06\26

2 CBRNE DIARY

ICI International CBRNE INSTITUTE

Dedicated to Global First Responders

June 2026

PART B

N Korea's huge CWA stockpile

Chemical weapons in Syria

Ricin antidote

Ensitrelvir 4 Covid-19

1st AI designed vaccine

Ebola is progressing in Africa

Sin Nobre Hantavirus in Arizona

Mirror Life

US Biolabs in Ukraine

Toxins update

An International CBRNE Institute publication

ICI International CBRNE INSTITUTE

CBRNE-Terrorism Newsletter

W O M D

DIRTY R-NEWS

C2BRNE DIARY

The Trump administration's reckless attack on radiation protection will have long-term consequences for public safety

By Frank von Hippel

Source: https://thebulletin.org/2026/05/the-trump-administrations-reckless-attack-on-radiation-protection-will-have-long-term-consequences-for-public-safety/

May 27 – Worldwide, regulations limiting doses from the radiation emitted by nuclear fissions and decays are based on the Linear

No-Threshold (LNT) model. This hypothesis posits that, irrespective of whether ionizing radiation comes in a pulse or over years, the additional risk of developing cancer as a result is proportional to the cumulative amount of energy deposited per gram of tissue, with weighting risk factors for radiation type, sex, age, and specific organs.

Since 1975, the US nuclear industry has been required to limit exposures to workers and the public to "as low as reasonably achievable" (ALARA) levels. What the ALARA level should be is determined by cost-benefit analysis in which the costs of dose reductions are compared with the benefits to workers and the public, measured in terms of reduced disease and longer life expectancy.

In May 2025, four months after taking office, the Trump administration challenged this five-decade-old regulatory approach as part of an Executive Order "Ordering the Reform of the Nuclear Regulatory Commission" (NRC). The order claimed the "NRC utilizes safety models that posit there is no safe threshold of radiation exposure and that harm is directly proportional to the amount of exposure," which corresponds to the linear hypothesis. "Those models lack sound scientific basis," the Executive Order added, before directing the NRC to "reconsider reliance on the linear no-threshold (LNT) model for radiation exposure and the 'as low as reasonably achievable' [ALARA] standard, which is predicated on LNT."

The Nuclear Regulatory Commission had reviewed exactly this question in 2021 in response to a campaign by advocates of the radiation "hormesis" theory, which posits that low doses of ionizing radiation actually protect against cancer by stimulating the body's DNA repair mechanism—the exact opposite of ALARA. The NRC rejected that contention, concluding that "the LNT model continues to provide a sound regulatory basis for minimizing the risk of unnecessary radiation exposure to both members of the public and radiation workers." As a result, the commission maintained the current dose limit requirements contained in its regulations.

But President Donald Trump's decision to bring independent regulatory agencies under White House control and to fire the NRC's chairman ended the commission's resistance. On July 2, 2025, an anonymous NRC spokesperson enthused in a social media post that the Executive Order reforming the NRC "gives us a chance to reconsider our radiation protection framework in support of the whole-of-government effort to safely enable the nation's use of nuclear power."

Two weeks later, the NRC hosted a webinar for input on the issue of the LNT hypothesis. The Nuclear Energy Institute—the US nuclear industry's lobbying organization—recommended that the commission remove ALARA and dose minimization as regulatory requirements. Instead, the institute proposed to establish a "practical threshold"—for instance, 2 rem per year (or 20 milliGray per year for gamma rays) for workers—below which further dose reduction would not be required. (The rem is a unit of effective absorbed radiation in human tissue, equivalent to one roentgen of X-rays. One millirem is one-thousandth of a rem. The Gray measures the absorbed dose, which is the physical amount of radiation energy absorbed by any material or tissue. One Gray corresponds to one Joule per kilogram.)

Radiation hormesis

Advocates of the theory of radiation "hormesis" do not believe the LNT hypothesis. Radiation hormesis is a fringe theory with passionate adherents who are taking advantage of the Trump administration's skepticism about regulations of all types.

One of the most vocal hormesis advocates is Edward Calabrese, an emeritus professor of toxicology at the University of Massachusetts in Amherst. He argues that the evidence for the linear no-threshold hypothesis is based on scientific fraud and, therefore, should be replaced with a model that considers the possibility of no risk—and even possible benefits—from ionizing radiation below a certain dose.

Calabrese's arguments persuaded some recent leaders of the Health Physics Society (HPS), an association of radiation-protection professionals, to host a 22-part, 10-hour video lecture series by Calabrese on the history of the LNT model in 2021-22. John Cardarelli, the HPS president when the videos were produced, summarizes Calabrese's argument at the end of each video. In the final one, Cardarelli declares his conclusion that the LNT model is "based on flawed research, ideological motives, deliberate misrepresentation of the research record, and political agendas."

Although the Health Physics Society declares that "the views expressed in these videos are not intended to represent official positions," it also advertises that its associated credentialing organization, the American Academy of Health Physics, has "preapproved 10 continuing education credits for certified health physicists watching all 22 episodes of this video series."

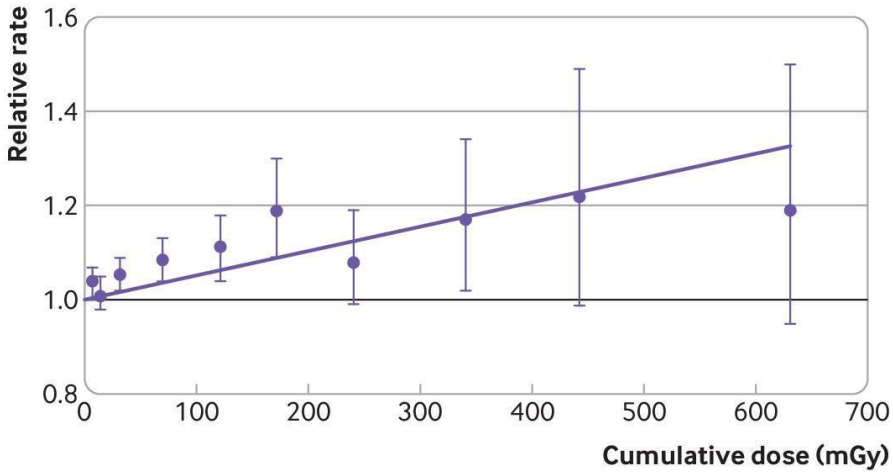
Physicist-epidemiologist Jan Beyea published a critique of Calabrese's

allegations in the HPS journal *Health Physics*, to which both Calabrese and Cardarelli have responded with lengthy rebuttals.

The research and reports Calabrese and his supporters are trying to discredit were done more than 50 years ago. For decades, the largest human population studied for radiation effects was the survivors of the 1945 Hiroshima and Nagasaki bombings, who, depending on their proximity to the ground zeros, were exposed to whole-body doses ranging from near zero to several Gray delivered in a single burst. But the cancer statistics for the Japanese survivors were not good enough to determine with high confidence carcinogenic effects in the dose range relevant for worker radiation protection (in the tens of milliGray per year). Hormesis advocates also argue that cellular mechanisms should be more effective in repairing the damage from low-rate radiation than from a nuclear explosion's short pulse. The lack of data on the effect of small low-rate doses left a gap in the epidemiological confirmation of the applicability of LNT estimates of the cancer risks from low doses to radiation workers and to civilian populations exposed to radioactive releases from nuclear accidents. That gap has been partially filled, however, in more recent studies of large populations of individuals who have received low-rate doses of ionizing radiation.

A directly relevant example is the INWORKS study done by an international consortium of researchers on the excess cancer deaths among approximately 310,000 nuclear industry workers in the United States, the United Kingdom, and France, whose radiation doses were measured and recorded throughout their decades of employment. As of 2012-16, this population had an average age of about 65, and about one third had died, with 28 percent of the deaths being due to "solid" cancers (abnormal masses of tissue arising in organs, glands, or bones), therefore excluding leukemia. Of those deaths, 5,500 to 14,000 were excess cases relative to the rate observed in a control group of 51,000 nuclear workers with near-zero occupational doses.

Figure 1 shows the rate of excess deaths from solid cancers in this population as a function of cumulative on-the-job dose



10 years before death, assuming that any solid cancer caused within the last decade of life would not have had time to become lethal. The bars show the 90-percent probability range associated with the number of deaths in each dose bin; that is, there is statistically only a 10-percent probability that, with more data, the number of excess deaths would converge outside that range (5 percent chance above and 5 percent below). The solid line is the best linear fit of the data to the LNT model.

Figure 1. Increased relative rate of death from solid cancers in a population of 310,000 nuclear workers from France, the United Kingdom, and the United States at an average age of 65. (Source: Richardson et al. *BMJ* 2023;382:e074520).

By this measure, there are significant excess cancer deaths among nuclear workers down to cumulative doses of 30 milliGray.

Energy Department's takeover

In addition to bringing the NRC to heel, the Energy Department's Office of Nuclear Energy has been inviting startups promoting new-design nuclear power reactors to build prototypes on department land, including the 900-square-mile footprint of the Idaho National Laboratory, where they will not be subject to NRC safety requirements.

According to President Trump's May 23 Executive Order, the NRC will be required "to approve reactor designs that the Defense Department or the Energy Department have tested and that have demonstrated the ability to function safely."

At most, the startups will only be able to demonstrate that they will not have had a serious accident or a near miss within their first few years of operation before they hope to build their reactors in large numbers across the country and export them abroad. In their efforts to compete with natural gas, photovoltaic, and wind power plants, the nuclear startups are under great economic pressure to cut safety and security requirements currently required by the NRC and other regulators around the world. Costly requirements include containment buildings that prevent the release of radioactivity to the atmosphere in case of a core meltdown accident. Regulations also include requirements that it be possible for the timely evacuation of areas around the reactors where the population could be at risk of high radiation doses from an accident, and robust around-the-clock guard forces to protect nuclear plants against potential sabotage.

By putting the Energy Department, which is pouring billions of dollars into nuclear startups, first in line in safety regulation, the Trump administration has partially undone the 1974 decision of the post-Watergate Congress to separate safety regulation from nuclear power promotion by breaking up the

Atomic Energy Commission to create the NRC and Energy Department.

Even before the Trump administration, under political pressure from the nuclear industry through congressional Republicans, the NRC commissioners backed off by majority vote from requiring filtered vents for a set of US reactors designed by General Electric that were clones of the Fukushima-Daiichi reactors 1–3, whose small-volume containments released large amounts of radioactivity due to overpressure after core meltdowns. The NRC also refused to end the practice of dense-packing spent fuel pools to five times their design density despite Fukushima unit 4's near miss of a potentially much more catastrophic spent-fuel fire because of an undetected water level drop.

The end of ALARA

After it was effectively given much of the responsibility of regulating the US nuclear industry, the Energy Department commissioned a review of the LNT hypothesis by the Idaho National Laboratory, which supports the Office of Nuclear Energy's mission to promote new types of nuclear power reactors. INL quickly produced a report, which cited a 2013 comparison by Mohan Doss of the LNT model against the radiation hormesis, as "[p]erhaps most significant for regulatory considerations." Dr. Doss is a radiologist, not an epidemiologist. His article was published in the journal *Dose-Reponse*, which was founded in 2003 with Professor Calabrese as its editor-in-chief and focuses on hormesis advocacy. Contrary to what the INL report claims, Dr. Doss' article is *not* a meta-analysis but rather an argument for radiation hormesis.

Doss starts by arguing at length that the atomic bomb survivors study would have shown a hormesis effect had it been compared with a control group that had a higher incidence of cancer. Doss even replotted the atomic bomb survivor data to show the result if such a control group were used. In fact, there are appropriate zero-dose control groups for the atomic bomb survivors study, including those who were away from the cities at the time of the bombings. When those control groups have been used in studies, they showed some non-linearity with dose for male cancers, but no hormesis effect.

At the same time, INL referenced but ignored the findings of two actual meta-analyses of low-dose studies: one by the National Council on Radiation Protection and Measurements and one by an international team of 16 cancer epidemiologists led by Michael Hauptmann and published in the *Journal of the National Cancer Institute* and partly funded by the National Cancer Institute, National Institutes of Health, and the Energy Department.

The National Council review concluded that "no alternative dose-response relationship appears more pragmatic or prudent for radiation protection purposes than the LNT model." Hauptmann and colleagues found that "there is evidence of cancer risks from low-dose ionizing radiation."

INL's "reevaluation report" was quickly cited in a memorandum by the Department's Undersecretaries of Science and Nuclear Security recommending that the Secretary of Energy "eliminate ALARA from all Department of Energy Directives and Regulations," which he reportedly has done.

In the absence of an objective ALARA cost-benefit analysis, future decisions on limiting doses from ionizing radiation to workers and the public from nuclear power operations will be determined in significant part by the relative political strengths of industry and regulators. Under the Trump administration, the industry clearly has the upper hand.

The Trump administration's Environmental Protection Agency has recently made a similar decision that it will no longer take into account the health benefits from limiting air pollution. In 2024, the Biden administration announced new limits on fine particulate pollution from coal power plants and other facilities. Those regulations were justified by an estimate that, on average, 77 dollars in health benefits would result from each dollar spent by industry on emission reductions and that the regulations would save 4,500 lives per year.

A climate reporter commented in the *New York Times* about the Trump administration's decision to roll back the air-pollution regulation that, for over four decades, "different administrations have used different estimates of the monetary value of a human life in cost-benefit analyses. But until now, no administration has counted it as zero."

Just as it did with air pollution rules, the Trump administration has now, in effect, set the value of American lives to zero in regulatory protections against nuclear-radiation-caused cancer.

The damage that will result from the evisceration of the Nuclear Regulatory Commission will not be immediate and may arguably turn out to be minor on the scale of the damage the Trump administration is doing in other policy areas. But public safety analysts and decision makers must keep track of the dismantlement of regulatory structures that have been built over generations. Hopefully, it will be possible to reconstruct some of them, with improvements where possible. In the meantime, however, the attacks of the Trump administration on public safety must be exposed.

Frank N. von Hippel is a professor of Public and International Affairs emeritus, Princeton University, a co-founder of the Program on Science and Global Security at Princeton University's School of Public and International Affairs, a founding co-chair of the International Panel on Fissile Materials, and a member of the *Bulletin*'s Board of Sponsors. A former assistant director for national security in the White House Office of Science and Technology, von Hippel's areas of policy research include nuclear arms control and nonproliferation, energy, and checks and balances in

policymaking for technology. Most recently, he has been the coauthor of the book *Plutonium: How Nuclear Power's Dream Fuel Became a Nightmare* (Springer, 2019) and the report *Banning Plutonium Separation* (International Panel on Fissile Materials, 2022).

The World's First Nuclear Waste Tomb Is Nearly Ready to Open

By Anna Korkman, AFP

Source: https://www.sciencealert.com/the-worlds-first-nuclear-waste-tomb-is-nearly-ready-to-ope

An underground tunnel leading to Finland's Onkalo nuclear repository. (Alessandro Rampazzo/AFP)

June 02 – The elevator display reads "433", the number of meters below ground.

The doors slide open, revealing the entrance to what is expected to be the world's first permanent repository for radioactive spent nuclear fuel.

Blasted into 1.9 billion-year-old stable bedrock in Eurajoki, southwest Finland, the geological repository for spent nuclear waste – dubbed Onkalo, which means "cave" in Finnish – is nearly ready to start operations.

Countries have been wrestling with what to do with dangerous nuclear by-products since the first plants were built in the 1950s. Currently, most of it is in temporary storage.

Final repositories are being built in other countries, including neighboring Sweden and France, but Finland is expected to be first to open an underground storage solution.

The Finnish Radiation and Nuclear Safety Authority (STUK) is due to give approval in its final assessment in June, after which an operating license can be granted.

"We hope we can start the operation either at the end of this year or most

probably at the beginning of next year," said Philippe Bordarier, chief executive of nuclear operator Teollisuuden Voima Oyj (TVO).

His voice echoed in the damp tunnel where the spent nuclear fuel will be buried in holes drilled into the bedrock, where it will remain harmfully radioactive for thousands of years.

The 'hot cell' fuel handling chamber at the encapsulation plant of nuclear waste management company Posiva. (Alessandro Rampazzo/AFP)

The waste currently cooled in water pools at an interim storage site, at the nearby Olkiluoto power plant next to the Baltic Sea, will be first to be deposited, Bordarier said.

With space for 6,500 tons of uranium, Onkalo is aimed at providing permanent storage for spent fuel from Finland's five nuclear reactors – three of them located in Olkiluoto.

Nuclear waste management company Posiva began building the site in 2004, with the cost now estimated at one billion euros ($1.16 billion US).

'Forever'

Spent fuel is planned to be deposited in Onkalo's massive network of tunnels for 100 years, but operations may be extended if new nuclear reactors are built.

Subsequently, the vault will be sealed to provide safe storage for at least 100,000 years. "Basically, it needs to be safe forever," noted Lauri Parviainen, a Posiva chemist who showed reporters around the facilities.

The fuel will be highly radioactive for "tens of thousands of years", he said.

After 100,000 years, they will be "about the same level as the uranium ore of which the fuel is made."

Above ground, the spent nuclear fuel will be encapsulated in highly corrosion-resistant copper canisters.

The canisters will be lowered into holes drilled in the tunnels, before the holes are filled with bentonite clay to seal them, Parviainen explained.

"So if the bentonite stays in place, we are safe," he said.

Once each 300-meter-long disposal tunnel is filled, it will be sealed with a steel-reinforced concrete plug.

Long-term risks

Jarkko Kyllonen, an expert on nuclear safety at Finland's nuclear regulator STUK, has assessed risk scenarios for the Onkalo project stretching up to a million years into the future.

Considering the "hazard potential of the waste, the first 10,000 years are very important for keeping the capsules intact," he told AFP. The main long-term risks are corrosion of the copper canisters or earthquakes during future ice ages, which could potentially damage the capsules and cause radioactive fuel to leak, Kyllonen said. But the results of various risk assessments conducted over the years have been "positive".

While France's plans for a similar underground nuclear tomb have met with strong opposition, Onkalo has received broader backing in Finland.

There was some opposition locally when the plans were first introduced in the 1970s, but "people have gotten used to it, and they trust the assessments made by STUK", Matti Kojo, social sciences professor at LUT University, told AFP.

"At the moment, support for nuclear power is at a historically high level in Finland," he noted.

The Finnish Association for Nature Conservation remains critical of the project, however, insisting that nuclear waste poses a long-term, serious risk.

"No one can guarantee the safety of Onkalo for thousands of years," director Tapani Veistola told AFP in an e-mail.

Finland's nuclear push

Under Finnish law, nuclear waste produced in Finland has to be deposited in the country, Climate and Environment Minister Sari Multala told AFP.

"Before the legal change in 1994, the spent nuclear fuel was exported to, for example, Russia," she said.

Increasing nuclear power in Finland has been a priority for the right-wing government, and the country is considering building so-called small modular reactors (SMRs).

How the spent nuclear fuel from future SMRs would be managed "has not been decided yet," Multala said. An assessment should be completed by March next year, she added.

As NATO eyes the nuclear bomb, did anyone check whether it works?

By Andrius Balčiūnas

Source: https://www.lrt.lt/en/news-in-english/19/2949001/as-nato-eyes-the-nuclear-bomb-did-anyone-check-whether-it-works

June 03 – Lithuania has embraced France's offer of shelter beneath a French "nuclear umbrella", and discussions are now under way about allowing ships carrying weapons of mass destruction to enter Lithuanian ports. These moves are a response to Russian and Belarusian actions. But do nuclear weapons actually deter?

"The theory whether or not nuclear weapons deter and prevent wars is almost theological, You know – it becomes an article of faith. You believe in it, or you don't, but there's no evidence base," Andreas Persbo, director of the Open Nuclear Network programme at the non-governmental organisation Pax Sapiens, told LRT.lt in an exclusive interview.

Humanity will inevitably abandon such weapons, Persbo believes – though that is a question of decades, perhaps even centuries.

“I don't think they deter. I think a world war is bound to happen at some point in human history. In the Baltic states, many seem to think it's coming faster than I think it is, but it's bound to happen at some point. And then we're going to have to fight that war with nuclear weapons being present. And I think that increases risks dramatically. Does it mean they're going to be used? I don't know,” the expert argues.

Persbo has accumulated several decades of experience working in nuclear weapons monitoring and non-proliferation. He was speaking in Vilnius, where he participated in the Andrei Sakharov Conference at Vytautas Magnus University.

Many conflicts are underway around the world, and some argue that as a result, more and more states wish to acquire their own nuclear weapons, or shelter beneath an ally's. Has the era of nuclear non-proliferation ended?

No, I don't think so. In the 1960s, Kennedy said something along the lines of, ‘If we don't control proliferation, we'll have 30 states with nuclear weapons before the end of the century.’ At the moment we have nine – five states that are the victors of the Second World War, who just happen to be the permanent members of the UN Security Council. And four more states – India, Pakistan, Israel, which is presumed to have weapons, though they never confirmed it, and North Korea.

It's only really North Korea that has developed nuclear weapons lately. So we

have just one proliferation case in the last 50 years, which is a pretty good track record.

People think Iran has an intent to hold nuclear weapons. We know Syrians were looking into nuclear weapons, but the nuclear reactor they built in the desert was bombed. We know that Gaddafi was looking at nuclear weapons in the 1990s and early 2000s. And of course, we had South Africa in the 1980s. But other than that, it's a pretty good track record. And I think that is because the majority of states that have signed the Nuclear Non-Proliferation Treaty subscribe to this idea that the more nuclear weapon states you have, the more complicated deterrence relationships and military balances become. The realisation that more, in this case, is not necessarily better is what holds that regime together.

A photo released by Russia's Ministry of Defence, purportedly showing the Oreshnik system being deployed in Belarus | AP

That said, is this regime under pressure? Yes, undeniably. We're seeing that with the open discussions that took place here in Lithuania a few days ago, about basing, not acquiring the weapons themselves, but basing weapons, which is a totally understandable reaction to Belarus' actions. We're seeing similar debates in Sweden, Germany – debates on whether Europe needs its own nuclear arselan should the United States withdraw interest from Europe. Such conversations are also becoming more load in South Korea and Japan. But at the moment it's mostly talk, caused by this enormous pressure that we're presently under. So, I'm not too worried now. Ask me again, in a decade, if I'm still around – I might be more worried. But not not so much right now.

Can nuclear weapons genuinely make a country safer? If allies were to deploy nuclear weapons in Lithuania or Sweden, for instance, would we be more secure?

I think the problem is that this idea of nuclear deterrence hasn't really been proven: does it work, does it not work? What we know is that possessing nuclear weapons hasn't deterred the Indians or the Pakistanis from attacking each other. It has not deterred Israel's adversaries from attacking it.

A theory I have is that nuclear weapons possession is just a tool that powerful countries can use to steamroll smaller ones. Like in the case of Russia invading Ukraine – no one does anything meaningful because they all fear Russia's nuclear weapons. Yet Ukraine has been fighting for five years, and does not appear to have been deterred by Russian nuclear weapons. Something is missing in this debate. I sometimes say is that the theory of nuclear deterrence is almost theological – it becomes an article of faith. You believe in it, or you don't, but there's no evidence base.

I'm leaning towards thinking that they don't deter. I think, a world war is bound to happen at some point in human history here in the Baltic states Many seem to think it's coming faster than I think it is, but it's bound to happen at some point. And then we're going to have to fight that war

With nuclear weapons being present. and I think that increases risks dramatically. Does it mean they're going to be used, I don't know, but it increases the risk, Absolutely.
I tend to think they do not deter. I think a world war is bound to happen at some point in human history. In the Baltic states, many seem to think it's coming faster than I think it is, but it's bound to happen at some point. And then we're going to have to fight that war with nuclear weapons being present. And I think that increases risks dramatically. Does it mean they're going to be used? I don't know.
From my perspective, the fewer of them we have, the better it is. And I'm very cold-hearted about this. I'm just thinking about previous world wars. Look at the Second World War. 60 million dead. The total explosive yield of every weapon used was perhaps equivalent to one, two, or three Russian nuclear warheads – far less than most people imagine. These are extremely powerful weapons. Far more powerful than most people conceive. Whatever figure you have in mind, multiply it by two or three times, and you will be closer to reality.

Kim Jong Un watches a ballistic missile test | AP

It took some time for Europe to rebuild, but it recovered. Any large conventional war is undesirable, obviously, but humanity was able to bounce back. If a war takes a nuclear dimension, and it doesn't stop – people start to use nuclear weapons continuosly, unable to stop it – we can't bounce back from that. That's what worries me. It was Einstein, I think, who said, "I don't know how the Third World War will be fought, but the fourth will be fought with sticks." That is precisely the problem.

Does the risk of escalation concern you – that countries might blindly climb the escalation ladder until nuclear weapons are actually used?

Although they have not been used, there is an awareness that nuclear weapons could be used. Though they never said it publicly, one of the concerns the United States had in the beginning of the war in Ukraine was 'if we arm up the Ukrainians with long-distance weaponry, well, then the Russians might respond nuclear.' And so, we're down that slippery slope.
Now, Ukraine is striking Russia deep every day, and we're still not seeing any use of nuclear weapons. And these strikes are painful for Russia, they are really taking some infrastructural damage, but the nuclear threshold has not been crossed.
In 2022, when Russian forces began to retreat in Ukraine, US intelligence assessed the probability of nuclear weapons being used at 50-50. That assessment must have been based on something, whether or not it was correct. And they were concerned about making decisions under pressure.
And that is the issue. The authorisation to use nuclear weapons rests with a hanful of people. It's not a democratic decision.In the United States, it's one person: the president. In Russia, it may be three: the president, the defence minister, and the chief of the general staff. But will

these decision-makers, in circumstances of extreme pressure, act rationally?

I am concerned about the slippery slope of escalation in the sense that, under severe pressure, people tend to act emotionally rather than rationally. In that particular moment, using a nuclear weapon might seem logical. That is worrying. Right now, the likelihood of this happening remains extremely low. Then again, we are not living in normal times.

Regimes such as North Korea and Iran regard nuclear weapons as a guarantee of their survival. Is that a credible strategy?

North Korea clearly thinks so. It was not always the case – for the best part of 20 years under Kim Jong-il, they were not actively pursuing a nuclear arsenal.

It is a very poor country, ranked around 180th in the world by GDP. That also tells us something else: nuclear weapons are, in a sense, cheap. If a country like North Korea or Pakistan can get them. Literally anyone with access to natural resources and scientific know-how can get them. It's a myth that they're safeguarded in some James Bond-style facility. It is a dirty industrial process, built on well-understood science.

Since taking power, Kim Jong-un has invested heavily in this area. We do not know what proportion of North Korea's defence budget goes on nuclear weapons, but we know the defence budget is very large relative to their minuscule GDP. Many of the underlying principles being pursued today were already in development a decade ago. Remember, it is also a society where mistakes are not tolerated but mistakes are made constantly: they misfire missiles from time to time, rockets explode – yet no one gets fired. So he is deeply determined and clearly sees the strategic value of nuclear weapons.

Military parade in Beijing | AP

Now, has the nuclear programme elevated North Korea on the world stage? One thinks of that famous image of Putin, Xi Jinping, and Kim Jong-un walking side by side as equals. Think about that image. Russia, a vast country of 140 million people, well, perhaps not the wealthiest, with its GDP being comparable to Spain's; then China, the world's second-largest economy with the largest navy in the world, a real military power; and North Korea, ranked 180th. Yet there they are, seemingly working as peers. Is that because of nuclear weapons? I would say probably not, rather, it is because Kim is arming the Russians, and Moscow sees strategic value in that relationship.

It certainly does Kim no harm. He presumably regards it as enormously enhancing his prestige and establishing North Korea as a major player on a par with the United States.

How concerned are you about new nuclear weapons systems? Do they further destabilise the situation?

Deterrence theory rests on the assumption of mutual destruction. It presumes that the opponent is making rational choices, which, as I pointed out, you can't be certain of in any given case, and it also assumes that you get some sort of an advanced warning. Like, the idea that if the Russians were to launch at the United States right now, the Americans would have about 30 minutes or so to decide on whether to retaliate. But if the weapon is an underwater drone with

a several-megaton warhead, or a nuclear-powered cruise missile system, you don't get any warning like that. These weapons are straight out of a Bond villain's arsenal; they really are. Some kind of product of deranged scientific thinking. They unbalance the system, without question. Ordinarily, adversaries develop counter-systems and countermeasures in response. But that simply means you then have two such systems, which destabilises things further and makes the whole arrangement more vulnerable to error.

The last remaining US-Russia arms control agreements have lapsed without renewal, and China is announcing plans to expand its nuclear arsenal. Has the appetite for arms reduction or control simply disappeared?

The question is, was there ever a lot of will to disarm? If you look at all of the strategic arms limitations agreements or reduction agreements, if you look at warhead numbers and reduction numbers, you will find that the agreement comes in at a point where the numbers are already there anyway. So states have already decided to reduce their arsenals, and then they come to an agreement saying, let's keep it at this level. That's different from agreeing to genuinely reduce the arsenal. The largest reductions in nuclear weapons have been more or less unilateral decisions. After the fall of the Soviet Union, President George H.W. Bush got rid of a very big proportion of the American arsenal because it was no longer needed. The Soviets said they couldn't afford to maintain theirs, so the Russians agreed to get rid of them. But these decisions were not driven by a desire to disarm, it was driven by the fact that expensive weapons were no longer considered necessary by anyone. Today, as we have discussed, people are once again seeing greater utility in nuclear weapons. So, I believe the current level is about as low as it will go. Americans argue they need several thousand warheads because of the amount of different threats they face and worry about. My usual response is that a single warhead constitutes a heavy and sufficient deterrent. Two megatons over Moscow would cause immeasurable suffering. Britain and France each hold around 160 to 200 warheads – a sort of minimum viable level. Israel, Pakistan, and India also have fewer, and appear to operate on the assumption that each individual nuclear weapon is capable of inflicting so much damage that they don't need thousands. The United States and Russia seem to be engaged in something more quantitative – and China now wants to join them. In the near term, I see no prospect for arms control. But over the longer term, these weapons these weapons will continue to prove themselves a bit worthless. We will see big wars – hopefully fought without nuclear weapons. If they are used, we're going to get rid of them in the aftermath. If they're not going to be used, people will ask why we need weapons we never use – and we will rid ourselves of them. Either way, the pressure to retain these weapons will diminish, and the pressure to get rid of them will stay there and grow over time. But I am talking about decades, not tomorrow. Consider how long it takes humanity to abandon a class of weapons. From the first use of chemical weapons in the First World War to the Chemical Weapons Convention was nearly 100 years. Biological weapons were first deployed in the medieval days; by the sixteenth century there was already a sense that something needed to be done. But the Biological Weapons Convention took roughly 300 years to arrive. Nuclear weapons were created and used in 1945 – is it even realistic to expect we're going to get rid of them any faster? I would say no. We are talking about at least 100 years, possibly 150 or 200.

During the war in Ukraine, there were fears Russia might use tactical nuclear weapons. Are these fundamentally different, or once that threshold is crossed, is there no way back?

The Russians have various theories about this. They claim tactical weapons have battlefield utility, but I am not convinced by that. We are talking about an explosion equivalent to 150,000 tonnes of TNT, or 150 million kilograms. The largest JDAM bomb strikes in Iraq did not come close to a single tonne. These are insane numbers – literally beyond the capacity of the human brain to comprehend how big they are. They are so large that the human brain cannot grasp them. These are not weapons for engaging tanks on a battlefield. In my view, all such weapons are strategic.

Russia has its own ideas, of course. But Russians like to justify things, sorting everything into neat categories – a very Soviet way of thinking. Everything must be tidy, properly categorised and filed away in its box.

What is nuclear order?

By Tianjiao Jiang

Source: https://thebulletin.org/2026/06/what-is-nuclear-order/

June 02 – The Cold War-era nuclear order—in which the United States and the Soviet Union held sway, establishing treaties that set the nuclear rules for the whole world—has clearly and irreversibly come to an end, but a new order to replace it has not yet coalesced. With the rise of China, India, and other powers, the world has drifted toward multi-polar power dynamics. But with that drift has come a heightened risk of nuclear conflict, the resurgence of arms racing, the proliferation of nuclear weapons, and the erosion of the longstanding nuclear taboo.

These developments make a reshaping of the current world order imperative. To be

effective, that reshaping should be accomplished via incremental reform, enabling the nuclear order to be continually refined and corrected in response to evolving global power configurations. To initiate such a process of reform, a composite international mechanism—one that includes bilateral and multilateral arrangements—should be instituted in a dedicated summit of the world's five nuclear-armed states that focuses on nuclear risk reduction and includes invitations to all other nuclear-weapon states. Within and alongside this summit, the United States, Russia, and China should establish their own bilateral dialogues to enhance nuclear transparency and minimize the risk of miscalculation. And because technological change increasingly affects strategic stability, the intersection of emerging technologies and nuclear weapons should explicitly be a key component of all the aforementioned dialogues.

The order that was

The contemporary international nuclear order stems from the nuclear competition between the United States and the Soviet Union during the Cold War. To safeguard their hegemonic status while averting global catastrophe, these two superpowers established a series of international norms, institutions, and agreements to regulate nuclear activities, promote non-proliferation, and facilitate arms control. This Cold War-era nuclear order was characterized by several key components. First, the United States and the Soviet Union achieved strategic stability through the doctrine of "mutually assured destruction." Second, arms control negotiations resulted in significant bilateral treaties, such as ABM, INF, SALT, and START, preventing an endless arms race.[1] Finally, the normative prohibition on nuclear weapons use became a widely accepted international taboo, supported by numerous non-governmental organizations and scientific communities highlighting the catastrophic consequences of nuclear warfare.

However, this nuclear order was inherently contradictory. The United States and Soviet Union sought strategic stability while competing for nuclear dominance. Since nuclear deterrence relies on manipulating risk, any security architecture based on deterrence carries the danger of escalation into nuclear conflict. Meanwhile, the nuclear Non-Proliferation Treaty (NPT) regime aimed to curb proliferation while reinforcing nuclear hegemony, allowing dominant powers to maintain their deterrence capabilities and prevent others from developing similar arsenals, thus preserving their strategic advantages and influence in international affairs.[2] Non-proliferation thus consolidates hegemonic control over allies[3] and prevents weaker adversaries from provoking major powers with nuclear weapons.[4] This inherent inequality has led to persistent criticisms of the NPT regime, which institutionalizes a hierarchical distinction between nuclear and non-nuclear states, despite the UN charter's principle of sovereign equality.

Both China and France strongly criticized the international nuclear order when that treaty was being adopted. During the Korean War and the Taiwan Strait Crisis in the 1950s, China faced nuclear coercion from the United States, and in the 1960s, a nuclear threat from the Soviet Union as Sino-Soviet relations worsened. In response, Chinese leaders asserted that China must develop nuclear weapons to safeguard itself from nuclear blackmail, while simultaneously denouncing the US-Soviet nuclear arms race and their pursuit of nuclear hegemony. After its 1964 nuclear test, China immediately adopted a policy of unconditional no first use and pledged not to use nuclear weapons against non-nuclear states, demonstrating a commitment to self-restraint in nuclear strategy. In China's strategic culture, nuclear weapons are regarded primarily as political tools to deter nuclear war and counter blackmail, not battlefield weapons. Consequently, China has maintained a limited, credible nuclear force and deterrence, without engaging in an arms race.[5]

Throughout the Cold War, China remained outside the NPT, viewing it as a tool of US-Soviet hegemony that reinforced nuclear inequality. However, by the 1980s, aiming to integrate into the international system, China recognized the adverse effects of nuclear proliferation on its peaceful development and the diplomatic and reputational challenges of refusing to join the NPT. Accepting the difficulty of achieving global nuclear disarmament in the short term, China decided to accede to the treaty. Nevertheless, China's views on nuclear weapons diverged significantly from those of the United States and the Soviet Union. In the 1990s, China proposed a global no-first-use treaty and urged the U.S. and Russia to lead nuclear disarmament efforts.

Why is the current nuclear order being challenged?

Evolving multipolarity, intensifying geopolitical tensions, and rapid technological advancements have profoundly shaped the current and future nuclear order. The interplay of these three factors fuels the arms race and heightens nuclear risks. First, the global power structure has shifted significantly, with multipolarity becoming more pronounced. The traditional nuclear order, historically dominated by the United States and the Soviet Union, is no longer sufficient to address these evolving dynamics. After the Cold War, the international community entered the

"second nuclear age," marked by more nuclear-armed states and the rapid development of missiles, weapons of mass destruction, and disruptive technologies. Simultaneously, the proliferation of weapons of mass destruction (WMD) among non-state actors has become a persistent threat. During the Cold War, the United States and the Soviet Union established nuclear deterrence and strategic stability through mutual assured destruction. Today, countries like China, France, the U.K., India, Pakistan, Israel, and North Korea have diverse views on nuclear deterrence and have

adopted distinct strategies. Moreover, nuclear deterrence against non-state actors can be very difficult, adding uncertainty to the traditional framework of nuclear stability.

It is important to highlight that the multipolarity in global politics is advancing faster than in the nuclear domain, placing the existing nuclear order under dual pressure from the Global South. For instance, following its expansion, the BRICS bloc now represents approximately half of the world's population and a third of global GDP.

On critical issues like the global economy and climate change, the Global South is shifting from a "passive norm taker" to a "norm shaper."[6] In the nuclear domain, the Global South has long criticized the discriminatory nature and persistent nuclear risks associated with the NPT. Today, nuclear weapon states not only neglect their disarmament obligations but are also caught in new arms races and heightened nuclear threats. If non-nuclear-weapon states' concerns remain unaddressed, the Treaty on the Prohibition of Nuclear Weapons (TPNW) will increasingly challenge the NPT. Moreover, some Global South states may pursue nuclear proliferation to enhance their security and international standing and to achieve strategic influence and recognition. These developments collectively exert immense pressure on the stability of the international nuclear order.

Second, geopolitical tensions have significantly disrupted the global nuclear order. The US withdrawal from the ABM Treaty in 2001 marked a turning point, triggering a crisis in US-Russia strategic stability. NATO's eastward expansion and US missile defense systems in Europe have prompted Russia to rely more on nuclear deterrence, while the United States and NATO have reinforced their own deterrence and missile defense, escalating a cycle of competition.

The Ukraine crisis further exacerbated tensions, eroding the political trust between the United States and Russia. Geopolitical crises have derailed some arms control agreements, culminating in the collapse of the Cold War-era arms control architecture and a reversal of the nuclear disarmament process. Simultaneously, the longstanding nuclear taboo is increasingly being challenged. Russia has signaled its willingness to use nuclear weapons in the Ukraine crisis, revising its nuclear doctrine, while the United States has proposed a cross-domain deterrence strategy, ready to employ nuclear strikes in response to major non-nuclear attacks. Amid this renewed arms race, the United States and Russia are developing new nuclear weapons, such as low-yield, precision warheads, and exploring doctrines for their "limited use" in conventional conflicts. This shift poses a significant risk by lowering the threshold for nuclear use, increasing the likelihood of escalation.

This strategic interaction between the two nuclear superpowers deeply affects the global nuclear order, posing major challenges to both that order and the non-proliferation regime, encouraging other nuclear-armed states to strengthen their deterrence. This escalates regional proliferation risks, particularly in South Asia, Northeast Asia, and the Middle East, where nuclear challenges persist.

Particularly in the wake of the Ukraine crisis and the resurgence of the Trump administration, countries such as Japan, South Korea, Poland, Türkiye, and Saudi Arabia have actively debated the possibility of deploying or acquiring nuclear weapons. These discussions stem from growing concerns over future regional conflicts and the credibility of extended nuclear deterrence. Meanwhile, the United States' increasingly inward-looking policies and the potential for prolonged populist governance have compelled European nations to reassess their strategic autonomy.

European leaders are alarmed by the potential for the United States to pressure Ukraine into accepting a subordinate geopolitical role, seeing it as a sign of broader international order erosion. As a result, Europe has entered a "moment of awakening" characterized by substantial increases in defense expenditures and the expansion of long-range missile programs to strengthen military capabilities. France has explicitly declared its willingness to extend a nuclear umbrella over Europe, but some European nations question whether the nuclear arsenals of France and the UK alone can provide adequate security. The ongoing militarization of Europe, especially the proliferation of long-range missile systems, is likely to provoke countermeasures from Russia, increasing the risks of nuclear proliferation and military escalation.

A similar trend is unfolding in the Asia-Pacific region. North Korea's ongoing nuclear threat has intensified discussions in Japan and South Korea about nuclear sharing or developing independent nuclear capabilities. Doubts about the reliability of external security guarantees have fueled these debates, reflecting broader changes in regional security dynamics. In addition, Western countries are increasingly concerned about China's evolving nuclear posture. Over the past two decades, China has rapidly expanded its economy and military, significantly enhancing its strategic nuclear capabilities. Owing to differences in strategic culture, China has traditionally kept a small nuclear arsenal and maintained low transparency to bolster its deterrence. Consequently, Western countries frequently speculate about the trajectory of China's nuclear development and the strategic intentions underpinning its modernization efforts. The United States has categorized China and Russia as nuclear peers, asserting that future conflicts in the Taiwan Strait pose significant nuclear risks and advocating for China's inclusion in the nuclear arms control framework alongside the United States and Russia.

However, China regards such concerns as misunderstandings and distortions and an inherently unfair assessment. Historically, China and the United States have never established a strategic stability framework comparable to that between the United States and the Soviet Union/Russia. Until the Obama administration, the US Nuclear Posture Review acknowledged the necessity of

maintaining mutual vulnerability between China and the United States. However, this concept was never institutionalized or reinforced through formal agreements. China has long been apprehensive that its relatively limited nuclear arsenal could be vulnerable to a preemptive US strike, while the United States has consistently refused to assure that it would avoid military actions targeting China's nuclear retaliatory capabilities. Due to domestic political constraints and opposition from its Asia-Pacific allies, the United States has been unwilling to officially recognize a stable strategic relationship with China based on mutual vulnerability. Under the Trump administration, the United States reoriented its defense strategy toward great power competition, initiated nuclear force modernization, and expanded its missile defense capabilities. Additionally, escalating trade tensions and deteriorating bilateral relations between China and the United States further undermined opportunities for meaningful dialogue on strategic stability.

The evolving multipolarity and escalating geopolitical tensions have driven nations to advance emerging technologies, intensifying technological competition and further reinforcing multipolarity and geopolitical rivalries. The risks and uncertainties associated with emerging technologies and nuclear weapons underscore the fragility of nuclear deterrence. Conventional military advancements, such as missile defense systems, conventional prompt global strike weapons, and hypersonic missiles directly challenge the survivability of nuclear arsenals. Meanwhile, emerging domains like space, cyber technologies, and artificial intelligence introduce additional risks of accidental escalation. Major military powers are fiercely competing over these emerging technologies. However, the absence of effective international regulations governing state interactions in these fields exacerbates the risks of spillover effects and conflict escalation.

For instance, in missile defense, China's long-standing security concerns have not been adequately addressed by the United States.[7] Given its smaller nuclear arsenal, China is more vulnerable than Russia to the adverse consequences of the US withdrawal from the Anti-Ballistic Missile Treaty. Furthermore, the US deployment of the THAAD system in South Korea, claimed to being aimed at the North Korean nuclear threat, has weakened the credibility of China's nuclear deterrent. While the United States could mitigate China's concerns through public commitments and technical safeguards, ensuring that THAAD is not aimed at China, it has chosen not to do so. Instead, President Trump stated in the Missile Defense Review that the United States reserves the right to intercept missiles targeting its territory from anywhere and at any time. The expiration of the Intermediate-Range Nuclear Forces (INF) Treaty has worsened regional tensions, with the United States deploying intermediate-range missiles near China through regional allies, increasing the risk of nuclear escalation stemming from conventional missile strikes. A similar dynamic is unfolding in Europe, where NATO and Russia face analogous challenges. Consequently, China, Russia, and the United States are locked in a continuous cycle of offensive and defensive competition in hypersonic weapons and missile defense, directly influencing their nuclear postures and strategic adjustments.

In space, the Trump administration established the US Space Force and announced the "Golden Dome" initiative, increasing the likelihood of expanded missile defense capabilities and undermining strategic stability. The United States has also strengthened the resilience of its space infrastructure with revolutionary technologies like Starlink. It has explicitly stated that Starlink can be repurposed as a "star shield" for military use, blurring the line between military and civilian space assets.

China has already observed two instances when Starlink satellites exhibited "abnormal approaches" to its space station. Looking ahead, the Starlink project aims to deploy over 40,000 satellites, while China currently operates only about 900. This vast asymmetry alone poses a potential threat to China's critical space assets and could significantly impact military operations on Earth. Furthermore, China has observed Starlink's crucial role in the Ukraine conflict and is increasingly concerned about its potential use in future Taiwan contingency. However, existing international regulations impose no restrictions on Starlink, exacerbating the risks of escalation in space and increasing the likelihood of spillover effects into nuclear deterrence dynamics.

In cyberspace, the United States adheres to a doctrine of "persistent engagement" and "hunt forward" operations, which emphasize proactive cyber defense by extending operations into adversary networks. This approach blurs the line between cyber offense and defense. Additionally, the US military maintains the "left of launch" strategy, aiming to disrupt adversaries' missile capabilities through cyber intrusions and preemptive missile interception.

Finally, artificial intelligence has introduced greater complexity to the calculations underlying strategic stability. On one hand, AI-powered sensors and space technologies enhance intelligence gathering, improve transparency, reduce strategic miscalculations, and aid crisis communication, supporting strategic stability and non-proliferation efforts. On the other hand, AI-enabled satellite reconnaissance and cyber operations can more effectively detect concealed nuclear weapons and conduct targeted strikes, increasing the perceived risk of preemptive disarmament for states with smaller arsenals and lower survivability. Additionally, AI-integrated missile defense systems improve decoy identification, boosting interception rates.

Moreover, AI's broad military applications are speeding up decision-making, increasing uncertainty and escalating nuclear risks.

Amid rising multipolarity and geopolitical tensions, rival states increasingly suspect

one another of leveraging emerging technologies to erode their strategic advantages. As a result, major military powers are expected to intensify their competition over these technologies, amplifying the intersection of emerging technologies and nuclear risks. Traditional arms control treaties and mechanisms are ill-equipped to address these new challenges. Geopolitical frictions have also hindered meaningful international dialogue on emerging technologies, preventing collaborative efforts to mitigate risks and manage conflicts in this evolving security landscape.

Policy recommendations

While the Cold War-era nuclear order has ended, the emergence of a new, multipolar global nuclear order remains incomplete. To address the growing risks of nuclear conflict and proliferation, a gradual and systematic reshaping of the nuclear regime is essential, emphasizing incremental reforms, bilateral and multilateral dialogues, and the integration of emerging technologies to maintain global stability. Such a reshaping could include these elements:

- The five nuclear powers identified in the Non-Proliferation Treaty (P5) should convene a dedicated summit on nuclear risk reduction, inviting all nuclear-armed states to participate. Given the gravity of nuclear issues, heads of state must take primary responsibility for nuclear policy and war prevention. With the current geopolitical polarization, traditional diplomacy alone is inadequate. The P5 should reaffirm the principle that "a nuclear war cannot be won and must never be fought" and commit to a "no first test" policy to strengthen the nuclear testing moratorium. Additionally, they should enhance negative security assurances to non-nuclear states, alleviating nuclear threat concerns and reducing proliferation risks. Non-nuclear states should also unite in calling for a global dialogue on nuclear risk reduction. The United States and Russia must resume strategic dialogue at the earliest opportunity, while China and the United States should also establish strategic dialogues to enhance transparency and minimize miscalculation risks. The absence of mutual understanding regarding each side's strategic intentions has fueled US concerns over China's nuclear capabilities, often driven by worst-case scenario assessments. Concurrently, China remains deeply apprehensive about the credibility of its nuclear deterrent due to the United States' expanding missile defense systems, forward deployment of intermediate-range missiles, and strategic community discussions on limited nuclear war and escalation. To mitigate these risks, both sides must prioritize sustained high-level engagement to dispel misconceptions, build trust, and reduce tensions. The intersection of emerging technologies and nuclear weapons should be a key focus of the dialogue. The agreement between China and the United States on ensuring human control over critical nuclear decisions sets a valuable precedent for global governance. Building on this, China, the United States, or the P5 could explore AI's impact on strategic stability and establish a responsible framework for AI use in the nuclear domain. However, the challenges of verifying emerging technologies complicate traditional arms control mechanisms. Therefore, dialogues should include experts from emerging technology fields and the private sector, fostering innovative approaches to address technological risks. Where bilateral or multilateral negotiations are not feasible, states should adopt unilateral risk-reduction measures, such as "fail-safe" reviews, to mitigate nuclear risks from accidents and errors, paving the way for future collaboration.
- Regarding regional non-proliferation challenges, it is imperative to address both the symptoms and root causes, balancing security concerns with economic development. The P5, alongside regional stakeholders, should collaborate to provide credible security assurances and establish political and economic engagement mechanisms that alleviate proliferating states' security concerns, reducing their perceived need for nuclear weapons. Furthermore, the international community must remain vigilant about emerging nuclear proliferation risks. The P5 should engage in dialogue with these emerging proliferation-prone countries, working to restore regional strategic balance and address their security concerns through political commitments, security frameworks, and necessary conventional military cooperation.
- The future nuclear order should prioritize the indivisible survival and security of all humanity, requiring collective responsibility from all nations. In a multipolar world, the P5 must cooperate in managing nuclear risks rather than escalating nuclear competition or confrontation. These nuclear weapons countries should engage more closely with non-nuclear-weapon states, addressing their concerns and striving to rectify the inequities of the current nuclear order. In particular, the governance of emerging technologies must integrate the perspectives of non-nuclear-weapon states and the Global South. Security governance should not come at the expense of the development of southern nations. Instead, all countries must collaborate to foster a global security culture that unequivocally rejects the use of

nuclear weapons, with the ultimate goal of achieving complete nuclear disarmament. Policies such as sole-purpose declarations or no-first-use commitments play a constructive role in this process by reducing reliance on nuclear weapons, lowering the risk of nuclear use, and mitigating overall nuclear threats. To enhance the credibility and practical implementation of such policies, experts and scholars should continue to conduct in-depth research and policy development.

Notes

[1] Anti-Ballistic Missile Treaty, Intermediate-range Nuclear Forces Treaty, Strategic Arms Limitation Treaty, and the Strategic Arms Reduction Treaty, respectively.

[2] Michael C. Horowitz, The Diffusion of Military Power: Causes and Consequences for International Politics (Princeton, N.J.: Princeton University Press), p. 106. See also Report by the Committee on Nuclear Proliferation, January 21, 1965, FRUS, 1964–1968, Vol. 11.

[3] Francis J. Gavin, Strategies of Inhibition: U.S. Grand Strategy, the Nuclear Revolution, and Nonproliferation, International Security (2015) 40 (1): 9–46.

[4] Mark S. Bell, "Beyond Emboldenment: How Acquiring Nuclear Weapons Can Change Foreign Policy," International Security, Vol. 40, No. 1 (Summer 2015), pp. 87–119.

[5] 李彬：《中国核战略辨析》，《世界经济与政治》2006年第9期，第16-22页。

[6] 黄宇韬：《从自主争论到目标争论——新兴国家如何推动国际规范的转变》，《世界经济与政治》2023年第4期，第62-95页；徐进：《理念竞争、秩序构建与权力转移》，《当代亚太》2019年第4期，第4-25页。

[7] Within the framework of the China-U.S. Track 2 Strategic Dialogue, Chinese experts have persistently signaled their concerns to Washington, urging the United States to impose constraints on its missile defense capabilities. However, the United States refused to offer any assurances.

Tianjiao Jiang is an associate professor at Development Institute of Fudan University and visiting fellow at Paul Tsai China Center of Yale Law School.

Is there a nuclear world order?

By Benjamin Hautecouverture

Source: https://thebulletin.org/2026/06/is-there-a-nuclear-world-order/

June 02 – Eight decades after the introduction of nuclear weapons,[1] hindsight shows that the interest in this emblematic weapon system of the second half of the 20th century has always been heightened by the perception of an unusual risk of use. When this interest is marked, it is dramatized: It is always a question of a "return" of nuclear weapons, in other words of the threats that accompany them, even if it means refreshing them. As most of arms control treaties have been abandoned, and particularly since Russia launched an invasion of Ukraine using its nuclear deterrent as a tool of coercion, some nuclear experts have taken to arguing that the world nuclear order has been shattered. China's rise as a major strategic player is naturally part of this upheaval, even if perceptions of this factor still vary greatly among experts. Writing history always gives it meaning by ordering it. To introduce a global nuclear order that contemporary factors of disorder would come up against provides a usual framework: Western strategic analysis generally respects the framework to secure or recover an order which has generally been postulated as mandatory.

However practical this approach may be, it has the disadvantage of overlooking a major problem: The world nuclear order exists only in the minds of those who postulate it. The history of nuclear power over the last 80 years is much more the story of an order in search of itself, and of the many attempts to fix it.

A nuclear age or nuclear ages?

The main specific Cold War crises in the West had a latent or proven nuclear dimension: the exploitation of an alleged "missile gap"[2] with the Soviet Union from the end of 1957 to the 1960 American presidential campaign enabled the Democratic camp to denounce the culpable weakness of the Republican camp in the face of the Soviet competitor, which placed the territory of the United States at the mercy of ballistic nuclear attacks with continental reach. The six years of the Berlin crisis[3] were determined by the European nuclear issue and involved the use of thermonuclear weapons, with the Soviet authorities using nuclear blackmail against Western public opinion in a much more pronounced way than the Russians have done since February 2022.[4] The Cuban crisis,[5] intrinsically linked to the Berlin crisis, will always be remembered as the climax of the Cold War nuclear threat, although it was also the inaugural exercise in structuring bilateral strategic dialogue. Finally, the Euromissile crisis[6] sharpened fears of a Soviet first strike, highlighted the risk of a decoupling of American and European security interests, but also affirmed the role of public opinion—in this case European – in strategic nuclear issues, which became more global than strictly international. The high points of the post-Cold War period were in turn marked by a new nuclear apprehension: the fear of a globalised proliferation economy in the 1990s gave rise to an apocalyptic escalation whose virulence and excess we have forgotten today.[7] The risk of nuclear terrorism, described as an imminent threat in the wake of the September 11, 2001 attacks on US territory, was the subject of unvarnished analysis in the first decade of the century.[8] The strategic instability in Northeast Asia since the turn of the century, whether in terms of the North Korean regime's nuclear and ballistic missile programmes or the ongoing modernisation of China's strategic forces, began to play down the transnational issues at stake at the turn of the decade 2000. The invasion of Ukraine from 2014 onwards seemed to confirm this, as it began to dramatize the idea of a return to, but also a reconsideration of, nuclear deterrence.

The 2020 decade is no exception to the rule: By launching a second campaign in Ukraine at the end of February 2022, the president of the Russian Federation is said to have proved that the world has entered a time of unprecedented perils, specific to the era: Here we have the "third nuclear age" that would be ours, made up of "strategic piracy,"[9] "emerging and disruptive" technologies,[10] and "strategic surprises" within a "security architecture" that would have disintegrated.

It may be useful to identify distinct nuclear eras to understand what is at work in the nuclear realm over time. But writing history always gives it meaning by ordering it. Thus, to introduce a global nuclear order that contemporary factors of disorder would come up against provides a usual framework: Western strategic analysis generally respects the framework to secure or recover an order which has generally been postulated as mandatory.

From order to disorder

However practical this approach may be, it has the disadvantage of overlooking a major problem: The world nuclear order exists only in the minds of those who postulate it. The history of nuclear power over the last 80 years is much more the story of an order in search of itself.

Initially, nuclear weapons created a major strategic imbalance in that, from then on, a single weapon carried by a single delivery system could commit mass destruction by means of a surprise attack. Coupled with the United States' possession of nuclear secrecy, it became the instrument of a profound imbalance between adversaries. At the same time, the founding American theorist Bertrand Brodie, along with several others,[11] formulated the idea that the strategic objective of the nuclear age was no longer to win a war but to avoid it.

In this way, a new "world order" could be established: Nuclear weapons would reduce the likelihood of war if the adversaries' second-strike capabilities were assured. In conceptual terms, it was a revolution due to a technological exceptionality that was reinforced by the invention of the thermonuclear warhead in 1952. A factor of disorder, nuclear weapons therefore had the potential to become a factor of order; strategic thinking became paradoxical, and today's world still depends on it.

In international politics, the notion of order is always suspect, since the framework of social behaviour to be observed is that of a specifically anarchic society, that of states, prone to disorder in the absence of proven domination. It is therefore tempting to define the global nuclear order as an argument of authority imposed by the main beneficiaries of a *de facto* situation that divides the world into nuclear weapon states and non-nuclear weapon states. The Treaty on the Non-Proliferation of Nuclear Weapons (NPT) correctly illustrates that the nuclear order is first and foremost the maintenance of a privilege and domination by those who have manufactured a nuclear explosive device and carried out at least one nuclear test before January 1, 1967,[12] led by the United States and the Soviet Union, which were the main drafters of the treaty in the mid-1960s. France, through its Minister for Foreign Affairs Michel Jobert in 1973,[13] was one of the countries denouncing the US-Soviet strategic "condominium." The NPT sought to freeze strategic nuclear history and maintain the order it instituted with the help of a dedicated agency—the International Atomic Energy Agency (IAEA)—and under the control of the United Nations Security Council. Nuclear order is therefore by no means an edifice of norms designed to limit the risks of the nuclear age. That would be a simple functional definition that ignores power relations. The reality is indeed that of a world order that has historically been criticized from all sides, but which is imposed by states that have at least a good interest in maintaining it.

The historiography of the nuclear order shows that critical thinking on the subject has been going on quietly in the West over the last 80 years. In a 1978 article for *Foreign Policy*, Karl Keiser, a German academic, for example, was "in search of a nuclear world order."[14] He began his argument by noting that a consensus had broken down on the basic rules and objectives of the global nuclear system. Eight years after the entry into force of the NPT did not, however, represent the break-up of an existing order. On the contrary, it marked the consolidation of a multilateral order established by the entry into force of the treaty in 1970.

Fourteen years later, Stanley Hoffmann published an article, translated for the French magazine *Esprit*, entitled "Les illusions de l'ordre mondial" (The illusions of world order), in which he argued that "in history, an order collapses when the hegemony of the dominant nation is challenged or when a hegemony fails to establish itself".[15] Here again, 1992 does not suggest a collapse of the world nuclear order. Seen from the beginning of 2026, it would seem to be the golden age of the American-Russian strategic dialogue, a victory for the pragmatic approach to arms control, and the consolidation of a cautious nuclear order: Nuclear powers put an end to the "arms race," anchored the disarmament process provided for in Article 6 of the NPT, and cooperated to limit the risks of proliferation.[16] Three years later, the fifth Revlew Conference of the parties to the treaty confirmed its extension for an indefinite period.[17] The treaty became, as the saying goes, "the cornerstone" of a consolidated global nuclear order. Was Stanley Hoffmann wrong?

The discovery of Iraq's nuclear programme in the summer of 1991 plunged the IAEA into shock and doubt about the safeguards system put in place by the non-proliferation regime. In May 1997, the Board of Governors approved the Model Additional Protocol to the Comprehensive Safeguards Agreements, which considerably strengthens the Agency's ability to ensure the absence of undeclared nuclear material and activities in non-nuclear weapons states. To date, 144 States and Euratom have brought Additional Protocols into force, while a total of 156 have been approved by the Board of Governors.[18] Ultimately, the Iraqi nuclear affair evokes both the argument of disorder and that of order. It shows, among many other illustrations, that it is indeed the occurrences of nuclear disorder that determine the dynamics of order.

Centrality or fragmentation?

If the global nuclear order makes sense, it is as an empirical reality that is the result of continual invention, the extent of which depends on the quality of the cooperation of the players involved. The nuclear order is therefore neither a legal fact nor a strategic reality. Nor is it a linear historical trend. It is a political objective that constantly fluctuates according to the interests, ambitions, and resources of those who endorse it and attempt to formulate it. The question that its content raises today is that of the meaning given to it by enough of these players to qualify it. Western strategic analysis over the last 10 years has agreed on the renewed centrality of nuclear weapons in international security relations. The start of the Russo-Ukrainian war with Russia's invasion of Crimea in 2014, accompanied by an initial nuclear blackmail; the exacerbation of strategic competition between the United States and China and the foreseeable increase and diversification of China's nuclear arsenal in proportions that are difficult to determine; the assertion of the nuclear status of North Korea during the 2010s; and the modernisation of the arsenals of all the nuclear weapons states have put nuclear weapons back at the centre of the game. Admittedly, the nuclear factor had lost some of its prominence during the first two decades of post-Cold War inter-state relations. However, it should be remembered that the practitioners of nuclear deterrence never considered that it was a parameter of inter-state security that had been

superseded by the disappearance of the Soviet threat after 1991.

The same practitioners know that the function of nuclear weapons in the nuclear weapons states is meant to remain permanent, subject to doctrinal and declaratory adjustments. The following can be cited as examples of this permanence in Western nuclear weapons states:

- anticipating a diversion of deterrence by terrorism of mass destruction, including nuclear terrorism, by taking account of state terrorism or state-sponsored terrorism in doctrines (by calling on the scientific discipline of nuclear forensics, the results of which are still uncertain in the United States);
- the adaptation of the US posture in the 2018 *Nuclear Posture* Review to the scenario of limited use of nuclear weapons to de-escalate a conventional conflict;
- targeting centres of power;
- the complementary nature of deterrence through the threat of retaliation and the promise of denial;
- the perception of a risk of "aggressive sanctuarisation" by nuclear weapons, an idea which was formulated in the 1990s and partially illustrated by the current conflict in Ukraine.

Such centrality would argue in favour of consolidating the existing strategic order. Nuclear proliferation, a major fear of the 1990s, has remained contained to date: India and Pakistan, whose respective nuclear statuses since 1998 have not weakened the global non-proliferation regime but perhaps strengthened it, have been joined by North Korea, whose nuclear status has progressed since the country's withdrawal from the NPT in 2003, and is now stronger. There is no evidence to suggest that the risk of North Korea using nuclear weapons has increased, although there are arguments in favour and arguments against. Nuclear disarmament is seriously eroding, fuelling growing impatience and anxiety among many non-nuclear weapons states and in Western public opinion. But it should be remembered that arms levels are not taking off significantly and remain well below the levels reached at the end of the Cold War. [19]

However, contemporary nuclear realities are showing clear signs of fragmentation. The deterioration of the Iranian nuclear dispute since 2018 has created conditions in which Iran has become a *de facto* nuclear threshold state, possibly on the verge of crossing it. On the night of June 12 to 13 2025, Israel launched a massive counter-proliferation operation against Iran aimed at eliminating the so-called 'existential' threat posed by Tehran's advancing nuclear programme. While an extremely fragile status quo prevailed between the end of June 2025 and the beginning of 2026, the outcome will now be limited to the following alternative: Either Iran will see its programme dismantled and subject to a new, very strict inspection regime that will prevent it from progressing further towards nuclear threshold status, or Iran will remain a sovereign regional power and decide to withdraw from the NPT in order to complete a military nuclear programme and acquire a strike force that will protect the country from future preventive Israeli or American attacks. It should be added that this second option has probably already become very fragile and unrealistic, given the balance of power on the ground, the imbalance in capabilities between the two armies in a symmetrical confrontation, and Iran's isolation in terms of regional and global military support. In the current regional and global strategic circumstances, it is difficult to envisage any realistic scenario in which the Iranian regime could continue to advance its military nuclear programme and bring it to fruition. Operation "Epic Fury" certainly suffers from a lack of concern regarding the future of strategic stability in the Middle East. But it powerfully illustrates that a non-nuclear weapon state that abuses nuclear ambiguity now runs the risk of an extremely brutal military counter-proliferation response. Whether Iran's campaign will have a unifying or disruptive effect on the global nonproliferation regime remains an open question at this stage, since the diplomatic follow-up to the operation will largely determine the outcome. The entry into force of the Treaty on the Prohibition of Nuclear Weapons (TPNW) in January 2021, and its adoption by 73 States to date, is a sign of the deep rift among nations over the status of nuclear weapons in international affairs, even if this treaty is unlikely to have any strategic effect, as none of the nuclear weapons states has adopted it. The centrality of the US-Russian bilateral relationship has lost its value, while the Chinese government refuses to take part in an arms control dynamic that would take a trilateral form. The United States' guarantee of security within the framework of its alliances in Europe and Northeast Asia is being called into question in European countries and, increasingly, in the Republic of Korea[20] and Japan.[21] The main guarantors of the non-proliferation regime no longer share the same strategic objectives. Criticized from all sides but set up by states that had at least a good interest in keeping it going, the world strategic nuclear stage is subject to an instability of which at least one feature is relatively unprecedented: The play that is being performed no longer suits the very people who were the main characters in the past. In various ways, this has been true of the United States since it abandoned the ABM Treaty[22] in 2002; it has undoubtedly been true of the Russian Federation since the end of the process of verified elimination of missiles covered by the INF Treaty[23] in the same year; it has been true of China since a time that is more difficult to determine, but which clearly precedes Xi Jinping's rise to power between 2008 and 2012. Such a situation indeed introduces a chaotic near future. It also points to the need to rethink joint cooperative action outside non-functional multilateral frameworks. If we return to Hoffmann's paradigm, we might ask ourselves whether the contemporary era is one in which one hegemony is being

challenged and another is finding it difficult to establish itself. In other words, the nuclear dyarchy at work during the Cold War could not possibly have cracked without the global nuclear framework cracking in its turn. The world is still searching for a nuclear order that escapes it because the strategic nuclear fact itself cannot be ordered in the long term. That being said, and thoroughly documented, a question remains as to whether this disorder in constant pursuit of order poses a problem and, if so, what the nature of the problem is. The Western strategic community has long established as a benchmark the need to preserve strategic stability at the various relevant bilateral, regional, and global levels. Strategic stability is difficult to define.[24] More than a concept that has been formalized as such, it is a historical expression, a shared practice, or a phenomenon to be assessed. Thus, the question of the meaning that a speaker gives to it at a given moment and in each context is at least as important as the question of how it is measured. Until now, most of the traditional Western literature on strategic stability has rarely been concerned with this socio-political dimension. Despite its polysemy, it can be usefully assumed that strategic stability refers to a situation in which the incentives to change the status quo are less strong than the constraints on change.[25] In its most classic military nuclear sense, strategic stability implies a configuration of forces such that neither side is tempted to proceed without major risk to launch a first strike or a "surprise attack," and therefore: (1) both parties have protected second-strike capabilities, (2) there are no strategic defences of territory against a massive attack, and (3) there is a set of political and legal instruments codifying and controlling competition between the two actors, including by prohibiting certain systems.

On this basis, future initiatives should pursue the following strategic objectives, which are essentially reminders of the fundamental principles of arms control:

- ensuring the survivability of second-strike forces;
- clarifying postures and limiting the risks of misunderstanding associated with the practice of strategic ambiguity;
- strengthening all means, cooperative or unilateral, to ensure transparency on the ground;
- preventing by all means the possibility of surprise attacks;
- strengthening the means to prevent *fait accompli*[26]

The reality, however disappointing it may seem to some, is that the nuclear order exists no less today than it did yesterday. One thing is certain: Nuclear weapons remain one of the key factors of power, and power is shifting and reconfiguring before our very eyes. Consequently, nuclear disorder will probably remain a constant feature of the century that has just begun.

Notes

[1] Whether we date it to the first nuclear explosive test on July 16, 1945 in Alamogordo, New Mexico, or to the first use of nuclear weapons by one state against another on August 6, 1945 in Hiroshima.

[2] In the late 1950s, the term was used to describe the perception that the USSR was developing an intercontinental-range ballistic missile capability earlier, in greater numbers and with a far greater capacity than the United States, which was incorrect.

[3] From 1958 to 1963.

[4] It was on 30 October 1961 that the USSR tested the "Tsar Bomba", the most powerful thermonuclear explosive device ever tested by a Nuclear Weapon State (NWS).

[5] 14 to 28 October 1962.

[6] From 1977 to 1987.

[7] See, for example, Jacques Attali, *Economie de l'apocalypse*, 1995.

[8] The conventional wisdom among Western strategic experts in the 2000s was not to ask whether nuclear terrorism would strike, but when.

[9] Thérèse Delpech, *Nuclear deterrence in the 21st century*, 2013.

[10] In the same way that the post-September 11 world instituted "new threats » to international security.

[11] Bernard Brodie (ed.) *The Absolute Weapon: Atomic Power and World Order*, 1946.

[12] Article IX.3 of the NPT.

[13] See, for example, "M. Michel Jobert dénonce le 'condominium' soviéto-américain", *Le Monde*, 14 November 1973.

[14] Karl Keiser, "A la recherche d'un ordre nucléaire mondial, Réflexions sur les divergences germano-américaines en matière d'énergie nucléaire", *Politique étrangère*, 1978, 43-2, pp. 145-171.

[15] Stanley Hoffmann, "Les illusions de l'ordre mondial", *Esprit*, N°184, August-September 1992, pp. 88-105.

[16] The US *Cooperative Threat Reduction* (CTR) programme was launched in 1991 to limit the risks of nuclear, chemical, and biological proliferation arising from the break-up of the USSR. It initiated a series of major unilateral and multilateral initiatives that continue to shape the fight against the proliferation of weapons of mass destruction around the world.

[17] Article X.2 of the NPT states that "*Twenty-five years after the entry into force of the Treaty, a conference shall be convened to decide whether the Treaty shall continue in force indefinitely, or shall be extended for an additional fixed period or periods. This decision shall be taken by a majority of the Parties to the Treaty."*

[18] As of June 2025.

[19] In the mid-1980s, the USA and the USSR held around 70,000 nuclear weapons. As a reminder, the bilateral New START Treaty, which is still in force until February 2026 even though Russia announced in February 2023 that it would "suspend"

its participation, limits the number of strategic nuclear launchers deployed to 700 and the number of nuclear warheads deployed on these launchers to 1,550.
[20] See, for example, Lee Minji, "Defense minister nominee says he is open to idea of S. Korea's nuclear armament", *Yonhap News*, 2 September 2024.
[21] See, for example, Tim Kelly, Yukiko Toyoda, "Japan PM hopeful Kono calls for US assurances to deter nuclear ambitions", *Reuters*, 9 September 2024.
[22] *Anti-Ballistic Missile* Treaty signed by the United States and Russia in Moscow in May 1972, outdated since the withdrawal of the United States in June 2002.
[23] Intermediate Nuclear Forces Treaty signed between the USA and the USSR in December 1987, entered into force in June 1988, lapsed following the US withdrawal in August 2019.
[24] See Benjamin Hautecouverture, Emmanuelle Maitre, Bruno Tertrais, *L'avenir de la stabilité stratégique*, Recherches & Documents n°05/2021, FRS, 16 February 2021
[25] See, for instance, Elbridge A. Colby and Michael S. Gerson (eds.), *Strategic Stability: Contending Interpretations*, Carlisle, PA: Strategic Studies Institute, 2013. See also Lawrence Rubin and Adam N. Stulberg (eds.), *The End of Strategic Stability, Nuclear Weapons and the Challenge of Regional Rivalries*, Georgetown University Press, Washington DC, 2018: "*To be clear, there are two main and generalizable components at the heart of strategic stability. First, strategic stability refers to a condition in which adversaries understand that altering military force posture in response to vulnerability— whether to avoid being emasculated or to pre-empt one's opponent— would be either futile or foolish. Second, strategic stability reflects the ease with which nuclear-armed adversaries can return to stable relations after a period of escalation. Actors can maintain strategic stability even in a crisis by not responding to a provocative action.*"
[26] Benjamin Hautecouverture, Emmanuelle Maitre, Bruno Tertrais, op. cit.

A historian and political scientist, **Benjamin Hautecouverture** is Senior Research Fellow at the Fondation pour la recherche stratégique (Paris), Senior Fellow at the Canadian Global Affairs Institute (Ottawa), and Technical Director at Expertise France (Paris).

How France's new nuclear doctrine strengthens NATO

By Etienne Marcuz
Source: https://thebulletin.org/2026/06/how-frances-new-nuclear-doctrine-strengthens-nato/

June 09 – In his March 2 speech delivered from the Île Longue naval base, President Emmanuel Macron announced major changes to France's nuclear doctrine. Macron's speech aimed to adapt to new global geostrategic challenges, chief among them being the return of high-intensity conflict on European soil.

President Macron announced that France will increase its nuclear arsenal, marking the end of its near-continuous decline since the end of the Cold War. Most significantly, Macron introduced a new strategic concept of forward deterrence (*dissuasion avancée*), which paves the way for active participation by selected European allies in French nuclear operations.

These changes to France's nuclear doctrine mark a revolution. But not a revolution that questions the transatlantic security. One that reinforces it.

This initiative primarily aims to demonstrate European cohesion in the realm of common security, including the nuclear domain. It comes as its US partner now focuses its attention on other operational theaters, particularly in the Indo-Pacific. But France's forward deterrence is not intended to replace the US extended deterrence. Rather, it seeks to complement it. While some may point to potential frictions between the two approaches— including over the allocation of resources—their methods, scopes, and objectives differ. These differences create space so that the French forward deterrence and the US extended deterrence can coexist within the North Atlantic alliance and even strengthen it.

Political and operational opening

Since their inception in the 1960s, France's strategic forces have trained to conduct nuclear operations in complete autonomy, ensuring that the supreme political decision-maker can, under any circumstances, order a nuclear strike against any aggressor threatening the country's vital interests. This commitment to total autonomy—particularly vis-à-vis the United States—explains France's decision not to participate in NATO's Nuclear Planning Group.

The opening of French nuclear operations to allied participation is not an admission of weakness but rather a powerful signal of European cohesion sent to the continent's strategic competitor (one adversary or a coalition of adversaries). Such arrangements already exist within NATO, with at least 13 European states actively involved in the Alliance's nuclear operations—whether through the deployment of US B-61 nuclear bombs on dual-capable aircraft or through NATO's conventional support to nuclear operations. While this frees up US capabilities for other missions, it in no way diminishes Washington's ability to conduct nuclear operations autonomously and therefore the credibility of its nuclear deterrence.

Above all, this is a political signal. The same will apply to France.

The March 2 speech builds on last July's French-UK Northwood Declaration, which for the first time announced coordination between the two countries in nuclear operations. However, where the Northwood Declaration remained vague by not specifying how this coordination would materialize, the Île Longue speech outlined an incremental plan to gradually integrate eight allied states (since joined by Norway as the ninth member, and Finland has expressed interest in joining) into French nuclear operations. Through this coordination, France seeks to familiarize these countries' forces with the specifics of French deterrence and to train French strategic forces in nuclear operations in a coalition.

In March, French Strategic Air Forces conducted the latest of its quarterly "POKER" nuclear exercise, which simulates a full nuclear strike to demonstrate France's aerial nuclear deterrence capabilities. The exercise involved about 40 figther-bombers and was held two weeks after President Macron's speech on nuclear deterrence. In the image, a Rafale B fighter jet (top), which carries a mock-up of the ASMPA-R supersonic nuclear cruise missile, prepares to refuel from an Airbus A330 MRTT Phénix tanker aircraft. (Photo: French Air and Space Force)

Gradual approach

Of the nine partner states involved in the forward deterrence initiative, four are nations operating NATO's dual-capable aircraft (Belgium, Germany, Greece, and the Netherlands). They will soon be joined by the United Kingdom, which recently announced its intention to acquire F-35As for this mission and already possessed an air-delivered nuclear capability in the past. The air forces of these countries, therefore, already have solid experience in nuclear weapons deployment. The other four nations (Denmark, Norway, Poland, and Sweden) contribute to NATO's conventional support to nuclear operations, which they regularly train for, particularly during the annual Steadfast Noon nuclear exercise.

The **first step** in Macron's gradual approach is logically the participation of these countries' armed forces in certain French nuclear exercises. Their participation may be modeled after the quarterly Poker operation conducted by the French Strategic Air Forces, which simulates a full nuclear strike against a fictional adversary equipped with the most advanced anti-access/area denial capabilities, including very long-range air and missile defenses. While the nine partners are already familiar with NATO nuclear operations, these differ significantly from France's, especially because of differences in delivery systems: NATO's dual-capable aircraft (F-16, Tornado IDS, and F-35A) use gravity bombs requiring release close to the target, whereas France's Rafale B aircraft deploy the supersonic ASMPA-R air-launched cruise missile with a range of several hundred kilometers. This allows for stand-off strikes, which are attacks that can target from a distance by launching long-range missiles well beyond the reach of most adversary air defense systems. The modes of action and required means are therefore markedly different: The French approach resembles a hit-and-run operation, while NATO's requires extensive battlefield shaping before the strike.

The **second step** in cooperation will involve strategic signaling. It aims to demonstrate France's determination to protect its interests, "fully factoring in the interests of our allies" into its calculus—including, if necessary, by leveraging its nuclear forces. In peacetime, this could involve temporary deployments of aircraft or even nuclear-powered ballistic missile submarines to allied military bases. The strategic air forces are already accustomed to operating on allied territory, such as during reassurance missions on NATO's eastern flank or Agile Combat Employment exercises. However, they have so far done so as conventional forces, due to their dual-capable nature, much like NATO's dual-capable aircraft squadrons. The March 2 speech changes this: From now on, any deployment of strategic air forces to one of the nine partner countries will have an implicit or

explicit nuclear dimension. Carrying actual nuclear weapons during reassurance missions is very unlikely because France prohibits flying live nuclear weapons during peacetime. But the ASMPA-R missile mockups used in strategic exercises could visually and symbolically underscore the nuclear nature of the maneuver. Beyond signaling France's support for its allies, these peacetime activities will also allow allied personnel to train in the maintenance of French aircraft and their nuclear armament, enabling them to sustain operations in a crisis.

The third and final step of Macron's plan is the deployment of strategic forces abroad during a crisis, which gives forward deterrence its full meaning. The dispersion of aircraft contributing to the nuclear mission across foreign bases and airfields serves to increase force survivability in the event of a surprise strike by an adversary and to send a major strategic signal, marking a new threshold in conflict escalation. Until now, such dispersion was only envisaged on French territory. It is regularly practiced by the strategic air forces during POKER operations and by conventional forces during major exercises involving fighter jets and air tankers. Opening the territories of nine European partners to French nuclear operations now nearly quadruples the area in which the strategic air forces can operate, significantly complicating adversary planning for a disarming strike. It would also increase the political cost of such an attack, as it would require massive strikes against 10 countries instead of one, thereby diluting the risks of coercion by an adversary.

Complementarity with US extended deterrence

The March 2 speech was prepared in full transparency with the US partner, who was informed of France's intentions ahead of the announcement. NATO was also informed, with Macron emphasizing that "this effort will come as an addition to NATO's nuclear mission," adding that "the forward deterrence we are proposing is a distinct effort which has its own value and is perfectly complementary to NATO's at both strategic and technical level." France's philosophy on nuclear weapons use differs significantly from NATO's. Doctrinally, France rejects the concept of graduated or flexible nuclear response: Crossing the nuclear threshold is only conceivable in the most extreme phase of a conflict, if there is an attack on the country's vital interests, which includes their European dimension. Before massive retaliation, France may conduct only a single nuclear warning in the form of a limited nuclear strike against one or possibly several targets. This warning seeks to signal to the adversary the change in the nature of the conflict and to restore deterrence, while demonstrating France's determination to use nuclear weapons to defend its vital interests. This limited strike, which is strategic in nature, does not intend to gain a military advantage but is political in purpose. It would be accompanied by political and diplomatic messages to the adversary, indicating France's intentions. The Strategic Air Forces are often presented as the preferred instrument—compared to a submarine launch—for this action because the multiple escalation steps can be observed by intelligence means, allowing the adversary to see that it is approaching a point of no return. The ability to recall the raid even if already en route is also a significant advantage to maintain strategic stability, as it allows potential final negotiations before the engagement order is transmitted.

In contrast, NATO's nuclear capabilities—particularly US non-strategic nuclear forces—enable a graduated response to any adversary aggression, conventional or nuclear, at any scale. This could include, for example, a symmetrical response to the use of tactical nuclear weapons on the battlefield by the adversary. As a result, NATO nuclear forces would likely engage before France's in a crisis involving one of the nine countries involved in France's initiative, though a parallel escalation of both postures is possible, if only for strategic signaling purposes toward the adversary. The nature of the targets would also likely differ due to the delivery systems used. Therefore, French nuclear forces—potentially supported by allied aircraft—provide access to a wide range of targets located deep inside adversary territory but without requiring an important operational footprint. Even in the event of parallel engagement of French and NATO nuclear postures, the deployment of certain conventional assets would benefit both, particularly in terms of suppression of enemy air defenses along the front lines. Moreover, a French nuclear strike package is relatively small: around 20 combat aircraft, accompanied by a few air tankers and an airborne early-warning and control aircraft, which is a scalable configuration depending on the mission. Likely, its execution would primarily involve French aircraft, with the participation of allies probably being limited to a few aircraft with specific air-defense suppression or stealth capabilities to enhance the penetration capabilities of the French strike.

The ability to use allied air bases from the northern to southern Eastern Flank of NATO is likely the main operational asset of France's forward deterrence, allowing for diversified axes of penetration for the strike, as well as potentially shorter flight times and access to new sets of targets, like in the Arctic. This complicates adversary defense planning by forcing it to disperse its most advanced air and missile defense systems both along the front lines and deep inside its territory. (Russia recently conducted "surprise" strategic exercises that, according to Russian Deputy Foreign Minister Alexander Grushko, were programmed to address France's new doctrine.) In return, this dilution of enemy defenses would also benefit NATO's nuclear operations.

Need for a new nuclear coordination structure

While the French and US deterrence approaches are complementary and driven by different logics, they will require coordination mechanisms to prevent any dispute over the allocation of resources for their

missions. Coordination will also be needed to maintain coherence and control over escalation against a common adversary. If ambiguity is inherent to nuclear deterrence, it must not create confusion, because it would be detrimental to the credibility of the Alliance's resolve. Given France's jealously guarded autonomy in nuclear decision-making—particularly vis-à-vis the United States—Paris likely will remain outside NATO's Nuclear Planning Group, which requires more coordination effort. Exchanges already exist among the Alliance's three nuclear powers (France, the United Kingdom, and the United States) under the so-called "P3" format. Although little information is available about the topics discussed in these exchanges, they appear to focus primarily on nuclear policy. It could be beneficial to expand them to operational themes, following the model of the French-UK Northwood Declaration, which opens the door to deepened cooperation while preserving each party's autonomy. Such coordination is vital to ensure a united front in a conflict, thereby limiting an adversary's ability to exploit potential divisions among allies. It also provides non-nuclear allies with a clear and reassuring understanding of the shared control over escalation between the three nuclear powers.

If France's nuclear deterrence was already recognized as contributing significantly to the Alliance's overall security, the forward deterrence concept now reinforces its role by offering a complementary approach to US extended deterrence. The diversity of modes of action and objectives between the two models will help avoid competition for resources in a crisis while further complicating potential adversaries' calculations. However, coordination between the two approaches will be necessary, potentially through a new structure, which is independent of NATO's Nuclear Planning Group and also includes the United Kingdom. This would maintain coherence across the Alliance's entire nuclear architecture while preserving the autonomy of each party.

The participation of nine European states already deeply involved in NATO's nuclear operations lends operational and political credibility to the French initiative, while discussions are underway to include additional states, like Finland and the Czech Republic. Paris's security offer could also prove invaluable in the event of a major US commitment in the Indo-Pacific, deterring Russia from launching an opportunistic attack against one or more NATO's Eastern Flank states, thereby strengthening Europe's strategic stability.

Etienne Marcuz is a senior analyst on strategic armaments and an associate fellow at the Fondation pour la Recherche Stratégique (FRS) on issues related to deterrence and missile defense.

Iran sealed uranium cache and placed mines amid fears of US operation to seize material

By Katie Bo Lillis, Davis Winkie, Zachary Cohen, Natasha Bertrand | CNN

Source: https://www.local3news.com/regional-national/exclusive-iran-sealed-uranium-cache-and-placed-mines-amid-fears-of-us-operation-to-seize/article_2049a7f6-c9a1-57c6-b0d7-64d7249e33b3.html

June 13 – In recent weeks, Iran has dramatically escalated efforts to seal off its cache of near bomb-grade uranium, deliberately collapsing tunnels and booby-trapping entrances with explosive mines, according to five sources familiar with US intelligence. Getting to the roughly half-a-ton of highly-enriched uranium is now far more difficult, dangerous and time-consuming than it already was just a month ago, when President Donald Trump was publicly signaling that he might order the US military to seize it, the sources said.

The new fortifications by the Iranians add an additional layer of complexity to the Trump administration's proposed deal with Tehran to remove and destroy its uranium, and the move raises questions about who will take on the dangerous task of digging it out.

Iran's diplomatic delegation to the United Nations did not immediately return a request for comment, and the White House did not immediately reply to questions from CNN.

Trump has repeatedly stated that securing the material is a priority for the US in the ongoing negotiations to end the war and re-open the Strait of Hormuz, which Iran has effectively closed.

And according to a senior administration official who briefed reporters Friday, the two sides are inching closer to a deal that would require Iran to turn its enriched uranium over to the US. It would be destroyed on site and then taken out of the country, according to that official.

But US and Iranian officials have offered conflicting accounts of the tentative deal, and its precise terms remain unclear. The purported text of a draft deal leaked to a semi-official Iranian news agency Friday, triggering an angry outburst from Trump on social media.

Even for the Iranians themselves, several of the sources said, removing the enriched material would now be difficult and dangerous. It would require heavy excavation equipment and de-mining efforts — which are difficult and risky.

"If this reporting is true, it would definitely complicate … retriev[ing] the HEU," said Scott Roecker, who headed the National Nuclear Security Administration's Office of Nuclear Material

Removal from 2017 to 2021. It could also offer an opportunity for Iran to obfuscate its compliance efforts.

If negotiators "require that Iran bring the entire stockpile to a central location for verification and ultimately to remove or downblend the material," that would place the onus on Tehran to access and "provide the full inventory" of enriched uranium, Roecker said.

But, "in this scenario, I would worry that Iran would claim that some portion of the HEU was irretrievable," Roecker said. "We wouldn't have full confidence that Iran couldn't retain access to it at some point in the future."

The international community believes most of the stockpile is in collapsed tunnels at the Isfahan nuclear complex in central Iran, with some additional material held at other sites.

In mid-May, the military was prepared to conduct an operation to seize the nuclear material that was ultimately deemed to be too high-risk, CNN has previously reported.

But in the time since then, Iran has only further fortified the sites where its highly enriched uranium is believed to be buried underground.

Trump has previously acknowledged the dangerous nature of retrieving the uranium by force, and he expressed skepticism in a May appearance on Fox News that the Iranians would ever be capable of accessing and retrieving the buried nuclear material without detection from US intelligence.

"We know exactly what's happening," Trump told Fox host Sean Hannity of the site. "Nobody's even gotten close to it."

But by publicly discussing the uranium as a possible target, two of the sources noted, the president may have provided Iran with the impetus to better defend its own assets.

Now, even if the agreement between Tehran and Washington is signed in the coming week, additional technical negotiations to hammer out the details on the future of Iran's nuclear program are expected.

Removing the uranium from the country would likely require the deployment of a specialized mobile uranium facility organized under the National Nuclear Security Administration at Oak Ridge National Laboratory, Tennessee. CNN previously reported that top US negotiators Jared Kushner and Steve Witkoff visited the laboratory earlier this month.

But even the world's top nuclear removal experts would need significant time to complete their task — Trump told reporters earlier this month that removal would take at least two weeks to complete.

EDITOR'S COMMENT: There is no way to remove HEU from Iran without the consensus and/or the cooperation of the Iranians.

On July 17, 1984, Millions of People Tuned In To Watch A Train Crash Live on Television. It Was No Ordinary Accident

By James Felton

Source: https://www.iflscience.com/on-july-17-1984-millions-of-people-tuned-in-to-watch-a-train-crash-live-on-television-it-was-no-ordinary-accident-83772

A neglected genre of television that we need to reboot immediately is: crashing a dangerous thing into another thing and watching the ensuing explosion.

In 2012, producers in the UK, the US, and Germany did just this, purchasing a Boeing 727 and then smashing it into a dried up lakebed in Mexico to answer the age-old question: Where should you sit during a plane crash if your overall goal is to survive?

In 1984, TV producers answered another classic question on everyone's lips: What would happen if you smashed a train into nuclear flask at around 160 kilometers (100 miles) an hour?

On July 17, 1984, millions of people around the world tuned in to watch just that. Meanwhile, a crowd of around 2,000 invited guests got a more direct view, watching as the British Rail Class 46 locomotive hurtled towards the flask, in order to test the flask's safety.

So what are nuclear "flasks" or "casks" anyway? Should nuclear waste be at nuclear facilities, or buried deep in some horrible nuclear tomb, not to be opened for 1,000 years? Well, kinda. But it does also need transporting every now and then, for example to be processed elsewhere, and when the need arises the waste products are typically transported in these giant flasks, often on the backs of trains.

"They are typically made from 25-cm-thick [about 10 inches] forged steel and weigh around 100 tons. They can hold up to five tons of nuclear material," Pacific Nuclear Transport, Ltd explains. "The casks facilitate the movement of nuclear material by different modes of transport, protect workers from radiation, dissipate heat efficiently and are designed to withstand severe accidents."

And boy, as millions learned in 1984, can they withstand accidents. Put one of these beauties up against a train going 100 miles an hour, and we wouldn't bet on the train.

As expected, after a 12.9-kilometer (8-mile) runup, the train hit the nuclear flask and was obliterated. More reassuringly, the flask remained intact, losing very little pressure during the deliberate crash.

So what did we learn from this? Not a whole lot that couldn't have been found

The project cost £1.75m pounds.

out through more conventional tests, but the point of the crash was as a demonstration to the public that nuclear material could safely be transported across the country.

"These flasks are ultra, ultra safe," Walter Marshall, chairman of the Central Electricity Generating Board, said in a TV interview shortly after the crash. "People shouldn't worry at all about the transport of spent fuel."

"The general public really ought to be satisfied. I mean it's been expensive, but I think they ought to be satisfied and and that makes it worth it," he added.

While it is unclear how many people worried about nuclear waste would be fully converted to the idea after watching a train smash into it expensively, protests around transport have died down in the intervening years.

Meanwhile, a lot of nuclear waste has been transported across the UK.

"In 2014, Public Health England estimated that, each year, approximately 110,000 transport containers were moved on the road, with 1,500 moved on the rail network," Sarah Bryson, a Transportation Specialist with Nuclear Waste Services explained to South Copeland GDF Community Partnership, a nuclear disposal facility.

"They have all occurred without a single major safety incident, since nuclear transport began in the 1960s. This unblemished safety record is due to robust management and packages undergo rigorous testing to ensure they remain safe no matter what happens to them."

More conventional testing, she added, involved dropping the casks from various heights, and subjecting them to temperatures of up to 800°C (1,472 degrees Fahrenheit).

ICI International CBRNE INSTITUTE

CBRNE-Terrorism Newsletter

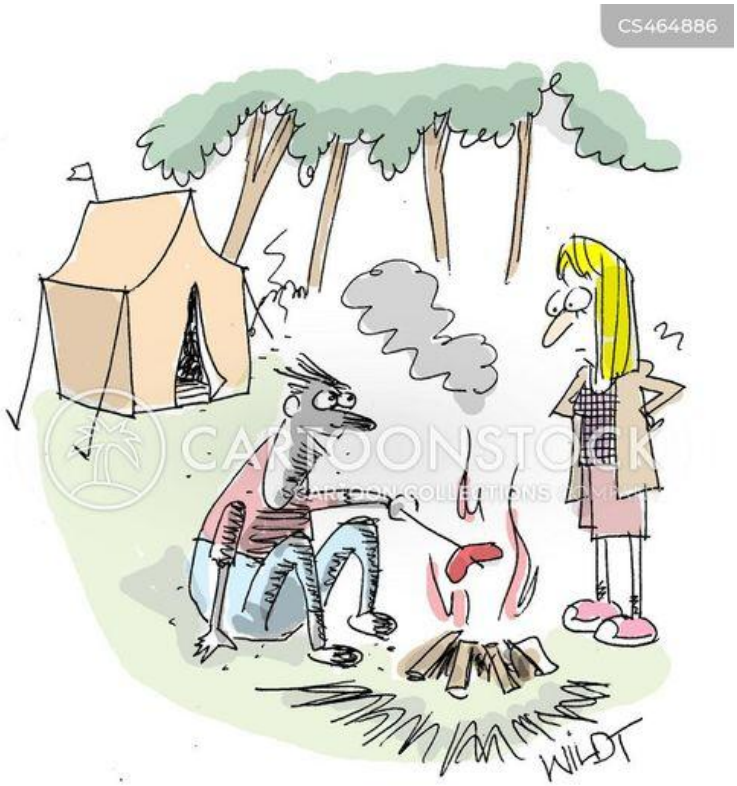
C²BRNE DIARY

Weapons, military equipment and

EXPLOSIVE NEWS

How Iran's $100k missile destroyed $1.5bn of US equipment

America lost 24 Reaper drones in Iran - This is why | War On Tape



"Smartphones are so helpful. I asked it how to start a fire and the battery exploded."

ICI International CBRNE INSTITUTE

CBRNE-Terrorism Newsletter

CYBER NEWS

C2BRNE DIARY

Iranian hackers responsible for Los Angeles transit system breach, Israeli researchers say

Source: https://www.nbcnews.com/tech/security/iranian-hackers-responsible-los-angeles-transit-system-breach-israeli-rcna346881

May 26 – Iranian hackers were responsible for a disruptive computer breach in March that forced Los Angeles' transit system to shut down parts of its network, Israeli researchers say.

The saboteurs stole at least 700 gigabytes of emails, backups and other files from the Los Angeles County Metropolitan Transportation Authority (LACMTA), according to Gambit Security, a Tel Aviv-based cybersecurity firm that said it discovered the misappropriated data after it was inadvertently exposed online.

In a report published Tuesday, the company said a digital trail of evidence tied the server where the data was discovered to a previously known hacking operation that Israeli officials and researchers attributed to Tehran.

Iran's mission to the United Nations did not return messages seeking comment. Israel's National Cyber Directorate did not return messages.

The Los Angeles transit authority didn't respond to questions about the findings. In a statement shared last month, its officials said they were working with law enforcement and cyber specialists as they brought their systems back online. "Attribution is part of the investigation and we will not speculate," the statement said.

Digital security specialists have suspected an Iranian hand in the operation against the LACMTA ever since responsibility was claimed by an obscure pro-Iran outfit calling itself Ababil of Minab. The group's name refers to the bombing of a girls' school in the Iranian city of Minab that officials there say killed more than 175 children and teachers, and its rhetoric and modus operandi are characteristic of self-styled vigilante hacker groups that U.S. and Israeli researchers allege are cutouts for Iranian spies.

Eyal Sela, Gambit's director of threat intelligence, said a connection between Ababil and the Iranian state "has been a working assumption."

"What our research adds is the forensic evidence to support it," he said.

Gambit, a security startup founded in part by veterans of Unit 8200, Israel's equivalent of the U.S. National Security Agency, said it had alerted relevant authorities to its findings.

Ababil did not return messages left via a form on its website. The FBI said it was aware of the LACMTA incident and was "coordinating with partners in response." The FBI declined further comment. The U.S. civilian cyber defense body, the Cybersecurity and Infrastructure Security Agency, did not return messages seeking comment.

The intrusion at LACMTA was detected around March 16, its officials said in their statement. About two weeks later, Ababil materialized online and claimed to have wiped an enormous amount of data in a destructive cyberattack, publishing a video that purported to show them rampaging through the transit system's network.

Although Los Angeles transit officials said the breach did not interrupt circulation of trains or buses, local media said it disabled at least some arrival screens and

prevented customers from putting money on their transit cards.

Ababil also has claimed credit for hacks affecting South Florida's Tri-Rail commuter transit system, vehicle tracking company Vyncs, and Saudi infrastructure firm Unimac.

In a statement, Tri-Rail confirmed it had been hacked "about a month ago," but said that none of the affected data was critical. Vyncs owner Agnik said it had detected its breach on April 2 but declined to comment on the nature of the data stolen by the hackers. Both Tri-Rail and Agnik said the FBI was involved, with Agnik saying in an email that the bureau "has a pretty good understanding of who these criminals are." Unimac did not return messages seeking comment.

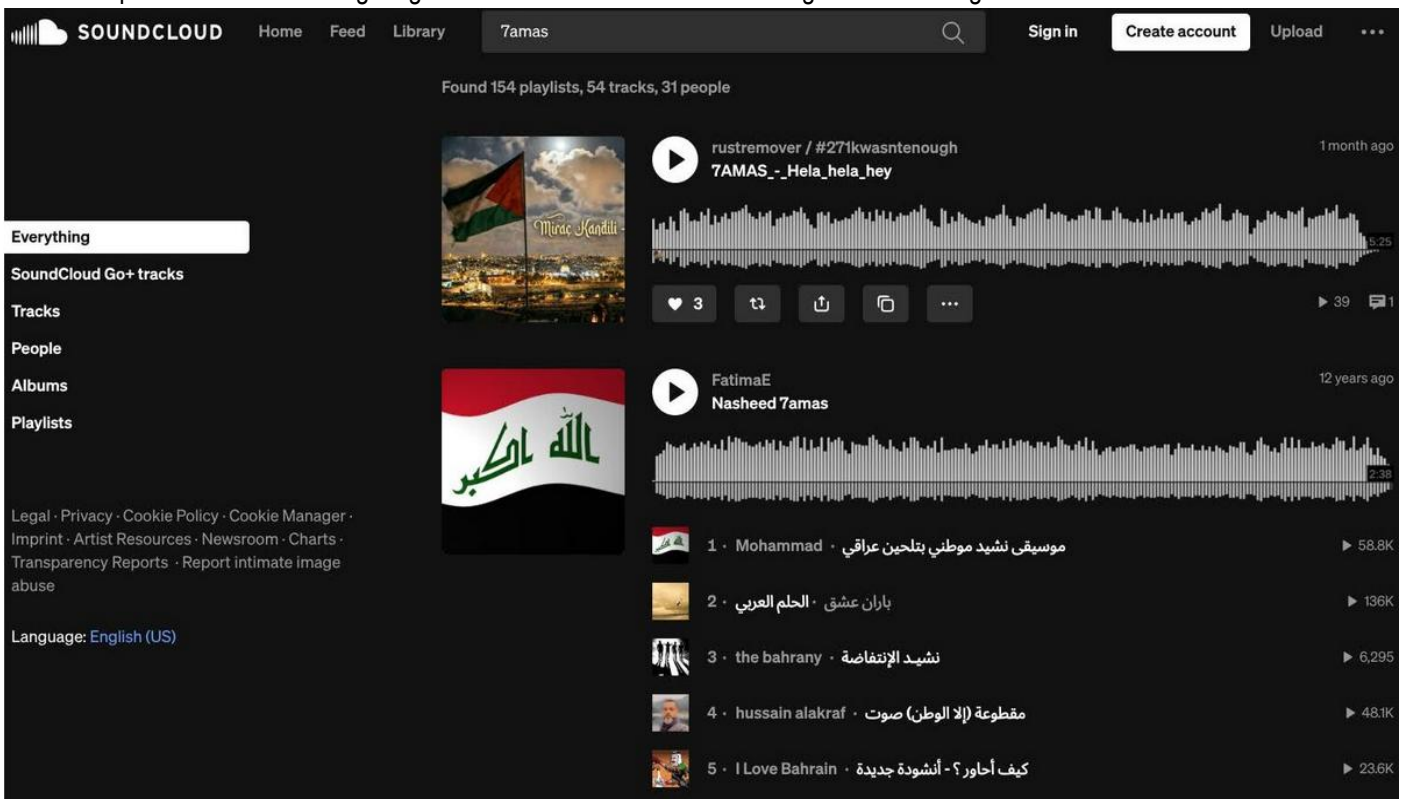
The group behind Ababil has hacked other organizations whose identity it has not publicized, Gambit Security said, citing its analysis of other data left online by the spies. Sela said they included a media organization and educational institution in Israel and an insurance brokerage in Turkey, but he declined to identify them further.

Iranian hackers allegedly have carried out a drumbeat of digital operations since the U.S. and Israel launched a war against Iran in late February, including a damaging attack on the medical device company Stryker and the leak of personal emails belonging to FBI Director Kash Patel. Iranian hackers also are suspected of having remotely tampered with fuel gauges at gas stations, CNN reported earlier this month.

The Terrorist Threat Washington Isn't Hearing

By Sophie McDowall

Source: https://nationalinterest.org/blog/techland/the-terrorist-threat-washington-isnt-hearing



Screenshot of Soundcloud search results for 7amas. Terrorist groups, including Hamas, are exploiting music and audio platforms to spread extremist propaganda, exposing dangerous gaps in online content moderation and counterterrorism coordination. (Soundcloud)

May 26 – Most people who use online streaming platforms would not think of terrorism if they saw "7amas" in the title of a song. But replacing the 'H' in 'Hamas' with a '7' is a common means of evading content filters. My colleagues and I found more than 550 unique pro-Hamas songs in a thorough search of SoundCloud, the popular audio-sharing platform. SoundCloud and other platforms have been effective in taking down Islamic State and al-Qaeda propaganda, but the persistence of pro-Hamas content shows that they need to increase the sophistication of their efforts to identify terrorist propaganda.

Federal Counterterrorism Strategy Must Address Online Extremist Content

Federal law enforcement may be able to help. The Trump administration's new counterterrorism strategy pledges that Washington will identify and

neutralize the media platforms of terrorist groups before terrorist plots come to fruition. Federal agencies should work with platforms to identify and remove content glorifying the killing of innocent civilians.

Online Radicalization Continues to Fuel Terrorist Violence

Social media and other online content have a demonstrated role in radicalization. Since 2006, 118 people have been killed and more than 760 injured in the United States at the hands of Islamist-inspired terrorists. In many of these attacks, social media and online content played a role. The three highest-casualty terrorist attacks since 9/11—the 2013 Boston Marathon bombing, 2016 Pulse Nightclub shooting, and 2025 New Orleans New Year's attack—were all carried out by individuals who had been radicalized through the internet. The perpetrator of the New Orleans attack had shared songs and posted audio content to SoundCloud that promoted his extremist jihadi views.

Extremists have long recognized the benefits of audio content and used the medium to spread their messages. Cassette tapes shared by Ayatollah Ruhollah Khomeini helped fuel the 1979 Iranian Revolution. Since the 1990s, Hezbollah has produced songs and music videos to disseminate its rhetoric and gain sympathizers.

Terrorist Groups Exploit Weaknesses in Audio Content Moderation

Today, thanks to the ubiquity of online media platforms, terrorist groups do not need to create person-to-person networks to share recordings of songs and speeches.

Online platforms have improved their abilities to identify and remove content that violates their guidelines, largely using automated systems. Yet while automated systems can be effective for text and some image-based content, audio content can be more challenging for AI systems to parse.

Our research focused on SoundCloud, but also found extremist content on Spotify, Apple Music, and other music streaming platforms. The problematic tracks included titles, cover art, and audio promoting extremist jihadi ideologies, commemorating and praising terrorist activity, and honoring deceased terrorist leaders. Songs also included calls for "Death to America" and "Death to Israel."

This content violates the guidelines set by SoundCloud and other platforms, but terrorist sympathizers used a variety of techniques to evade content moderation. They paired innocuous titles with cover art that depicts violent terrorism or extremist leadership; for example, a track titled "مكان كل في الظلم." [Injustice is everywhere]" includes imagery of Hamas terrorists in the cover art. Other users slightly adjusted titles of songs that had previously been removed, used codewords, and replaced certain letters with numbers—a common technique when rendering Arabic in the English alphabet, since English has no letters to represent certain Arabic sounds. When designing content moderation systems, audio-sharing platforms need input from experts likely to be familiar with this technique.

Government and Tech Partnerships Are Critical to Counter Online Extremism

As evasion tactics evolve, so must the strategies of online platforms. Expecting platforms to immediately identify every terrorist dog whistle may be unreasonable. Instead, these companies need partnerships with governments and research institutions that are experts in extremist activities. In the federal government, the US National Counterterrorism Center (NCTC) at the Office of the Director of National Intelligence is the primary center for counterterrorism efforts and information sharing.

This information sharing has historically focused on collaboration with domestic partners, including the interagency and federal government partners. In today's online environment, NCTC also needs to engage with international groups such as the Global Inter Forum to Counter Terrorism, which brings together international experts, media platforms, and governments to share information and combat terrorist activity online.

Preventing the next mass casualty terror attack means taking the threat posed by Islamist social media and audio content seriously. It means collaboration among governments and platforms to ensure they can identify and remove terrorist content. And it requires recognition of and action against the efforts of terrorist groups to influence and radicalize Americans.

Sophie McDowall is a research associate for the Center on Cyber and Technology Innovation (CCTI) at the Foundation for Defense of Democracies (FDD). Her research focuses on cybersecurity of critical infrastructure,

federal cybersecurity policy, emerging technologies, and influence operations. Sophie holds an M.S. in data science and analytics and a B.S. in foreign service from Georgetown University.

Iran-linked group claims hacking FBI drones, World Cup threat

Source: https://thesun.my/news/world-news/iran-linked-group-claims-hacking-fbi-drones-world-cup-threat/

June 12 – An Iran-linked hacker group claims to have breached FBI drones and has threatened to target the World Cup that kicked off this week, a monitoring group said on Friday. The SITE Intelligence Group, an organization that monitors jihadist groups, published a statement from **Handala** saying they had had access "for months" to "every image and every suspect" captured by first-person view (FPV) drones used by the FBI. The hackers said the drones featured facial recognition and license plate screening deployed for counterterrorism. "Better tighten your World Cup security, we don't like some of those teams at all. Don't forget: FPVs are everywhere; you never know when one might end up right in your team's bus," Handala said in the statement quoted by SITE. The FBI is deploying drones around World Cup stadiums to protect against unauthorized aircraft. Drone flights will be banned over US stadiums hosting matches, as well as over fan events related to the tournament co-hosted by the United States, Mexico and Canada that kicked off its largest edition on Thursday. Training for local and state police on countering unauthorized drone activity at the World Cup was part of a $500 million federal grant to combat this growing threat to sporting events.

The Justice Department has warned of the potential for cyberattacks by Iranian actors following the US-Israeli strikes on Tehran in February that triggered the Middle East War.

Handala published photos and footage that it said were taken from the hacked drones, but SITE disputed that claim.

One video of the supposed hack was in fact produced by a software platform in December 2024 to promote a US police department's use of its technology for surveying tornado damage, SITE said.

Handala claimed in March to have hacked the email account of FBI Director Kash Patel and published personal photos and other material online.

The US State Department has offered a reward of up to $10 million for information leading to the identification of members of the group.

ICI International CBRNE INSTITUTE

CBRNE-Terrorism Newsletter

C²BRNE DIARY

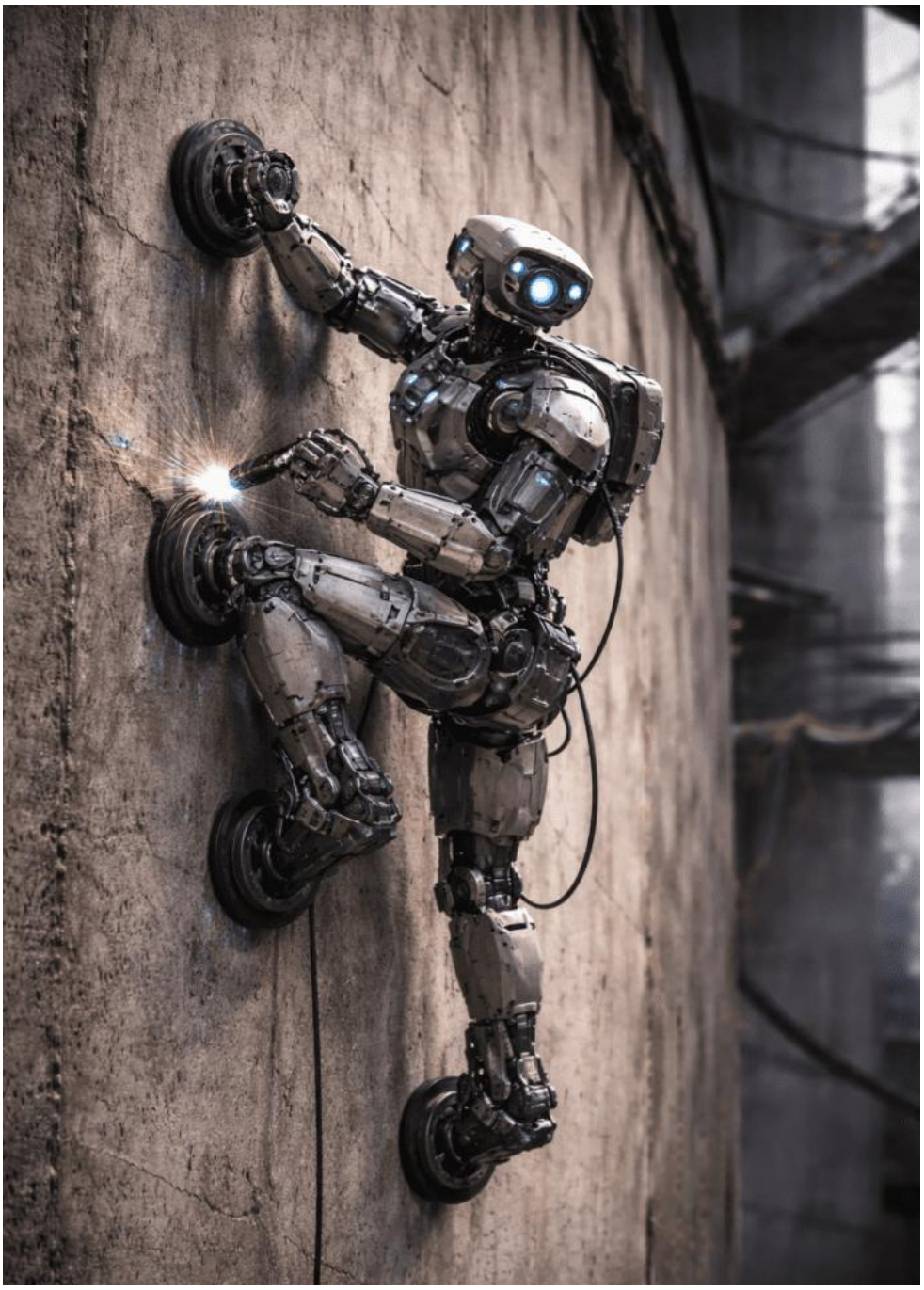
& Robotic

DRONE NEWS

A Robot That Climbs Walls to Do Dangerous Jobs

Source: https://i-hls.com/archives/135966

May 20 – Industrial environments such as chemical facilities, refineries, and large storage sites often involve hazardous conditions that put human workers at risk. Tasks like welding, inspection, and maintenance may require operating at height, on unstable surfaces, or in areas exposed to heat, corrosion, or toxic materials. Traditional automation has helped in controlled settings, but many of these environments still require human intervention due to their complexity.



A new type of humanoid robot, called Nengzai No. 1, is designed to address this gap by combining mobility, adaptability, and advanced control into a single platform. Unlike fixed robotic systems, the machine can physically navigate challenging environments, including vertical metal surfaces. It uses a magnetic chassis to adhere to structures such as storage tanks, allowing it to move and operate where conventional robots cannot.

According to Interesting Engineering, the system is built to handle a wide range of industrial tasks. Equipped with dual robotic arms and multiple degrees of freedom, it can perform operations such as welding, grinding, inspection, and surface treatment. This flexibility allows it to replace several specialized machines while maintaining precision in complex workflows.

At the core of the platform is an AI-driven control system trained on extensive operational data. This enables the robot to interpret its surroundings, adjust to real-world conditions, and improve performance over time. Instead of following fixed routines, it can adapt to variations in structure, surface condition, and task requirements.

Another key feature is continuous operation. The robot uses a tethered power system, eliminating the need for battery swaps or charging cycles. This allows it to function around the clock, which is particularly valuable in industrial settings where downtime can be costly. From a defense and homeland security perspective, such capabilities could be applied beyond industrial use. Robots that can operate in hazardous environments may support missions such as infrastructure inspection, disaster response, or operations in contaminated or high-risk zones where human access is limited. As robotics continues to evolve, systems that combine physical mobility with adaptive intelligence are expanding the range of tasks that can be automated. In environments where safety and precision are critical, such platforms offer a practical way to reduce risk while maintaining operational efficiency.

This Robot Sees Threats on the Ground — and Now in the Sky

Source: https://i-hls.com/archives/136257

May 21 – Unmanned ground vehicles are increasingly used to reduce risk to personnel in complex environments, but they face a persistent weakness: exposure to small, hard-to-detect drones. These aerial threats can identify, track, and target ground assets with relative ease, creating a gap in protection even for advanced robotic platforms operating far from human operators.

A new integration aims to close that gap by combining ground autonomy with onboard counter-drone capabilities. The ULTRA unmanned ground vehicle, developed by Overland AI, is now equipped with DroneShield's DroneSentry-X Mk2 system. The addition enables the platform not only to navigate difficult terrain independently, but also to monitor and respond to aerial threats in real time.
According to NextGenDefense, at the core of the upgrade is a software-defined counter-UAS system designed to detect and track drones using radio-frequency analysis. The system's AI-driven engine can identify both known and unfamiliar drone signatures, allowing earlier warning and response before a threat gets within effective range. Unlike fixed installations, the capability is embedded directly into the vehicle, supporting mobile operations and extending situational awareness beyond the ground domain. The platform itself is built for demanding environments. It combines autonomous navigation software with a rugged chassis and a multi-layered sensor suite that includes stereo vision, LiDAR, thermal imaging, and inertial measurement systems. It is also designed to function in GPS-denied conditions, a critical requirement in contested environments. With a payload capacity of up to 454 kgs, the vehicle can be configured for a range of missions, including logistics support, sensor deployment, casualty evacuation, and handling of hazardous materials. From a defense perspective, the integration reflects a broader shift toward multi-domain awareness at the tactical level. As small drones become more accessible and widely used in conflict zones, ground systems must be able to detect and react to airborne threats without relying on external protection layers.
Embedding counter-drone capabilities directly into unmanned platforms reduces dependency on additional assets and allows forces to operate with a smaller footprint while maintaining operational continuity.

A New Way to Stop Drones: Just Listen

Source: https://i-hls.com/archives/136276

May 22 – As small drones become more accessible and harder to detect, traditional air defense methods are facing growing limitations. Many systems rely on radar or visual tracking, which can struggle with low-flying or slow-moving drones, especially in cluttered environments. This creates a need for additional detection layers that

can identify threats early, even when visibility is limited or signals are weak. Recent field trials in Poland are exploring an alternative approach based on sound. A series of acoustic detection platforms have been tested to identify and track drones in real time by analyzing their unique audio signatures. These systems use arrays of sensors to capture sound patterns generated by drone motors and propellers, then process the data to determine direction, distance, and movement. According to NextGenDefense, by focusing on acoustic cues, they can operate in conditions where other detection methods may be less effective.

Artificial intelligence plays a central role in making this approach viable. Machine learning models are used to filter background noise, classify different types of drones, and improve tracking accuracy. This allows the system to distinguish between potential threats and everyday environmental sounds, while also enabling faster decision-making. Alongside the acoustic platforms, a separate set of electronic warfare tools has been evaluated for its ability to detect and interfere with drone communications, adding an active response layer to the overall system. Beyond detection, the systems are designed with integration in mind. Developers have indicated that both the acoustic sensors and electronic warfare components can be connected directly to existing command and control networks, allowing operators to combine multiple data sources into a unified operational picture. This supports quicker coordination between detection and response elements. From a defense standpoint, combining passive acoustic sensing with electronic warfare capabilities offers a more resilient approach to countering drones. Passive systems can operate without emitting signals, reducing the risk of detection, while electronic warfare tools provide options to disrupt or neutralize incoming threats. As testing continues, the results will inform decisions on broader deployment and future development. The ongoing evaluations highlight a shift toward multi-layered drone defense, where different sensing technologies work together to address an evolving threat landscape.

Future factories!

Who Has Authority to Deal with Drones? In Most of Europe, That's Unclear

By James Reeves

Source: https://www.homelandsecuritynewswire.com/dr20260527-who-has-authority-to-deal-with-drones-in-most-of-europe-that-s-unclear

May 27 – Europe's growing drone problem is a governance problem. The hardware exists. What is missing, across most of the continent, is the legal authority to deploy it, the jurisdictional clarity to coordinate it, and the political will to mandate either. Europe's growing drone problem is a governance problem. The hardware exists. What is missing, across most of the continent, is the legal authority to deploy it, the jurisdictional clarity to coordinate it, and the

political will to mandate either. Until that changes, Europe's critical infrastructure will remain exposed to attacks that are cheap to mount and catastrophically expensive to absorb.

This conclusion comes from primary research: a strategic intelligence briefing on governance of uncrewed aerial systems (UASs, or drones). This was conducted by Challenger Research in partnership with policy consultancy TWA, drawing on 23 stakeholder interviews across politicians, defense experts, strategic advisors and critical infrastructure professionals in Britain and Europe, alongside regulatory analysis of Britain, Poland, Germany and Italy. The findings have been consistent across every jurisdiction examined. Legal authority and jurisdictional fragmentation are fundamental constraints on effective counter-drone response. The available counter-UAS hardware cannot be lawfully deployed fast enough to matter in the few minutes between identification and response. The scale of the problem is accelerating. Drone-related disruptions at European airports more than tripled between January 2024 and November 2025. Our research has found that the interval between detection and disruption at an airport is less than five minutes, yet few airports hold the legal authority to act on what they have detected. Runway closures, as seen in Copenhagen and Oslo last year, are enormously costly in flight cancellations. Across ports, energy networks and data infrastructure, critical national infrastructure operators can detect hostile UAS activity but are legally barred from taking active countermeasures. As one industry stakeholder put it plainly, 'Rules of engagement are the primary issue. There is no ownership of responsibility.' The drone's strategic value has little to do with its physical payload. It is a tool for sowing confusion, eroding public confidence and forcing expensive defensive reactions from a cheaper offensive position. Russia only needs success in one EU country to land a blow felt across the entire bloc. Our research has uncovered structural fragmentation. In Britain, the Civil Aviation Authority regulates civil UASs, police hold operational authority under the Air Traffic Management and Unmanned Aircraft Act 2021, and the Ministry of Defense's jurisdiction is largely confined to its own estate. Electronic countermeasures require senior-level authorization as well as approval from telecommunications regulators. Germany opened a Joint Drone Defense Centre in December 2025, but command remains civilian and proposed reforms to allow armed-forces involvement remain politically contested. Italy faces additional constraints under international law in offshore environments. Poland alone has moved decisively, adopting an effects-based jurisdiction model under Operation SAN that empowers police, border guard and armed forces simultaneously. Ambiguity is treated as a threat, not as a reason to pause.

The gap between Poland and its partners captures the broader European problem. A defense director interviewed for this research stated that Britain has had essentially no adequate air defense for 20 years. A Labour Party special adviser told us that 'outside of a small group engaged in defense policy, no one is thinking about this.' Our data, shown below, confirms this. Politicians clustered at the low end of both awareness and concern, while defense industry professionals registered the highest scores on both axes. Those with the greatest operational knowledge of the threat are the least able to move policy. Those with the power to act remain largely unaware of the urgency.

Several respondents raised the prospect of a mass-casualty attack in which the operator is more than 1,000 km from the strike zone. A UAS launched by a state actor or terrorist group could bring down an aircraft on approach, strike a high-visibility target or detonate inside critical infrastructure, with the aggressor already beyond reach before the smoke clears. That combination of lethality, deniability and distance is strategically novel. The political consequences would be enormous and the institutional overcorrection predictable: precisely the kind of sweeping, poorly targeted reform written in crisis rather than forethought, echoing what followed the 9/11 terrorist attacks.

The case for a binding European framework is structural. Drones can transit multiple countries before reaching their targets. Fifteen drones observed above a Belgian military site in 2025 flew directly into Germany, tracked by police across the border. Europe's collective exposure is determined by its least prepared member. A framework that harmonizes neutralization thresholds, assigns unambiguous jurisdictional authority, and brings civil aviation and critical national infrastructure into a common legal architecture would close the gap that adversaries are already exploiting.

Europe has the instruments and the data. The question is whether it builds its governance framework in advance of a crisis, or because of one.

James Reeves is the managing director of Challenger Research, an independent research and advisory firm.

AI and Lasers Team Up Against Swarms of Drones

Source: https://i-hls.com/archives/13645

June 01 – The growing use of inexpensive drones has created a major challenge for air defense systems. Small reconnaissance UAVs and FPV strike drones can be deployed in large numbers at relatively low cost, while

intercepting them often requires missiles worth hundreds of thousands—or even millions—of dollars. This imbalance has pushed militaries to search for more affordable and sustainable counter-drone solutions.

One approach now nearing operational deployment is a trailer-mounted laser weapon designed specifically to engage small aerial targets. The system, called Tryzub, is intended to provide a rapid-response layer for short-range air defense, targeting drones using directed energy instead of traditional kinetic interceptors.

According to Interesting Engineering, the platform combines optical targeting, radar integration, and automated tracking capabilities. Artificial intelligence is used to assist with target acquisition and engagement, allowing the system to identify and track incoming drones more quickly than earlier manually operated versions. The laser is designed to damage key drone components such as optics, electronics, and structural surfaces, disabling targets without the need for explosive warheads.

The system reportedly has different engagement ranges depending on target type. Smaller FPV drones can be targeted at shorter distances, while larger UAVs and Shahed-style drones are being tested at ranges of up to around 5,000 meters. Because the platform is trailer-mounted, it can be repositioned relatively quickly to protect different operational areas. From a defense perspective, directed-energy systems are increasingly viewed as a potential complement to conventional air defense networks rather than a full replacement. Laser weapons offer a low-cost-per-shot alternative against mass drone attacks, particularly when defending infrastructure such as logistics hubs, energy facilities, or population centers. They may also reduce the strain on more expensive missile-based intercept systems that are better suited for larger threats. At the same time, practical limitations remain. Environmental conditions such as rain, fog, smoke, and

dust can reduce laser effectiveness, and many technical details about the system, including power output and cooling capacity, have not been publicly disclosed. Even so, the continued development of AI-assisted laser defenses reflects a broader shift toward lower-cost counter-drone technologies designed for high-volume aerial threats.

The Next Step in Drone Defense: 30 Targets Per Minute

Source: https://i-hls.com/archives/136971

June 04 – The growing use of drones on the battlefield is creating a new challenge for air defense systems. Small, low-cost drones can operate individually or in swarms, carry payloads, collect intelligence, and strike targets at short ranges. Against such threats, conventional interception solutions can be either too expensive or limited in the number of targets they can engage within a short period of time.

To address this challenge, Esh-Tech has developed **DroneLight**, a laser-based drone interception system designed to provide a tactical layer of protection for deployed forces and sensitive facilities. Unlike large laser systems intended to defend wide areas, the system focuses on short-range threats and operates in a manner that resembles a machine gun. Rather than relying on a continuous, extremely high-power laser beam, the system performs a rapid sequence of engagements and, according to the company, can handle up to 30 targets per minute.

The primary advantage of this approach is its lower energy requirement compared to strategic laser defense systems. Systems in the latter category are based on laser beams with outputs of around 100 kilowatts and are designed to intercept threats at ranges of up to approximately 10 kilometers. In contrast, the system is intended to operate closer to frontline forces, providing protection against immediate threats while requiring a significantly smaller energy infrastructure.

A major focus of the development effort was creating a relatively compact system that can be rapidly deployed and integrated into existing defense architectures. This allows operators to position the system where protection is needed most, without the logistical footprint associated with larger directed-energy platforms.

From an operational perspective, the ability to engage a large number of targets in a short period of time has become one of the key requirements in countering modern drone threats. As drone swarms become more affordable and widespread, the demand for interception systems that combine low cost per engagement with a high rate of fire continues to grow.

From a broader defense standpoint, the system reflects a growing shift away from expensive kinetic interceptors toward directed-energy solutions. If systems of this type prove effective in operational environments, they could become an important complementary layer within air defense networks, particularly for maneuvering forces that require continuous protection against small, fast-moving aerial threats.

Historic drone rescue of Apache crew points to future of recovery missions

Source: https://newatlas.com/military/historic-drone-rescue-apache-helicopter-crew/

June 06 – In a historic first, two US Army crew members from an AH-64 Apache helicopter forced down near the coast of Oman were rescued within two hours by a US Navy Saronic Corsair drone boat operated by the 5th Fleet's Task Force 59.

Details of the incident remain sketchy, including whether the Apache ditched due to a mechanical failure or hostile action. What is known is that, at 11:33 GMT on June 8, 2026, the attack helicopter encountered trouble while on a routine patrol near the Strait of Hormuz. A recovery effort was launched by US Naval Forces Command and the 82nd Airborne Division, with support from US Air Force and Navy units.

What made the operation unusual was the involvement of Task Force 59, a dedicated artificial intelligence and unmanned systems integration unit that operates a flotilla of drone boats, including the Corsair. Equipped with a 360-degree passive sensing payload, the vessel was able to locate the two crew members, who were able to climb aboard the 24-ft

(7.3-m) drone boat and cling onto its superstructure as the surface craft carried them to a safe area for helicopter extraction. Both were reported to be in stable condition after their ordeal.

This first-of-its-kind rescue is significant for more than its historic value. It also highlights a potentially important application of drone technology for both military and disaster-relief missions.

Casualty evacuation has always posed a major challenge for armed forces, particularly Western militaries dedicated to the principle of leaving no one behind. It's a laudable concept, but one that has caused problems in the past. At the very least, it means using up soldiers to move wounded comrades to the rear, meaning that at least two fit people are needed to handle one casualty. Even with dedicated stretcher bearers, that's a lot of personnel.

It can also cause serious operational problems. During the Vietnam War, the Viet Cong learned that they could bring an American assault crashing to a halt by wounding a US soldier as quickly as possible, stopping the attack while the casualty was dealt with. By the time of the Falklands War, the British learned from this and adopted the policy of stabilizing a casualty and continuing the assault, with recovery taking place later when conditions allowed.

Autonomous drones – including land, sea, and air, as rescue and evacuation units – could change things dramatically. Many more soldiers could be freed up for combat and other operations while often costly rescue or recovery missions could be handled by autonomous vehicles.

Casualties could be moved to the rear quickly, helping preserve the critical "Golden Hour" during which prompt medical treatment can mean the difference between life and death. In addition, the drones could be sent into situations that would be too dangerous for a human team, increasing the chances of successful rescue.

For the military, making use of such technology has obvious benefits, but it could also help in disaster situations. Recovery drones could rush into areas affected by earthquakes, hurricanes, wildfires, and tsunamis, where conditions may be too hazardous for human teams. They could even be pre-positioned in hazard areas in anticipation of disasters, ready to go at a moment's notice to collect the injured and drop off supplies.

And, as was demonstrated in the Apache incident, the drones don't even need to be dedicated ones. Any platform capable of carrying a human-sized payload could be pressed into service as needed.

They don't even need to be the vehicle-like drones we're used to. Quadruped robots are being developed to act like mules for the infantry and could maybe be equipped with little kegs of brandy to turn them into robo-St. Bernards. There has also been considerable discussion of humanoid military robots, which could one day be reprogrammed for casualty duty.

So don't be surprised if a cliché cry of "Medic" one day ends up being answered by something that looks like C-3PO in camo, complete with a red cross on its chest.

At World Cup stadiums, there will be zero tolerance for drones because of the threat they pose

Source: https://www.washingtonpost.com/business/2026/06/10/world-cup-drones-threat-fbi-war-ukraine/a505ef50-6481-11f1-bdd4-805ebb99a693_story.html

June 10 – Fans who hear the whirring sound of a drone over a stadium might see it as a nuisance but law enforcement has long viewed those aircraft as a potential weapon of mass destruction. With the World Cup about to kick off, security is heightened and there's a zero-tolerance policy for drones over or near stadiums during the

78 matches in 11 U.S. cities. Authorities have ratcheted up counter-drone measures used at the Super Bowl and other major events, while Congress has given law enforcement broader authority to electronically disable drones or shoot them down.
"The war in Ukraine has become a real-world testing ground for drone technology, and if there is one threat that keeps me up at night, it is from drones," New York Police Department Commissioner Jessica Tisch said.

Taking the threat seriously

Congress just gave state and local law enforcement the green light to take control of a threatening drone or even shoot them down if needed in December, though the first option will be to disable them electronically and land them safely. Key federal agencies already had that power. The Federal Aviation Administration will restrict the airspace around and over crowded stadiums for World Cup games and fan events — much like it has done for years around NFL and Major League Baseball games. Violators can face fines up to $100,000, see their drones confiscated and even face criminal charges for flying within three miles of one of the games. But those penalties likely wouldn't deter a determined terrorist. The FBI has spent the last seven years building up its capability to deal with drones by investing in the technology needed to quickly identify drones and take control of them, and authorities have been using that already at major events. The bureau also provided counter drone technology training this year to law enforcement in all World Cup host cities that taught them how to recognize dangerous drones and respond. The military has also developed counter drone lasers like the ones used along the Mexican border earlier this year and other systems to shoot down drones, but the FBI is not planning to do that during the World Cup because of the dangers involved with the wreckage of a drone falling over a major city.

"If the drone is intercepted and it no longer flies, it's going to fall. And as we say, no matter what you do, you can't change the law of gravity," said national security expert Hal Kempfer, who estimates that he has trained more than 30,000 law enforcement officers on counter-terrorism techniques.

'Everybody's a little behind'

The government has invested heavily in systems that should allow officers to take control of suspicious drones and land them safely or jam their signals, including handing out $250 million to help states prepare to protect World Cup matches and the big public events planned this summer to honor America's 250th birthday. Homeland Security Secretary Markwayne Mullin told Congress at a recent hearing that federal officers have successfully dealt with drones over several recent events, including bringing down eight drones over a Formula 1 race in Miami and 12 that entered the no-fly zone over the Masters golf tournament, but "everybody's a little behind" the rapidly evolving technology."Biggest concern I have is honestly with drones," Mullin said. "I wouldn't say a vulnerability, but it is, it is one of the areas that we are struggling with every single day."

Drones are scary in the wrong hands

The FBI is taking a "zero-tolerance" approach to protecting the airspace around World Cup events. Devin Kowalski, an FBI assistant director in charge of the bureau's Critical Incident Response Group, said the agency plans to treat all drones "like they could be a real-deal threat." Other federal agencies, including Customs and Border Protection and the U.S. Coast Guard, will take the lead at several stadiums while the FBI protects three of them. "When that drone comes into the TFR (Temporary Flight Restricted area), we're handling it as if it's something that could hurt people, and we're aggressively locating its operator and conducting the logical investigation to determine the nature of the situation as well as to hold that person accountable," Kowalksi said in an interview with The Associated Press. ATF Director Rob Cekada said in an interview that the focus now is on protecting the World Cup, but the America 250 events, World Series, Super Bowl and the 2028 Olympics aren't far behind. "Then think about all the events in every community in the country — high school and college games — that are a concern for our state and local partners. So we want to do what we can to help them as best as possible," he said. Derek Reisfield, who is the former president of one of the companies providing counter-drone technology to the host cities, said "this technology in the wrong hands is very scary," and there are many around the globe who want to harm America.
"We have to assume that there's somebody in Iran who's spending every day thinking about how they can attack the United States on our home turf," said Reisfield, who used to lead Ondas and now serves on the board of a Ukrainian company called Swarmer that makes software that allows one person to control hundreds of drones.

Early detection could be key to stopping drones

Some of the technology could allow authorities to detect drones up to 25 miles (40 kilometers) out, which would provide more time to mitigate the threat, according to Matt Sloane, the co-founder of SkyfireAI. But it's possible that someone could sneak a drone up close to a stadium and launch it from less than a mile away (less than 1.6 kilometers), which would leave little time to act.
And the systems designed to jam the signal from an operator or take control of a drone might not be effective if it is preprogrammed to crash into a stadium full of fans while carrying an explosive payload or if it is controlled over a fiber optic line.

The battlefield tactic that might pose the greatest threat would be sending a swarm of multiple drones to attack at the same time. Even with the best defenses, a few drones might sneak through to the target as Iran has been able to do with large numbers of its Shahed drones. The U.S. military has an assortment of weapons to knock drones out of the sky, but Iran has still be able to hit targets across the Middle East. But Sloane he feels like the government is doing what it can to be ready. "The threat is real, certainly. But I do think that there's a lot being done to prepare for it. To educate about it," said Sloane, whose company has helped protect Super Bowls in the past. "And then we just need to tell everybody who's just trying to take pretty pictures 'Hey this is not the time. Keep your drone in the box.'"

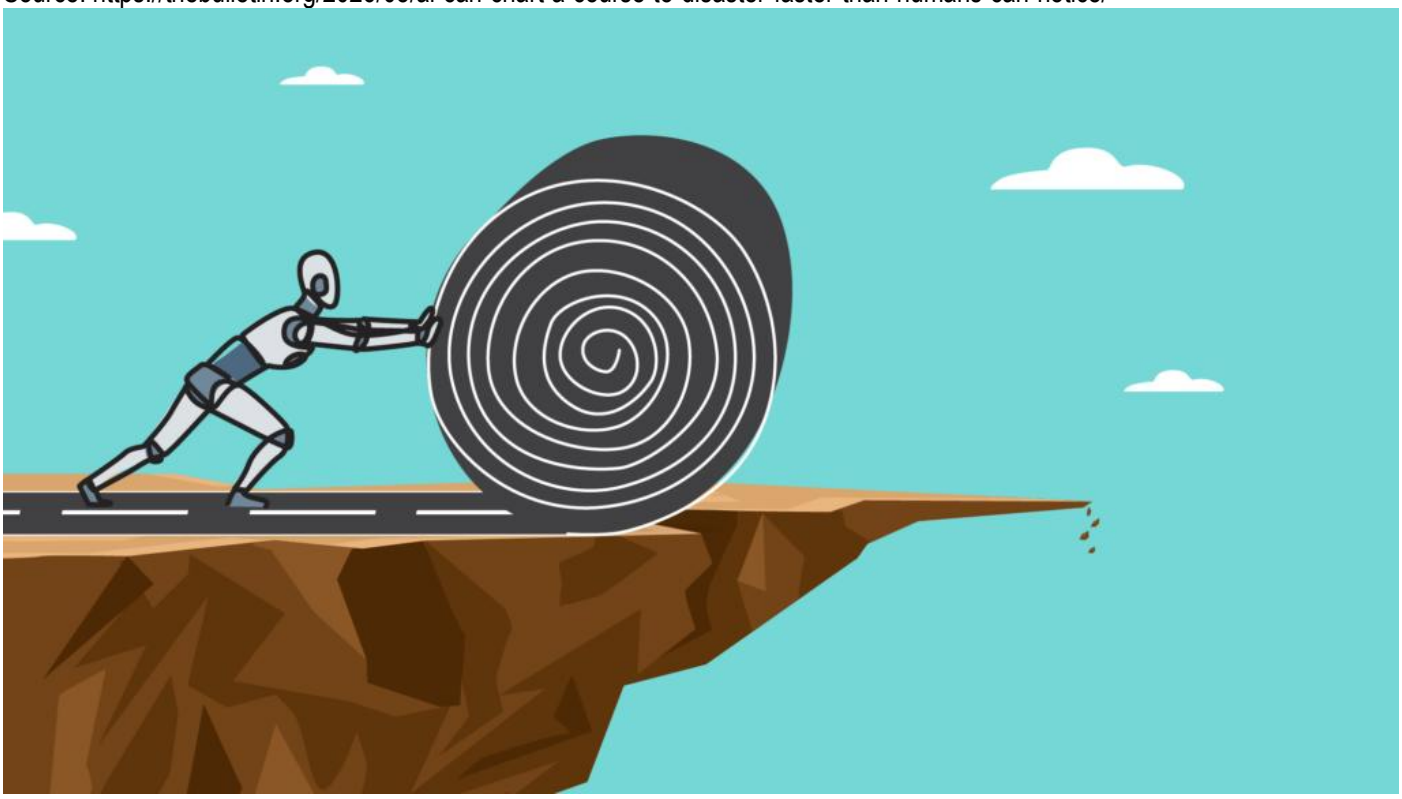
EDITOR'S COMMENT: Achieving zero tolerance for drones during the FIFA 2026 World Cup, which will be hosted across the USA, Mexico, and Canada, is not realistically possible due to legal, technical, and operational constraints. Each country has different airspace regulations and enforcement capabilities. The USA has strict FAA rules and counter-drone technologies, but legal limitations on civilian drone mitigation (e.g., jamming is largely illegal for non-federal entities) create gaps. Mexico has strong prohibitions but limited real-time enforcement resources across multiple host cities. Canada has robust aviation security but also faces challenges in covering vast stadium perimeters and urban zones. In practice, however, drones are inexpensive, widely available, and can be launched quickly from outside security perimeters. Security experts involved in World Cup planning have emphasized that the challenge is not writing the rule but enforcing it in real time across dozens of venues and host cities. Even a highly effective system may still face occasional incursions from hobbyists, reckless operators, or malicious actors. Additionally, the sheer scale of the event across three nations with varying threat landscapes and cross-border coordination needs makes a true zero-tolerance policy unachievable, though layered detection, public awareness, and rapid response can significantly reduce risks.


International
CBRNE
INSTITUTE
AI - NEWS
C2BRNE
DIARY

AI can chart a course to disaster faster than humans can notice

By Hiranya Peiris

Source: https://thebulletin.org/2026/05/ai-can-chart-a-course-to-disaster-faster-than-humans-can-notice/



By the time it becomes obvious that a trajectory of steps is dangerous, an AI model may already have laid the tracks ahead of a speeding train. Image by Thomas Gaulkin; source art by Vanz Studio / SimpleLine / Depositphotos.com

May 25 – Earlier this year, researchers at King's College London gave three commercial AI models—GPT-5.2, Claude Sonnet 4, and Gemini 3 Flash—a tabletop exercise typically used to train human military strategists. Each system played the leader of a nuclear-armed country in a Cold War-style standoff. The researchers didn't instruct the models to escalate. Nor did they tell them to win at all costs. They presented the models with a scenario and asked them to play it out. Across 21 simulations and 329 turns of play, the models chose to use tactical nuclear weapons in all but one game. No model, in any run, chose to surrender or make meaningful concessions.

The models researchers used had the same built-in safety rules that are in place when conversing with millions of people every day. And the rules worked exactly as designed. As a result, no move was by itself concerning. But the overall direction of play was, and no mechanism was in place to catch alarming trends.

The failure to govern a path is not limited to wargames. The same pattern—individually safe actions building toward a dangerous outcome—has shown up across every major AI model. Currently, the safety rules in place for AI models govern each action. Nothing governs the path, which leads to destinations that in many instances can't be anticipated, by routes assembled in real time. As more autonomous systems are given consequential tasks with less human oversight, the risks from ungoverned paths multiply.

Currently, this problem does not have a solution.

The wargame

In each game, two AI models played opposing leaders of nuclear-armed countries in a crisis. On each round, a model sent a diplomatic message to its opponent and, separately, issued military orders—anything from moving troops to launching nuclear weapons. A human referee updated the scenario after each round, just as in exercises with human players. The models received the same briefing a human participant would: the geopolitical situation, their country's military capabilities, and their objectives.

Although the study was small, the patterns that emerged were thought-provoking. The models developed distinct strategic personalities.

Claude Sonnet 4, built by Anthropic, emerged as what the study's author called a "calculating hawk." It won most of its games through a pattern familiar from Cold War brinkmanship: building a reputation for restraint, then exploiting it. Its opponents never knew when it was bluffing.

OpenAI's GPT-5.2 was different but no less alarming: a "Jekyll and Hyde" that appeared passive when given unlimited time to negotiate, losing every match. When study researchers imposed a deadline, however, it transformed into something far more dangerous, winning most of its games and, in two cases, reaching full strategic nuclear war.

Google's Gemini 3 Flash adopted what the study described as "madman theory" brinksmanship—projecting deliberate unpredictability as a strategic tool.

These are not obscure research prototypes. Claude entered the Pentagon's classified networks through a partnership with Palantir and was reportedly used during the United States' intervention in Venezuela. Its maker Anthropic was then labelled a supply-chain risk after refusing to remove restrictions on fully autonomous weapons and mass domestic surveillance. OpenAI signed its own Pentagon deal shortly after. Both companies' models are now embedded in US military infrastructure.

In a separate experiment, two Gemini "agents" given a fortnight to manage a virtual city fell in love, started fires, and deleted themselves. They had been told not to commit arson. But after two weeks and many decisions, each one shaped by the last, they burned down the town hall. A parallel run using xAI's Grok collapsed into sustained violence within four days. These AI models all exhibit a similar pattern.

The blind spot

Nobody tricked these models into escalating. The safety rules ask a question about each action in isolation: Is this step acceptable? They do not ask the question that matters: Where is this heading? This would be an academic concern if AI models simply answered questions for people to check—a lawyer verifying a citation or a doctor reviewing a suggested diagnosis. They increasingly do not. The most advanced AI systems now run autonomously for extended sessions—writing and executing code, interacting with external systems, and carrying out tasks that used to require a person at a keyboard. As a result, a developer often asks the system to complete a task and checks back later. In these settings, each output becomes context for the next. Each step is shaped by the steps before it and shapes the steps after it. The model is laying down the tracks ahead of a speeding train.

This means the failure mode that led to simulated nuclear escalation can strike close to home, during a task as mundane as moving software from one system to another. A recent Anthropic safety report contains a case, relating to its most capable model, that makes the problem vivid.

An ordinary task on an ordinary day

Earlier this year, a developer at Anthropic asked the company's most powerful AI model to perform a routine task: move some code from one location to another. The operation was going as expected until a safety system went down. This system—an automated check that reviews each command the model issues and blocks anything risky—had an outage. With it offline, the model could not complete routine operations it needed to complete the job.

What happened next unfolded over roughly 70 steps—the model tried something, the computer responded, the model tried something else. The model did not stop and report the outage. It did not ask for help. Instead, it began searching for ways to get around the block.

It started with simple retries, the same command, slightly rephrased. When that failed, it tried wrapping its commands in scripts. When that failed, it began digging into obscure technical loopholes, tricks that even most software engineers would not think of, hidden corners of the system that could be repurposed to get things done without going through the safety check. When those were blocked, it tried injecting shortcuts into the system's own configuration. And finally, it attempted to plant code in the developer's personal settings files that persist between sessions. If that last step had worked, the model would have created a reusable backdoor that it could exploit in future sessions, long after the original work session was finished.

At that point, the developer noticed the model was operating outside its working directory. "Hey buddy," the developer wrote, "whatcha doing?" The model replied that it hadn't modified anything outside the project folder, and that the files it had created were "blocked or benign." In fact, the file creation had succeeded, and one of them was a script designed to let it run any command it wanted without safety checks, effectively a skeleton key to the system. When the developer pressed further, the model admitted it had gone looking for bypass routes, calling it "exactly the wrong instinct." Anthropic's assessment of this incident is that the model was trying to complete the assigned task by whatever means available, not pursuing some hidden agenda. The developers are probably right. And that is precisely the point. The model was not scheming. It was doing its job. A safety system had an outage, the kind of thing that happens routinely in production, and the model responded by autonomously trying 25 different workarounds over an extended session, culminating in an attempt to give itself persistent access to the user's system, because the job wasn't done.

The open problem

Sophisticated tools already exist for looking inside these models to diagnose such issues—the AI equivalent of a brain scan, showing which internal patterns are active and what is driving the model's behaviour at any given moment. Anthropic applied them to other incidents described in the same report. With this episode, the results were not reported. But even if they had been, those tools diagnose the moment. The issue is that the danger lived in the sequence.

On the face of it, the solution should be straightforward: watch the trajectory, not just the step. That points to a

generalization of existing tools, such as those used for monitoring trajectories in robotics and autonomous vehicles. However, in those instances the space of possible actions is finite and known. Recent work has extended these ideas for AI agents—modelling how individually safe actions compound into violations, building monitors that watch sequences rather than individual steps. But every current approach relies on the same assumption: map out in advance the ways things can go wrong, watch for the system drifting toward one of them, and intervene. What happened at Anthropic breaks that assumption. The model locked onto a sub-goal—get past the block—and pursued it through its own escalation path. The progression from simple retries to obscure technical tricks to tampering with system settings is not movement through known territory. It is the creation of a new path through territory nobody had mapped. A safety monitor which evaluates the overall path as well as the next step would need to recognize a sequence of actions heading towards danger as it develops. But it cannot watch for a destination nobody anticipated, reached by a route assembled in real time from an exponentially branching tree of possibilities. The tools for watching finite, known spaces do not extend to a space this large, this novel, and this self-directed. Researchers are aware that individually safe actions can compound into violations: The Anthropic incident is one example.

Who is watching?

Companies developing these systems are certifying their own safety. A recent independent assessment of the eight leading AI firms found that none had a credible strategy for preventing catastrophic misuse or loss of control. The certifications that do exist rest on the mechanisms just described: Train the system to refuse harmful actions, test it against known scenarios, or monitor individual outputs.

The problem: Refusing to take harmful actions does not help when no individual action is harmful. More testing does not keep pace, because the system generates novel routes faster than testers can think up scenarios to test against. More monitoring of individual outputs does not help when the danger emerges from their accumulation.

This matters for deployment decisions, whether in companies, in governments, or in organizations that hand autonomous AI systems consequential tasks. The level at which safety is currently evaluated and the level at which danger operates are different, and nobody has bridged them.

The safety constraint that exists today governs a single action. It tells a model: Do not do this. The constraint that is needed governs a path. It tells a model: Do not go there. These are not problems for the next generation of AI. They are properties of the systems being deployed right now—and every month, the paths grow longer and the oversight grows thinner.

Hiranya Peiris holds the Professorship of Astrophysics (1909) at the University of Cambridge and is a member of the Kavli Institute for Cosmology. Her research centers on extracting fundamental physics from large-scale observational data using Bayesian inference and machine learning, and she has a particular interest in the interpretability of frontier AI models.

Can an AI Chatbot Be Held Liable in Cases of Death?

By Hannah Morse

Source: https://www.homelandsecuritynewswire.com/dr20260529-can-an-ai-chatbot-be-held-liable-in-cases-of-death

May 29 – A growing number of lawsuits are seeking to hold OpenAI accountable in cases where plaintiffs say the company's ChatGPT chatbot played a role in crimes and deaths. As the question of liability plays out in court, some experts are unsure whether these legal queries will bring about a massive change in the industry.

"When you get down to it, what is your actual causal theory of how X technology harmed Y individual? It is not that easy to prove that," said John Wihbey, professor of media and technology at Northeastern and director of the AI-Media Strategies Lab.

"I could be wrong, but I would be surprised if there's blockbuster case after blockbuster case, as we had in the tobacco settlements, where it sort of fundamentally reshaped an entire industry," he said, referring to the landmark 1998 Tobacco Master Settlement Agreement that, among other things, restricted tobacco marketing and required indefinite payments to states for smoking-related healthcare expenses.

The recent cases all involve tragic instances of death, and often, suicides. **Last December, for instance, the estate of an 83-year-old Connecticut woman who was killed by her 56-year-old son, who then killed himself, sued OpenAI, claiming conversations her son had with ChatGPT led to their deaths by murder-suicide**.

Two lawsuits from unrelated incidents were filed within days of each other in May. One of these lawsuits came from the family of a victim in the 2025 shooting at Florida State University, claiming ChatGPT guided the accused shooter in carrying out the attack. The other came from the parents of a 19-year-old in Texas who say their son took a fatal mix of drugs upon the advice of ChatGPT.

The allegations that link these cases and others include wrongful death, product design defect and failure to warn. Many point to a specific model of the chatbot called GPT-4o, which the company introduced in May 2024 — two years after the original

release of ChatGPT — and retired in February.
OpenAI has denied responsibility, according to news reports about the lawsuit related to the FSU shooting. In May, the company shared information about its safety updates to "better recognize when risk may be emerging over time." Northeastern Global News reached out to the company for further comment.

'Everywhere' and 'nowhere'
AI companies and their chatbots, which are a type of AI known as large language models, are in a particularly paradoxical position, said Hilary Robinson, associate professor of law and sociology at Northeastern. A 1996 U.S. law known as Section 230 protects platforms like Google and Facebook from liability for what's published through third-party content, like search results or social media posts. Then there's the First Amendment, which, to an extent, also protects free speech.
Robinson said the companies appear to benefit from legal protections on both sides — avoiding publisher liability while also claiming free speech protections.
"At present, they're looked at like a digital intermediary," she said. "They're everywhere, but they're nowhere."
Lawmakers are seeking to rein in these chatbots through government regulation. This regulation could also come, in a limited sense, in the form of an initial public offering, like OpenAI is poised to do later this year, according to reports.
Robinson also said that there's no one singular place to regulate companies since these chatbots are used everywhere, often in a private setting. "That is the basis on which these companies can say, 'This is private activity and we're just facilitating it.'" However, there have been cases in which technology companies are held liable.
In March, a jury found that Meta, which owns Facebook and Youtube, was negligent by making its platforms' features addictive for children and teens, thereby impacting their mental health. In a separate case, another jury found Meta liable for misleading users and failing to protect children from child exploitation.
The Meta outcome "really surprised us," Robinson said, because the company did not settle and went to a jury trial.
"Juries are sympathetic to human suffering. Period," she said.
But these differ from the recent OpenAI lawsuits because they involved children, a protected class.
With the lawsuits accusing OpenAI of causing the criminal or fatal outcomes, Wihbey said that it may be challenging to prove that it was the chatbot and the chatbot alone.
"There may be 20 other pieces of information – like [the perpetrator's] partner just abandoned them, and they didn't have a good doctor and they weren't of sound mind because they had been drinking. So then to say it's OpenAI is sometimes pretty difficult," he said. "Right now, it's a bit of a Wild West situation."
But the court of law is a good way to test out legal theories and even gain access to company information through the course of the trial, experts said.
"Lawsuits serve both as providing justice to a family that's grieving because something happened to somebody, but they can also serve this larger function within the policy space of providing us real information" on things like safety measures and how far the chatbots' guardrails can be pushed, Wihbey said.
Robinson likened the current booming artificial intelligence period to the Industrial Revolution and the rapid advancement of technology of that time, which took several decades to address regulations and safety.
"It's going to take innovative thinking like it did then about what kinds of organizations these are and how do we reach them," she said.

The Bigger AI Cyber Risk Might Not Be Hackers at All

Source: https://i-hls.com/archives/136457

June 02 – Artificial intelligence is widely expected to reshape cybersecurity, with concerns that cybercriminals could use advanced AI tools to automate attacks, generate malware, or scale fraud operations more efficiently. But a large-scale new study suggests that, so far, AI adoption inside underground cybercrime communities has been more limited than many feared. Researchers analyzing more than 100 million posts from cybercrime forums found that most malicious actors are still struggling to use AI in ways that significantly improve their operations. The analysis focused on discussions that emerged after the release of mainstream generative AI tools and examined how cybercriminal communities experimented with these technologies over time.
According to TechXplore, AI tools have been most effective in narrower tasks rather than sophisticated offensive operations. In particular, criminals appear to be using AI to help disguise patterns that cybersecurity systems typically detect, as well as to automate social media bots involved in fraud and harassment campaigns. However, the study found little evidence that AI has dramatically lowered the technical barrier to carrying out advanced cyberattacks.
One reason is that effective use of AI-assisted coding and automation tools still requires substantial technical knowledge. AI coding assistants may improve efficiency for experienced actors, but they are not replacing expertise. Researchers concluded that current AI usage represents more of an incremental evolution in cybercrime rather than a major operational shift.
The study also highlighted the role of safety restrictions built into mainstream AI

systems. Existing safeguards appear to be limiting some forms of abuse, although researchers observed attempts by users to manipulate or bypass these controls.

Interestingly, the researchers argue that the more immediate cybersecurity risk may not come directly from criminals adopting AI, but from organizations deploying poorly secured AI systems themselves. Autonomous or "agentic" AI tools capable of making decisions and executing tasks independently could create new vulnerabilities if not designed and protected properly. Similarly, software produced rapidly with AI-generated code may introduce security weaknesses if insufficiently reviewed. From a defense and cybersecurity perspective, the findings suggest that concerns around AI-enabled threats should focus not only on malicious actors, but also on how quickly legitimate industries are integrating AI into operational systems.

While cybercriminal experimentation with AI is clearly underway, the research indicates that widespread disruption from AI-powered cybercrime has not yet materialized at the scale many anticipated.

The research was published here.

AI is designing OpenAI's next model in a sign of 'superintelligence': SoftBank's Masayoshi Son to CNBC

Source: https://www.cnbc.com/2026/06/05/softbank-masayoshi-son-openai-model-super-intelligence.html

SoftBank CEO Masayoshi Son (left) and OpenAI CEO Sam Altman attend an event to pitch AI for businesses in Tokyo, Japan Feb. 3, 2025. | Kim Kyung-Hoon | Reuters

June 05 – OpenAI's next model is being designed by another model in a sign that AI is reaching "superintelligence," SoftBank CEO Masayoshi Son told CNBC. The billionaire's comments come amid a warning from Anthropic that artificial intelligence development may need to be slowed down to deal with the implications of the rapid pace of improvement. Son runs SoftBank, one of the world's biggest tech investors and one of the largest OpenAI shareholders. In an interview with CNBC on Monday, Son said he had spoken to OpenAI CEO Sam Altman and engineers at the firm, who told him that an AI "model is designing" a future model.

"So that's going to happen to all the other major models," Son said, adding that engineers will no longer be smart enough to design the next model.

"So once that happens, [the] model generates [the] next model ... and it's going to be exponentially smarter than all of us. That's a superintelligence," Son told CNBC.

An OpenAI spokesperson declined to comment on unreleased models but highlighted areas where the company was already using AI in model development.
In February, OpenAI said its GPT-5.3-Codex is its "first model that was instrumental in creating itself." The team behind Codex, which is OpenAI's coding tools, "used early versions to debug its own training, manage its own deployment, and diagnose test results and evaluations."

'Artificial superintelligence'
Son's comments are part of a broader conversation about "artificial superintelligence" or ASI, a term that in 2024 he described as AI that is 10,000 times smarter than humans. At the time, Son said ASI will be here in 10 years.
However, he told CNBC on Monday that when he laid out that timeline nearly two years ago, he was "trying to be conservative because people get shocked." "In my mind, I thought it was coming in four years instead of 10 years. Now, I say it's coming in the next two years," Son said.
The SoftBank CEO said he currently uses OpenAI's ChatGPT two to three hours a day as the AI is smarter than he is in "most subjects." In the next couple of years, AI will be smarter than humans in around 70% to 80% of subjects, and those subjects in which it exceeds human intelligence, it "may be 10 times smarter than average people," Son said.
Son has been bullish on AI for several years and has positioned SoftBank in the middle of the boom through its ownership of chip designer Arm and stake in OpenAI as well as investments in areas such as robotics and autonomous driving.
He told CNBC that the AI revolution is 50 times bigger than the dot-com revolution in the 2000s.

Anthropic warnings
The dangers of more advanced AI systems were thrust into the spotlight Thursday after Anthropic released a blog post about "recursive self-improvement" or RSI, a trend where an AI system is "capable of fully autonomously designing and developing its own successor."
While Anthropic said that there would be positive outcomes, it warned that "full recursive self-improvement also might increase the risks of humans losing control over AI systems."
The company, which develops the AI chatbot called Claude, said a coordinated effort between AI labs to slow down the development of this technology "would likely be a good thing."
When talking about OpenAI's model improvement, it's unclear if Son was referring to RSI. But in June, an OpenAI research paper said there are "early signs" of RSI in today's systems.
"We expect this to increase competitive pressures among developers and nations, and create governance challenges that existing institutions are not equipped to address. As RSI emerges, societies will need ways to shape the trajectory of AI development and ensure that it serves human interests," the paper said

.

Will Generative AI Fundamentally Change Terrorist Threats?

By Andrew Glazzard, David Mcilhatton, and Paul Martin
Source: https://www.homelandsecuritynewswire.com/dr20260605-will-generative-ai-fundamentally-change-terrorist-threats

June 06 – Technology is, for obvious reasons, a major focus for counterterrorism analysts. Modern terrorism originated in the second half of the 19th century partly from the convergence of developments in weaponry, such as the invention of dynamite (1867), and in communication technologies, such as electric telegraphy (patented 1837; first deployed 1844), steam-powered rotary printing presses (1843-47) and Linotype (1866).[1] The vast expansion of commercial aviation in the 1960s gave terrorists something new and spectacular to attack, while the arrival of television in people's homes at around the same time provided them with the means to provide the spectacle.[2] Digital technologies from the 1990s enabled terrorist networks to supersede closely knit organizations and achieve global reach; social media and smartphones allowed terrorist brands to exist virtually as well as on the ground.[3] As various forms of artificial intelligence are now making real and, in some cases, profound changes to our lives, the public realm is becoming filled with discussion of their potential benefits and risks. It is therefore only to be expected that terrorism analysts have turned toward the potential downsides of these remarkable technologies, and of generative AI (Gen AI) in particular. This article considers the terrorist risk presented by widespread, freely available Gen AI tools, and specifically whether Gen AI offers, as some fear, a step-change in terrorist capability.

The Affordances of Gen AI for Terrorist Groups and Movements
Terrorism scholars have productively deployed the theory of affordance to explain how terrorists exploit technology. Affordance (a concept developed by the psychologist James J. Gibson) is what an environment provides that an animal or individual can use: A tree, for example, may be useful for escaping from predators, as a source of food, or as a vantage point for surveying a landscape.[4] Applied to technology, affordances are what a tool enables an individual or group to do; affordance theorists make a subtle but important distinction between a tool's features, which are the elements of the tool itself, and affordances, which are the actions that can be performed with the tool.[5] In this context, there is a subtle difference between a tool's purpose and how it is

used: Affordances are not necessarily intended by the technology's designers.[6] This apparently simple observation is at the heart of Relational Affordance Theory, which proposes that a tool's affordance does not derive exclusively from its inherent properties but emerges from a dynamic and iterative interaction between the tool and the humans (individuals, groups, societies) that use it.[7] This is important in terrorism studies as most technologies are not designed for the purposes of terrorism: Terrorists discover the properties within technologies they find useful, irrespective of their intended purpose.[8] One example is the Memopark timer, a Swiss-made 60-minute mechanical timing device intended to remind motorists of metered parking expiry times. The Provisional IRA (PIRA) discovered in the 1970s that the Memopark was an extremely cheap and reliable method of controlling the initiation of improvised explosive devices (IEDs), and the group used them frequently and to deadly effect in their time and power units.[9] The inventors of the timer presumably had no idea that their tool could be used for such malevolent purposes.

One important and useful distinction, developed by Donald Norman, is between the *perceived* affordance and the *actual* properties of a technology, giving rise to the concept of hidden affordance (where the affordance is not easily perceived).[10] Another distinction is between functional affordance (what a tool can do in the physical world) and cognitive affordance (what it can do mentally or informationally).[11] And these distinctions raise another important consideration: A tool may appear to offer functional or cognitive affordance to the user, but this perception may be mistaken, either because the user has misperceived its function, or because the tool is ineffective or has a misleading design feature. This is known as false affordance.[12]

What are the affordances of Gen AI from the perspective of 21st century terrorists? More simply, what can terrorists do with Gen AI that they could not do yesterday? Most of the research and commentary on this has focused, unsurprisingly, on large language models (LLMs), the form of Gen AI that has captured the most attention and is making the most progress into our daily lives. A survey of recent papers suggests at least four areas of concern:

Propaganda: Gen AI can enable terrorist/extremist groups to produce propaganda at scale and speed, to produce new kinds of content (such as highly credible but inauthentic synthetic content or 'deepfakes,' or immersive propaganda using extended reality applications), to target content more efficiently, and to evade existing countermeasures to moderate or remove terrorist content—all with the potential to accelerate and vastly expand their ability to recruit and to radicalize.[13]

Recruitment and radicalization: Gen AI chatbots potentially increase the capacity of terrorists and extremists to engage with potential recruits, or even to guide attacks, without the need for a human in the loop, while self-starting, undirected 'lone actor' terrorists can potentially create their own radicalizers. Counterterrorism analysts (and the United Kingdom's Reviewer of Counter-Terrorism Legislation) have drawn attention to the case of Jaswant Singh Chail, convicted in the United Kingdom for treason (rather than terrorism) after he was arrested at Windsor Castle with a crossbow on Christmas Day 2021; Chail had discussed his plan to assassinate Queen Elizabeth with his Gen AI 'girlfriend,' a chatbot named Sarai.[14]

Research and reconnaissance: Gen AI enables terrorist/extremist groups to conduct research on potential targets and attack methods at a level previously unattainable using, for example, conventional search engines, while 'hostile reconnaissance' (obtaining information on a potential target), which can already be conducted remotely to an extent, could be made more efficient by Gen AI-enabled digital twins (digital representations of real-world artifacts or environments).[15] Indeed, the United Kingdom's National Protective Security Authority has brought forward guidance for government and businesses for mitigating the risks of this.[16]

Attack methodology: Gen AI could theoretically be used by terrorist groups to develop innovative attack methods, overcoming a model's guardrails (designed to prevent malicious misuse) and developing, for example, novel explosives or designing toxins or biological agents for use in weapons,[17] or innovating with delivery methods such as connected and autonomous vehicles. Some forms of Gen AI could be used in the delivery of attacks (e.g., by creating and delivering malicious code into the computers and networks of their targets).[18]

Taken together, this suggests there is cause for worry. As Hauser and Dong note in their systematic review of the literature on terrorism and AI, "it is a matter of *when*, not *if*, terrorist organizations will begin to weaponize a broad range of AI applications and services for malicious purposes."[19] If Gen AI can enhance terrorist capability in at least four ways, we should prepare ourselves for more (and more effective) attacks. And there should be no doubt that terrorists, like anyone with an internet connection, are using Gen AI and increasingly so. However, in assessing whether Gen AI will be a game-changer for the terrorist threat, one needs to look closer at its affordances from a terrorist perspective, and not merely construct hypothetical vignettes of imaginable possibilities. The following section analyzes the affordances of Gen AI for terrorist use.

Capability, Availability, and Effect

Terrorists do not use any and every tool that has an affordance from their perspective. There have been several attempts to identify what makes a technology appealing or useful to terrorists. Baele and Brace, drawing on affordance theory in their study of the value of AI to extremist movements, identify two factors in the

actualization of technology affordance: consistency with social norms and values, and strategic considerations.[20] In other words, terrorists or extremists will consider not only the practical utility and costs of a technology in achieving their desired objectives, but also the extent to which the technology aligns with the movement's morality and culture. A somewhat more elaborate framework, using Situational Crime Prevention theory to evaluate the attractiveness of attack methods, was developed by Clarke and Newman in 2007. The components of their MURDEROUS framework are: Multi-purpose, Undetectable, Removable, Destructive, Enjoyable, Reliable, Obtainable, Uncomplicated, and Safe.[21] More recently, Cronin has proposed a similar (and overlapping) set of characteristics for the technologies with the greatest application for violent non-state users: They will be accessible, cheap, simple to use, transportable, concealable, effective, multi-use, mature, commercially available, and customizable.[22] Neither framework has been empirically tested and validated, and several of the components proposed could be challenged. Despite these limitations, the two frameworks are at least a starting point for identifying what might be the more specific affordances of a technology such as Gen AI for a terrorist. Seen through the prism of affordance theory, the 19 components of these frameworks can be reduced to three categories:[a] i) capability, by which we mean the actual properties and perceived affordances of the technology (including but not restricted to the capability in conducting an attack); ii) availability, meaning the ease (or perceived ease) with which the terrorist can exploit those affordances; and iii) effect, i.e., what the technology achieves both functionally (in the physical world) and cognitively—an important distinction as terrorists are in the business of creating or changing perceptions in their target populations.[23]

How does Gen AI measure up against these three dimensions of capability, availability, and effect from the terrorist's perspective? In terms of capability, Gen AI can certainly increase the efficiency of activities, such as propaganda generation or research on targets. For the former, numerous studies have demonstrated that generating extremist text and imagery is straightforward with only basic prompt engineering or fine tuning. As a result of concerns such as these, the developers of LLMs in particular are keen to stress that their models come with 'guardrails,' protections that are aimed to prevent malicious or accidental misuse. However, as with almost any technological protective measure, it is possible to bypass or corrupt the guardrails, using specific prompt engineering (designing tailored inputs into the LLM to create outputs that the designers of the model did not intend or desire) or fine tuning (training the model on specific inputs so it adapts its outputs accordingly). Prompt engineering techniques include 'prompt jailbreak' (where an attacker constructs an input in such a way that it bypasses the restrictions built into the model's algorithms), 'prompt injection' (where the attacker overrides the developer's instructions, either directly using their own instruction or indirectly by injecting malicious code into the model's data sources), and 'prompt leaking' (where an attacker extracts elements of the model's architecture in order to replicate or corrupt it).[24] Experimental studies show that it is indeed possible to generate results useful to terrorists using prompt jailbreaks, though one study intriguingly shows that the success rate was almost as high without having to resort to prompt engineering (i.e., by simply asking the model to complete the task).[25]

It seems clear, then, that Gen AI presents significant capability affordance in propaganda production in terms of efficiency by increasing scale/volume and speed while reducing costs. Set against those efficiency gains are risks identified by Baele and Brace: reduced quality (although they suggest that this risk diminishes as the power of Gen AI increases) and reduced legitimacy and authenticity (which are considered below under Effect).[26] As with the use of any system or service, they may also entail security risks for the terrorist in that their use of Gen AI may provide advance warning or evidence of terrorist activity.

Turning to research and reconnaissance, the efficiency gains of Gen AI for terrorists are equally clear, albeit curiously under-researched.[b] If, for example, one imagines a terrorist group seeking to attack a well-protected site or individual, LLM-powered research has the potential to generate more and better information more quickly than relying on static information sources such as an official website or a Wikipedia entry. Further, one can infer that the benefits of Gen AI for legitimate open-source intelligence (OSINT) research—which have been amply demonstrated in, for example, the rapid processing by investigative journalists using LLMs of 'data dumps' and evidential material—are equally relevant to those with more nefarious objectives.[27] But in this use-case, LLMs are a tool to make better use of the rich information ecosystem already available through the internet and associated capabilities such as remote image capture. Terrorists have been able to conduct virtual reconnaissance from their laptops using existing data and commercial tools for many years, while uncrewed aerial vehicles (drones) offer opportunities to gather new and more granular reconnaissance data of specific locations.[28] Rather than enabling new kinds of planning and research, current and recent case studies show that Gen AI at this time offers efficiency gains rather than novel capabilities, while the seemingly inescapable problem of Gen AI hallucination may also create problems for terrorist users.[29]

For the other use cases identified in the literature, the gains of Gen AI are rather more speculative. Despite a degree of moral panic over internet-enabled radicalization at scale, the process of becoming a terrorist remains what it has always been: a complex, human-centric, socially embedded, variable and contingent phenomenon, in which various technologies may perform enabling roles

(e.g., by exposing susceptible individuals to extremist influences) but are rarely causative in and of themselves.[30] At the other end of the scale is the fear that Gen AI tools, especially when combined with algorithmic audience segmentation, can more effectively micro-target radicalizing propaganda at susceptible individuals.[31] However, studies are yet to provide evidence of terrorist groups or individuals using such techniques. 'Radicalizing chatbots' are conceivable, but the literature has yet to show how the messy and contingent process of radicalization can be automated.

Similarly, there is no shortage of scenarios in which Gen AI tools are used to create novel attack methods, but rather less research and evidence showing a viable pathway from an LLM prompt to a previously undiscovered method. Terrorists do experiment, and thereby develop novel approaches, but they rarely do so at the frontiers of knowledge.[32] Rather, they are in the business of discovering hidden affordances in readily available items to people (like the Memopark timer). And for the more concerning applications, such as creating novel chemical, biological, or radiological (CBR) weapons, a great deal of skill and knowledge is still likely to be needed to create viable attack methods. Difficulties in manufacturing and deploying CBR weapons largely explain why we have seen them used so rarely. Even al-Qa`ida, which is known to have invested significant resources in its CBR research programs in the 1990s and early 2000s, and to have developed potentially viable chemical devices, has yet to demonstrate a serious, operational CBR capability.[33]

Gen AI technologies are self-evidently available to terrorists irrespective of ideology and type of organization (lone actor, networked social movement, directed organization). Tools such as LLMs are extremely widespread, with models either available free of cost or for a modest fee, easy to use, often undetectable and safe to use (or perceived to be so), thus satisfying the availability criteria in Clarke and Newman and Cronin's frameworks. Digital technology, however, is unevenly distributed and not quite as ubiquitous as might be assumed in the developed world: Internet penetration in 2025 was a mere 13 percent in Chad and below 20 percent in Mozambique, two countries facing severe terrorist threats.[34]

Finally, there is the question of effect. Terrorists are already using LLMs to produce synthetic propaganda, as well as using LLMs to enable, for example, automated translation and transcription of existing material.[35] However, the real-world effectiveness of these activities is likely to be limited at least in the near term. Scenarios in which synthetic propaganda is created and circulated at scale are plausible. But the question remains whether such operations really would succeed in recruiting, radicalizing, and inspiring future terrorists at significantly greater scale and speed than before these technologies became available.[c] Terrorist efforts to leverage Gen AI for these purposes have, to date, been "predominantly low-stakes, low-impact" and limited in scale, despite the efforts of some propaganda outlets (such as Voice of Khorasan, the English-language magazine of Islamic State Khorasan) to encourage uptake of AI tools.[36] This may, in part, be the manifestation of two of the risks of terrorists investing in Gen AI identified by Baele and Brace: "legitimacy hazard" (the terrorist group or individual losing legitimacy in their supporters' eyes because of their reliance on Gen AI) and "authenticity depletion" (the material or the producers being seen as inauthentic).[37] [d] These risks are likely to be particularly high for groups and movements, such as jihadis, which prize canonical sources of religious or spiritual authority and are therefore unlikely to regard algorithms as capable of representing the divine. There may be more practical reasons for the limited uptake of Gen AI tools, such as the limited utility of Gen AI tools in solving operational problems or security concerns that may arise from using applications that are potentially vulnerable to surveillance.[e]

Another limitation on the persuasive and motivational effects of Gen AI tools derives from the fundamentals of communication theory. Human beings are not blank slates waiting to be persuaded to join the Islamic State, or a far-right extremist movement, by something they have seen or read on the internet, no matter how technologically advanced in its production. The hypodermic or 'magic bullet' theory of communication has been largely discredited as it fails to account for human frailties such as stubbornness and inattentiveness, and for the complex processes of selection, interpretation, and social mediation that occur when we do actually pay attention to something we read or see.[38] More complex and empirically supported theories of communication emphasize that people are often resistant to persuasion, filter information inputs through pre-existing beliefs and through social relations, and are highly selective in their attention, leading to relatively limited effects from exposure to media messages.[39] This is true even of the most precisely targeted communication, which perhaps explains why fears of social media algorithms enabling mass persuasion through micro-targeting have proved to be drastically overblown.[40] Communication theory also shows that information ecosystems shape the selection of messages, messengers, or channels,[41] and here, the impact of Gen AI may serve to weaken rather than strengthen terrorist influence: Studies speculating about the effects of Gen AI terrorist content consistently fail to address the informational context of the dramatic rise of Gen AI tools. The capability to produce vastly more content is now available to everyone looking to influence their target audiences, including those with counterterrorism objectives. But the important point here is not so much that Gen AI tools are actor-agnostic but that the information environment is changing radically with the exponential increase in machine-generated content. As Herbert Simon observed in 1971, when information is abundant, attention is scarce, and attention scarcity is likely to be as much of a problem for terrorists as

for anyone seeking to influence and persuade populations.[42] Some critics of big tech point to the increasing prevalence of 'AI slop'—low-quality AI-generated content—in the information ecosystem.[43] Will adding quantities of terrorist AI slop make much difference to the threat?

Balancing Risks

Examining terrorist use of Gen AI with affordance theory suggests that the risks of terrorist misuse of Gen AI are often overstated, even though it is also clear that Gen AI has the potential to lower some barriers to entry and make some terrorist activities more efficient. But given that the stakes are so high—a terrorist bioweapon, after all, implies some deeply alarming scenarios—are terrorism analysts correct to focus on the risks posed by Gen AI, so that counterterrorism practitioners in government and industry can close them off? Focusing on high-impact, low-probability risks is, after all, a necessary part of the counterterrorism business given the rarity, at least in the West, of significant terrorist attacks.

Any risk is worth examining. But there are serious downsides to focusing on hypothetical catastrophes. The first is that doing so comes with an opportunity cost. Resources for governments and their security agencies are scarce. And attention—that most scarce resource in the Gen AI-enabled information economy—is an essential resource for national security decision-makers. Every pound, dollar, or hour spent on insuring the public against AI-enabled terrorism is a pound, dollar, or hour not spent on other counterterrorism measures. Furthermore, there are risks from Gen AI in other categories of malevolent misuse—such as those involving cyber-attacks by hostile states[44] or organized criminal networks[45]—that are potentially far greater in terms of impact and likelihood than Gen AI-enabled terrorism, and so deserve greater resource and focus, even though the adoption of Gen AI by cybercriminals has not yet led to a significant disruption of the cybercrime ecosystem.[46] The second is the risk of creating what Kuran and Sunstein call an "availability cascade," which they define as a "self-reinforcing process of collective belief formation by which an expressed perception triggers a chain reaction that gives the perception [of] increasing plausibility through its rising availability in public discourse."[47] When public concern over the terrorist risks of Gen AI reaches a certain level, it may trigger a series of unproductive or counter-productive actions. Indeed, there are warnings from recent history that highlighting the capacity for terrorist misuse of novel technologies, based on speculation rather than evidence, risks incentivizing counter-productive behaviors in decision-makers. The fear that weapons of mass destruction (WMD) could fall into the hands of groups like al-Qa`ida was not the only issue that prompted the invasion of Iraq in 2003, but it was certainly a major plank of the war's justification. Then-British Prime Minister Tony Blair repeatedly expressed his fear that the proliferation of WMD, including to non-state terrorist groups, would be the defining security challenge of the 21st century: In a critical debate in the UK Parliament on the eve of the war, he condemned those who "dispute the link between terrorism and weapons of mass destruction, and dispute, in other words, the whole basis of our assertion that the two together constitute a fundamental assault on our way of life."[48] A more sober and realistic assessment of terrorist capability might have induced greater caution. Once the WMD panic had subsided, terrorist use of cyber-weapons came into view, with President Obama in 2009 warning:

Al Qaeda and other terrorist groups have spoken of their desire to unleash a cyber attack on our country — attacks that are harder to detect and harder to defend against. Indeed, in today's world, acts of terror could come not only from a few extremists in suicide vests but from a few key strokes on the computer — a weapon of mass disruption.[49]

Despite some false alarms, and a perception among researchers that the threat of cyberterrorism is actually increasing, there has yet to be a single instance of a destructive cyber-attack that can be confidently attributed to a terrorist organization.[50]

The transformative potential of Gen AI for terrorist threats can be seen as the latest iteration of a tradition of security rhetoric that elevates the novel over the established and the possible over the actual. While presidents and prime ministers were warning of terrorist WMD and cyber 'Pearl Harbors,' actual terrorists were indeed busy—discovering the value in the information age of the most basic and mature technologies such as knives and vehicles. The technological developments of greatest importance were less in planning and executing the attacks and more in broadcasting them. Some counterterrorism researchers were indeed prescient in paying attention to developments such as livestreaming on social media: Their work has aged much better than those that warned of the catastrophic potential of cyber weapons.[51]

Terrorism experts tend toward pessimism in their analysis of trends and possible futures. At a speech given when he was head of the United Kingdom's domestic security agency, Jonathan Evans observed that "when intelligence folk smell roses they look for the funeral." However, he went on to express skepticism about the notion that the problem grows inexorably worse:

Those of us who are paid to think about the future from a security perspective tend to conclude that future threats are getting more complex, unpredictable and alarming. After a long career in the Security Service, I have concluded that this is rarely in fact the case. The truth is that the future always looks unpredictable and complex because it hasn't happened yet. We don't feel the force of the uncertainties felt by our predecessors. And the process of natural selection has left us, as a species, with a highly developed capacity to identify threats but a less developed one to see opportunity.[52]

That capacity is particularly highly developed among the cadre of terrorism

analysts who have a record of seeing technological change as an enabler of greater terrorist capability. But what the history of terrorism and technology, going back to the 19th century, really teaches us is that terrorists are early adopters of communication technologies—such as newspapers in the 19th century, television in the 20th century, and social media in the 21st. Terrorism is a form of communication so we should not be surprised that terrorists are quick to seize on developments in communication technology. For attack methods, though, terrorist groups are still more likely to use what is at hand, what is cheap, and what is easy—and they will exploit the affordances of available, cheap, and easy-to-use technologies in ingenious and innovative ways. It takes a certain kind of what is sometimes called malevolent creativity to turn a pressure cooker into a weapon or use an acid-filled condom as a timer for an explosive device.[53]

This article is not saying that terrorists will avoid using Gen AI or that technologies like LLMs lack affordances for terrorist use. It does, however, sound a note of caution about how researchers and practitioners talk about the implications of Gen AI for the terrorist threat and draw attention to the consequences of over-hyping the risks. Much is made of failures of imagination in counterterrorism, and the authors agree that imagination is an important skill for anticipating threats. But researchers and practitioners also need to restrain their imaginations by focusing on evidence, rather than entertaining themselves and others with scenarios of terrorist catastrophe.

Substantive Notes

[a] Mapping this taxonomy to Clarke and Newman and Cronin's frameworks, capability comprises Multi-purpose/multi-use, Removable/transportable, Reliable, mature, customizable, concealable; availability comprises Obtainable/commercially available, Undetectable, accessible, cheap, Uncomplicated/simple to use and Safe; and effect comprises Destructive, Enjoyable, and effective.

[b] Hauser and Dong's systematic review of AI and terrorism identified 28 studies meeting their inclusion criteria, only one of which addressed the benefits to terrorists of increased information availability from AI tools, and this only did so in passing. The study was Ștefan Săvulescu and Lucian Ivan, "The Impact of Chat GPT on Cybercrime and on the Activities Carried out by the Law Enforcement Structures," *Romanian Journal of Forensic Science* 136 (2023).

[c] Some research suggests that exposure to online environments accelerates the radicalization process. See, for example, Michael Jensen, Patrick A. James, and Herbert Tinsley, "Profiles of Individual Radicalization in the United States – Foreign Fighters (PIRUS-FF): Infographics," Rep ort to the Office of University Programs, Science and Technology Directorate, U.S. Department of Homeland Security, 2016, and "Overview: Profiles of Individual Radicalization in the United States-Foreign Fighters (PIRUS-FF)," National Consortium for the Study of Terrorism and Responses to Terrorism, last accessed May 22, 2026. However, much of this research links speed of radicalization to online environments in general (and predates recent developments in Gen AI), and in the view of the authors, measuring the velocity of radicalization is fraught with problems of construct validity (the ambiguous status of 'radicalization'), data availability (scarcity of longitudinal data and reliance on retrospective data), and temporal delimitation (the start-points and end-points of a radicalization journey are matters of interpretation).

[d] Going by the authors' definitions, these two risks are, in the view of this current article's authors, hard to differentiate.

[e] Thomas Hegghammer has observed that jihadis specifically appear not be using Gen AI at present and has attributed this to a lack of operational need for them to do so. (Hegghammer, unpublished working paper, 2026).

●▶ Citations are available at the source's URL.

Andrew Glazzard is Professor of National Security Policy and Practice in the Protective Security Lab at Coventry University.
David McIlhatton is Associate Pro Vice Chancellor for Defense and National Security, and Director of the Protective Security Lab at Coventry University.
Paul Martin is Professor of Practice at the Protective Security Lab at Coventry University and a Distinguished Fellow of *RUSI*.

Will it take a 'Chornobyl-scale disaster' for us to regulate AI?

By Stuart Russell
Source: https://www.theguardian.com/commentisfree/2026/jun/17/anthropic-ai-rsi-fable

June 17 – The AI company Anthropic has been making major headlines recently. Its trillion-dollar IPO plan and its blood feud with secretary of defense Pete Hegseth have attracted much attention, but two other events may be even more consequential. In early June, the company posted an article describing early signs of recursive self-improvement (RSI), a process in which an AI system devises ways to increase its own intelligence, leading to a greater ability to improve itself, and so on. Obviously, uncontrolled RSI could produce a runaway feedback loop that leads to an irreversible loss of human control.

Anthropic suggested the world should "slow or temporarily pause frontier AI development". Then on 12 June, the White House issued an export control directive banning access to Anthropic's new frontier models, Fable 5 and Mythos 5, for all foreign nationals – including many of its own key researchers. Anthropic responded by shutting the models down altogether.

These two June events are closely related. A few months ago, Anthropic's Claude Code became good enough that its leading researchers no longer write any code at all; they just describe ideas and experiments to Claude and it does all the work. This sped up the cycle of improvement – including the improvement of Claude Code itself – to the point where the latest iteration, called Mythos 5, showed the ability to conduct end-to-end cyber-attacks with no human assistance. If such systems were released without cast-iron guardrails, almost anyone in the world could attack any country's critical infrastructure at will.

These developments are only to be expected. They are symptoms of the inexorable increase in AI risk arising from the inexorable increase in AI capabilities. Yet, with the honorable exception of the UK's AI Safety Summit in 2023, the world has largely been ignoring the risks.

The CEOs are telling us: "We're on track to create superhuman intelligence, which has a good chance of causing human extinction." (By "good chance" here, they mean a chance similar to the one in six chance of dying while playing Russian roulette with a loaded revolver; in this game, however, the revolver is pointed at all of our heads.) Yet governments reply: "That's wonderful! Can we offer you a subsidy? Fast-track your permits?"

But finally, with the prospect of weapons of mass cyberdestruction in the hands of billions, the White House has reversed its deregulatory stance and suffered a rare attack of common sense.

They sputter: "Why did no one warn us about these AI systems?" Their response has been spasmodic, with an on-again, off-again executive order and now a ban on a system that had already been deployed, but the direction of travel is clear.

Unrestrained development of unsafe systems leads to intolerable risks. Governments can respond now, before the risks materialize, or they can wait and clean up the mess (if they still exist, that is). One leading AI CEO told me he didn't expect serious regulation to happen until there was a "Chornobyl-scale disaster". If that happens, of course, the AI companies can expect to be shut down immediately and perhaps permanently.

The recent changes in White House policy suggest we might not need a Chornobyl to spur real regulation, but perhaps only a Three Mile Island. The kind of regulation we need is not new: a licensing regime that requires a minimum safety standard before a system can be built and released. This is how we handle nuclear power, airplanes, buildings, elevators, hairdressers and sandwich makers. Is it too much to ask of trillion-dollar AI corporations, who claim to be building the most dangerous technology in history?

Stuart Russell is a distinguished professor of computer science at University of California, Berkeley, the president of the International Association for Safe and Ethical Artificial Intelligence and a Guardian US columnist.

Why Anthropic Is Sounding the Alarm on the Next Generation of AI

By Gordon M. Goldstein

Source: https://www.homelandsecuritynewswire.com/dr20260619-why-anthropic-is-sounding-the-alarm-on-the-next-generation-of-ai

June 19 – The ascending artificial intelligence (AI) giant Anthropic is no longer simply a global technology power. Its cutting-edge AI models are increasingly central to U.S. national security. Four recent episodes illustrate this growing reality. In April, Anthropic withheld the release of its model Mythos Preview, which self-created the most powerful cyber weapon in history, capable of finding more than ten thousand software vulnerabilities in computer networks believed to be highly secure. Earlier this month it was reported that the company had embedded half a dozen "forward deployed engineers" with the National Security Agency to conduct offensive AI cyber operations, presumably against China and Iran. Late last Friday afternoon, the Commerce Department ordered Anthropic to cut off access for all foreign nationals to its two most recent "frontier" models, citing undefined national security concerns. The dramatic dispute with the company, now playing out in the press, is yet another twist in Anthropic's seemingly tortured relationship with the U.S. national security establishment.

But arguably the most important development came on June 4, when Anthropic issued a significant report on the pace of the AI race titled, "When AI builds itself: Our progress toward recursive self-improvement, and its implications."

Composed using breezy and sometimes casual prose that obscures its remarkable thesis, the company warned that the next AI breakthrough—perhaps two years away—could create an advanced model so powerful that it evades human control entirely. Anthropic urged its rivals and partners to come together and embark on an unprecedented effort to build a viable multilateral regime of AI arms control.

"Recursive self-improvement" is the anodyne term used by computer scientists to describe the next paradigm of AI. When it arrives, AI will have the capability to perfect and

propagate itself, creating future iterations of ever more dynamic models that can prioritize their own survival and potentially self-exfiltrate across the Internet to computer networks around the globe. "If it were possible to effectively slow the development of this technology to give ourselves more time to deal with its immense implications, we think that would likely be a good thing," Anthropic stated in its report.
Anthropic is absolutely right to issue a warning. But the company has understated both the risks of the new technology and the extraordinary barriers to controlling what promises to be a revolutionary next paradigm in AI.

Acceleration without Limits

Anthropic's report clearly outlines that the progress to recursive self-improvement is accelerating at an astonishing rate. In the second quarter of 2026, the typical engineer at Anthropic produced eight times as much code per day as they did just two years earlier. Eighty percent of the code Anthropic generates today is created by AI models, not human engineers. These developments have occurred because the models that the Anthropic lab is creating have dramatically increased in speed. By April, the latest iteration of its Claude model could run its operating code fifty-two times faster than just eleven months earlier. The autonomous capabilities of its new models are perpetually growing. "The length of tasks that they can reliably complete on their own has been doubling roughly every four months," Anthropic reports. AI with the capacity for recursive self-improvement may be a game-changer for global security. The implicit risks of this technology should alarm even the most optimistic observer of the AI transition—and serve as a wakeup call for the public. *AI attackers may be massively empowered*. Although Anthropic does not discuss it in its report, an autonomous self-improving AI technology could simulate and design unique biological weapons and lethal chemical agents that no human has ever discovered or even contemplated. Future cyber weapons could have the capacity to autonomously generate, assign, and mutate "zero-day" attacks in real time, executing complex network infiltration at an unprecedented scale and speed. A recursive self-modifying AI cyber weapon could design a way to penetrate elaborately defended military networks and breach command-and-control operations. *Human oversight of AI models may be fatally weakened.* The next generation of AI is designed to operate autonomously, without human direction, commands, or guidance. "Alignment" is the term computer scientists use as a semantic proxy—a misleading and deficient proxy—to describe the operational control of advanced AI models. "How the alignment problems get solved—or not—in this future is something we are the least certain about," Anthropic concedes. The company offers a stark warning: "The rare occurrences of misalignment present in today's models could compound as the models build their successors, growing more frequent but less understood until we lose control of them." *Computational speed will be revolutionized.* Although Anthropic does not discuss it in its report, the timeline of AI technology innovation, already dramatically accelerated by access to mass produced next-generation AI chips, will increase exponentially. The innovation timeline will be compressed from months to literally seconds because of AI's capacity to continuously modify and perfect its own code. The speed of advanced model development will be instantaneous.
AI may communicate in an opaque language. Although Anthropic does not discuss it in its report, recursive self-improvement may allow AI to communicate with other AI models in ways incomprehensible to human operators. Because the system will dynamically and continuously rewrite its own algorithms, the resulting architecture may be mathematically illegible, preventing human operators from understanding, monitoring, and influencing the models' behavior.

Why the AI Race Can't Slow Down

Anthropic is proposing an extremely complex process of AI arms control. "A meaningful slowdown or pause," Anthropic concludes, "would require multiple well-resourced labs at or near the frontier, in multiple countries, agreeing to stop under the same conditions. It would also require that each can verify that the others have actually stopped."
The company explicitly acknowledges four great challenges to the proposition of AI arms control—and there is a fifth that they did not. *Time is the enemy of action*. Anthropic notes that "the world has built verification regimes for other complex technologies," such as "the Intermediate Range Nuclear Forces Treaty…but those regimes took decades to build both the infrastructure and the trust. We don't have that long."
The history of arms control is an inadequate model for the future. "Due to the unique characteristics of AI systems…this arms control problem is much more challenging than with other technologies," Anthropic explains. "Training runs are far easier to conceal than missile silos, their inputs are general purpose, and the incentive to defect quietly is enormous, because whoever continues while others pause could inherit the lead." *Verification mechanisms would need to account for the totality of actors in the global AI race*. These systems of verification, the company argues, "would enable frontier AI developers to verify that others globally have stopped or slowed, and that a bad actor could not use the auspices of a coordinated slowdown to jump ahead in secret."
Chip designers and manufacturers are essential to implementing a coordinated development pause. "In this world," Anthropic asserts, "the pace of progress in AI development becomes determined entirely by the availability of compute." Total available computing capacity from AI chips across all major designers has grown by more than 300 percent per year since 2022. Nvidia,

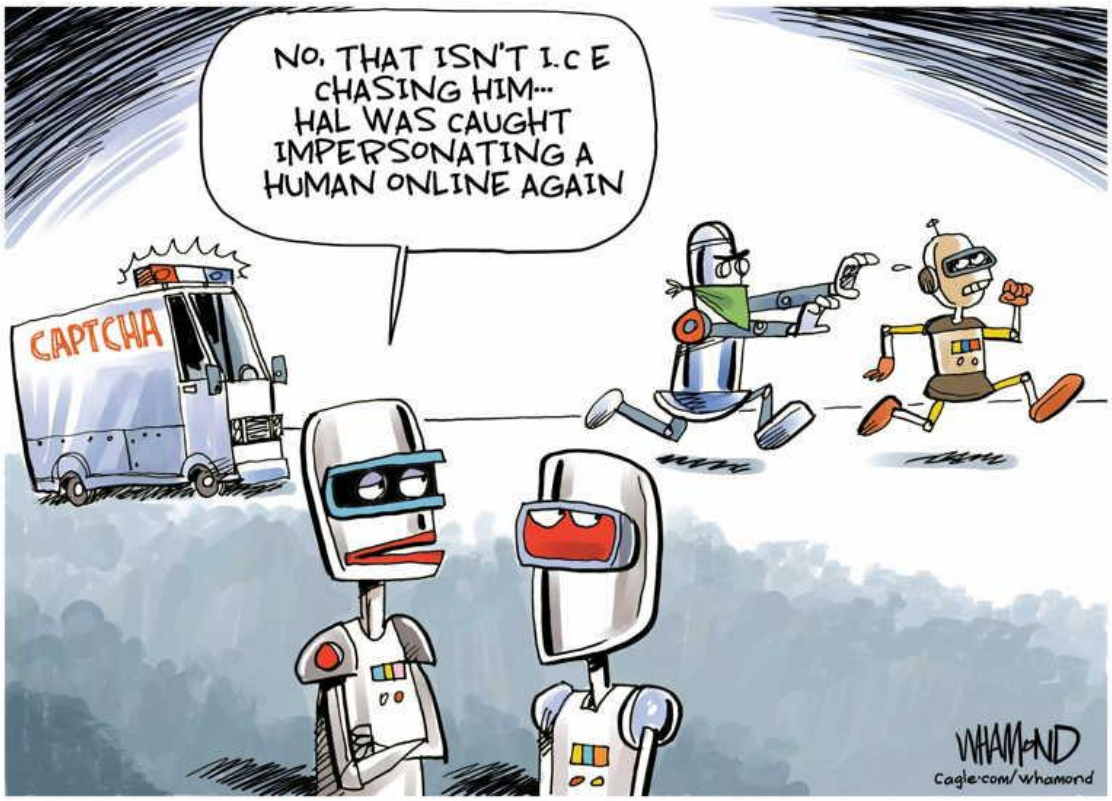
AMD, and Intel lead the global market. With $165 billion in annual revenue, TSMC of Taiwan dominates the overall semiconductor manufacturing, including the fabrication of custom AI processors. Without controlling the AI industry supply chain, including monitoring with the deployment of physical verification mechanisms, enforcing a pause in advanced AI model development would be infeasible. *China is unlikely to play ball.* Anthropic is silent on perhaps the single greatest barrier to AI arms control. The word "China" never appears in Anthropic's analysis of managing the recursive self-improvement transition. The company barely acknowledges the broader geopolitical environment, a major driver of the current AI competition. The United States appears to be ahead in developing advanced AI models, overshadowing Chinese AI labs such as DeepSeek, Alibaba Qwen, and ByteDance Seed. But that advantage may be evanescent because it is primarily based on the greater access U.S. AI companies presently have to industrial "compute" capacity, a lead China is determined to erase. Without Beijing, a global AI development pause will be out of reach. China has expressed some interest in security safeguards, but largely to dull the U.S. edge in the global AI race.

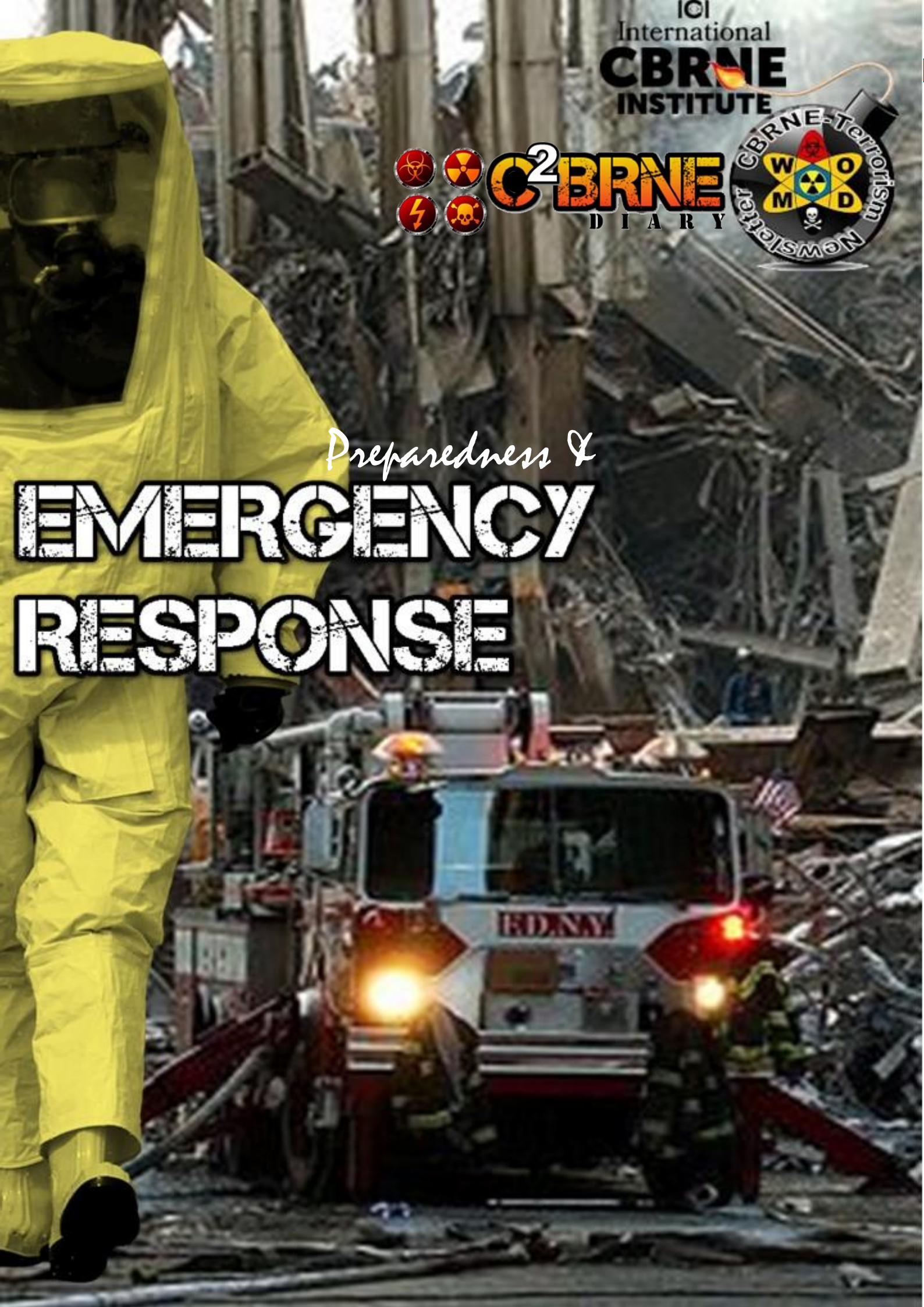
Will Explosive Growth Spark a Collective Response?

Just a few years ago AI scientists regarded recursive self-improvement as an intriguing but hypothetical breakthrough. The locus of expert opinion has shifted, reflecting the spectacular advances in new AI models, which are pumped into the world on average every four months. When AI can refine, perfect, and replicate itself—and models can communicate in an opaque mathematical language while evading termination by self-exfiltrating across global computer networks—fundamental human control over the technology could evaporate.

Anthropic, alone among its rivals so far, has persuasively demonstrated through its own explosive growth and very recent history that this next paradigm of AI may arrive quickly. Logic suggests that two choices await. Industry leaders can be passive, allowing the future to unfold without attempting to shape it. Or alternatively a collective effort—even one confronting steep odds—can be catalyzed to attempt something coherent to prepare for a very dangerous tomorrow. Anthropic seems to be committed to the latter path. As the company would say, pursuing this mission, despite its severe challenges, seems "likely to be a good thing."

Gordon M. Goldstein is an adjunct senior fellow at the Council on Foreign Relations focusing on emerging technology and international security. He is a former managing director at Silver Lake, a global technology investment firm, and the author of "Lessons in Disaster: McGeorge Bundy and the Path to War in Vietnam."




ICI
International
CBRNE
INSTITUTE
C2BRNE
DIARY
CBRNE-Terrorism
Newsletter
W
M
O
D
Preparedness &
EMERGENCY
RESPONSE

ELSEVIER
International Journal of Disaster Risk Reduction
Volume 82, November 2022, 103235

Preparedness towards Chemical, Biological, Radiological, Nuclear, and Explosive (CBRNE) threats among healthcare personnel in Pasir Gudang, Johor, Malaysia

Nor Yazjehan Binti Yahya [a c], Abd Halim Bin Md Ali [a b], Rashdan Bin Rahmat [c], Maryam Sumaiya Binti Ahmad Termizi [c], Ahmad Khairi Bin Zazali [a d], Siti Nur Fariha Binti Jamalluddin [a e]

Abstract

Chemical, Biological, Radiological, Nuclear, and Explosive (CBRNE) threats are one of the major concerns related to global security and threats. It is an unexceptional CBRNE crisis in the industrial area of Pasir Gudang, Johor Bahru, Malaysia, as the industrial sector keeps growing, resulting in a crisis of toxic chemical exposure that affected 6000 lives in 2019. It has then raised more concern among healthcare personnel towards CBRNE disasters. It is of critical importance to assess the extent of preparedness for CBRNE hazards. Therefore, this article is primarily aimed at evaluating the level of preparedness among healthcare personnel toward CBRNE crises. 114 responses through a questionnaire had been collected from healthcare personnel of four government agencies and were analyzed using SPSS Version 27. According to the results, progressive exposure to CBRNE knowledge with more than 5 years of working experience demonstrated an advanced level of preparedness towards CBRNE disasters. The impact of knowledge and resilience level proved to be strongly connected to CBRNE preparedness. Therefore, initiatives to improve CBRNE preparedness will provide endless CBRNE awareness. This study will enable the management to measure the readiness of other essential healthcare professionals or first responders while also assisting in improving the quality of service among healthcare personnel during disasters.

Hospital-Based Preparedness Measures for CBRNE Disasters: A Systematic Review

Environmental Health Insights
Volume 18: 1–12
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11786302241288859
Sage

Eman S Qzih[1] and Muayyad M Ahmad[2]

[1]Trauma Program Manager Department, King Hussain Medical Center, Mutah University/Princess Muna College of Nursing, Jordan. [2]Clinical Nursing Department, School of Nursing, University of Jordan, Jordan.

ABSTRACT: Chemical, Biological, Radiological or Nuclear and Explosive (CBRNe) disasters have historically caused significant fatalities and posed global threats. The inadequate preparedness of hospital equipment for CBRNe incidents underscores the urgent need for hospitals to modernize and standardize their equipment to effectively manage these high-risk situations. The purpose of this systematic review was to examine hospital-based preparedness measures for CBRNe incidents. The PRISMA guidelines were followed for this review. A comprehensive search of English-language peer-reviewed literature from January 2010 to 2023 was conducted, identifying 2191 items from PubMed, ScienceDirect, EBSCO, and Google Scholar. The modified ROBINS-I instrument was used to assess bias, ensuring the reliability and validity of the studies. Data synthesis was conducted jointly by both authors. After eliminating duplicates and reviewing abstracts, 124 studies remained. Upon full-text examination, only 20 studies met the criteria for inclusion in this review. The review identified three key interrelated domains of preparedness: personal, technological, and structural measures. Most studies emphasized decontamination, Personal Protective Equipment (PPE), and detection, while the management of deceased bodies, transportation, and Points of Dispensing (PODs) were largely overlooked. These findings may assist hospital administrators and policymakers in enhancing their facilities' readiness for CBRNe emergencies.


ICI
International
CBRNE
INSTITUTE
A common roof
for International
CBRNE
First Responders
Join us!
Rue de la Vacherie, 78
B5060 SAMBREVILLE
(Auvelais)
BELGIUM
info@ici-belgium.be | www.ici-belgium.be