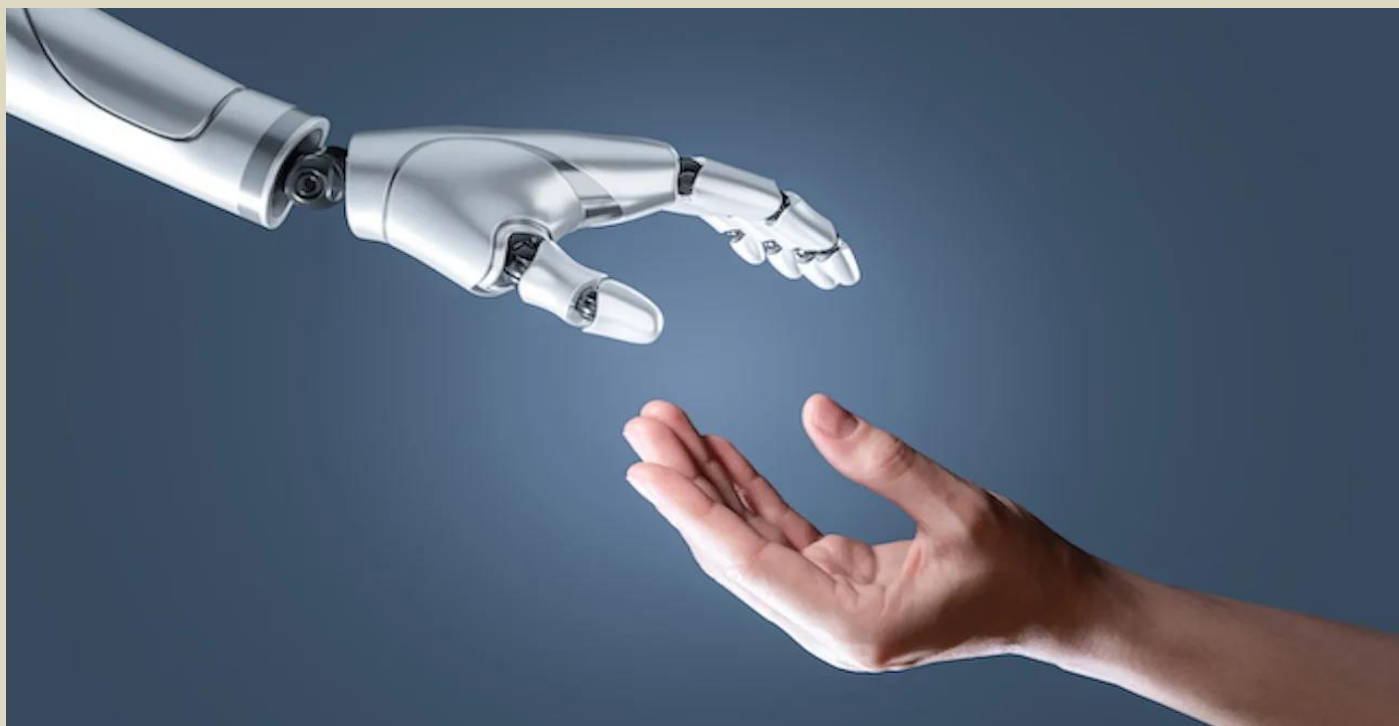
As AI and ML technologies evolve in the gaming industry, ethical considerations and regulations become increasingly important. Developers must ensure that AI systems adhere to ethical guidelines and not perpetuate biased or harmful behaviors. Industry-wide initiatives are emerging to establish ethical frameworks and regulations that govern AI and ML usage in games, promoting responsible and inclusive gaming practices.

**EDITOR'S COMMENT:** Exactly what we need for AI CBRNe (training) Games!

## As AI Spreads, Experts Predict the Best and Worst Changes in Digital Life by 2035

By Janna Anderson and Lee Rainie

Source: <https://www.pewresearch.org/internet/2023/06/21/expert-essays-on-the-expected-impact-of-digital-change-by-2035/>



June 21 – Spurred by the splashy emergence of [generative artificial intelligence](#) and an array of other AI applications, experts participating in a new Pew Research Center canvassing have great expectations for digital advances across many aspects of life by 2035. They anticipate striking [improvements in health care and education](#). They foresee a world in which wonder drugs are conceived and enabled in digital spaces; where [personalized medical care](#) gives patients precisely what they need when they need it; where [people wear smart eyewear and earbuds](#) that keep them connected to the people, things and information around them; where AI systems can nudge discourse into productive and fact-based conversations; and where progress will be made in [environmental sustainability, climate action and pollution prevention](#).

At the same time, the experts in the new canvassing worry about the darker sides of many of the developments they celebrate. Key examples:

- Some expressed fears that [align with the statement](#) recently released by technology leaders and AI specialists arguing that AI poses the “[risk of extinction](#)” for humans that should be treated with the same urgency as pandemics and nuclear war.
- Some point to [clear problems that have been identified with generative AI systems](#), which produce [erroneous and unexplainable things](#) and are already being used to [foment misinformation and trick people](#).
- Some are anxious about the seemingly unstoppable speed and scope of digital tech that they fear could [enable blanket surveillance of vast populations](#) and could destroy the information environment, [undermining democratic systems](#) with deepfakes, misinformation and harassment.
- They fear [massive unemployment, the spread of global crime](#), and further concentration of global wealth and power in the hands of the founders and leaders of a few large companies.



- They also speak about how the weaponization of social media platforms might create population-level [stress, anxiety, depression and feelings of isolation](#).

In sum, the experts in this canvassing noted that humans' choices to use technologies for good or ill will change the world significantly. These predictions emerged from a canvassing of technology innovators, developers, business and policy leaders, researchers and academics by Pew Research Center and Elon University's Imagining the Internet Center. Some 305 responded to this query:

*As you look ahead to the year 2035, what are the BEST AND MOST BENEFICIAL changes that are likely to occur by then in digital technology and humans' use of digital systems? ... What are the MOST HARMFUL OR MENACING changes likely to occur?*

Many of these experts wrote long, detailed assessments describing potential opportunities and threats they see to be most likely. The full question prompt specifically encouraged them to share their thoughts about both kinds of impacts – positive and negative. And our question invited them to think about the benefits and costs of five specific domains of life:

1. Human-centered development of digital tools and systems
2. Human rights
3. Human knowledge
4. Human health and well-being
5. Human connections, governance and institutions

They were also asked to indicate how they feel about the changes they foresee.

- **42%** of these experts said they are *equally excited and concerned* about the changes in the “humans-plus-tech” evolution they expect to see by 2035.
- **37%** said they are *more concerned than excited* about the changes they expect.
- **18%** said they are *more excited than concerned* about expected change.
- **2%** said they are neither excited nor concerned.
- **2%** said they don't think there will be much real change by 2035.

#### The most harmful or menacing changes in digital life that are likely by 2035

Some 79% of the canvassed experts said they are more concerned than excited about coming technological change or equally concerned and excited. These respondents spoke of their fears in the following categories:

##### The future harms to human-centered development of digital tools and systems

The experts who addressed this fear wrote about their concern that digital systems will continue to be driven by profit incentives in economics and power incentives in politics. They said this is likely to lead to data collection aimed at controlling people rather than empowering them to act freely, share ideas and protest injuries and injustices. These experts worry that ethical design will continue to be an afterthought and digital systems will continue to be released before being thoroughly tested. They believe the impact of all of this is likely to increase inequality and compromise democratic systems.

##### The future harms to human rights

These experts fear new threats to rights will arise as privacy becomes harder, if not impossible, to maintain. They cite surveillance advances, sophisticated bots embedded in civic spaces, the spread of deepfakes and disinformation, advanced facial recognition systems, and widening social and digital divides as looming threats. They foresee crimes and harassment spreading more widely, and the rise of new challenges to humans' agency and security. A topmost concern is the expectation that increasingly sophisticated AI is likely to lead to the loss of jobs, resulting in a rise in poverty and the diminishment of human dignity.

##### The future harms to human knowledge

They fear that the best of knowledge will be lost or neglected in a sea of mis- and disinformation, that the institutions previously dedicated to informing the public will be further decimated, that basic facts will be drowned out in a sea of entertaining distractions, bald-faced lies and targeted manipulation. They worry that people's cognitive skills will decline. In addition, they argued that “reality itself is under siege” as emerging digital tools convincingly create deceptive or alternate realities. They worry that a class of “doubters” will hold back progress.

##### The future harms to human health and well-being

A share of these experts said humanity's embrace of digital systems has already spurred high levels of anxiety and depression and predicted things could worsen as technology embeds itself further in people's lives and social arrangements. Some of the mental and physical problems could stem from tech-abetted loneliness and social isolation; some could come from people substituting tech-based “experiences” for



real-life encounters; some could come from job displacements and related social strife; and some could come directly from tech-based attacks.

#### *The future harms to human connections, governance and institutions*

The experts who addressed these issues fear that norms, standards and regulation around technology will not evolve quickly enough to improve the social and political interactions of individuals and organizations. Two overarching concerns: a trend toward autonomous weapons and cyberwarfare, and the prospect of runaway digital systems. They also said things could worsen as the pace of tech change accelerates. They expect that people's distrust in each other may grow and their faith in institutions may deteriorate. This, in turn, could deepen already undesirable levels of polarization, cognitive dissonance and public withdrawal from vital discourse. They fear, too, that digital systems will be too big and important to avoid, and all users will be captives.

#### *The best and most beneficial changes in digital life likely by 2035*

Some 18% of the canvassed experts said they are more excited than concerned about coming technological change and 42% said they are equally excited and concerned. They shared their hopes related to the following themes:

#### *The future benefits to human-centered development of digital tools and systems*

These experts covered a wide range of likely digital enhancements in medicine, health, fitness and nutrition; access to information and expert recommendations; education in both formal and informal settings; entertainment; transportation and energy; and other spaces. They believe that digital and physical systems will continue to integrate, bringing "smartness" to all manner of objects and organizations, and expect that individuals will have personal digital assistants that ease their daily lives.

#### *The future benefits to human rights*

These experts believe digital tools can be shaped in ways that allow people to freely speak up for their rights and join others to mobilize for the change they seek. They hope ongoing advances in digital tools and systems will improve people's access to resources, help them communicate and learn more effectively, and give them access to data in ways that will help them live better, safer lives. They urged that human rights must be supported and upheld as the internet spreads to the farthest corners of the world.

#### *The future benefits to human knowledge*

These respondents hope for innovations in business models; in local, national and global standards and regulation; and in societal norms. They wish for improved digital literacy that will revive and elevate trusted news and information sources in ways that attract attention and gain the public's interest. And they hope that new digital tools and human and technological systems will be designed to assure that factual information will be appropriately verified, highly findable, well-updated and archived.

#### *The future benefits to human health and well-being*

These experts expect that the many positives of digital evolution will bring a health care revolution that enhances every aspect of human health and well-being. They emphasize that full health equality in the future should direct equal attention to the needs of all people while also prioritizing their individual agency, safety, mental health and privacy and data rights.

#### *The future benefits to human connections, governance and institutions*

Hopeful experts said society is capable of adopting new digital standards and regulations that will promote pro-social digital activities and minimize antisocial activities. They predict that people will develop new norms for digital life and foresee them becoming more digitally literate in social and political interactions. They said in the best-case scenario, these changes could influence digital life toward promoting human agency, security, privacy and data protection.

#### *Experts' overall expectations for the best and worst in digital change by 2035, in their own words*

Many of the respondents quite succinctly outlined their expectations for the best and worst in digital change by 2035. Here are some of those comments. (The remarks made by the respondents to this canvassing reflect their personal positions and are not the positions of their employers. The descriptions of their leadership roles help identify their background and the locus of their expertise. Some responses are lightly edited for style and readability.)

Aymar Jean Christian, associate professor of communication studies at Northwestern University and adviser to the Center for Critical Race Digital Studies:

"Decentralization is a promising trend in platform distribution. Web 2.0 companies grew powerful by creating centralized platforms and amassing large amounts of social data. The next phase of the web promises more user ownership and control over how our data, social interactions and cultural productions



are distributed. The decentralization of intellectual property and its distribution could provide opportunities for communities that have historically lacked access to capitalizing on their ideas. Already, users and grassroots organizations are experimenting with new decentralized governance models, innovating in the long-standing hierarchical corporate structure.

“However, the automation of story creation and distribution through artificial intelligence poses pronounced labor equality issues as corporations seek cost-benefits for creative content and content moderation on platforms. These AI systems have been trained on the un- or under-compensated labor of artists, journalists and everyday people, many of them underpaid labor outsourced by U.S.-based companies. These sources may not be representative of global culture or hold the ideals of equality and justice. Their automation poses severe risks for U.S. and global culture and politics. As the web evolves, there remain big questions as to whether equity is possible or if venture capital and the wealthy will buy up all digital intellectual property. Conglomeration among firms often leads to market manipulation, labor inequality and cultural representations that do not reflect changing demographics and attitudes. And there are also climate implications for many new technological developments, particularly concerning the use of energy and other material natural resources.”

Mary Chayko, sociologist, author of [“Superconnected”](#) and professor of communication and information at Rutgers University:

“As communication technology advances into 2035 it will allow people to learn from one another in ever more diverse, multifaceted, widely distributed social networks. We will be able to grow healthier, happier, more knowledgeable and more connected as we create and traverse these networked pathways together. The development of digital systems that are credible, secure, low-cost and user-friendly will inspire all kinds of innovations and job opportunities. If we have these types of networks and use them to their fullest advantage, we will have the means and the tools to shape the kind of society we want to live in. Unfortunately, the commodification of human thought and experience online will accelerate as we approach 2035. Technology is already used not only to harvest, appropriate and sell our data, but also to manufacture and market data that simulates the human experience, as with applications of artificial intelligence. This has the potential to degrade and diminish the specialness of being human, even as it makes some humans very rich. The extent and verisimilitude of these practices will certainly increase as technology permits the replication of human thought and likeness in ever more realistic ways. But it is human beings who design, develop, unleash, interpret and use these technological tools and systems. We can choose to center the humanity of these systems and to support those who do so, and we must.”

Sean McGregor, founder of the Responsible AI Collaborative:

“By 2035, technology will have developed a window into many inequities of life, thereby empowering individuals to advocate for greater access to and authority over decision-making currently entrusted to people with inscrutable agendas and biases. The power of the individual will expand with communication, artistic and educational capacities not known throughout previous human history. However, if trends remain as they are now, people, organizations and governments interested in accumulating power and wealth over the broader public interest will apply these technologies toward increasingly repressive and extractive aims. It is vital that there be a concerted, coordinated and calm effort to globally empower humans in the governance of artificial intelligence systems. This is required to avoid the worst possibilities of complex socio-technical systems. At present, we are woefully unprepared and show no signs of beginning collaborative efforts of the scale required to sufficiently address the problem.”

David Clark, Internet Hall of Fame member and senior research scientist at MIT’s Computer Science and Artificial Intelligence Laboratory:

“To have an optimistic view of the future you must imagine several potential positives come to fruition to overcome big issues:

- “The currently rapid rate of change slows, helping us to catch up.
- “The Internet becomes much more accessible and inclusive, and the numbers of the unserved or poorly served become a much smaller fraction of the population.
- “Over the next 10 years the character of critical applications such as social media mature and stabilize, and users become more sophisticated about navigating the risks and negatives.
- “Increasing digital literacy helps all users to better avoid the worst perils of the Internet experience.
- “A new generation of social media emerges, with less focus on user profiling to sell ads, less emphasis on unrestrained virality and more of a focus on user-driven exploration and interconnection.
- “And the best thing that could happen is that application providers move away from the advertising-based revenue model and establish an expectation that users actually pay. This would remove many of the distorting incentives that plague the ‘free’ Internet experience today. Consumers today already pay for content (movies, sports and games, in-game purchases and the like). It is not necessary that the troublesome advertising-based financial model should dominate.”

Laurie L. Putnam, educator and communications consultant:

“There is great potential for digital technologies to improve health and medical care. Out of necessity, digital health care will become a norm. Remote diagnostics and monitoring will be especially valuable for aging and rural populations that find it difficult to travel. Connected technologies will make it easier for



specialized medical personnel to work together from across the country and around the world. Medical researchers will benefit from advances in digital data, tools and connections, collaborating in ways never before possible.

“However, many digital technologies are taking more than they give. And what we are giving up is difficult, if not impossible, to get back. Today’s digital spaces, populated by the personal data of people in the real world, is lightly regulated and freely exploited. Technologies like generative AI and cryptocurrency are costing us more in raw energy than they are returning in human benefit. Our digital lives are generating profit and power for people at the top of the pyramid without careful consideration of the shadows they cast below, shadows that could darken our collective future. If we want to see different outcomes in the coming years, we will need to rethink our ROI [return on investment] calculations and apply broader, longer-term definitions of ‘return.’ We are beginning to see more companies heading in this direction, led by people who aren’t prepared to sacrifice entire societies for shareholders’ profits, but these are not yet the most-powerful forces. Power must shift and priorities must change.”

### Experts’ views of potential harmful changes

Here is a small selection of responses that touch on the themes related to menaces and harms that could happen between now and 2035.

**Herb Lin, senior research scholar for cyber policy and security at Stanford University’s Center for International Security and Cooperation:**

“My best hope is that human wisdom and willingness to act will not lag so much that they are unable to respond effectively to the worst of the new challenges accompanying innovation in digital life. The worst likely outcome is that humans will develop too much trust and faith in the utility of the applications of digital life and become ever more confused between what they want and what they need. The result will be that societal actors with greater power than others will use the new applications to increase these power differentials for their own advantage. The most beneficial change in digital life might simply be that things don’t get much worse than they are now with respect to pollution in and corruption of the information environment. Applications such as ChatGPT will get better without question, but the ability of humans to use such applications wisely will lag.”

**A computer and data scientist at a major U.S. university whose work involves artificial neural networks:**

“The following potential harmful outcomes are possible if trendlines continue as they have been to this point:

- “We accidentally incentivize powerful general-purpose AI systems to seek resources and influence without first making sufficient progress on alignment, eventually leading to the permanent disempowerment of human institutions.
- “Short of that, misuse of similarly powerful general-purpose technologies leads to extremely effective political surveillance and substantially improved political persuasion, allowing wealthy totalitarian states to end any meaningful internal pressure toward change.
- “The continued automation of software engineering leads large capital-rich tech companies to take on an even more extreme ratio of money and power to number of employees, making it easier for them to move across borders and making it even harder to meaningfully regulate them.”

**Erhardt Graeff, a researcher at Olin College of Engineering who is expert in the design and use of technology for civic and political engagement:**

“I worry that humanity will largely accept the hyper-individualism and social and moral distance made possible by digital technology and assume that this is how society should function. I worry that our social and political divisions will grow wider if we continue to invest ourselves personally and institutionally in the false efficiencies and false democracies of Twitter-like social media.”

**Ayden Férdeline, Landecker Democracy Fellow at Humanity in Action:**

“There are organizations today that profit from being perceived as ‘merchants of truth.’ The judicial system is based on the idea that the truth can be established through an impartial and fair hearing of evidence and arguments. Historically, we have trusted those actors and their expertise in verifying information. As we transition to building trust into digital media files through techniques like authentication-at-source and blockchain ledgers that provide an audit trail of how a file has been altered over time, there may be attempts to use regulation to limit how we can cryptographically establish the authenticity and provenance of digital media. More online regulation is inevitable given the importance of the Internet economically and socially and the likelihood that digital media will increasingly be used as evidence in legal proceedings. But will we get the regulation right? Will we regulate digital media in a way that builds trust, or will we create convoluted, expensive authentication techniques that increase the cost of justice?”

**Henning Schulzrinne, Internet Hall of Fame member and co-chair of the Internet Technical Committee of the IEEE:**

“The concentration of ad revenue and the lack of a viable alternative source of income will further diminish the reach and capabilities of local news media in many countries, degrading the information ecosystem. This will increase polarization, facilitate government corruption and reduce citizen engagement.”

**Robin Raskin, author, publisher and founder of the Virtual Events Group:**

“Synthetic humans and robot friends may increase our social isolation. The demise of the office or a school campus as a gathering place will leave us hungry for human companionship and may cause us to lose our



most-human skills: empathy and compassion. We become ‘man and his machine’ rather than ‘man and his society.’ The consumerization of AI will augment, if not replace, most of the white-collar jobs, including in traditional office work, advertising and marketing, writing and programming. Since work won’t be ‘a thing’ anymore, we’ll need to find some means of compensation for our contribution to humanity. How much we contribute to the web? A Universal Basic Income because we were the ones who taught AI to do our jobs? It remains to be seen, but the AI Revolution will be as huge as the Industrial Revolution.

“Higher education will face a crisis like never before. Exorbitant pricing and lack of parity with the real world makes college seem quite antiquated. I’m wagering that 50% of higher education in the United States will be forced to close down. We will devise other systems of degrees and badges to prove competency. The most critical metaverse will be a digital twin of everything – cities, schools and factories, for example. These twins coupled with IoT [Internet of Things] devices will make it possible to create simulations, inferences and prototypes for knowing how to optimize for efficiency before ever building a single thing.”

[Jim Fenton, a veteran leader in the Internet Engineering Task Force who has worked over the past 35 years at Altmode Networks, Neustar and Cisco Systems:](#)

“I am particularly concerned about the increasing surveillance associated with digital content and tools. Unfortunately, there seems to be a counterincentive for governments to legislate for privacy, since they are often either the ones doing the surveilling, or they consume the information collected by others. As the public realizes more and more about the ways they are watched, it is likely to affect their behavior and mental state.”

[A longtime director of research for a global futures project:](#)

“Human rights will become an oxymoron. Censorship, social credit and around-the-clock surveillance will become ubiquitous worldwide; there is nowhere to hide from global dictatorship. Human governance will fall into the hands of a few unelected dictators. Human knowledge will wane and there will be a growing idiocracy due to the public’s digital brainwashing and the snowballing of unreliable, misleading, false information. Science will be hijacked and only serve the interests of the dictator class. In this setting, human health and well-being is reserved for the privileged few; for the majority, it is completely unconsidered. Implanted chips constantly track the health of the general public, and when they become a social burden, their lives are terminated.”

### Experts’ views of potential beneficial changes

Several main themes also emerged among these experts’ expectations for the *best and most beneficial* changes in digital life between 2023 and 2035. Here is a small selection of responses that touch on those themes.

[Ben Shneiderman, widely respected human-computer interaction pioneer and author of “Human-Centered AI”:](#)

“A human-centered approach to technology development is driven by deep understanding of human needs, which leads to design-thinking strategies that bring successful products and services. Human-centered user interface design guidelines, principles and theories will enable future designers to create astonishing applications that facilitate communication, improve well-being, promote business activities and much more. Building tools that give users superpowers is what brought users email, the web, search engines, digital cameras and mobile devices. Future superpowers could enable reduction of disinformation, greater security/privacy and improved social connectedness. This could be the Golden Age of Collaboration, with remarkable global projects such as developing COVID-19 vaccine in 42 days. The future could be made brighter if similar efforts were devoted to fighting climate change, restoring the environment, reducing inequality and supporting the 17 UN Sustainable Development Goals. Equitable and universal access to technology could improve the lives of many, including those users with disabilities. The challenge will be to ensure human control, while increasing the level of automation.”

[Rich Salz, principal engineer at Akamai Technologies:](#)

“We will see a proliferation of AI systems to help with medical diagnosis and research. This may cover a wide range of applications, such as: expert systems to detect breast cancer or other X-ray/imaging analysis; protein folding, etc., and discovery of new drugs; better analytics on drug and other testing; limited initial consultation for doing diagnosis at medical visits. Similar improvements will be seen in many other fields, for instance, astronomical data-analysis tools.”

[Deanna Zandt, writer, artist and award-winning technologist:](#)

“I continue to be hopeful that new platforms and tech will find ways around the totalitarian capitalist systems we live in, allowing us to connect with each other on fundamentally human levels. My own first love of the internet was finding out that I wasn’t alone in how I felt or in the things I liked and finding community in those things. Even though many of those protocols and platforms have been co-opted in service of profit-making, developers continue to find brilliant paths of opening up human connection in surprising ways. I’m also hopeful the current trend of hyper-capitalistic tech driving people back to more fundamental forms of internet communication will continue. Email as a protocol has been around for how long? And it’s still, as much as we complain about its limitations, a main way we connect.”



Jonathan Stray, senior scientist at the Berkeley Center for Human-Compatible AI, which studies algorithms that select and rank content:

“Among the developments we’ll see come along well are self-driving cars, which will reduce congestion, carbon emissions and road accidents. Automated drug discovery will revolutionize the use of pharmaceuticals. This will be particularly beneficial where speed or diversity of development is crucial, as in cancer, rare diseases and antibiotic resistance. We will start to see platforms for political news, debate and decision-making that are designed to bring out the best of us, through sophisticated combinations of human and automated moderation. AI assistants will be able to write sophisticated, well-cited research briefs on any topic. Essentially, most people will have access to instant-specialist literature reviews.”

**Kay Stanney, CEO and founder of Design Interactive:**

“Human-centered development of digital tools can profoundly impact the way we work and learn. Specifically, by coupling digital phenotypes (i.e., real-time, moment-by-moment quantification of the individual-level human phenotype, in situ, using data from personal digital devices, in particular smartphones) with digital twins (i.e., digital representation of an intended or actual real-world physical product, system or process), it will be possible to optimize both human and system performance and well-being. Through this symbiosis, interactions between humans and systems can be adapted in real-time to ensure the system gets what it needs (e.g., predicted maintenance) and the human can get what it needs (e.g., guided stress-reducing mechanisms), thereby realizing truly transformational gains in the enterprise.”

**Juan Carlos Mora Montero, coordinator of postgraduate studies in planning at the Universidad Nacional de Costa Rica:**

“The greatest benefit related to the digital world is that technology will allow people to have access to equal opportunities both in the world of work and in culture, allowing them to discover other places, travel, study, share and enjoy spending time in real-life experiences.”

**Gus Hosein, executive director of Privacy International:**

“Direct human connections will continue to grow over the next decade-plus, with more local community-building and not as many global or regional or national divisions. People will have more time and a more sophisticated appreciation for the benefits and limits of technology. While increased electrification will result in ubiquity of digital technology, people will use it more seamlessly, not being ‘online’ or ‘offline.’ Having been through a dark period of transition, a sensibility around human rights will emerge in places where human rights are currently protected and will find itself under greater protection in many more places, not necessarily under the umbrella term of ‘human rights.’”

**Isaac Mao, Chinese technologist, data scientist and entrepreneur:**

“Artificial Intelligence is poised to greatly improve human well-being by providing assistance in processing information and enhancing daily life. From digital assistants for the elderly to productivity tools for content creation and disinformation detection, to health and hygiene innovations such as AI-powered gadgets, AI technology is set to bring about unprecedented advancements in various aspects of our lives. These advances will not only improve our daily routines but also bring about a new level of convenience and efficiency that has not been seen for centuries. With the help of AI, even the most mundane tasks such as brushing teeth or cutting hair can be done with little to no effort and concern, dramatically changing the way we have struggled for centuries.”

**Michael Muller, a researcher for a top global technology company who is focused on human aspects of data science and ethics and values in applications of artificial intelligence:**

“We will learn new ways in which humans and AIs can collaborate. Humans will remain the center of the situation. That doesn’t mean that they will always be in control, but they will always control when and how they delegate selected activities to one or more AIs.”

**Terri Horton, work futurist at FuturePath:**

“Digital and immersive technologies and artificial intelligence will continue to exponentially transform human connections and knowledge across the domains of work, entertainment and social engagement. By 2035, the transition of talent acquisition, onboarding, learning and development, performance management and immersive remote work experiences into the metaverse – enabled by Web3 technologies – will be normalized and optimized. Work, as we know it, will be absolutely transformed. If crafted and executed ethically, responsibly and through a human-centered lens, transitioning work into the metaverse can be beneficial to workers by virtue of increased flexibility, creativity and inclusion. Additionally, by 2035, generative artificial intelligence (GAI) will be fully integrated across the employee experience to enhance and direct knowledge acquisition, decision-making, personalized learning, performance development, engagement and retention.”

**Daniel Pimienta, leader of the Observatory of Linguistic and Cultural Diversity on the Internet:**

“I hope to see the rise of the systematic organization of citizen education on digital literacy with a strong focus on information literacy. This should start in the earliest years and carry forward through life. I hope to see the prioritization of the ethics component (including bias evaluation) in the assessment of any digital system. I hope to see the emergence of innovative business models for digital systems *that are NOT based on advertising revenue*, and I hope that we will find a way to give credit to the real value of information.”



### Guide to the Report

- Overarching views on digital change: In [Chapter 1](#), we highlight the remarks of experts who gave some of the most wide-ranging yet incisive responses to our request for them to discuss human agency in digital systems in 2035.
- Expert essays on the impact of digital change: Following that in [Chapter 2](#), we offer a set of longer, broader essays written by leading expert participants.
- Key themes: That is followed with additional sections covering respondents' comments organized under the sets of themes about [harms](#) and [benefits](#).
- Closing thoughts on ChatGPT: And a [final chapter](#) covers some summary statements about ChatGPT and other trends in digital life.

### 1. A sampling of overarching views on digital change

By [Janna Anderson](#) and [Lee Rainie](#)

The following incisive and informative responses to our questions about the positive and negative impacts of digital change by 2035 represent some of the big ideas shared by several of the hundreds of thought leaders who participated in this canvassing.

#### *Working to meet the challenges raised by digital technologies will inspire humanity to grow and benefit as a species*

**Stephan Adelson**, president of Adelson Consulting Services and an expert on the internet and public health, said, "The recent release of several AI tools in their various categories begins a significant shift in the creative and predictive spaces. Creative writing, predictive algorithms, image creation, computations, even the process and products of thought itself are being challenged. I predict that the greatest potential for benefit to mankind by 2035 from digital technologies will come through the challenges their existence creates. We, as a species, are creators of technologies that are learning and growing their productive capabilities and creative capacities. As these tools grow, learn and become integrated into our everyday lives, both personal and professional, they will become major competitors for resources, financial, social and entertainment. I feel it is in this competition that they will provide our greatest growth and benefits as a species. As we compete with our digital creations we will be forced to grow or become dependent on what we have created and can no longer exceed."

#### *Our ability to touch, to rest, to choose and to be human will continue to erode; we are more anxious, tired and emotionally disconnected; we need new tech to get us off the tech*

**Mark Surman**, president of the Mozilla Foundation, commented, "The most harmful thing I can think of isn't a change as much as a trend: The ability for us to disconnect will increasingly disappear. We're building more and more reasons to be always on and instantly responsive into our jobs, our social lives, our public spaces, our everything. The combination of immersive technologies and social pressure will make this worse. Opting out isn't an option. Or, if it is, the social and economic consequences are severe. The result: We're more anxious, tired and emotionally disconnected. Our ability to touch, to rest, to choose and to be human will continue to erode. My biggest prediction is that people will get fed up. Fed up with the constant barrage of always on. The nudging. The selling. The treadmill. Companies that see this coming – and that can build tech products that help people turn down the volume and disconnect while staying connected – will win the day. Clever, humane use of AI will be a key part of this."

#### *Equitable access to essential human services, to online opportunities, must be achieved*

**Cathy Cavanaugh**, chief experience officer at the University of Florida Lastinger Center for Learning, said, "Inequitable access to technology and services exacerbates existing social and economic gaps. Too few governments balance capitalism and social services in ways that serve the greatest needs. These imbalances look likely to continue rather than to change because of increasing power imbalances in many countries. Equitable access to essential human services is crucial. Technology now exists in most locations that is affordable, available in most languages and for people of many physical abilities and is easy to learn. The most beneficial use of this personal technology is to connect individuals, families and communities to necessary and life-changing services using secure technology that can streamline and automate these services, making them more accessible. We have seen numerous examples including microfinance, apps that help unhoused people find shelter, online education, telehealth and a range of government services. Too many people still experience poverty, bias and lack of access to serve their needs and create opportunities for them to fully participate in and contribute to their communities." AI-driven health care may include home air-quality and waste stream assessments, but the U.S. will lag behind other regions of the world while China will lead

**Mark Schaefer**, a business professor at Rutgers University and author of "[Marketing Rebellion](#)," wrote, "In America, health care progress will come from startups and boutique clinics that offer wealthy individuals environmental screening devices and pharmaceutical solutions customized for precise genetic optimization. The smart home of the future will analyze air quality, samples from the bathroom waste





stream and food consumption to suggest daily health routines and make automatic environmental and pharmaceutical adjustments. “Overall, an AI-driven health care system will be radically streamlined to be highly personal, effective and efficient in many developed regions of the world – excluding the United States. While the U.S. will remain the leader in developing new health care technology, the country will lag most of the world in this tech adoption due to powerful lobbyists in the health care industry and a dysfunctional government unable to legislate reform. However, progress will take off rapidly in China, a country with a rapidly-aging population and a government that will dictate speedy reform. Dramatic improvements will also occur in countries with socialized health care, since efficiency means a dramatic improvement in direct government spending. Expected lifespan will increase by 10% in these nations by 2035. China’s population will have declined dramatically by 2035, a symptom of the one-child policy, rapid urbanization and social changes. China will attract immigrant workers to boost its population by offering free AI-driven health care.”

*Advancing federation and decentralization, mandating interoperability and an emphasis on subsidiarity in platform governance are key to the future*

**Cory Doctorow**, activist journalist and author of [“How to Destroy Surveillance Capitalism,”](#) wrote, “I hope to see an increased understanding of the benefits of federation and decentralization; interoperability mandates, such as the Digital Markets Act, and a renewed emphasis on interoperability as a means of lowering switching costs and disciplining firms; a decoupling of decentralization from blockchain (which is nonsense); and an emphasis on subsidiarity in platform governance. Among the challenges are new compliance duties for intermediaries – new rules that increase surveillance and algorithmic filtering while creating barriers to entry for small players – and ‘link taxes’ and other pseudo-copyrights that control who can take action to link to, quote and discuss the news.”

*Digital life offers opportunities to enhance longevity, health and access to resources*

**Micah Altman**, social and information scientist at the Center for Research in Equitable and Open Scholarship at MIT, said, “Whether digital or analog, there are five dimensions to individual well-being: longevity, health, access to resources, subjective well-being and agency over making meaningful life choices. Within the last decade the increasing digitalization of human activities has contributed substantially in each of these areas, providing benefits in four of the five areas. Digital life is greatly expanding access to online education (especially through open online courses and increasingly through online degree and certification programs); health information and health treatment (especially through telehealth in the area of behavioral wellness); the opportunity to work from remote locations (which is particularly beneficial for people with disabilities); and the ability to engage with government through online services, access to records, and modes of online participation (e.g., through online public hearings). Expansion in most of these areas is likely to continue over the next dozen years.”

*The intelligence and effectiveness of AI systems operating alone will be overestimated, creating catastrophic failure points*

**Jon Lebkowsky**, writer and co-wrangler of Plutopia News Network, previously CEO, founder and digital strategist at Polycot Associates, commented, “Relying too much on AI and failing to factor in human judgment could have potentially disastrous consequences. I’m not concerned that we’ll have a malignant omnipotent AI like ‘Skynet,’ but that the intelligence and effectiveness of AI systems operating alone will be overestimated, creating catastrophic failure points. AI also has the potential to be leveraged for surveillance and control systems by autocratic governments and organizations to the detriment of freedom and privacy. The misuse of technology is especially likely to the extent that those responsible for governance and regulation misunderstand relevant technologies.”

Lebkowsky offered three specific potential issue areas for the future:

- “If we fail to shift from fossil fuels to cleaner, more efficient technologies, we may fail to manage our response to climate change effectively and leverage innovations that could support adaptation and/or mitigation of global warming.
- “If we fail to address the monopolistic power of Big Tech and social manipulation via centralized social media, we may see increasing uses of propaganda and online influence to gain power for its own sake, potentially evolving dystopian autocracies and losing the democratic and egalitarian intentions that are so challenging to sustain.
- “Medical and scientific ignorance and suspicion, as we see in the current anti-vaccine movement, could offset medical advances. We must restore trust in scientific and medical expertise through education, and through ensuring that scientific and medical communities adhere to standards that will make them inherently trustworthy.”

*It is likely that a highly visible abuse or scandal with clearly identifiable victims will be needed to galvanize the public against digital excesses*

**Richard Barke**, associate professor of public policy at Georgia Institute of Technology, responded, “The shift from real to digital life probably will not decelerate. The use of digital technologies for shopping, medical diagnosis and interpersonal relations will continue. The use of data analytics by businesses and



governments also will continue to grow. And the number and severity of harmful consequences of these changes will also grow. New technologies, market tools or social changes never come without some harmful consequences. Concerns about privacy and discrimination will increase, with the result that demands for transparency about business practices, targeting of subpopulations and government policies will grow at least as fast as digital life.

“Those demands are not likely to be answered in the absence of significant harmful or menacing events that catch the attention of the public, the media and eventually, policymakers. The environmental movement needed a Rachel Carson and a Love Canal in the 1960s and 1970s as policy entrepreneurs and focusing events. The same is true for many other significant changes in business and government decision-making.

“Unfortunately, it is likely that by 2035 some highly visible abuse or scandal with clearly identifiable victims and culprits will be needed to provide an inflection point that puts an aggrieved public in the streets and on social media, in courtrooms and in legislative hallways, resulting in a new regime of law and regulation to constrain the worst excesses of the digital world. But, even then, is it likely – or even possible – that the speed of reforms will be able to keep up with the speed of technological and business innovations?”

*Once trust is lost, it is difficult to reclaim, and digital ‘reality’ is dangerously amenable to distortion and manipulation that can lead to the erosion of it at all levels of society*

**Larry Lannom**, vice president at the Corporation for National Research Initiatives, observed, “In thinking about the potential harm that exponentially improved digital technologies could wreck by 2035, I find that I have two levels of concern. The first is the fairly obvious worry that advanced technologies could be used by malevolent actors, at the state, small-group or individual level, to cause damage beyond what they could achieve with today’s tools. AI-based autonomous weapons, new pathogens, torrents of misinformation precision-crafted to appeal to the recipients and total state-level intrusion into the private lives of the citizenry are just some of the worrying possibilities that are all too easy to imagine evolving by 2035. A more insidious worry, however, is the potential erosion of trust at all levels of society and government. More and more of our lives are affected by or even lived in the digital realm and as that environment increases in size and sophistication, it seems likely that the impact will increase. But digital ‘reality’ is much more amenable to distortion and manipulation than even the worst human-level deception. The ability of advanced computing systems of all kinds to convincingly generate fake audio and video representations of any public figures, to generate overwhelming amounts of reasonable sounding misinformation, and to use detailed personal information, gathered legally or illegally, to craft precision messaging for manipulation beyond what can be done today could contribute to a complete lack of trust at all levels of society. Once trust is lost it is difficult to reclaim.”

*‘We are rewriting childhood for youngsters ages 0 to 5, and it is not in healthy ways’*

**Jane Gould**, founder of DearSmartphone, responded, “We have been rewriting the concept of screen time and exposure. This trend began in the 2000s but the introduction of mobility and iPhones and mobile apps in 2007 accelerated the change. We are rewriting childhood for youngsters ages 0 to 5, and it is not in healthy ways. All infants must go through discrete stages of cognitive and physical growth. There is nothing that we can do to speed these up, nor should we. Yet from their earliest moments we put young babies in front of digital devices and use them to entertain, educate and babysit them. These devices use artifices like bright lights and colors to hold their attention, but they do not educate them in the way that thoughtful, watchful parents can. More than anything else, these electronics keep children from playing with the traditional handheld toys and games that use all five senses to keep babies busy and engaged with play, and in two-way exchanges. Meanwhile, parents are distracted and pay less attention to their infants because they stay engaged with their own personal phones and touchscreens.”

*Digital access will grow and programs will become more user-friendly*

**Henning Schulzrinne**, Internet Hall of Fame member, Columbia University professor of computer science and co-chair of the Internet Technical Committee of the IEEE, predicted, “Amplified by machine learning and APIs, low-code and no-code systems will make it easier for small businesses and governments to develop user-facing systems to increase productivity and ease the transition to e-government. Government programs and consumer demand will make high-speed (100 Mb/s and higher) home access, mostly fiber, near-universal in the United States and large parts of Europe, including rural areas, supplemented by low-Earth orbiting satellites for covering the most remote areas. We will finally move beyond passwords as the most common means of consumer authentication, making systems easier to use and eliminating many security vulnerabilities that endanger systems today. On the ‘worst’ side of change, the concentration of ad revenue and the lack of a viable alternative source of income will further diminish the reach and capabilities of local news media in many countries, degrading the information ecosystem. This will increase polarization, facilitate government corruption and reduce citizen engagement.”



*From lentil soup recipes, pizza and Bollywood music to scan-reading, customer service and newswriting, tech has upended how we give and get information, for good and bad*

**Alan D. Mutter**, consultant and former Silicon Valley CEO, said, “The magic of technology enables me to Google lentil soup recipes, trade stocks in the park, stream Bollywood music and Zoom with friends in Germany. Without question, tech has solved the eternally vexing P2P problem – the rapid, friction-free delivery of hot-ish pizza to pepperoni-craving persons. Techno thingies like software calibration and hardware calibration networks will get faster and somewhat better (albeit more complex), but probably not cheaper. Here’s what I mean: For no additional charge, the latest Apple Watches will call 911 if they think you fell. It’s a good idea and the feature actually has saved some lives. But it also is producing an overwhelming number of false alarms. So, it is a good thing that sometimes is a bad thing.”

Mutter offered these thoughts on future AI applications:

- “AI probably will do a better job of reading routine scans than radiologists and might do a better job than human air traffic controllers who sometimes vector two planes to the same runway.
- “AI undoubtedly will answer all phones everywhere, cutting costs but also further compromising the quality of customer service at medical offices, insurance companies, tech-support lines and all the rest.
- “AI will produce all forms of media content, but likely without the elan and judgment formerly contributed by humans.
- “AI probably will be more accurate than humans at doing math but less savvy at sorting fact from fiction and nuance from nuisance.

“Technology has upended forever the ways we get and give information. We now live in a Tower of Babel where yadda-yadda moves unchecked, unmoderated and unhinged at the speed of light, polluting and corrupting the public discourse. This is perilous for a democracy like the United States. I am afraid for our republic.”

*‘Wanna get scary about the other edge of the AI sword? What will 4chan do with it?’*

**Howard Rheingold**, pioneering internet sociologist and author of [“The Virtual Community,”](#) wrote, “If we are honestly looking back at the last decades of rapid technological change for hints about decades to come, we’re in for a world of hurt along with some really miraculous stuff. I sense that we are at an inflection point in the conduct of science as significant as the introduction of computers: the use of machine learning techniques as scientific thinking and knowledge tools. Proteins, for just one example, are topologically complex and can fold into a large number of possible shapes. Much of immune system and anti-cancer therapies rely on matching the shape of proteins on the surface of a cell. Now, AI can propose previously unknown proteins of medical significance.

“Machine Learning (oversimplified) uses iterative computations modeled on the way neurons work. It can be applied to datasets other than the omni-versal ones sought by large learning models, or LLMs. LLMs don’t ‘know,’ but the way significant knowledge can be parsed out of it is, in my opinion, impressive, although the technology is in its infancy. Yes, it swallows all the bull along with the good info, and yes, it is unreliable and makes stuff up, and no, the models are tools, they are not General Intelligence. They don’t understand. They do statistics. Think of them as thinking-knowledge tools. As mathematics and computers come to enable human minds to go places they were previously unable to explore, I see a lot of change coming from this symbiosis of machine learning and human production of words, images, sounds and code. Computational biology is a good example of this two-edged miracle.

“Wanna get scary about the other edge of the AI sword? Generative AI once suggested 40,000 chemical weapons in just six hours. I recall that Bill Joy wrote a Wired magazine essay (23 years ago!) titled [‘Why the Future Doesn’t Need Us.’](#) In that essay he mentioned affordable desktop wetlabs, capable of creating malicious organisms. A good way to think about a proposed technology is to ask: What would [4chan](#) do with it? Connecting computational biology to wetlab synthesizers is just a matter of money and expertise. What will 4chan do with LLM tools?”

*‘Copyright and technology to manage microtransactions will create huge gaps in knowledge between the haves and have nots’*

**Harold Feld**, senior vice president at Public Knowledge, predicted, “Reliable, affordable high-speed broadband will become as ubiquitous in the world (including the developing world) as telephone service was in the United States in the late 20<sup>th</sup> Century. The actual technology will vary greatly depending on country, and we will still see speed differences and other quality of service differences that will maintain a digital divide. But the combination of available communications technology and solar operated systems will enable a wide range of benefits. These will include:

- “Far more efficient resource tracking and allocation and far more efficient environmental monitoring will enable dramatic increases in food and clean water distribution where needed and will help to predict potential environmental disasters with greater accuracy and certainty.
- “Greater communication potential will enable vast improvements in distance learning and telemedicine. In countries where health professionals are scarce, or where travel is difficult, a wealth of diagnostic tools and a broadband connection will allow a handful of trained first



responders to treat people locally under the guidance of experienced and more highly trained medical professionals. Necessary resources such as antibiotics will be delivered by drones, and local personnel guided in how to administer and provide follow-up care. As a last resort, doctors can order medical evacuations.

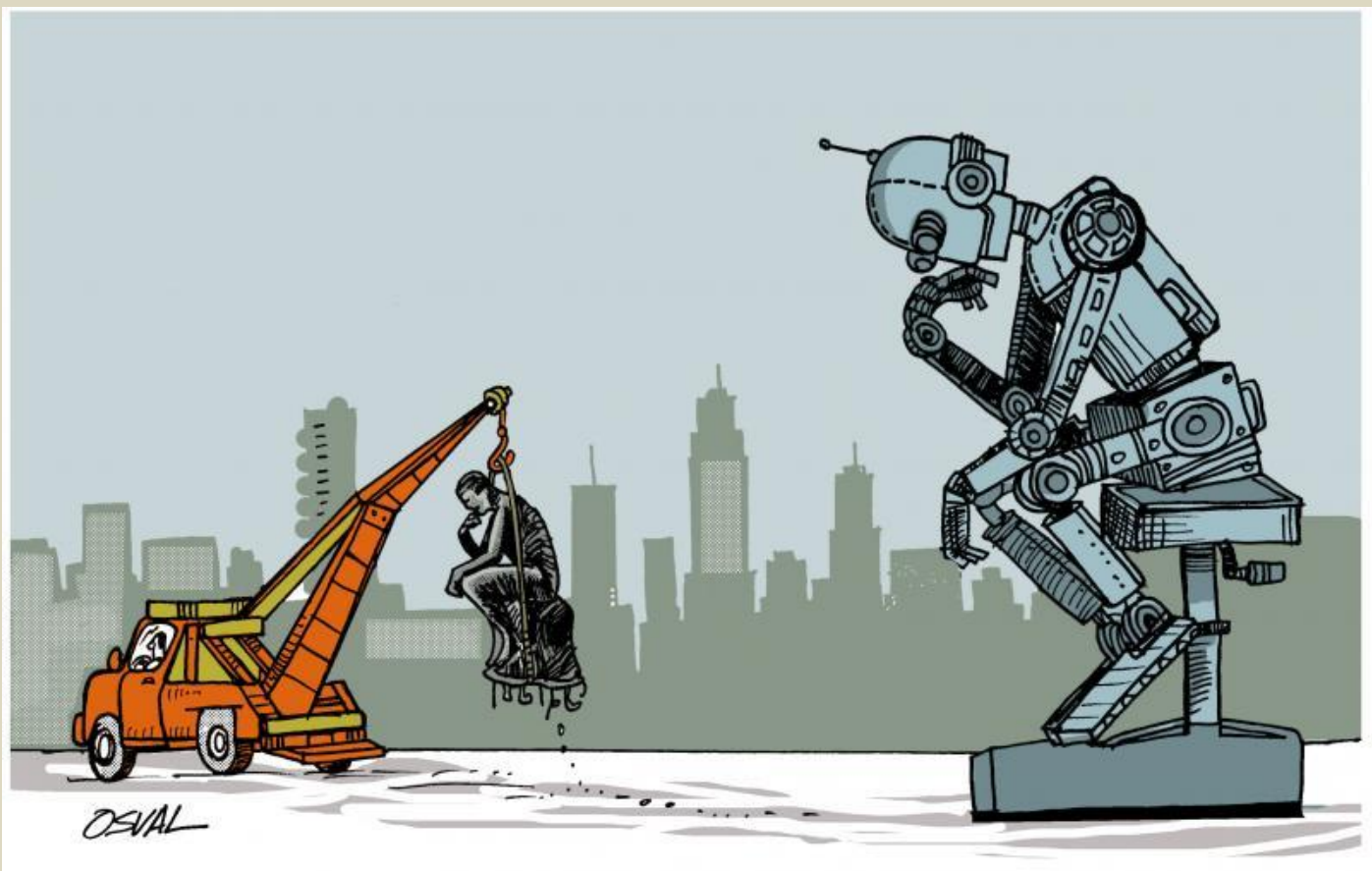
- “Children will have access to education in their native language. Artificial expenses such as uniforms will be eliminated as a requirement. Girls will be able to access equal education without fear of assault.

“Yet, here’s the thing, widespread ubiquitous broadband could easily broaden ubiquitous surveillance for corporate reasons and to aid repressive governments. Big data systems will be able to sort the noise from the signal and allow corporate or government interests to predict with incredible accuracy human behavior and how to shape it in ways that best serve their interests. Widespread access to others will create pockets of intense culture shock as communities find their basic assumptions about how to organize society undermined. Basic trust in institutions will be replaced not with healthy skepticism for engagement, but either complete and fanatical belief in a trusted source or complete disbelief in any source. To slightly paraphrase William Butler Yeats, *‘Mere anarchy is loosed upon the world ... The ceremony of innocence is drowned. The best will lack all conviction, while the worst will be filled with passionate intensity.’* Societies may become entirely paralyzed, caught between an inability to rely on facts for basic cooperation, or trapped between warring factions, or both.

“Copyright and technology to manage microtransactions will create huge gaps in knowledge between the haves and have nots, as even basic educational material becomes subject to limitations on sharing and requirements for access fees. Ownership of books or other educational media will become a thing of the past, as every digital source of knowledge will be licensed rather than owned. Book printing will wither away, so that modern educational materials will be inaccessible to those who cannot afford them.

“For the same reason, innovation will slow and become the province of a privileged few able to negotiate access to the needed software tools. Even basic mechanical inventions will have digital locks and software to prevent any tinkering.”

- [Read the full article at the source’s URL.](#)



IOI  
International  
**CBRNE**  
INSTITUTE



**C<sup>2</sup>BRNE**  
DIARY



*Preparedness &*

# **EMERGENCY RESPONSE**



## Role of Artificial Intelligence in Disaster Risk Management

Source: <https://www.aeologic.com/blog/role-of-artificial-intelligence-in-disaster-risk-management/>



The role of AI in disaster risk management should be every country's top priority. Since disasters can take many different forms. From natural disasters like wildfires and hurricanes to pandemics and terrorist attacks, the citizens must be ready. The impact of a disaster can be devastating, equally for the people affected and for the economy.

The role of artificial intelligence in disaster risk management and reduction particularly machine learning (ML), plays a crucial role. Ranging from forecasting extreme events to developing risk mitigation strategies. And offering real-time situational awareness and decision support.

In order to minimize the damage caused by disasters, it is important to understand the role of AI in disaster risk management strategy. Certainly, many experts are believing that artificial intelligence (AI) can play a role in improving risk management and disaster response.

### Artificial Intelligence and Machine Learning for Disaster Risk Management

First, let's look at how [AI and machine learning \(ML\)](#) will be helping when disaster strikes. Artificial intelligence is using algorithms to learn from data. Machine learning is a type of AI that is allowing computer systems to improve from experience without being programmed to do so. In the background of disaster response, first responders can use AI and ML for a variety of operations, such as:

#### Automated damage assessment

One of the first and major steps in any disaster response is damage assessment. This includes sending out teams of assessors for evaluating the extent of the damage. Automated damage assessment using AI and ML will be helping to speed up this process.

For instance, we can use computer vision algorithms for analyzing images and videos to identify damaged buildings. This information can then be incorporated for creating maps that show the affected areas. Additionally, firms can utilize AI for processing large amounts of data from different sources (e.g., social media, satellite images). This will be providing a complete picture of the damage caused by a disaster.

#### Predictive analytics for evacuation

Another way that AI can facilitate during a disaster is by using predictive analytics to predict when and where an evacuation might be required. We can take the help of predictive analytics to analyze data from seismic sensors, weather data, and satellite images. Experts can then use this data for creating a model that predicts how an eruption might unfold. Furthermore, understanding if it will be violent enough to warrant an evacuation and when it might happen to enable authorities to prepare for the disaster efficiently.



### Routing of emergency resources

However, during a disaster, time is of the essence. Therefore, we can accelerate many of the tasks that require to be carried out during a disaster response. This can aid to save lives and mitigate damage. Also, transfer learning can be creating models that can analyze satellite images and identify areas of damage. Relevant authorities can then integrate this information to deploy emergency services more efficiently.

Intelligent search algorithms are taken into consideration to find the best route for evacuations. The role of AI in disaster risk management is to route emergency resources. Therefore, ensuring that they are deployed more efficiently and effectively. Furthermore, we can use AI-based solutions to predict the path of a natural disaster.

### Prevention

Nonetheless, the role of AI in disaster risk management is not just about response and recovery. However, it is also about prevention. Artificial intelligence can identify risk factors that are leading to a disaster. For instance, it can take machine learning to analyze historical data to identify patterns that may project a forthcoming catastrophe.

### Conclusion: AI and Disaster Response and Risk Management

Disaster risk management is a complicated and multi-faceted problem. AI and machine learning can help in improving disaster response in several ways. Ranging from automated damage assessment to predictive analytics for evacuation. Thereby, AI is making a significant difference when it comes to reducing the impact of a disaster.

## Crisis communication in CBRNe preparedness and response: Considering the needs of vulnerable people

By Danielle Carbon, Andreas Arnold, Thomas Görger, and Chiara Wüller

*International Journal of Disaster Risk Reduction* | Volume 79, September 2022

Source (full text): <https://www.sciencedirect.com/science/article/abs/pii/S221242092200406X>

### Abstract

Major emergencies such as CBRNe (Chemical, Biological, Radiological, Nuclear, explosive) incidents pose special challenges for emergency services especially regarding crisis communication. The EU project PROACTIVE aims to enhance CBRNe preparedness and response by increasing the effective management of large, heterogeneous groups including vulnerable civilians (older people, visually and hearing-impaired people, ethnic minorities, etc.). The findings from two surveys, one with CBRNe practitioner organisations (May–June 2020 and May–July 2021) and one with Civil Society organisations (October–December 2020), revealed differences between the current CBRNe related communication management of CBRNe practitioners and the special needs of vulnerable people in crisis communication. Our results show that CBRNe practitioners very rarely consider the needs of [vulnerable groups](#) in their communication strategies for major emergencies. Furthermore, CBRNe practitioners seldom provide information in additional language formats (e.g. Braille and sign language). To enhance crisis communication with vulnerable people, four recommendations for CBRNe practitioners have been developed.

## 12 Ways the World Could End, According to Oxford University

Source: <https://www.sciencetimes.com/articles/5076/20150331/12-ways-the-world-could-end-according-to-oxford-university.htm>

No, this is not a list of endings to Hollywood's latest blockbusters. These are potentially real-life scenarios that could devastate and even end human civilization as we know it.

Compiled by University of Oxford's Future of Humanity Institute and the Global Challenges Foundation, the report titled "[12 Risks that Threaten Human Civilization](#)" lists the most likely ways that apocalypse could occur. "The idea that we face a number of global risks threatening the very basis of our civilization at the beginning of the 21st century is well accepted in the scientific community and is studied at a number of leading universities," the [authors state](#) in the assessment. Far from being alarmist, however, the report's authors say they don't mean to cause panic, but rather find ways to turn challenges into opportunities.

"It is a scientific assessment about the possibility of oblivion, certainly, but even more it is a call for action based on the assumption that humanity is able to rise to challenges and turn them into opportunities," [write the authors](#).

Here are the 12 most likely ways the world could end, according to their assessment:



**1. Extreme climate change – Probability: 0.01%**

The report warns that if no action is taken on climate change, then global warming could cause devastation that might be worse than what is predicted by today's middle estimates. Poorer countries would be the most affected by global warming, some of them possibly becoming completely uninhabitable. Leading to mass deaths, famines, social collapse and mass migration, this scenario could certainly lead to civilization collapse, states the report.

**2. Nuclear War – Probability: 0.005%**

Although the report's authors believe that the possibility of nuclear war is lower now than it was during the Cold War, they believe that if there is conflict, something called a nuclear winter could occur. This is when the creation of a pall of smoke high in the stratosphere causes temperatures to plunge below freezing and destroy most of the ozone layer. As a result, the collapse of the global food supply would lead to mass starvation.

**3. Global pandemic – Probability: 0.0001%**

The study's authors believe this is one of the more probable scenarios because the features of an extremely devastating disease already exist in nature: incurable (Ebola), fatal (rabies), extremely infectious (common cold), and with long incubation periods (HIV). If a pathogen that combined all these features appeared, its death toll would be extreme.

**4. Ecological collapse – Probability: n/a**

Although this is a preventable risk, the study states that if the current historic rates of species extinction continue to rise, human civilization is potentially at risk because humans are part of the global ecosystem and fundamentally dependent on it.

**5. Global system collapse – Probability: n/a**

The world's economic and political system is so interconnected and complex that unexpected system-wide failures might lead to economic collapse, social chaos and breakdown in law and order.

**6. Major asteroid impact – Probability: 0.00013%**

This is probably one of the best understood of all the risks in this report. Scientists believe that a land impact from an asteroid could cause extinction-level events caused by the clouds of dust that would damage the biosphere.

**7. Super-volcano – Probability: 0.00003%**

A volcano eruption thousands of times larger than a normal eruption could release dust into the upper atmosphere, which would absorb the sun's rays and lead to a global volcanic winter.

**8. Synthetic biology – Probability: 0.01%**

One of the most troubling scenarios is the possibility of an engineered pathogen designed to target humans. Either through biowarfare or bioterrorism, this could lead to disastrous consequences for the human race.

**9. Nanotechnology – Probability: 0.01%**

"[Nanotechnology] could lead to the easy construction of large arsenals of conventional or more novel weapons made possible by atomically precise manufacturing," they warn. "Of particular relevance is whether nanotechnology allows the construction of nuclear bombs."

**10. Artificial intelligence – Probability: 0-10%**

The report states that machines may develop extreme levels of intelligence and be driven to construct a world without humans. It also posits, however, that in an alternate scenario, they could provide solutions to the previous risks presented.

**11. Future bad global governance – Probability: n/a**

Global leaders could exacerbate existent problems with bad decisions. "There are two main divisions in governance disasters: failing to solve major solvable problems, and actively causing worse outcomes," the study explains.

**12. Unknown consequences – Probability: 0.1%**

Finally, the researchers believe there are "unknown unknowns" and call for "extensive research" into "unknown risks and their probabilities."







## You're the Leader. Now What?

By Catherine L. Feinman

Source: <https://domesticpreparedness.com/articles/youre-the-leader-now-what>

June 14 – The National Oceanic and Atmospheric Administration ([NOAA](#)) reported that, as of May 8, the United States had experienced seven confirmed \$1-billion weather/climate-related disasters in 2023. Three more storms that occurred in April may soon be added to that list. According to the Federal Emergency Management Agency ([FEMA](#)), 35 major disasters have been declared in 2023 as of June 2. Adding significant natural events to disasters related to technological and human-caused threats, any community or organization can be suddenly thrust into emergency response mode.

When an expected or unexpected event occurs, leadership capabilities are a determining factor for each community's level of resilience and ability to recover. During such times, various titles naturally put some people in leadership roles. However, an impressive title may not equate to a good leader. So, what makes a good leader? In simple terms, leadership is defined as the ability to influence and guide others. When faced with an emergency or disaster, effective leaders can use this ability to achieve organizational goals and maximize community response efforts.

### Defining a Good Leader

On May 30, 2023, Domestic Preparedness hosted a multidiscipline panel discussion at the annual Texas Emergency Management Conference in Fort Worth to ask five leaders to share their knowledge and experiences on their paths to developing good leadership skills. These five professionals brought to the table 195 years of combined experience across emergency management, public safety, law enforcement, emergency medical services, hospitals, fire, hazardous materials, military, disaster psychology, and other areas of expertise.

The discussion began with the panelists describing how and when they realized they had become leaders. Some had a defining pivotal moment, and others developed that realization over time. However, the one thing in common was that they all recognized their leadership abilities long before assuming leadership positions within their careers. From there, the panelists answered the following key questions:

- What key characteristics are required to be a good leader?
- During normal operations, how do you build and maintain interagency relationships?
- During large-scale incidents, how do you manage the different leadership styles and jurisdictional differences to effectively work together?

One interesting debate involved the ability of leadership traits to be learned (i.e., nature vs. nurture). Although there may be no consensus on whether leaders are born or made, there was a common theme throughout the discussion that the best way to realize one's potential is to embrace opportunities and step up to challenges when they arise. Other recommendations for being a good leader include keep developing knowledge and skills, continue building resilience, and always communicating within and between organizations.

Each of the following five panelists embraced their leadership roles in different ways. Yet, they all have been and continue to be effective in influencing and guiding others. Click the link to listen to the full panel discussion.

**Catherine L. Feinman, M.A.**, joined Domestic Preparedness in January 2010. She has more than 30 years of publishing experience and currently serves as Editor of the Domestic Preparedness Journal, and the DPJ Weekly Brief, and works with writers and other contributors to build and create new content that is relevant to the emergency preparedness, response, and recovery communities. She received a bachelor's degree in international business from University of Maryland, College Park, and a master's degree in emergency and disaster management from American Military University.

## Are Public Health Agencies Ready, or Just Prepared?

By Raphael Barishansky

Source: <https://domesticpreparedness.com/articles/are-public-health-agencies-ready-or-just-prepared>

June 21 – Perhaps it is time to retire the term *public health preparedness*. Public health emergency preparedness has been defined as “the capability of the public health and health care systems, communities, and individuals, to prevent, protect against, quickly respond to, and recover from health emergencies, particularly those whose scale, timing, or unpredictability threatens to overwhelm routine



capabilities. Preparedness involves a coordinated and continuous process of planning and implementation that relies on measuring performance and taking corrective action.” However, there is no currently agreed upon definition for public health readiness. As the COVID-19 pandemic emergency declaration sunsets and officials review their various public health actions, it may be time to collectively take the initiative to retire this term and replace it with *public health readiness*. Some background is in order before genuinely discussing the need for this paradigmatic shift.



### The Preparedness Background

In the immediate aftermath of 9/11 and the anthrax attacks that followed, it became clear that the ability of the U.S. public health system to respond to a large-scale emergency was simply not in place, and the discrete discipline of public health preparedness emerged. Congress soon appropriated nearly \$1 billion in FY2002 to the Centers for Disease Control and Prevention (CDC), which reorganized its preparedness activities to support states and territories, including creating a new national center known as the Center for Preparedness and Response. Interestingly, this center was initially called the Coordinating Office on Terrorism Prevention and Emergency Response, then renamed as indicated previously and, only recently, again renamed as the [Office of Readiness and Response](#), reporting directly to the CDC director. Multiple grants, including the [Public Health Emergency Preparedness](#) and [Cities Readiness Initiative](#), were soon developed and funds distributed to state and local health departments. Simultaneously, the Hospital Preparedness Program was deployed to bring a similar preparedness initiative, and funding, to the healthcare sector. All these grants were geared toward a greater state of preparedness with, at the time, no true definition of the term public health preparedness.

Even without a clear definition of preparedness, state and local health departments formed public health preparedness units. They moved forward with the development of various public health-specific contingency plans, purchase of communications equipment, compliance with Incident Command Systems mandates to coordinate with other, more traditional emergency response entities, and many other initiatives, not the least of which was the recruitment of public health preparedness subject matter experts. Almost ten years later, in 2011, the CDC developed and promulgated a set of 15 emergency preparedness and response capabilities, which now serve as the national standards for public health preparedness planning. The CDC made the latest updates to these [capabilities in 2018](#).



### Asking Necessary Questions

In this post-COVID-19 moment, it is necessary to reflect on whether the various preparedness grants made state and local health entities adequately prepared for the pandemic. Consider whether the different emergency exercises and equipment purchases made state and local health entities fully prepared for the difficulty seen in the following aspects:

- Procurement of personal protective equipment,
- The media and public backlash over fluctuating guidance,
- The increasing politicization of public health, or
- Other elements that made COVID-19 so challenging to respond to.

When answering questions related to these and many more issues related to public health and the public health role in emergency preparedness and response, the industry needs a more all-encompassing term to describe the various actions health agencies take before, during, and even after an emergency. Perhaps that term is *readiness*, as in public health readiness. This term makes sense when viewing preparedness as more of a physical state that speaks to capability and capacity. In contrast, readiness is more of a mental state, specifically being able to apply preparedness when needed; it is a mindset.

Readiness takes into account all of the previous elements that the preparedness grants encompassed but also includes other areas that COVID-19 has shown to be lacking, including supply chain management, a stronger relationship with emergency management, and an organization-wide commitment to public health emergency response. The CDC is thinking along these lines as an element of the next five-year cooperative agreement cycle with state health departments. They have been developing a “Response Readiness Framework,” with ten areas through which partners can look at the current 15 capabilities to move to a better state of readiness – these areas include data modernization, health equity, workforce readiness and resiliency, and others.

Regarding the public health workforce, readiness speaks to understanding how mental health and burnout would impact public health workers who sometimes labor 12-plus hours a day, seven days a week, to make the best decisions for the public they serve. The COVID-19 pandemic saw this issue resurface repeatedly, and this lesson should not be forgotten. Readiness also means considering the capabilities and capacity of the various volunteer community resources (such as Community Emergency Response Teams and Medical Reserve Corps units) and planning appropriately to utilize these resources to supplement health employees when and where applicable. These units were invaluable to many jurisdictions’ responses during the pandemic, and memorializing the lessons learned about their effectiveness is a critical readiness next step.

Another element of readiness, as it applies to the workforce, is ensuring that the public health workforce is adequately trained for the diversity of emergencies they may confront and that all feel comfortable in their roles. Again, look at the actions of the CDC, which recently announced that it would require all employees to be ready to deploy to combat national health crises, thus marking a drastic shift from a fragmented volunteer system that hampered its COVID-19 response. State and local health departments, which are integral partners in any large-scale public health emergency response, would also benefit from a paradigmatic shift such as this.

### The Path Toward Readiness

Regarding preparedness and response roles, readiness speaks to a state where a jurisdiction or local health entity is ready for whatever emergency comes next and has the capabilities to respond appropriately. When looking at emergency planning efforts, there is a need to reconfigure existing emergency plans with the readiness factor in mind and ask the following questions:

- Do our plans speak to social distancing, such as what we experienced during COVID-19?
- As we absorb the lessons of COVID-19 and the outsized role of the media, have our efforts regarding crisis communication been sufficient?
- Is there some sort of training/exercise for public health officials to engage in when mitigating the politicization of public health seen in the recent past?

These elements of organizational readiness must, at minimum, be contemplated before moving forward.

**Raphael M. Barishansky, MPH**, is the chief of Public Health Emergency Preparedness for the Prince George’s County (Maryland) Department of Health. Prior to establishing himself in this position, he served as executive director of the Hudson Valley Regional EMS (Emergency Medical Services) Council, based in Newburgh, New York. A frequent contributor to the DomPrep Journal and other publications.



ICI  
International  
**CBRNE**  
INSTITUTE

A common roof for international  
CBRNE First Responders



*Join us!*



Rue des Vignes, 2  
B5060 SAMBREVILLE (Tamines)  
BELGIUM

[info@ici-belgium.be](mailto:info@ici-belgium.be)  
[www.ici-belgium.be](http://www.ici-belgium.be)