

The Nuances of Responsibility of Artificial Intelligence for Irresponsible Space Activity

Anna Hurova

Ph.D. in Law, Institute of State and Law of National Academy of Science of Ukraine
(Kyiv, Ukraine)

E-mail: a.m.hurova@gmail.com

<https://orcid.org/0000-0003-4134-761X>

Hurova, Anna (2024) The Nuances of Responsibility of Artificial Intelligence for Irresponsible Space Activity. *Philosophy and Cosmology*, Volume 32, 27-33. <https://doi.org/10.29202/phil-cosm/32/2>

The article examines various aspects of responsibility for irresponsible space activities related to the use of artificial intelligence. It is stated that the existing legal instruments for regulating AI do not allow us to fully make space law effective and compatible with this technology, as well as the risks that this technology can bring. An analysis of the Convention on International Liability for Damage Caused by Space Objects (1971) is carried out, an interim conclusion is made that its norms apply to spacecraft that used AI technology. It is found that the “problem of many hands” inherent in space activity leads to difficulties in tracing the causal relationship between the action of a particular person in the chain of creation, maintenance, use of AI, etc. and the damage caused by an autonomous spacecraft. The article reveals the key contradictions of the current norms of AI responsibility for irresponsible space activities and suggests ways to solve them by developing and adopting comprehensive rules of conduct in space activities.

Keywords: responsibility, artificial intelligence, space activity, intent, guilt, inaction, harm, Liability Convention.

Received: 04 December 2023 / Accepted: 02 January 2024 / Published: 5 February 2024

Introduction

For objective circumstances, space activities have always been considered and are considered extremely dangerous. Therefore, it is clear that the issues of causing damage, its compensation and procedures for bringing to justice for improper or illegal space activities are relevant.

The concept of responsible behavior has a preventive accent, aims to mitigate misperception and miscalculations, and is oriented on the political mechanisms of space governance. The concept of liability deals with consequences (damage, unlawful act, failure of the objective standard of care).

The AI makes difference less visible according to characteristics of complexity, vulnerability, openness, and opacity. In the vulnerable digital system, the autonomy of machine learning, and “many hands engaged in operation” characteristics could be tended to regulation based on the responsibility concept.

According to the Liability Convention, liability depends on where the damage occurred. A launching state is absolutely liable for damage caused on the Earth’s surface or to aircraft in flight while, damage caused in outer space is based on fault (Convention, 1971). However, in both cases the mechanism is the same. The launching state bears liability for damage caused by its space object. Therefore, we could answer the question about coverage under the Liability Convention by means of analyzing the concepts “damage,” “space object” and “launching state” within the light of threat of AI usage.

Indeed, any unregulated activity carries risks, and even more so if such activities are related to outer space and artificial intelligence. Let’s briefly define these risks:

1. Commercial space activities, which account for about 70% of space-related activities (Inter-Agency, 2018), have increased ubiquitous impact on Earth and the benefits that space products and technologies provide in solving the problems and needs of humankind (Soroka, 2020; Gal et al., 2020).
2. New actors, together with emerging new technologies such as AI, develop new global business models driven by demand, such as satellite constellations, tourism, asteroid and lunar mining, in-situ resource utilization (ISRU), 5G, in-orbit servicing (IoS), 3D printing of satellite parts (e.g., solar panels, etc.), and commercial space station (Gal et al., 2020).
3. The use of AI is realized without a global regulator and global rules (Soroka et al., 2022).

All of the above determines the relevance and prospects of studying the legal aspects and risks of using IS in the space sector.

Theoretical prerequisites for the application of responsibility in the space sphere

The international debate over the establishment of a comprehensive regime of responsibility for inappropriate activities in outer space was a direct consequence of the space race between the United States and the USSR. Unfortunately, disagreements between the two great space nations and Cold War-era enemies prevented any agreement to establish an international regime of responsibility in space.

The beginning of the global regulation of space activities was the resolution 1963 (XVIII) of the UN General Assembly (Declaration, 1963), which proclaimed the legal principles of states’ activities in the exploration and use of outer space, it was also a declaration that included the principle of responsibility (paragraph 8). It was the forerunner of the first space treaty – Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies (Treaty, 1966), which was signed almost a decade after the launch of the first satellite in 1957 and the Convention on International Liability for Damage Caused by Space Objects (Convention, 1971). The latest document was the result of a consensus among countries that took place after the adoption of the 1963 resolution. The Liability Convention entered into force in September 1972 as *lex specialis* to Article VII of the Outer Space Treaty for those states that are states parties to the convention (Schmalenbach, 2023).

Taking into account the negotiation process in which the Outer Space Treaty and the Liability Convention were adopted, it is necessary to note their organic unity. But taking into account the current realities of space activities, they require certain clarifications.

Article 1 of the Liability Convention (Convention, 1971) prescribes that the term “damage” means “loss of life, personal injury or other impairment of health or loss of or damage to property” of States or persons, natural or juridical, or property of international intergovernmental organizations in turn, the Treaty on Outer Space (Treaty, 1966) also specifies the obligation of states to bear responsibility for damage caused by their space objects. A comparison of the adverse effects highlighted by both international acts lets us make the conclusion that they protect the same values, i.e., life, health, and property. However, neither the Tallinn Manual (The Tallinn, 2013), which is a piece of “soft law,” nor any source of international law have the mechanisms of imposing liability on as many States as the Liability Convention. Therefore, we can conclude that the consequences of irresponsible behavior of AI are covered by the term “damage” in the Liability Convention.

According to the Liability Convention, the “injurer” and “target” are the space objects. So, to answer the question about the applicability of the Liability Convention for the cases of -irresponsible behavior of AI it is necessary to understand if software is covered by the term “space object.” At the forefront, we need to take into consideration Article 31 of the Vienna Convention on the Law of Treaties, which sets out that a “treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its object and purpose” (Vienna, 1969). The term “space object” in the Liability Convention is not clear as to what is and what isn’t covered by this term.

The development of space activity and its consequences call for applying the norms of the Liability Convention through evaluative interpretation to the new cases that did not exist at the time of its adoption. For instance, there is a very common position that every piece of space debris is considered as a space object according to the Liability Convention (Kerrest, 2001). This includes functional or nonfunctional debris, including component parts of space objects as well as its launch vehicle and other parts. There is also a position that it is necessary to include into the legal category of “space objects” the objects having a movable or immovable character, directly constructed or manufactured in outer space during deep space exploration or space mining activity (Skopowska, 2017).

It should be noted that the boundary between evaluative interpretation and creating the new norm through interpretation is thin. The authors of the Cologne Commentary on Space Law (Cologne, 2013) argue that the term “space object” does not leave room for damage resulting from non-material origin. They argue that an interpretation based on the ordinary meaning of the relevant provision supports the view that the Liability Convention rather envisages the damages caused by impact, than the damage inflicted through activity, where satellite signals are related by complementary technology and equipment (Cologne, 2013: 12). That is concerning electromagnetic interference or interruption in satellite radio signals caused by means located beyond the satellites. On the one side, AI also has a non-material origin, but on the other side, what would have been left from the modern satellite if the software were not there?

According to the ordinary meaning the terms of the Liability Convention in their context and in light of its object and purpose, a space object is acknowledged as a man-made object that is launched or intended to be launched into outer space that needs to be protected from another space object. Since the adoption of the Liability Convention, space objects and mechanisms of their control have greatly changed, but regardless of this transformation of

technologies, the term “space object” has included the component parts of a space object, and nowadays software is an integral part.

Artificial intelligence adds its share of risks to the concept of responsibility for inappropriate behavior in space activities to the above problems. Since the irresponsible behavior of artificial intelligence can lead to the seizure of control over a space object and without causing direct damage. But, in the future, this behavior of artificial intelligence to seize control of a space object can lead to a deliberate collision with another space object. In this case, we cannot deny the application of the Liability Convention, but the problem of proving guilt in consequences in this aspect has become a very significant issue.

Content and types of liability

James Goudkamp conducted a very meaningful analysis of liability in tort law in his work “Rethinking fault liability and strict liability in the law of torts” (Goudkamp, 2023). He believes that most torts are a mixture of two forms of liability (intentional and negligent), as they usually have several elements of behavior, and only a few of these elements are combined with a claim of guilt. He also defined the following:

1. Liability is shifting loss from the person who originally suffered the damage (the victim) to a person who caused the damage (the tortfeasor).
2. Fault liability – in the base is wiliness of tortfeasor in the form of intent or negligence.
3. Vicarious liability – based on the generally ‘unfit’ nature in terms of personality or skills, or on the fact that someone failed to meet an objective standard of care.
4. Strict liability – do not require any kind of defect or mal-performance or wiliness of tortfeasor but based on causation (Goudkamp, 2023).

The challenges of AI in outer space – are determine and enforce a “meaningful human control” (MHR) over the entire technological cycle and system functioning.

The strict liability rules are often complemented with the extension of current tortious liability doctrines to tackle compensation gaps in accidents provoked by AI systems. Doctrines of tort law include duties of care and theories of agency, as well as procedural standards on burdens of proof, presumptions, etc.

The HLEG’s report on “liability for AI and other emerging technologies,” one of these ways regards the implementation of compensation schemes: “compensation funds may be used to protect” (Liability, 2019).

Scientists constate difficulties in attributing liability on the fault-based regime as well as exoneration based on the “gross-negligence” that is generally “associated to an action or omission part of a human activity under the Liability Convention. It is necessary to pay attention to the concept of a source of increased danger, which provides for liability without fault for the functioning of things that a person does not control, but which he or she has set in motion and must supervise.

Another author, Christian Wendehorst (Wendehorst, 2022), analyzing the various potential risks associated with artificial intelligence, examines liability issues in two main areas: security risks and fundamental rights risks. Based on this, the author claims that AI challenges existing modes of responsibility, and we agree with this.

Sources of increased danger in space law

It is necessary to pay attention to the concept of a source of increased danger, which implies responsibility without fault for the functioning of things that a person does not control, but which he has set in motion and should control.

Thus, the Montreal Declaration of 2018 defines the following principle of responsibility: “Only human beings can be held responsible for decisions stemming from recommendations made by AI, and the actions that proceed therefrom (§ 1); When damage or harm has been inflicted by an AIS, and the AIS is proven to be reliable and to have been used as intended, it is not reasonable to place blame on the people involved in its development or use (§5)” (Montréal, 2018).

That is, from the above norms, AI acts as a source of increased danger, for which procedures and rules for regulating liability in the event of harm have long existed at the national level.

As an example of the operation of this norm at the national level, we can cite Article 1187 of the Civil Code of Ukraine, which lists the circumstances that determine the specifics of liability for damage caused by a source of increased danger, they are: 1) the intent of the victim excludes the responsibility of the causer of harm; 2) gross negligence of the victim divides responsibility between the causer and the victim; 3) illegal possession of the source of increased danger by a third party (Civil, 2003).

Multiplication of the source of irresponsible behavior of AI:

- a) Multiple Launching States;
- b) Lack of requirement for authorization that proves the irresponsible behavior of AI components;
- c) Threats could have the origin of malfunctioning electronics; no resilient to cyberthreat and malfunctioning of digital ecosystems (supply chain confidence for protection of the backdoors); opacity and autonomy of the AI (“black box effect” – there is no explanation why an AI behaved in a particular manner in a given situation).

Thus, the reason for changes made in program settings often arises outside of the space objects, so the launching State could not be liable for such damage because it probably has no fault for the malfunctioning of the space object.

In this case, the Liability Convention is unlikely to be applied, but according to international space law, the States are not only liable for damage caused by their space objects or space objects of entities under their jurisdiction but also bear national responsibility for national activity in outer space.

Conclusions

Existing solidary tort liability is not the best solution, since this type of liability does not comply with either Article XVIII or any other provision of the Liability Convention, which substantive law is used to resolve the issue on the merits and to resolve the issue of compensation for damage. Thus, it is necessary to create universal global rules for the field and technologies that are constantly changing and characterized by spontaneity, while the developed and adopted legal norms should ensure a regulatory balance in the interaction of robotic and virtual technologies with people in space missions.

 References

- Civil Code of Ukraine* (2003) Law of Ukraine dated January 16 No. 435-IV. Verkhovna Rada of Ukraine official website. Available online: <https://zakon.rada.gov.ua/laws/show/435-15#Text>
- Cologne Commentary on Space Law Vol. 2* (2013) German Edition by Stephan Hobe, Bernhard Schmidt-Tedd, Kai-Uwe Schrogl. Heymanns Verlag GmbH.
- Convention on International Liability for Damage Caused by Space Objects* (1971) RES 2777 (XXVI) General Assembly 26th session. United Nations Office for Outer Space Affairs. Available online: <https://www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/introliability-convention.html>
- Declaration of Legal Principles Governing the Activities of States in the Exploration and Use of Outer Space* (1963) RES 1962 (XVIII) General Assembly 18th session. United Nations Office for Outer Space Affairs. Available online: https://www.unoosa.org/oosa/oosadoc/data/resolutions/1963/general_assembly_18th_session/res_1962_xviii.html
- Gal, George Anthony, Cristiana Santos, Lucien Rapp, Réka Markovich, and Leendert van der Torre (2020) *Artificial Intelligence in Space*. Available online: https://www.researchgate.net/publication/342377395_Artificial_intelligence_in_space
- Goudkamp, James (2023) From Strict Liability to Fault Liability in the Law of Torts. *Law Quarterly Review*. Available online: <https://search.informit.org/doi/10.3316/agispt.20230412086202>
- Inter-Agency Meeting on Outer Space Activities: 2018, Thirty-eighth session (2018) UNOOSA. Available online: <https://www.unoosa.org/oosa/en/ourwork/un-space/iam/38thsession.html>
- Kerrest, Armel (2001) Space debris, remarks on current legal issues. *3rd European Conference on Space Debris*. Available online: <https://conference.sdo.esoc.esa.int/proceedings/sdc3/paper/3/SDC3-paper3.pdf>
- Liability for artificial intelligence and other emerging digital technologies* (2019) Directorate-General for Justice and Consumers (European Commission). Office of the EU. Available online: <https://op.europa.eu/en/publication-detail/-/publication/1c5e30be-1197-11ea-8c1f-01aa75ed71a1/language-en>
- Malysheva, Natalia (2010) *Essays on space law*. Kyiv, Alerta.
- Montréal Declaration for a Responsible Development of Artificial Intelligence* (2018) Available online: https://monoskop.org/images/d/d2/Montreal_Declaration_for_a_Responsible_Development_of_Artificial_Intelligence_2018.pdf
- Schmalenbach, Kirsten (2023). Convention on International Liability for Damage Caused by Space Objects. In: *Gailhofer, P., Krebs, D., Proelss, A., Schmalenbach, K., Verheyen, R. (eds) Corporate Liability for Transboundary Environmental Harm*. Springer, Cham. https://doi.org/10.1007/978-3-031-13264-3_11
- Skopowska, Laura Rut (2017) *Is an object built in the outer space a “space object” under the Liability Convention?* (LLM thesis, University of Luxembourg) (on file with author).
- Soroka, Larysa (2020) *Administrative and Legal Mechanism for the Implementation of the Space Doctrine of Ukraine: Theory and Practice*. Private individual Chalchynska.
- Soroka, Larysa, Anna Danylenko and Maksym Sokiran (2022) Legal Issues and Risks of the Artificial Intelligence Use in Space Activity. *Philosophy and Cosmology*, Volume 28, 118-135. <https://doi.org/10.29202/phil-cosm/28/10>
- The Tallinn Manual (2013) Ccdcoe.org. Available online: <https://ccdcoe.org/research/tallinn-manual/>
-

- Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies* (1966) RES 2222 (XXI) General Assembly 21st session. United Nations Office for Outer Space Affairs. Available online: https://www.unoosa.org/oosa/oosadoc/data/resolutions/1966/general_assembly_21st_session/res_2222_xxi.html
- Vienna Convention on the law of treaties* (1969) Concluded at Vienna on 23 May 1969. United Nations. Available online: <https://treaties.un.org/doc/Publication/UNTS/Volume%201155/volume-1155-I-18232-English.pdf>
- Wendehorst, Christiane (2022) In *Liability for Artificial Intelligence*. The Cambridge Handbook of Responsible Artificial Intelligence Interdisciplinary Perspectives, 187-209. Cambridge University Press. <https://doi.org/10.1017/9781009207898.016>