IAC-22- D5.4 (x70406)

"Understanding Space Vulnerabilities: Developing Technical and Legal Frameworks for AI and Cybersecurity in Space"

Mr. Antonio Carlo[a]*, Dr. Nebile Pelin Mantı[b], Mr. Bintang A.S.W.A.M[c], Ms. Francesca Casamassima[d], Mr. Nicolò Boschetti[e], Dr. Paola Breda[f], Tobias Rahloff[g]

[a] *TalTech - Tallinn University of Technology, Tallinn, Estonia, ancarl@taltech.ee*
[b] *PIL Dept., Istanbul University Faculty of Law, Istanbul, Türkiye, np_manti@yahoo.com*
[c] *Space Generation Advisory Council, Depok town, West Java, Indonesia, baswam95@gmail.com*
[d] *IT & Security Governance, Dedalo GRC Advisory, Rome, Italy, fcasamassima8@gmail.com*
[e] *Johns Hopkins University, Baltimore, MD, USA, nbosche1@jhu.edu*
[f] *HyImpulse Technologies GmbH, Neuenstadt am Kocher, Germany, breda@hyimpulse.de*
[g] *DHBW, Hamburg, Germany, contact@tobiasrahloff.com*
* Corresponding Author

Abstract

Over the past decades, industries and governments have progressively been relying upon space data-centric and data-dependent systems. Consequently, this led to the emergence of malicious activities, also known as cyber-threats, targeting such systems. To counter these threats, new technologies such as Artificial Intelligence (AI) have been implemented and deployed. Today, AI is highly capable of delivering fast, precise and reliable command-and-control decision-making, as well as providing reliable vulnerability analysis using well-proven cutting-edge techniques. Nonetheless, this might not yet be the case when used for space applications. AI can also play a transformative and important role in the future of space cybersecurity, and it poses questions on what to expect in the near-term future.

Challenges and opportunities, deriving from the adoption of AI-based solutions to achieve cybersecurity and later cyber defence objectives in both civil and military operations, bring a new framework and new ethical requirements. In fact, most of these technologies are not designed to be used or to overcome challenges in space. Because of the highly contested and congested environment, as well as the highly interdisciplinary nature of threats to AI and machine learning technologies, including cybersecurity issues, a solid and open understanding of the technology itself is required, as well as an understanding of its multidimensional uses and approaches. This includes the definition of legal and technical frameworks, ethical dimensions and other concerns such as mission safety, national security, and technology development for future uses.

The continuous endeavours to create a framework and regulate interdependent uses of combined technologies such as AI and cybersecurity to counter "new" threats require the research and development of "living concepts" to determine in advance the vulnerabilities of the networks and the AI.

This paper will develop a cybersecurity risk and vulnerability taxonomy for the future applications of AI in the space security field. Moreover, it will assess to what extent a network digital twins' simulation can still protect networks against relentless cyber-attacks in space against users and ground segments. These concepts will be applied to the case study of Earth Observation (EO) operations, which allows for conclusions to be drawn based on the business impact (reputational, environmental, and social) of a cyber malicious activity. Since AI technologies are developing daily, a regulatory framework will be proposed using ethical and technical approaches for the technology and its use in space.

Keywords: Cybersecurity, Artificial Intelligence, Earth Observation, Cyber Risk, Emerging Disruptive Technologies

Acronyms/Abbreviations

AI	Artificial Intelligence
DDoS	Denial-of-service attack
DL	Deep Learning
DT	Digital Twin
EO	Earth Observation
EDTs	Emerging Disruptive Technologies
EVT	Experientable Virtual Twin
GNN	Graph Neural Network
IP	Internet Protocol
Geovis Spat Anal	Journal of Geovisualization and Spatial Analysis
ML	Machine Learning
MBSE	Model Based Systems Engineering
RF	Radio Frequency
RaaS	Ransom as a Service

1. Introduction

Over the last decade, the fourth industrial revolution has brought significant scientific and technological progress that has deeply affected spatial data-centric and data-dependent systems. Given the inherent criticality of the space sector, scientific and technological progress has also resulted in the emergence of new malicious capabilities targeting space systems. Among the countermeasures adopted to tackle them is the development and deployment of the so-called Emerging Disruptive Technologies (EDTs), such as Artificial Intelligence (AI), which are highly capable of delivering fast and reliable command-and-control decision-making, as well as providing reliable vulnerability analysis using well-proven cutting-edge techniques.

The significant advances made in the field of AI in the last decades has contributed to human progress in a wide range of scientific fields, such as robotics and machine learning. Moreover, it has substantially contributed to boosting current space efforts. AI is applied in many fields ranging from mission planning and designing, processing extensive amounts of data collected by satellites, assisting navigation systems as well as enhancing satellite imagery [1].

High-quality and precise satellite imagery is particularly important for Earth Observation (EO) and monitoring activities, defined as “the gathering of information about planet Earth’s physical, chemical, and biological systems via remote sensing technologies, usually involving satellites carrying imaging devices” [2]. Remote-sensing provides unique capabilities and advantages such as observing wide areas, contributing to the increasingly accurate development of early warning or weather detection systems, allowing for the collection of data without jeopardising national sovereignty, rapid measurement of acquired images, and ensure operational continuity in the use of sensors belonging to previous missions, thus, in long-term data collection [3].

In recent years there has been an increasing use of Machine Learning (ML) and AI for EO applications. In fact, the exponential growth of data collected by satellites, now on the order of several petabytes, requires the use of technologies for a quick and accurate analysis [4]. An example of this use is the PhiSat-1 (Φ-Sat-1) satellite, the first European satellite to use AI to efficiently send EO data back to Earth. More specifically, the hyperspectral camera collects a significant number of images, some of which have poor quality due to external factors, such as cloud coverage. Φ-Sat's artificial intelligence chip filters them to return only usable data, autonomously discarding those images that cannot be used [5].

The use of AI to support EO and monitoring activities has raised some challenges, more specifically deriving from the adoption of AI-based solutions to achieve cybersecurity and later cyber defence objectives in both civil and military operations. This includes the definition of legal and technical frameworks, ethical dimensions, and other concerns such as mission safety, national security, and technology development for future uses.

The objective of this paper is to propose a regulatory framework using ethical and technical approaches for the use of AI in space, specifically applied to EO and system health monitoring. To achieve the objective, the paper develops a cybersecurity risk and vulnerability taxonomy for the future applications of AI in space and assesses to what extent a network digital twins’ simulation can still protect networks against relentless cyber-attacks.

2. New technologies

New technologies are a set of applications of scientific knowledge that offer a significant improvement over an established technology for a given process. The definition of "new" is in a continuous redefinition as technology changes over time in a cyclical way. These new technologies are a focal point for the development of our society and are discussed frequently for their potential use in both the civil and military domain. This research paper will focus mainly on AI and Cybersecurity.

AI, machine learning, deep learning and others are disruptive technologies that have been at the centre of attention for their potential use in conflict, deterrence, assurance, and competition. AI is often used as an umbrella term for a large variety of disciplines.

Although its use is increasing, AI still doesn't have an universally accepted definition. Today there are many definitions, however the term Artificial Intelligence appeared for the first time in a workshop at Dartmouth University in 1956. John McCarthy, also known as the father of AI, defined AI as "the science and engineering of making intelligent machines" [6]. Referring to intelligent machines, computer scientist Elaine Rich regards artificial intelligence as "the study of how to make computers do things at which, at the moment, people are better" [7].

Further differentiation should be made on the type of AI – weak, strong and super [44]. The distinction is provided by the range of functions and capabilities that each of the three AI supports. Nowadays, most progress has been made in weak AI. This is specialised on a very narrow range of functions, such pre-programming assistance. Weak AI repeats similar codes that were predefined by their makers and classifies them accordingly. This kind of AI has entered the market and private homes. It is now widely used through smart devices such as smart-homes, phones and cars. On the other hand, strong AI aims to duplicate human intellectual abilities by copying them. While even more advanced, super AI seeks to outperform human intelligence with the increasing computational power that computers are able to elaborate [43]. With regards to the development of AI technologies, after a period of so-called "AI winter" referring to a decline in interest and funding in AI technologies, an era of "AI spring" has entered. In fact, only this technology raised an estimated US$ 6.9 billion in the first quarter of 2020, although covering all industries and not only space. [9].

It is important to distinguish between artificial intelligence and automation. Whereas automation refers to a "broad category describing an entire class of technologies rather than just one" [48] including robotics, AI can be regarded as a type of automation that replaces "human labour in tasks both physical and cognitive" [48].

3. Risk and Vulnerabilities

One of the most important particularities of space data is its "instrumental" nature and the fact that the data received from satellites needs to be converted into meaningful information. Therefore, specific AI methods to leverage advances in physical parameters extraction are needed and used. AI itself, on the other hand, can represent different uses, such as machine learning and deep learning methods, which are mainly used for image classification or object segmentation. The effective use of space data could require hybrid AI methods, encompassing mathematical models for the satellite orbit, the physics of electromagnetic propagation and scattering, signal processing, machine learning, or knowledge representation [27].

3.1 Overview of Cyber Risks Against AI Space Assets

It is not possible or reliable to estimate the probability of a cyberattack. A cyber risk can be defined as the product of threats, vulnerabilities and impacts divided by the possible mitigations [23].

The use of AI bears some risks varying from lack of AI implementation traceability, data sourcing and privacy violations, as well as black box algorithms and lack of transparency, which require the adoption of a system-focused policy to track, assess, prioritise, and control cyber-AI risks. Secondly, the use of AI can introduce program bias into decision making processes. As algorithms become considerably more complex, it is difficult to make a comprehensive overview of existing security vulnerabilities, as well as adopting cyber security measures to prevent any attack.

Other risks are data sourcing and privacy violations since unfettered access to satellite data creates privacy-related legal and ethical problems. Whether governmental or non-governmental entities or even civilians, in the wrong hands, can become a source of national security threats, like revealing the position of secret military bases and global peacekeeping operations [24].

As well as black box algorithms, lack of transparency is one of the major concerns related to AI. AI-based decision-making tools can become target and attacked by cyber means and unintended consequences of these can be the obsolescence of

existing controls, can cause complexity in operations, and the possibility of cascading errors, which take place when only one part of the system fails, and other parts must compensate for the failed component [25]. This in turn overloads these nodes, causing them to fail as well, prompting additional nodes to fail one after another.

3.2. Overview of AI Cyber Vulnerabilities

AI, in particular weak AI, is a cyber vulnerable technology. AI systems are not only embedded with traditional forms of cyber vulnerabilities, particularly the ones deploying machine learning, but also depending on how AI works and learns, existing attack surface composed of coding errors can be complemented by additional, and un-patchable ones, which can render the system using AI more open to attacks [37]. Attack codes to exploit vulnerabilities of AI systems have already proliferated in space by many States and agencies. On the one hand machine learning vulnerabilities futher enable hackers to manipulate systems' integrity (causing them to make mistakes), confidentiality (causing them to leak information), and availability (causing them to cease functioning), while AI cyber defensive techniques are limited and hard to keep up with new means.

The uses of ML algorithms can help to identify and defend against computer-based vulnerabilities [21] and threats by automating the detection of an attack and its response. On the other hand, offensive AI algorithms can render cyberattacks increasingly difficult to block or defend against by enabling rapid adaptation of malware to adjust to restrictions imposed by countermeasures and security controls [22].

In terms of AI cyber security, vulnerability refers to a weakness in hardware, software, or procedures. Risk on the other hand, refers to the potential for lost, damaged, or destroyed assets. Starting from the mission execution level to the data analysis, AI systems still have significant limitations and vulnerabilities, particularly in terms of predictability, verifiability, and reliability. Both AI systems and AI-enabled systems deployed in different contexts in space can be attacked. AI attacks are enabled by inherent limitations in the underlying AI, algorithms that currently cannot be fixed, therefore, they are different from traditional cyberattacks that are caused by "bugs" or human mistakes in codes. An attack can target security in the training algorithm (e.g., adversarial machine learning), or vulnerabilities in the training process (e.g., data poisoning attacks). On the other hand, vulnerabilities in the platform on which the AI system runs can also have an impact on the classification results. An example is a concrete proof-of-concept attack to prove the feasibility and impact of platform attack, or a higher-level qualitative analysis to reason about the impact of large vulnerability classes on AI systems [28].

4. Cybersecurity risk and vulnerability taxonomy

AI technologies are one of the enabling and innovative technologies that can both reduce and augment cybersecurity risks and vulnerabilities. A cyber taxonomy would help to align cybersecurity definitions and terminologies to enable the categorisation of potential risk and vulnerabilities. Understanding technical aspects will help to shape legal and policy aspects.

Even if one might intuitively think that space assets can be challenging to attack, they are prone to multiple risks, even of a cyber nature. Satellites are the core of many industrial sectors such as telecommunications and, in the case of navigation, are the elements that, if disabled or destroyed, completely prevent operations. In addition, the importance of cyber risks for the space sector stems from the fact that there are no common standards and regulations in this field; that supply chains are particularly complex to manage; and that often these types of attacks deliver significant benefits to a relatively low price and visibility.

Cyber threats can affect all segments of a space operation, so both space, link, and ground segments need to be monitored and protected [47]. If kinetic threats aim to destroy or physically harm targets, and electronic threats aim to intercept or disable RF communications, cyber threats target data directly. The complexity of an attack is relatively low. Private hacker groups or individuals with low budgets can pose a threat. Space cyber threats can be analysed under two main categories, thus as technical cyber threats and as social engineering cyber-Threats [39]. The former exploits the technical weaknesses of the various segments of space activities, while the latter exploits the deception or psychological manipulation of the victims in order to penetrate a system.

Technical cyber threats include a variety of attacks like signal hijacking, seizure of control, data corruption, data interception, Denial-of-service attack (DDoS) and Internet Protocol (IP) satellite attacks [39]. Protection against signal hijacking is particularly important in telecommunications satellites. Using an antenna connected to a computer, the attacker can

identify a free communication slot in a transponder and use the bandwidth capacity in excess. In this way, the attacked asset will be used to relay malevolent information, even if the actual risk consists of possible cross-talk interferences or denial of service. Another vulnerability is related to the Command and Control (C2) link which retrieves data from the subsystems. An intrusion into the C2 link of a satellite operator can make it possible for an attacker to seize control of the satellite. This could lead to an unintended change of orbit or a change of attitude to deteriorate optical instruments in an EO satellite. An intruder in the C2 link could also take control of the entire communication subsystem of the satellite leading to the interception of uplink data or the corruption of the downlink. [40] A Denial-of-Service attack to the ground segment and the C2 link could block the control of the satellite's operations and the data collection. Hacker groups can also detect IP addresses from satellites providing internet connectivity and then initiate a TCP/IP connection from a stolen IP address [38].

Social engineering cyber threats include phishing, pretexting,baiting attacks, quid pro quo attacks, tailgating. Such attacks are not addressed to the technology directly, but to the human operators. These practices involve different ways of manipulating the victim's behaviour and psychology. As an example, if phishing exploits human naivety or distraction, a baiting attack exploits human curiosity. Quid pro quo and tailgating (or piggybacking) involve the deception of the victim and camouflage.

In addition to the two categories described above, another way of targeting space systems by means of cyber attacks is to disable or infiltrate the systems that monitor their flight, position, and collision probability, in other words the Space Situational Awareness networks. These attacks have two main objectives; to prevent the observation of space traffic and promote traffic congestion, and to hide the presence of spacecraft from the eyes of a competitor.

4.1. Overview of EO-Assets Security Risk

With the rapid growth of internet services and dependence on the interconnected physical and digital technologies in the 21st century, cyber-physical security is persistently raised as one of the prevalent research in the modern digitalisation realm. Cyber-physical security addresses security concerns for physical systems used to maintain and implement cybersecurity solutions. At the same time the practical angle of AI is gradually emerging to contribute in the advancement of the automated and integrated cyber-physical systems using the ground-breaking AI techniques.

Recently, most space defence agencies customised the backbone of cyber-physical security by gradually augmenting the technical purpose of AI which consists of identifying, collecting, analysing, interpreting as well as neutralising and recovering from interference and intrusion, while constantly blocking doubtful actions on cyber-physical technologies including data communication protocol, data transmission bandwidth, and data management with secure protection [10].

The cyber-physical adversaries essentially can be expressed in terms of two necessary intrusion parameters: cyber-threats and cyber-attacks. It is hard to devise AI-based automatic tools consisting of well-operated techniques of threats and attacks. Furthermore, the intrusion parameters also elaborate AI, ML, and DL for probing the intrinsic features representation from the existing cyber-physical security big data set. They have been deployed to various real cyber-physical security cases, for instance identifying, predicting, and scrutinising particular sets of threats and attacks which occurred in the field of EO [11]. However, the cyber-physical systems which are supported by AI enable the development of transformative approaches to ensure effectiveness and optimality in such a way that achieve the desired outcomes [12].

On the other hand, in the context of space-based EO-assets, the most urgent need is to progressively develop sustained and trustworthy data handling including other core-technical capabilities, such as data-fetching, data-recognition, data-streamwise, and data-delivery in the form of image and/or non-image types. It leads to implicitly unlocking the long-term intersection research activities between EO-assets and space-security. Thus, there is the need to set up and maintain reliable statistical information for detailed multi-temporal and multi-spatial data provision so as to uphold continuous surveillance and mapping transformations.

As an example, the development of high-fidelity decision support tools referred to as digital twins (DT), allows to counteract the advanced persistent malicious threats and lethal attacks during incessant reconnaissance missions. Other use cases of digital twins are vulnerability detection through visual adversarial analytics, advanced real-time intrusion

monitoring, and resilience assessment on active cyber physical threat intelligence systems.

5. Technical Countermeasures

One of the countermeasures to optimise AI in space applications can be the application of digital twins in the combination with AI. The digital twin is a high-fidelity digital model of a physical system or asset that can be used to optimise operations and predict faults of the physical system and for space to understand different use cases of digital twin for its potential for cybersecurity incident predictions. The integration of digital twin technology and AI has significant effects in aerospace flight detection simulation, failure warning, aircraft assembly, and even unmanned flight, therefore the use of this technology for space provides a good discussion of the trend and the challenges of using DT in providing the technical benefits associated with cyber physical systems.

The notion of DT was firstly proposed by Michael Grieves and conceptualised as a subsidiary part of the strategic diagnostic and prognostic toolset in the context of product life-cycle management [13]. It is basically understood as the essential engineering advancement in the production and operation of technology, while it also offers digital representation of a real-world or physical object to the reformation of a virtual replica, including its process throughout its lifecycle and the required real-time and historical data [14]. The virtual replica can be used for further analysis which can deliver actionable insights in the form of the desired key-performance measurements. This allows to enhance both tangible and intangible products in terms of eight vital values: real-time remote monitoring-control; predictive maintenance-scheduling; scenario-risk assessment; synergy of abnormalities detection; informed decision support system; personalization of products and services; efficiency and safety; and documentation and communication [18]. Aside from that, the necessary key-terms from various thrived definitions of DT being constantly proposed and formally used can be simply characterised in terms of three elementary components: physical reality, a virtual replica, and the bi-directional data flow. The latter occurs in the form of information exchange using cutting-edge cognitive systems [15] between the physical reality and the virtual replica, which comprises data streamwise and actionable insights as depicted in Figure 1.

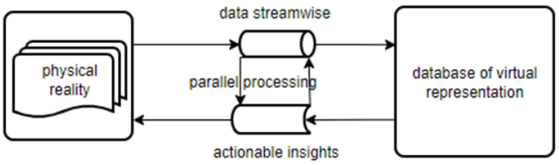

Figure 1. Essential building blocks of Digital Twin [41]

5.1. Conceptual Network Digital EO Twin as a Mitigating Measures

The first part of this analysis focuses on the overview of a cutting-edge network DT architecture proposed by one of the co-authors [41][42], which is compared with the technical white paper developed by a team of industrial practitioners from scalable Network Technologies enterprises [17]. The analysis includes the additional technical explanation on how to carry out the comprehensive idea of network DT in modelling, simulating, monitoring, and assessing the existing EO-assets threats taxonomy.

The second step of the analysis focuses on the breakthrough approach in formalising the main building blocks of network digital-EO twin for handling the RF intrusion and interference, as one of the existing technical suggestions for the upcoming remote-sensing EO activities conducted by authorised space-based research and development institutions, and space-military and defence enterprises.

By definition, network DT is the digital simulation-based model of the communication network integrated with its operating environment and the application of the traffic carried by it. To satisfy its intended goals, the network DT must have sufficient fidelity to accurately reflect and propagate the network dynamics due to the tangible interaction amongst the communication protocols, topology, traffic, and physical environment. A network DT can be further upgraded by incorporating cyber vulnerabilities and defences. The cyber-enhanced network DT can be used to verify and validate the cyber resilience of the simulated system in an adversarial environment, while analysing its behaviour and resilience under various collections of spiteful intrusion and interference scenarios [17]. The visualisation of network DT architecture developed by industrial practitioners and academic researchers is provided as depicted in the Figures.

5.2 Objectives

Practical use cases for the digital twin approach, and its use in EO imagery, communications and networking, referred to as network digital-EO twin.

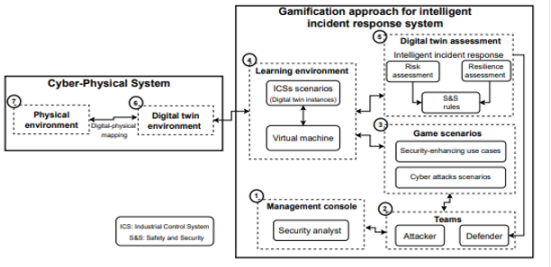

Figure 2. Network Digital Twin in the context of Intelligent Incident Cyber-Physical Response System [41, p.7]

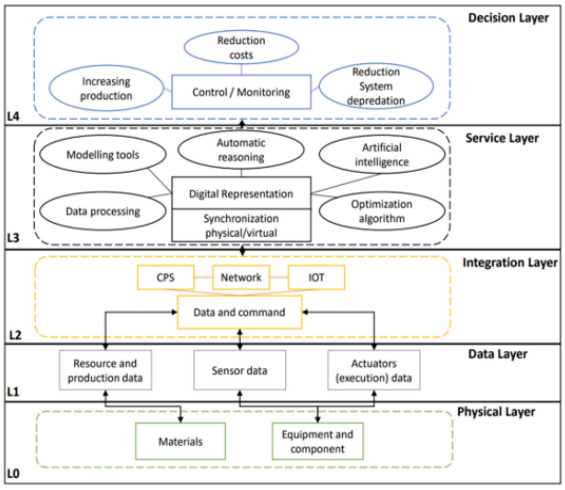

Figure 3. Network Digital Twin in Hierarchical Layer-by-Layer mode from the bottom level L0 to the top level L4 [42, p.7]

The main idea behind the development of network digital-EO twins is the elaboration between three essential pillars, which are Experientable Virtual Twin (EVT), reliable adversarial ML models, and advanced AI solver using Graph Neural Networks (GNN), as shown in Figure 4. GNN is strongly chosen as a neural network solver for developing a highly resilient, secure, and lightweight architecture model of data-driven networks, including the capability of detecting particular anomalies. These pillars can be identified as the essential building blocks for EVT for AI space systems in order to align with the latest scientific research and breakthrough cybersecurity solutions.

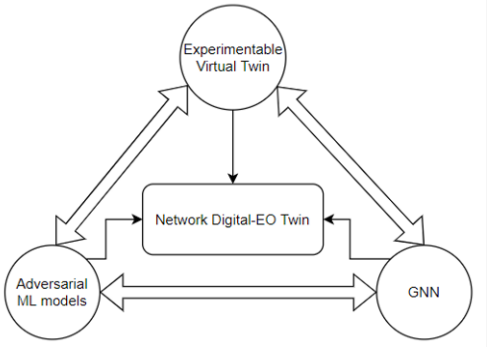

Figure 4. The Primary Building Blocks of Network Digital-EO Twin [41]

EVT basically combines the underlying notion among MBSE (Model Based Systems Engineering), simulation-based technology, and DT itself. Besides, it is created to comply with high-fidelity simulation based systems engineering processes for a variety of different applications, from the development of verification, training, optimization, testing, validation, up to the realisation of intelligent systems [19].

6. Legal and policy aspects of a cybersecurity risk and vulnerability taxonomy

Not only satellites but also satellite data have to be a priority subject of international dialogues on cyberlaw and international security. Most of the privacy-related legal and ethical problems are generated by unfettered access to satellite data, and can be a source of national security threats, like revealing the position of secret military bases and global peacekeeping operations. AI development and use in the space sector bears a regulatory vacuum, except for some national provisions as to technology, and cyber security and safety dimensions of AI use have not been regulated at all.

The expanding variety of space stakeholders and those able to use emerging technologies effectively in their system designs will create unique challenges for each actor, system and uses, that may not be applicable in other areas, therefore will require tailored regulations for cybersecurity and AI in space, in order to regulate both technology and the use of emerging technologies for cross-domain challenges, while watching implementation for compliance to the constituent values, to make the policy and law regulations germane to domain (space-cyber) and technology related cyber challenges.

6.1. Design of AI Cyber Policy for /in Space

The use of AI technology in space without adequate verification and acceptance tests in the engineering phase could cause a high level of risk. Therefore, imperatives for policymakers and legal designers are different. Policymakers have to focus on recognizing the problem, identifying vulnerable systems, and taking steps to mitigate risks before any undesired consequence, for the present uses of emerging technologies, not excluding the possible and future uses. For space, cyber security policy defines and documents any organisations' statement of intent, principles and approaches to ensure effective management of cybersecurity risks in pursuit of its strategic objectives. While law has a reactive approach and clearly defines, regulates and protects against violations of core values, cybersecurity policy, on the other hand, is more successful when it is proactive, not reactive, and when it answers to rapid technological changes and challenges. In terms of cybersecurity and AI, policymakers have to consider many parameters as to the security impacts of AI.

From the cybersecurity aspect, AI can have an impact on the national and international landscape in multiple forms, in physical and non-physical environments; will increasingly be used as a tool to help carry out cyberattacks or to defend against cyber-attacks by nations and private space actors for future missions. While generating new modes of informational warfare, its use will expand the threat landscape, and might contribute to the destabilisation and generation of new forms of weaponization for conventional and non-conventional actors. With the growth of AI, the intrusion caused by obtaining and retaining the data is not a fixed impact, but will vary according to the quality of data and what the scope of cyber intrusion will be, as analytic processes change and develop, and the legal and policy frameworks will have to catch up with them.

Embedded artificial intelligence in space systems, services, processes, and decision-making, is shifting attention on how the data is and will be used by the software, particularly by complex, evolving algorithms, and the consequences of developing uses. Security focused policies for AI underlines the importance of transparency, testing, and accountability for algorithms and their developers. However, operationalizing these policies in practice requires the establishment of legal responsibility for the occurrence of harmful consequences as a result of the use of artificial intelligence.

AI cyber security space policy, therefore, has to find a perfect balance between innovation and resiliency for all four segments (ground-link-user-space) [47], as well as space actors, including space vendors, contractors, and governments. Space cyber security is on the agenda of US and European actors [8] to set the framework for the urgency, tone and guidelines taking into consideration the particularities of the space and technology challenges and bring different actors together on a common set of principles.

In terms of policy and regulation, the main focus for agencies and governments will be to reduce the risk of attacks on AI systems, and to mitigate the impact of successful attacks. Therefore;

(a) The first consideration, before creating an AI cybersecurity policy and regulation, is the classification of AI systems on a risk basis and on the intended purposes, and in line with existing product safety legislations.
(b) The classification of AI depends not only on the function performed, but also on the specific purpose and modalities for which that system is used. Agencies and actors therefore, need to understand firstly the system particularities, and while reinforcing specific controls depending on the nature of the risk in technical terms, in legal terms AI must be regulated by “sets of harmonised rules for the development and use” of AI systems.
(c) The third phase will be creating regulations, sanctioning not only the cyber attempts and consequences.

The development of an effective space cybersecurity policy will require, firstly, designing cyber resilient systems and therefore, (1) adopting cybersecurity as a priority (in line with existing technical standards and regulations), not as an afterthought. Therefore, (a) defining security elements before defending ground-based systems, networks and space assets, and first minimising risks and vulnerabilities. Then, (b) following the adoption of cybersecurity best practices for both technologies used and their components (Cognitive Computing, Machine Learning, Deep Learning, Neural Networks, Natural Language Processing). Bearing in mind the critics directed at Space Policy Directive-5 [49], setting security frameworks require (2) designing actual risk management frameworks, on the basis of collaboration between governments, private initiatives and operators, for (3) designing security frameworks in

official documents can only be effective with (a) standardisation, (b) modernisation, (c) transformational initiatives, and (d) verifications through experience and understanding which tools/designs/policies are effective and which are not. In these terms, exercises and game-playing like wargames and hack-a-sats can help to open and develop dialogues, to set common grounds and principles, and clearly see 'how to's' in order to build a living, adaptable to evolving threats, up-to-date to be resilient, and reactive policy.

6.2. Liability in Terms of AI Cybersecurity for/in Outer Space

Cyber attacks and the other new technologies such as AI or blockchain were unknown during the adoption of the Liability Convention [50] and how the Convention will be able to cope with new challenges posed by harmful 'activities/interferences' committed by using these new technologies were as well unknown during the era of its adoption.

The presence of cybersecurity vulnerabilities in new and emerging technologies pose great risks. While providing consistency and time advantage to assess data, AI bears varying risks, and one of the important ones is "Unclear Legal Responsibility" for many aspects originated from the technology. Firstly, the Tallinn Manual, in Rule.11. reasons that only cyber attacks of sufficient "severity," "invasiveness," and "military character" amount to uses of force [33]. In terms of the Liability Convention, the 'injurer' and 'target' are the space objects, accordingly the first consideration as to the applicability of the Liability Convention for the cases of cyberattacks against software or software defined space assets is based on 'whether the software is covered by the term 'space object', and the answer is positive.

Another reason for uncertainty is the difficulty to foresee the final results of the implementation of AI and lack of precedents as to problems that will arise from the use of AI, and cases that are specific to incidents involving AI cyber security [24]. As to the liability under international space law, the use of AI and rise of cyber security and safety issues are significant concerns and challenges regarding the interpretation of Art. III of the Liability Convention, for the determination of 'fault' and the establishment of causal link between the fault and the damage.

In terms of cyber-attacks, compared to other types of interference targeting space assets, particularly low-intensity cyber-attacks are mainly physically non-destructive, with latent intervention, and have a low threshold to access [30]. However, both Art.30 of the Tallinn Manual [31] and Art.1 of the Liability Convention [32] require 'damage' to 'life, health, and property', and neither of these documents foresee mechanisms to impose liability for low-intensity cyber attacks, which can be considered as a legal vacuum for ongoing low-intensity cyber threats against space objects including software.

In terms of liability, famous Roman law maxim, "sic utere tuo ut alienum non laedas" principle, which states that "each must use his property in a way that does not cause injury to another's", can be a hint to understand and discuss possibilities for the realm in outer space and cyber security of space technologies. In order to strengthen international peace and security in space, within an unstable cyber environment, and minimise threats, application of cyber due diligence can be considered as one of the options.

A state is responsible for failing to take, either generally or with respect to the conduct of individuals, according to due diligence care as the particular obligation requires [34]. States are obliged under international law to exercise due diligence in preventing their territories from being used to perpetrate harmful conducts that will interfere with the rights of other states. The principle of due diligence would require states to set standards and norms in terms of their cyber infrastructure, cyber activity, and people engaged in cyber activities, however, both as to AI and space and cyber due to the lack of established international law, as well as the different features between cybersecurity and space security, and uses of technology, leaves victims navigating unknowns, since the wrongdoer is (often unknown), the types of wrongful acts (intentional human/State actions), the damage (personal data theft and damage to systems), and attribution (the difficulty of identifying those responsible).

As underlined in many occasions, such as The UN Group of Governmental Experts (GGE) indicated the importance of procedural obligations to prevent harm, and encourages states to cooperate "to mitigate malicious ICT activity emanating from their territory" [35], uses of emerging technologies, like AI require attention.

In conclusion, the future regulation of liability generated by cyber attacks/interferences against space technologies has to find and design a balance. Whether through national or international norms, addressing the attribution-response gap will be difficult. Therefore, in order to regulate the legal regime as to the use and the consequences of these uses for Emerging

Technologies, States and industries have to understand and redefine, as;

(a) "*'Harm' considering the technology used and the environment in which the technology used,*
(b) *The likelihood and the degree of the technology used that contributed to the harm;*
(c) *The risk/ known vulnerabilities within the technology and environment the technology used,*
(d) *The Informational asymmetry, the degree of ex-post traceability and intelligibility of processes within the technology that may have contributed to the cause;*
(e) *The degree of ex-post accessibility and comprehensibility of data collected and generated by the technology.*
(f) *The kind and degree of harm potentially and actually caused*" [36].

Even if the harm is caused/originated by a cyber-attack, the liability is conditional upon the intent of the perpetrator or negligence of the operator. In order to counter general expectations of reasonable care and regard for harms to sovereignty between States, due diligence can serve in the absence of a legal regime. However, the legal vacuum as to the non-state actors and targeting private space activities, as well as low-intensity cyber interference, remains.

A state can be liable for an act of transboundary harm, even if the activities giving rise to the harm were not in themselves breaches of international law. The Liability Convention rather envisages the damages caused by impact, than the damage inflicted through activity. Malicious transboundary cyber conduct committed by non-state actors can exceed that committed by states. The international legal regime is based upon the sovereignty equality of its member states, international law demands the existence of effective international legal rules that provide states with protection from non-state actors that commit malicious cyber conduct from the territory of other states.

7. Conclusions

We are transitioning into a new era in terms of policy making and legal regulation of responsibilities for governments and growing private space actors, as well as outsider adversaries using emerging technologies in and against space.

The AI technologies are expected to be used more extensively in future space missions, and augmented use of these new technologies and cybersecurity concerns as to the latter, brings more topics to discuss as to the security of future space missions and as to the applicability of existing norms for new technology driven challenges. However, cyber-attacks on space assets are different from the cyber-attacks targeting other kinds of critical infrastructure, because numerous States and now private actors are engaged in space activities, and considering the augmentation of services provided from space, the regulation of the new relationships require new discussions, beyond existing frameworks provided by existing international space law.

AI is enabling progress and innovation in the space sector and helps to provide robust solutions to the most relevant problems. Therefore, creating processes and frameworks to use AI technologies requires taking into consideration particularities of the technology, in the first place, in order to ensure clarity, in normative and policy grounds, and to respond to cyber security requirements timely. Neither existing space policy nor cybersecurity policy is prepared for the challenges created by the meshing of space, cyberspace and emerging technologies, especially designed for space assets and use of emerging technologies in space activities. In order to ensure adaptable/compatible use of emerging technologies with other technologies in complex environments, adoption of responsive universal principles and regulatory frameworks [20] becomes an important agenda for authorities, governments and industry. In the absence of dialogue and formal policy and regulations, it will become difficult to use emerging technologies, minimise and mitigate risks, develop and use technologies for future missions within a security framework and to build robust defences against emerging technological threats

Therefore, it is essential to note that there are important challenges such as the use of AI-enabled DT technologies with full performance. These challenges might depend on the scale and integration complexity of the applications, besides the uses for space missions. The main challenges to consider are issues related to data, including trust, privacy, cybersecurity, convergence and governance, acquisition and large-scale analysis [45]. While DT promises many advantages, this technology is under development and far from maturity in the near future. The existing limitations for more mature and complex implementations of DTs across all domains, including both space and cyber, will also require overcoming communication network related obstacles on the

technical aspect, which also creates another difficulty for the widespread adoption of this technology and makes accessibility difficult. Trust in technology is another challenge, since the information flowing from various levels of indicator systems presents a challenge for developing common policies and standards. Therefore, lack of standards, frameworks and regulations for DT implementations [46], is one of the grand challenges and has many aspects to consider. For complex implementations of this technology in specific environments, regulations will become more difficult in the future, considering the access related problems to sensitive data by private and military actors, and the adoption of uniform methodologies for data security and authenticity.

References

[1] SERRANO, I., How Artificial Intelligence is advancing Space efforts, 10 August 2021, https://www.geospatialworld.net/blogs/how-artificial-intelligence-is-advancing-space-efforts/#:~:text=Scientists%20use%20AI%20to%20control,communication%20between%20Earth%20and%20space, (*Accessed on 06.07.22*).

[2] Group on Earth Observation, What is Earth observation?, https://www.earthobservations.org/g_faq.html, (*Accessed on 06.07.22*).

[3] European Space Agency, Role Of EO In Understanding Climate Change, https://climate.esa.int/en/evidence/role-eo-understanding-climate-change/, (*Accessed on 05.07.22*).

[4] GEVAERT, C. M., Explainable AI For Earth Observation: A Review İncluding Societal and Regulatory Perspectives, *in* International Journal of Applied Earth Observation and Geoinformation, Vol. 112, 2022, pp. 1-11.

[5] EO Portal, PhiSat-1, https://directory.eoportal.org/web/eoportal/satellite-missions/p/phisat-1, (*Accessed on 04.07.22*).

[6] John McCarthy: father of AI, Intelligent Systems, October 2002, *in* IEEE Xplore, Vol.17(5), pp.84 - 85.

[7] LIAO Matthew, Ethics of Artificial Intelligence, Oxford: Oxford University Press, 2020, p. 3.

[8] CARLO, A,, Cyber Threats to Space Communications: Space and Cyberspace Policies. In: Froehlich A. (eds) Outer Space and Cyber Space. Studies in Space Policy, vol 33. Springer, Cham, 2021.

[9] ERTEL Wolfgang, Introduction to Artificial Intelligence, Cham: Springer, 2017, p. 2.

[10] TORRES, Martínez, COMESAñA, IGLESIA, J. & GARCIA-NIETO, C., P.J. Review: machine learning techniques applied to cybersecurity, *in* Int. J. Mach. Learn. & Cyber, Vol. 10, 2019, pp. 2823–2836. https://doi.org/10.1007/s13042-018-00906-.

[11] SALIH, A., ZEEBAREE, S. T. , AMEEN S. , ALKHYYAT, A., and SHUKUR, H. M., "A Survey on the Role of Artificial Intelligence, Machine Learning and Deep Learning for Cybersecurity Attack Detection," 2021 7th International Engineering Conference “Research & Innovation amid Global Pandemic" (IEC), 2021, pp. 61-66.

[12] How AI Can Ensure Better Cyber Security?, Innefu Labs, https://www.innefu.com/blog/how-ai-can-ensure-better-cyber-security (*Accessed on 10.07.22*)

[13] TREMBLAY, Matt, The Digital Twin: The Benefits of Taking An Incremental Journey, *in* The Maritime Executive, November 3, 2020, https://www.maritime-executive.com/editorials/the-digital-twin-the-benefits-of-taking-an-incremental-journey (*Accessed on 27.07.22*)

[14] Confronting Cyber Risk: An Embedded Endurance Strategy for Cybersecurity (textbook)

[15] APPLEGATE, Scott, A.Stavrou, Towards a Cyber Conflict Taxonomy, International Conference on Cyber Conflict, 2013.

[16] AGRAFIOTIS, Ioannis et.al, A Taxonomy of Cyber-Harms: Defining the Impacts of Cyber-Attacks and Understanding How They Propagate, in International Journal of Cybersecurity, 2018, 1-15, doi:10.1093/cybesec/tyy006

[17] Automated Creation of Network Digital Twins handbook, Scalable Network Technologies, © 2021 SCALABLE Network Technologies, Inc. All Rights Reserved PN MRL141217 QualNet and EXata are registered trademarks of SCALABLE Network Technologies, Inc

[18] Digital Twins for IoT Applications: A Comprehensive Approach to Implementing IoT Digital Twins, Sep 2019, [online] available: http://www.oracle.com/us/solutions/internetofthings/digital-twins-for-iot-apps-wp-3491953.pdf.

[19] SCHLUSE, M., PRIGGEMEYER, M., ATORF, L. and ROSSMANN, J., Experimentable Digital Twins–Streaming Simulation-based Systems Engineering for Industry 4.0, IEEE Trans.Ind. Informat, Vol. 14, April 2018, pp. 1722-1731,

[20] ČERKA, Paulius, GRIGIENĖ, Jurgita, and SIRBIKYTĖ Gintarė, Liability for Damages Caused by Artificial Intelligence, *in* Computer Law & Security Review, Vol. 31 (3), 2015, (pp.376-389), pp.383-384.

[21] YAMPOLSKIY, Roman “AI Is the Future of Cybersecurity, for Better and for Worse”, *in* Harvard

Business Review, May 8, 2017. https://hbr.org/2017/05/ai-is-the-future-of-cybersecurity-for-better-and-for-worse.

[22] SWAYNE, Matt, Researchers detail privacy-related legal, ethical challenges with satellite data, Pennsylvania State University, in Phys.org, July 12, 2019. https://phys.org/news/2019-07-privacy-related-legal-ethical-satellite.html?deviceType=mobile

[23] FALCO, Gregory & ROSSENBACH, Eric, Confronting Cyber Risk: An Embedded Endurance Strategy for Cybersecurity, OUP 2022, p.5.

[24] JONES, Kate, BUCHSER, Marjorie, WALLACE Jon, Challenges of AI, Chatham House, March 22, 2022, https://www.chathamhouse.org/2022/03/challenges-ai

[25] DU, P., BAI, X., TAN, K. *et al.,* Advances of Four Machine Learning Methods for Spatial Data Handling: A Review, *in* J Geovis Spat Anal, Vol. 4, 2020 p. 13. https://link.springer.com/article/10.1007/s41651-020-00048-5#citeas

[26] LOHN, Andrew, Hacking AI: A Primer for Policymakers on Machine Learning Cybersecurity, *in* CSET, December 2020, https://cset.georgetown.edu/publication/hacking-ai/ (*Accessed online on 05.07.22*)

[27] MAC WILLIAMS, Carmen, Earth Observation Big Data Challenges: The AI Change of Paradigm, in European AI Platform, EU AI Alliance, Futurium, European Commission, January 29, 2020. https://ec.europa.eu/futurium/en/european-ai-alliance/earth-observation-big-data-challenges-ai-change-paradigm.html

[28] ASHLEY Ashley Hyowon Kim, The Impact of Platform Vulnerabilities in AI Systems, *Ph.D.* Dissertation, Massachusetts Institute of Technology, 2020, https://dspace.mit.edu/handle/1721.1/129159

[29] HUROVA, Anna, Liability for a Cyber Attacks on a Space Objects, *in* Czech Yearbook of International Law, Vol. XXI, 2021, (pp.57-24), p.57.

[30] Tallinn Manual 2.0 On The International Law Applicable to Cyber Operations, Michael N. Schmitt (Ed.), 2nd Ed. 2017. http://csef.ru/media/articles/3990/3990.pdf

[31] Tallinn Manual 2.0, Art. 30. qualifies a cyber attack as cyber operation, 'whether offensive or defensive', "*if it is reasonably expected to cause* ***'injury or death to a person, damage or destruction to objects***"".

[32] Art.1 of the Liability Convention defines the term 'damage' as '**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 organisations,

[33] **Rule 11, Tallinn Manual** specifically notes that "[a] cyber operation constitutes a use of force when its scale and effects are comparable to non-cyber operations rising to the level of a use of force."

[34] Under the doctrine of state responsibility, states are responsible for "wrongful" acts that are (a) attributable to the state and (b) breaches of an international obligation. Int'l Law Comm'n, Draft Articles on Responsibility of States for Internationally Wrongful Acts, with Commentaries, Art. 2, Rep. of the Int'l Law Comm'n on the Work of Its Fifty-Third Session, U.N. Doc. A/56/io, at 68 (2001) [hereinafter Draft Articles on State Responsibility].

[35] UN, General Assembly, A/70/174, Developments in the field of information and telecommunications in the context of international security Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security, Section 17(e), https://ccdcoe.org/sites/default/files/documents/UN-150722-GGEReport2015.pdf (*Accessed on August 20, 2022*).

[36] KOCH, Bernhard A. "Liability for Emerging Digital Technologies: An Overview", in *Journal of European Tort Law*, vol. 11, no. 2, 2020, pp. 115-136. https://doi.org/10.1515/jetl-2020-0137

[37] PUPILLO,Lorenzo, FANTIN, Stefano, FERREIRA, Afonso, POLITO, Carolina, Artificial Intelligence and Cybersecurity: Technology, Governance and Policy Challenges, Report of a CEPS Task Force, May 28, 2021, p.36.

[38] FALCO, Gregory, "Cybersecurity principles for space systems", *in* Journal of Aerospace Information Systems, 16.2, 2019, pp. 61-70.

[39] FALCO, Gregory, BOSCHETTI, Nicolo, "A Security Risk Taxonomy for Commercial Space Missions", *in* ASCEND 2021, 2021-4241. https://arc.aiaa.org/doi/10.2514/6.2021-4241 .

[40] MANULIS, M., BRIDGES, C., HARRISON, R., SEKAR, V., and DAVIS, A., "Cyber security in New Space: Analysis of threats, key enabling technologies and challenges," *in* International Journal of Information Security, Vol. 20, 2021.

[41] SUHAIL, Sabah, ZEADALLY, Sherali, JURDAK, Raja, HUSSAIN, Rasheed, MATULEVİČİUS, Raimundas, SVETINOVIC, Davor, Security Attacks and Solutions for Digital Twins, 2022.

[42] da SILVA MENDONÇA R, de OLIVEIRA LINS S, de BESSA IV, de CARVALHO AYRES FA Jr., de MEDEIROS RLP, de Lucena VF Jr. Digital Twin Applications: A Survey of Recent Advances and Challenge, *in* Processes, 2022; vol. 10(4), p. 744. https://doi.org/10.3390/pr10040744

[43] CARLO, Antonio, Artificial Intelligence in the Defence Sector, MESAS20, Prague, Czech Republic, 2021, p. 269–278. http://dx.doi.org/10.1007/978-3-030-70740-8_17

[44] Artificial Intelligence, *in* Plato Stanford Encyclopedia of Philosophy, Jul 12, 2018. https://plato.stanford.edu/entries/artificial-intelligence/

[45] SINGH, M.; FUENMAYOR, E.; HINCHY, E.P.; QIAO, Y.; MURRAY, N.; DEVINE, D., Digital Twin: Origin to Future, in Appl. Syst. Innov. 2021, pp. 4, 36.

[46] HARRISON, R.; VERA, D.; AHMAD, B., A Connective Framework to Support the Lifecycle of Cyber–Physical Production Systems, in Proc. IEEE 2021, 109, pp. 568–581.

[47] CARLO, A.; LACROIX, L.; ZARKAN, L., The Challenge of Protecting Space-based Assets against Cyber Threats, IAC-20,E9,2.D5.4,11,x59386, 71st International Astronautical Congress, 2020.

[48] GAYNOR Michael, *Automation and AI sound similar, but may have vastly different impacts on the future of work*, Washington: Brookings, 29 January 2020. (*Accessed on 04.07.22*). https://www.brookings.edu/blog/the-avenue/2020/01/29/automation-and- artificial-intelligence-sound-similar-but-may-have-vastly-different-impacts-on-the-future-of-work/

[49] Executive Office of the President "Space Policy Directive-5: Cybersecurity Principles for Space Systems" (2020) 09 September. Available at https://www.federalregister.gov/documents/2020/09/10/2020-20150/cybersecurity-principles-for-space-systems

[50] Convention on International Liability for Damage Caused by Space Objects entered into force Oct. 9, 1973, 24 U.S.T. 2389, 961 U.N.T.S. 187