



Unintended Nuclear War

Karl-Hans Bläsius¹ · Jörg Siekmann²

Received: 23 January 2021 / Accepted: 25 January 2021
© The Author(s) 2021

Abstract

We want to use the 22nd of January 2021 as an opportunity to honor the “*Treaty on the Prohibition of Nuclear Weapons*”, TPNW, by this article, as the treaty will enter into force on this day.

By resolution 71/258, the General Assembly of the United Nations decided to convene a UN conference in 2017 to negotiate a legally binding instrument to prohibit nuclear weapons, leading towards their total elimination. And on July 7, 2017, 122 states voted in favour to this request (with one vote against and one abstention). The treaty prohibits the production, possession and use of nuclear weapons. On October 24, 2020 the 50th state ratified the treaty and now the TPNW will enter into force on January 22nd, 2021. However, since the nuclear powers have boycotted the conference and have not signed, the treaty is currently of political and symbolic value only. The arms race including atomic weapons still continues.

In 2017, the International Campaign to Abolish Nuclear Weapons (ICAN) received the Nobel Peace Prize for „*its work to draw attention to the catastrophic humanitarian consequences of any use of nuclear weapons and for its groundbreaking efforts to achieve a treaty-based prohibition of such weapons*” and called on nuclear powers to engage in serious disarmament negotiations.

This article deals with the problem whether decisions based on AI techniques in early warning systems can be useful and if these systems are then safer with regard to possible false alarms.

1 AI in Military early Warning and Decision-Support Systems

Securing the nuclear second-strike capability is the basis of the deterrence strategy that has so far prevented any potential attacker from launching a nuclear attack: “Whoever shoots first dies second”.

To be able to react when the second-strike capability is threatened, the nuclear powers have developed and installed sophisticated early warning and decision support systems with the aim of an in-time recognition of an attack in order to activate their own nuclear launchers before the destructive impact.

Such a strategy is called “launch-on-warning”.

These systems detect an attack on the basis of sensor data, such as radar, light or ultra sound, and rely on a computer-based interpretation of this data.

The structure and functioning of early warning systems became known primarily from the U.S. through various investigative reports and publications in the 1980s. An early command centre of the U.S. was NORAD (North American Aerospace Defence Command), that began operations as early as 1957 and by 1983 it contained about ten million lines of code. Since, these systems have been largely modernized to take account of the new arsenal of weapons and they are now substantially more complex.

Like any large combined software and hardware system, these enormous installations are susceptible to errors and this could lead to an accidental nuclear war.

Although the time required for the decision between a reported attack and the launch of missiles for the counterattack has fallen to a few minutes in recent years, the final decision—not least because of the susceptibility to errors of such systems—is still left to the commander in chief, i.e. the president of the United States.

✉ Jörg Siekmann
Joerg.Siekmann@dfki.de
Karl-Hans Bläsius
blaesius@hochschule-trier.de

¹ Fachbereich Informatik, Hochschule Trier, Schneidershof, 54293 Trier, Germany

² Saarland Informatics Campus, German Research Centre for Artificial Intelligence, Saarland University, Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany

Due to the increasing number of different types of sensors, satellites in space and monitoring systems, the available data and information increases disproportionately in a concrete situation. Thus, for the classification of sensor data and the evaluation of an alarm situation, more and more Artificial Intelligence Systems are installed for subtasks that prepare the final human decision.

The end of the INF Treaty (Intermediate Range Nuclear Forces) has led to a new arms race, also with hypersonic missiles, which shorten this time span even further. Therefore politicians and military personnel have the expectation that AI systems will be capable of making better decisions than military personal within this very short time for considerations for the counterattack.

1.1 False Alarms and the Political Context

Because of the uncertainty of the data, the military personal base their decisions also on contextual knowledge about the political situation and the assessment of the opponent. For example, the operating crew of the American early warning system decided that the alarm caused by several missiles heading for the USA on the 5th of October 1960 must be false and no retaliation made sense, because the Soviet head of state was on a state visit to New York at the time.

That is, even in a machine-based decision, contextual knowledge of the political world situation must be included in the evaluation of alarms, and this knowledge is also *uncertain, vague and incomplete*. The result of the analysis by an AI system is therefore correct only within the limits of a statistical probability.

1.2 Two Examples Pro and Contra a machine-based Decision

1.2.1 Example 1

In January 2020 the USA killed the Iranian General Soleimani with a drone attack and in retaliation Iran attacked American positions in Iraq a few days later. Shortly afterwards a Ukrainian airliner was accidentally shot down in Iran, since the operating crew came to the conclusion that the flying object could be an attacking cruise missile.

In this situation a computer might have made a better decision, because the pure facts, such as the size of the radar signal, would probably have been better interpreted and then judged against a cruise missile attack. In addition, a machine could have taken more information into account, such as civil flight plans, even within the short time available.

The wrong decision came about mainly because the crew had expected war and an attack by the US and obviously overestimated the political context.

1.2.2 Example 2

A satellite of the Russian early-warning system reports on September 26, 1983 five incoming Intercontinental Ballistic Missiles (ICBM). Since the correct function of the satellite was checked and established, the Russian officer on duty, Stanislav Petrov, should have sent this information on to his superior and to the Head of State of the Soviet Union, J.W. Andropow according to his regulations. However, he considered an attack by the Americans with only five missiles unlikely and decided, despite the available data, that it was probably a false alarm, thus preventing a catastrophe with nuclear strike and counterattack. The incident occurred during an unstable political situation: the modernisation with medium-range missiles was pending and a few weeks before the Soviets had accidentally shot down a Korean passenger plane over international waters. An AI-system would have assessed the attack more likely as real and would have initiated the counterattack based on these facts.

Petrov however had emotionally hoped for a false alarm and he did not want to be responsible for the millions of deaths of people and decided against his orders.

1.3 New Technologies and Tests

With new technologies there are always fears, justified or not. For example, fast train journeys were thought to be dangerous in the nineteenth century and currently the danger of autonomous vehicles is debated in the media. However, safety has usually increased with new technologies, and many experts believe that autonomous driving significantly reduces the risk of accidents. But it is the repeated cycle of “trial and error” that is important when developing innovations into reliable technologies. The misclassification of a truck tarpaulin as a “free road” in the earlier use of the Tesla, which led to a serious accident, has been eliminated and has resulted in better and more robust classification methods. In such cases the losses are within limits, whereas an all-out nuclear war cannot be confined to a local area of the world, since in combination with the subsequent nuclear winter would mean billions of deaths and possibly eliminate human life on earth. Moreover, the early warning systems can only be tested using simulation software, since obviously they cannot be tried out in reality.

1.4 The Atomic Doomsday Clock

In 1947 the *Atomic Doomsday Clock* was established to alert to the danger of an impending nuclear war. The clock is reset once a year by nuclear scientists and Nobel laureates and the reasons for the setting are published in the Bulletin of the Atomic Scientists. The first setting in 1947 was at seven minutes to twelve, it decreased to three minutes to twelve in

1984 because of the accelerated arms race at the time and in 1991—because of the disarmament agreement between Mikhail Gorbachev and Ronald Reagan—it was set back to seventeen minutes. Because of the recent worsening political and military situation, it was set in 2020 to an all-time low of 100 s. The 2021 reset announcement will be shown LIVE on January the 27th.

The reasons for this dramatic warning are that the nuclear powers currently modernize and even expand their nuclear arms, that most of the important treaties on limitation and mutual trust have been invalidated, and that the climate change and the resulting deterioration of living conditions may lead to conflicts.

2 Summary

While AI systems can be quite useful for political assessment and military reconnaissance, the final decision for a counterattack by a computer is not acceptable. The survival of humanity as a whole should never depend on the decision of a single person or a machine. Due to the uncertain and incomplete data, machines cannot with one hundred percent reliability evaluate incoming alarm messages and once the false counterstrike has been initiated there is no chance for correction.

With several colleagues we have set up and maintain a website.

“Atomkrieg aus Versehen”.

<http://www.unintended-nuclear-war.eu/>

3 Further reading

This article is based on our more detailed paper:

K.H. Bläsius, J. Siekmann; Computergestützte Frühwarn- und Entscheidungssysteme, 2020, which also contains all references.

K.H. Bläsius, J. Siekmann; Early Warning and Military Decision Support Systems, both in: <https://www.fwes.info/fwes-19-3.pdf> where you can switch between English and German.

K.H. Bläsius, J. Siekmann: Computergestützte Frühwarn- und Entscheidungssysteme, Informatik-Spektrum, Band 10, Heft 1, 24-39, 1987

K.H. Bläsius, J. Siekmann: Frühwarnsysteme und Cyberangriffe – gefährliche Wechselwirkungen möglich, Behördenspiegel August 2019, S. 44,

in: https://issuu.com/behoerden_spiegel/docs/2019_august

Bläsius, Siekmann: Künstliche Intelligenz in Frühwarnsystemen

in: Newsletter des Behördenspiegels Verteidigung, Streitkräfte, Wehrtechnik, am 19.12.2019, https://www.fwes.info/nl_defence_253_02.jpg

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.