UNIVERSITETET I STAVANGER

MASTERGRADSSTUDIUM I

RISIKOSTYRING OG SIKKERHETSLEDELSE

MASTEROPPGAVE

SEMESTER: 2022 V

FORFATTER: Rico Behlke

VEILEDER: Jon Tømmerås Selvik

TITTEL PÅ MASTEROPPGAVE:

How do emerging technologies for collision avoidance in Earth’s orbit impact risk management processes?

EMNEORD/STIKKORD:

Space debris, risk management, collision avoidance

SIDETALL: 76

STAVANGER, 20.06.2023

Contents

1. Abstract

Human society is strongly dependent on satellite-based services such as communication, Earth observation, navigation etc. A by-product of satellite operations is the generation of space debris which consists amongst others of launch stages, defunct satellites, tools or debris resulting from collisions. A collision between an operating spacecraft and space debris can severely damage the operating spacecraft, render it inoperable or in the worst lead to its destruction and many thousands of new pieces of debris. The commercialization of space operations has led to the introduction of thousands of new spacecraft into orbit which provides challenges for collision avoidance operations. Earlier, these have been manual operations, but the sheer number of objects in orbit around the Earth questions if human operators can handle this. Thus, autonomous collision avoidance systems including emerging technologies such as Artificial intelligence have been utilized.

Earlier, for manual collision avoidance operations, classical risk management procedures could be adopted, since the space environment resembled a rather static problem. Now, the orbital picture is much more dynamic with constant manoeuvrers and the change from human to AI operators. Thus, the question arises if the introduction of AI into collision avoidance requires risk management procedures to be revised and the inherent risk of AI operations to be addressed.

This thesis discusses these two issues and concludes that, although AI operations can significantly support collision avoidance features, emerging technologies require modernised risk management approaches with human feedback in order to quality-check the deliveries of the AI and to address the highly dynamic system the AI operates in. Otherwise, AI operations can lead to situations which should have been prevented from the beginning, namely collisions that created even more space debris. This could in the worst case render the orbit around the Earth inoperable.

2. Introduction

Space age started with the launch of Sputnik I in 1957 and with it the insertion of space debris in the Earth's orbit. Space debris comprise all man-made objects in orbit, such as remains of defunct satellites, old launch stages, tools lost by astronauts, remains from collisions or caused anti-satellite rocket tests. Given human kind's dependency on spaceborne technical solutions (e.g., GPS), space debris was gradually considered a threat, not only on ground but also to human space flight.

Until a few years ago, the launch of spacecraft was reserved to state agencies, such as the European Space Agency. This meant that few, but large spacecraft were launched, mostly for research, communication and Earth observation. However, a change of pace occurred a few years ago and the so-called "new space era" dawned. Now, private companies entered the market leading to revolution in spacecraft and launcher design. Suddenly, it was feasible to launch smaller spacecraft in larger numbers (in their thousands) and this was implemented as seen in Figure 1. However, space operations are characterized by the lack of regulations and traffic management, i.e., space can be considered a law-less area where each participant in the space race can act as pleased. This unregulated growth in spacecraft launches carried with it a significant increase in space debris, worsening the issue even further as seen with "pre-newspace" eyes. It has already happened that spacecraft have been damaged by space debris – or as a worst-case scenario – destroyed.

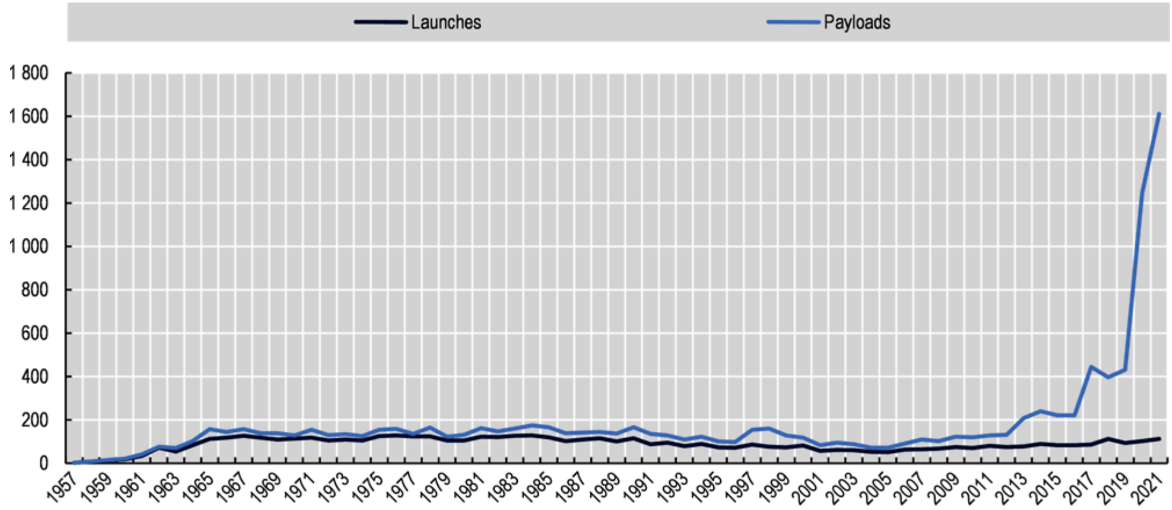


Figure 1: The significant increase in spacecraft in orbit, as depicted by the blue line (launched payloads). Adapted from (OECD, 2022).

This caused space debris to be attributed higher attention. Despite being located hundreds or thousands of kilometres above the Earth's surface, space debris can have a profound impact on modern life as we know, since we are highly dependent on spacebased infrastructure. Firstly, GPS has found its way into our private life but is also considered a crucial asset for the economy. Degradation or even failure of GPS services would hamper flight traffic, offloading of container ships etc etc. Secondly, satellites are widely used in Earth observation services such as climate change monitoring or weather forecasting. Thirdly, communication services such as satellite phone or spacebased global internet are becoming widely used. Fourthly, financial transactions depend on satellite communication. Further dependencies can be added to this list.

Also at a higher level, space debris has been acknowledged as a problem which needs to be dealt with. An example is the inclusion of space debris in the Danish Risk Picture (Beredskapsstyrelsen, 2022). Thus, space debris equates to other risk factors, such as flooding, nuclear accidents, terror etc. A further example is the OECD report on the economics of space sustainability which space debris is an important part of. This report considers space sustainability the next major societal challenge (OECD, 2022).

Despite the lack of a legal framework, reacting to space debris before the newspace era – meaning potential collision between an active spacecraft and space debris (or another active spacecraft) – was characterized by few operators (since mostly only state agencies were active) and few spacecraft. This changed dramatically by the introduction of private companies now launching hundreds if not thousands of spacecraft without any legal framework concerning the generation of space debris, responsibilities, traffic management in orbit or – at least – any established routines concerning the communication between operators. In addition, autonomous collision avoidance systems were introduced with little information to others than the actual operators on the details of their operative details.

One could compare this change with a motorway which originally was populated by a few single cars owned by a few state operators. If any cars got close to each other, there was enough overview and space to solve the problem. Now suddenly, a great wealth of other cars of many different sizes operated by many private owners poured into the motorway driving everywhere and in any direction without any rules. Crashes occurred littering the motorway with debris causing further crashes. To make things worse, autonomous cars were introduced at the same time with only the operators knowing of their operative details.

Collision avoidance processes are also risk management processes. This raises an interesting question: how has the entry of newspace companies changed risk management processes in space? The significant increase in number of spacecraft and space debris has lead to a very complex regime where small changes can lead to enormous changes. What role does spacecraft and space tracking play in this process? How is the lacking legal framework incorporated in the risk management process? Does the introduction of autonomous processes improve or worsen the situation and how do they impact risk management processes? How are collision probabilities calculated in this highly complex system?

The largest player in the newspace segment is Starlink which is owned by SpaceX. Starlink plans to insert more than 10000 spacecraft into orbit in order to provide satellite-based internet on Earth. Hence, Starlink is the ideal company to study in this thesis.

Thus, the topic of this master thesis is: **How do emerging technologies for collision avoidance in Earth's orbit impact risk management processes?**

The thesis has the following contents: Section 3 describes the underlying theory which is required to discuss the problem raised. The concept of space debris is introduced and an overview of the different parts of space debris and its development is given. Also, examples of effects of space debris both in space and on Earth are given and ways of tracking and modelling space debris are explained. Furthermore, an outlook of future developments is included. Subsequently, the concepts of risk, risk analysis and risk management are presented, since collision avoidance is considered a risk management process. Consequently, the legal framework of space operations is described, since this also has to be included in the risk analysis and management. As shown, satellite operations can be considered complex operations.

Section 4 describes the methodology used in this thesis. Section 5 presents the empiri. Section 6 discusses the facts in the light of presented theory. Section 7 concludes the thesis.

3. Theory

3.1. Space debris

3.1.1 What is space debris?

Space age started with the launch with Sputnik-1 on October 4, 1957, and has been associated with the introduction of debris into space, hence space debris (ESA, 2022). Awareness of this issue has not been raised before the early 1960s and culminated temporally with the postulation of the Kessler syndrome which claims that space flight might be hindered by an exponential increase of man-made objects in space (Kessler & Cour-Palais, 1978). The first conference on space debris was held in 1982 followed by the creation of the Inter-Agency Space Debris Coordination Committee (IADC) in 1993. The IADC published the *IADC Space Debris Mitigation Guidelines* (IADC, 2002). Although considered a major advance, the guidelines are non-binding and different nations, operators and manufacturers have introduced their own safety procedures. This missing standardisation represents a major challenge which in its turn has to be addressed by, e.g., the International Standards Organisation (ISO). In addition to the IADC guidelines, the United Nations' Committee on the Peaceful Uses of Outer Space (UNCOPUOS) has introduced the *Guidelines for the long-term sustainability of outer space activities* (UNCOPUOS, 2021).

Definition of space debris: "*Space debris* is defined as all artificial objects including fragments and elements thereof, in Earth orbit or re-entering the atmosphere, that are non-functional" (ESA, 2022, s. 10). Often space debris is characterized as either launch related or unidentified **(UI)**. For launch related space debris, the following subcategories are utilized:

- Payloads **(PL)**: All objects that are not engaged in launch processes, e.g., active spacecraft.
- Payload mission related objects **(PM)**: All objects that were related to the operation of a payload, e.g., astronaut tools.
- Payload fragmentation debris **(PF)**: space debris arising from explosions or collisions which can be traced back to a unique event.
- Payload debris **(PD)**: as for payload fragmentation debris, but without the possibility of a trace-back to an unique event.
- Rocket body **(RB)**: space debris that has been utilized during a launch operation.

- Rocket mission related objects **(RM)**: intentionally released with a designation for operation of a rocket body.
- Rocket fragmentation objects **(RF)**: space debris originating from a rocket body due, e.g., an explosion.
- Rocket debris **(RD)**: as for rocket fragmentation objects but without the possibility of a trace-back to a unique source. (ESA, 2022)

Space debris can also be characterized in terms of its existence in surveillance catalogues. A *catalogued object* is registered in a space surveillance catalogue. In contrast, an *asserted object* is not registered in such a catalogue, but is known to exist by its design. (ESA, 2022)

3.1.2. The temporal development of space debris

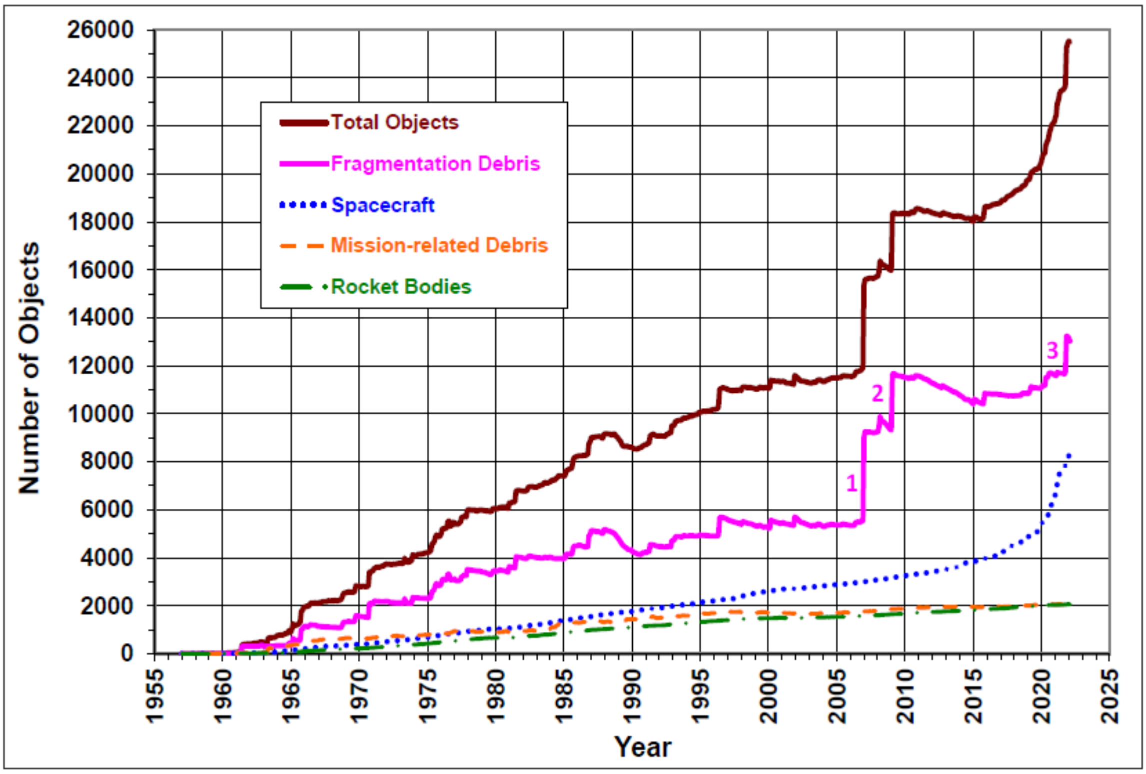


Figure 2: Temporal development of number of space debris objects. Note the sharp increases associated with (1) which is connected to the fragmentation of Fengyun-1 in 2007 and (2) which is related to the collision between Iridium-33 and Cosmos-2251 in 2009. Adapted from (NASA, 2022).

Figure 2 shows the temporal development of both of the number of launches and space debris in orbit around the Earth. The number of launches increases steadily until approximately 2020 when a sharp increase can be observed. This relates to the entry of newspace companies and the beginning of launches connected to Starlink and Oneweb constellations. Also, clearly visible are two sharp increases in the number of fragmentation space debris. The first relates to the fragmentation of the Fengyun-1 spacecraft which was targeted by a anti-spacecraft rocket during a test. The second corresponds to the accidental collision between the Iridium-33 and Cosmos-2251 satellites – a textbook example for the amount of space debris that can be created by a collision between spacecraft which in its turn leads to an increased risk for collisions between other spacecraft and debris created by this collision (NASA, 2022).. This is the development which ultimately could lead to the Kessler syndrome.

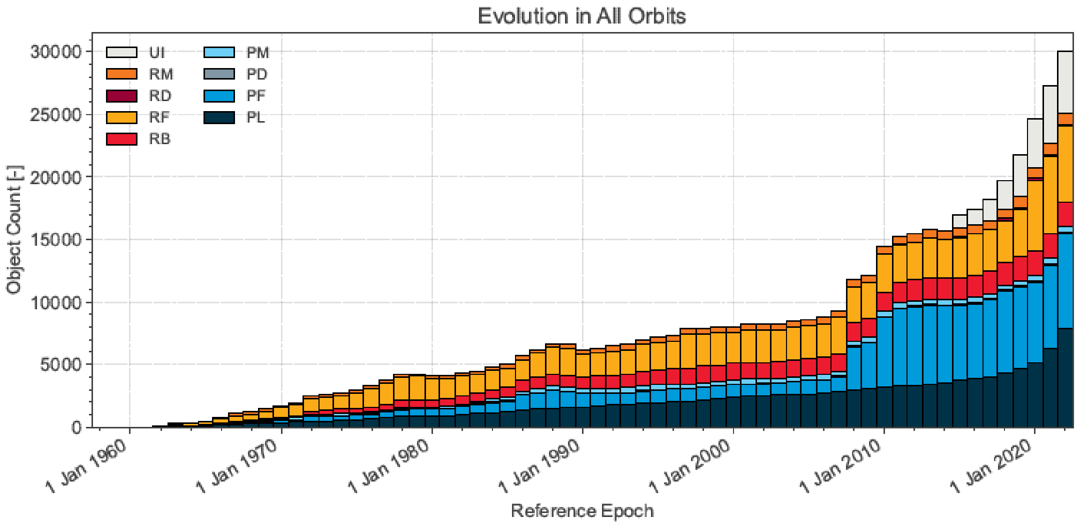


Figure 3: Development of number of space debris objects according to type of debris. Adapted from (ESA, 2022).

Figure 3 shows a different point of the development depicted in Figure 2, namely the characterization of different regimes of space debris. Here, the three largest contributors are payload related space debris (PL), payload fragmentation debris (PF) and rocket fragmentation debris (RF). This does however not describe the spatial of space debris in the different orbital shells.

Figure 4 describes this spatial distribution of space debris with respect to different orbital shells. The by far largest contribution of space debris is to be found in Low Earth Orbit (LEO), the orbital shell which Starlink populates.

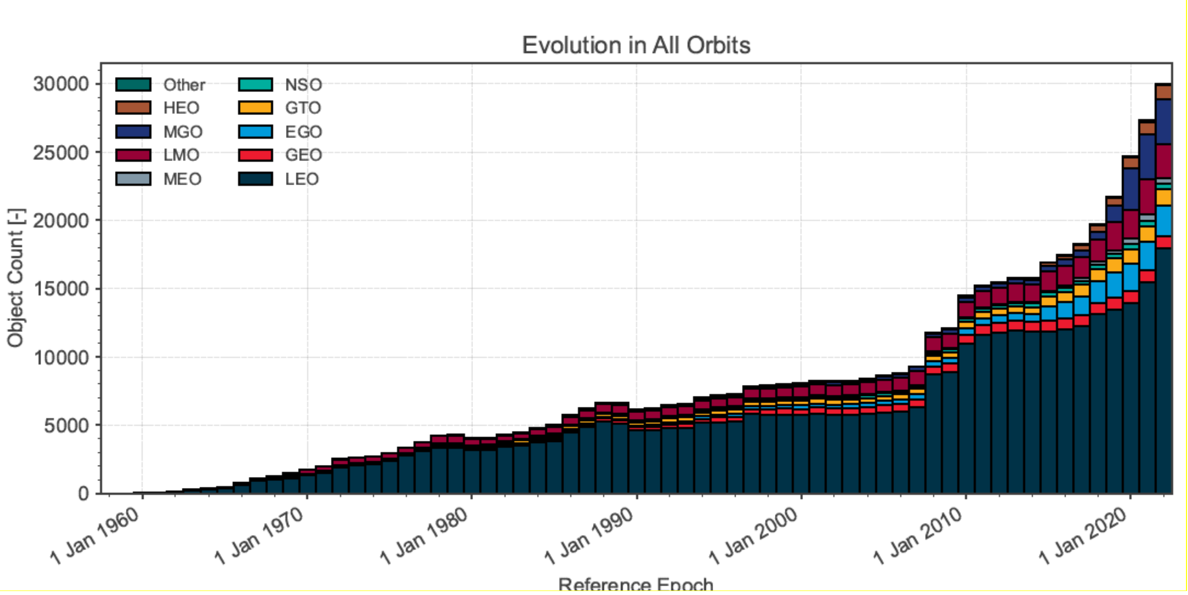


Figure 4: Development of number of space debris objects according to orbital shell. Adapted from (ESA, 2022)**.**

When talking about orbital shells, it is of course relevant to define the different shells, see Table 1.

Orbit	Description	Definition		
GEO	Geostationary Orbit	$i \in [0, 25]$	$h_p \in [35586, 35986]$	$h_a \in [35586, 35986]$
IGO	Inclined Geosynchronous Orbit	$a \in [37948, 46380]$	$e \in [0.00, 0.25]$	$i \in [25, 180]$
EGO	Extended Geostationary Orbit	$a \in [37948, 46380]$	$e \in [0.00, 0.25]$	$i \in [0, 25]$
NSO	Navigation Satellites Orbit	$i \in [50, 70]$	$h_p \in [18100, 24300]$	$h_a \in [18100, 24300]$
GTO	GEO Transfer Orbit	$i \in [0, 90]$	$h_p \in [0, 2000]$	$h_a \in [31570, 40002]$
MEO	Medium Earth Orbit	$h_p \in [2000, 31570]$	$h_a \in [2000, 31570]$	
GHO	GEO-superGEO Crossing Orbits	$h_p \in [31570, 40002]$	$h_a > 40002$	
LEO	Low Earth Orbit	$h_p \in [0, 2000]$	$h_a \in [0, 2000]$	
HAO	High Altitude Earth Orbit	$h_p > 40002$	$h_a > 40002$	
MGO	MEO-GEO Crossing Orbits	$h_p \in [2000, 31570]$	$h_a \in [31570, 40002]$	
HEO	Highly Eccentric Earth Orbit	$h_p \in [0, 31570]$	$h_a > 40002$	
LMO	LEO-MEO Crossing Orbits	$h_p \in [0, 2000]$	$h_a \in [2000, 31570]$	
UFO	Undefined Orbit			
ESO	Escape Orbits			

Table 1: Definition of different orbital shells according to their semi-major axis a, eccentricity e, inclination I, perigee height h_p and apogee height h_a. Units are presented in kilometres and degrees. Adapted from (ESA, 2022).

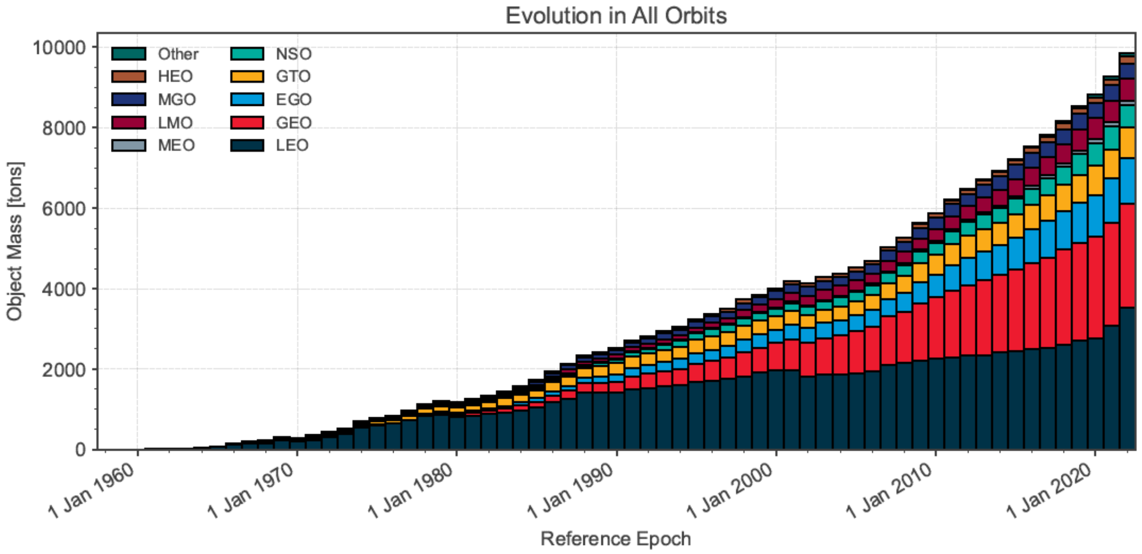


Figure 5: Development of the mass of space debris objects according to mass and orbital shell. Adapted from (ESA, 2022).

Figure 4 depicted the number distribution of space debris in the different shells, but space debris can range from remains of the order of a few centimetres to several metres. Therefore, it is also of interest to illustrate the mass distribution in each orbital shell. Figure 5 shows the mass distribution of space debris in each orbital shell. As expected from Figure 4, LEO contains a significant amount of space debris also in terms of mass, but Geostationary Orbit (GEO) also contains a great of space debris in terms of mass. This has its explanation in the fact that GEO is commonly used for Earth observation and communication satellites which usually are of larger size.

As it will be shown later, space debris created in one orbital shell can propagate into other orbital shells. Figure 6 and 7 depict the propagation of space debris into the LEO and GEO shell, respectively. Figure 6 shows clearly that most of the debris present in LEO also originates from LEO. This is however different for GEO shells. Here, only part of the space debris originates from GEO, other significant contributors are EGO, MGO and HEO.

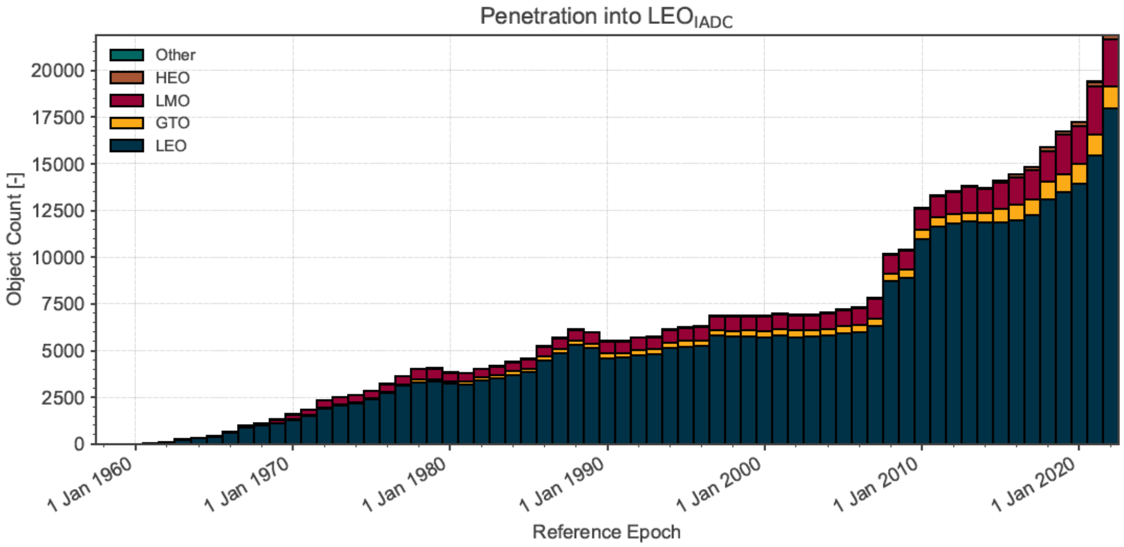


Figure 6: Development of number of space debris objects in protected LEO orbit according to origin. Adapted from (ESA, 2022).

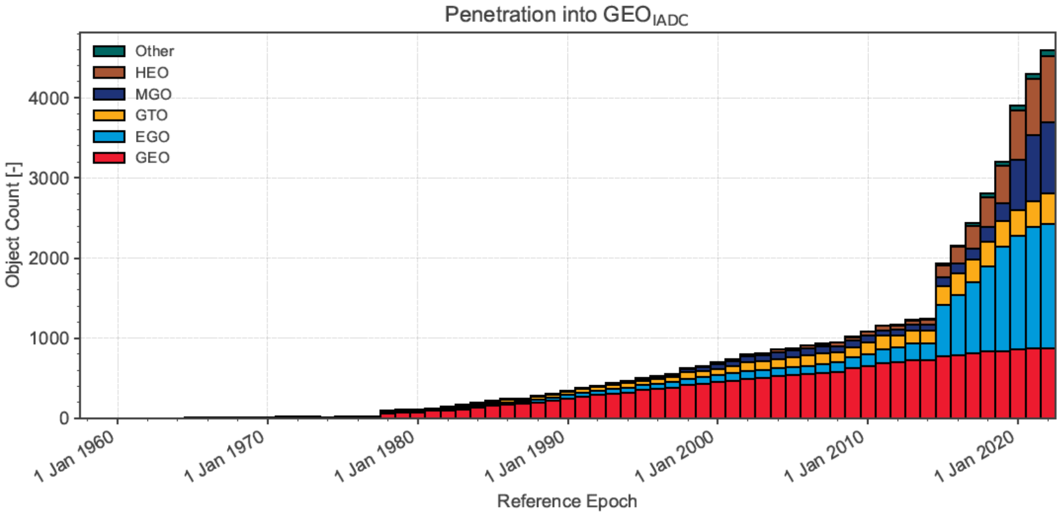


Figure 7: Development of number of space debris objects in protected GEO orbit according to origin. Adapted from (ESA, 2022).

Figure 1 showed the significant increase in payload traffic, especially during the last years, however without distinguishing between different orbital shells. Figure 8 and 9 show payload launch traffic in the shells between (1) 200 and 1750 kilometres and (2) 35000 and 36800 kilometres, respectively. These regions are not equivalent to the protected LEO and GEO regions, but they roughly resemble them. Comparing Figure 8 and 9, one clearly can understand that most of the significant increase of launch payloads was inserted into LEO, that is where amongst others Starlink operates.

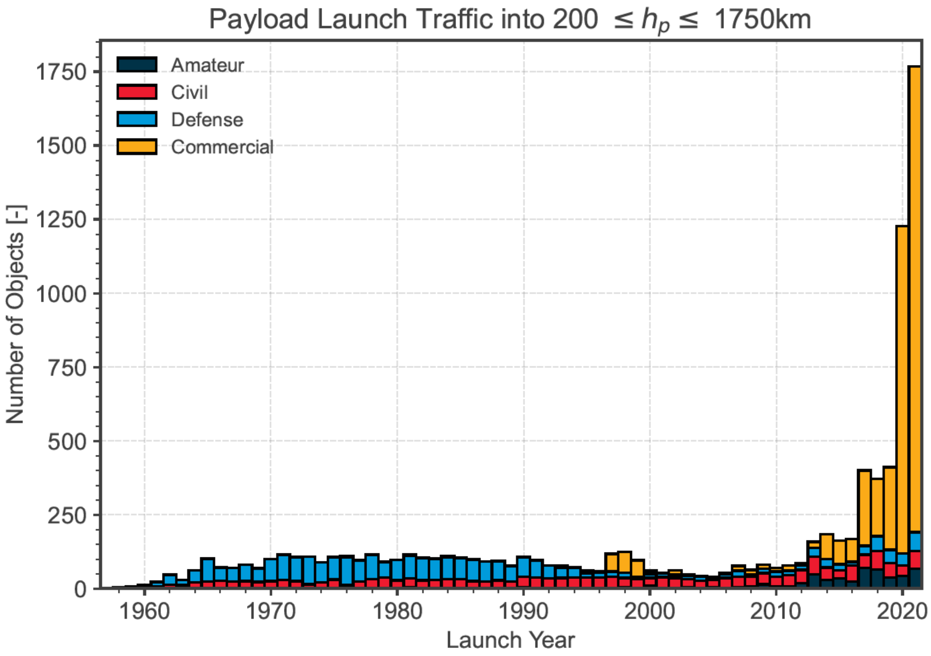


Figure 8: Development of launch traffic for perigee height between 200 km and 1750 km. Adapted from (ESA, 2022).

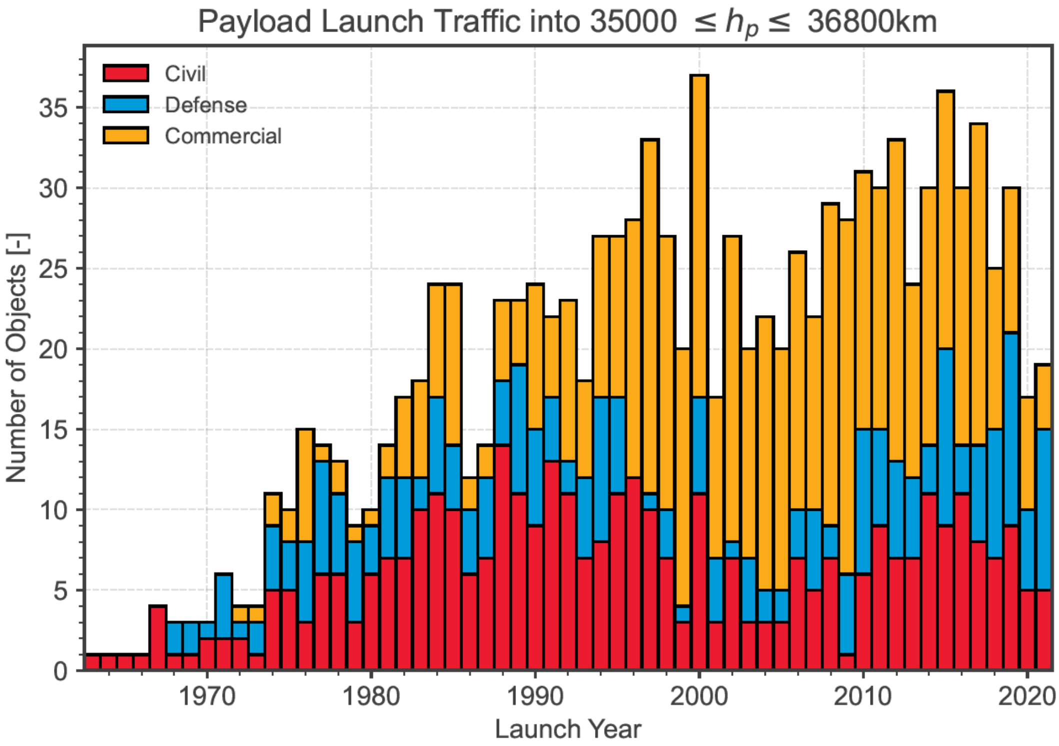


Figure 9: Development of launch traffic for perigee height between 35000 km and 36800 km. Adapted from (ESA, 2022).

The entry of newspace companies is associated with the change from large, non-constellation spacecraft to small, constellation satellites. Figure 10 clearly shows this change for LEO ranges.

Whereas launches in the time period 1960 to 2010 were associated with spacecraft with a mass larger than 1000 kg, this shifted clearly to spacecraft with a mass smaller than 1000 kg.

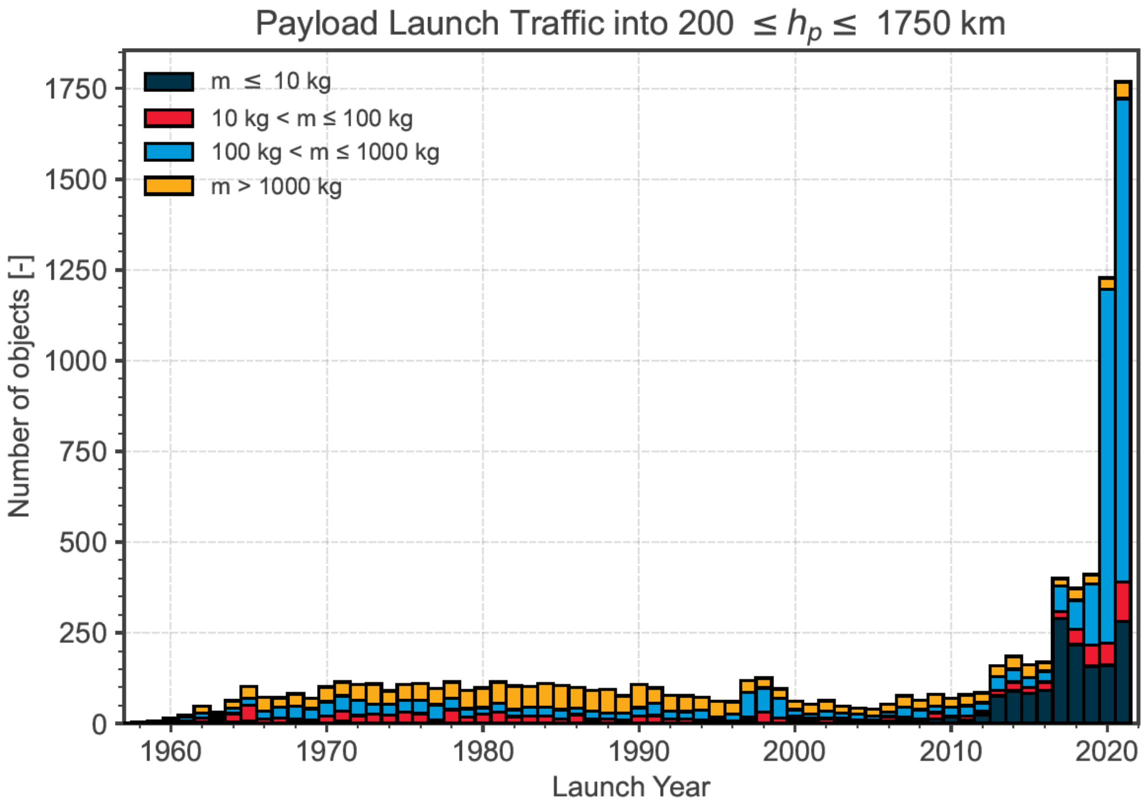


Figure 10: Development of launch traffic for perigee heights between 200 km and 1750 km according payload mass. Adapted from (ESA, 2022).

This is obviously not the case for GEO ranges, as shown in in Figure 11. Here the trend is the opposite compared with LEO. The explanation is that GEO spacecraft usually are Earth observation and communication spacecraft, being much larger by design. The requirements for these satellites include usually longer lifespan and more power required for different instruments such as optical and radar sensors or relay antennas.

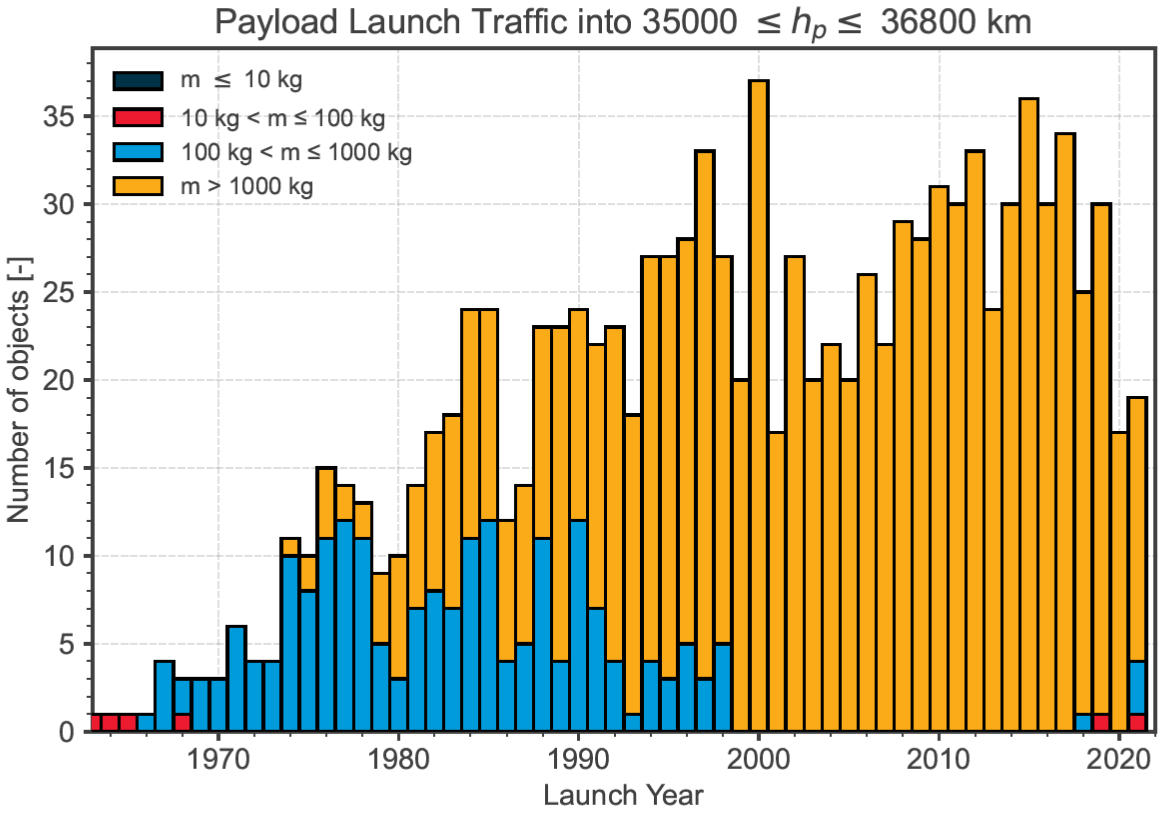


Figure 11: Development of launch traffic for perigee height between 35000 km and 36800 km according payload mass. Adapted from (ESA, 2022).

Figure 12 shows another trend: during early years, launches usually inserted a few spacecraft into orbit, here shown by the ranges of 0-1 or 1-5 payloads per launch. By 2020, a significant part of the launches is now characterized by multiple payloads, particularly in the range of 50-100 and 100-150 payloads per launch. This might seem surprising at first, but the reduction in size of the payloads allows of course for more payloads to be launched by the same rocket. In fact, Starlink for examples launches its satellites in batches up to 60 at a time (Starlink, 2022a).

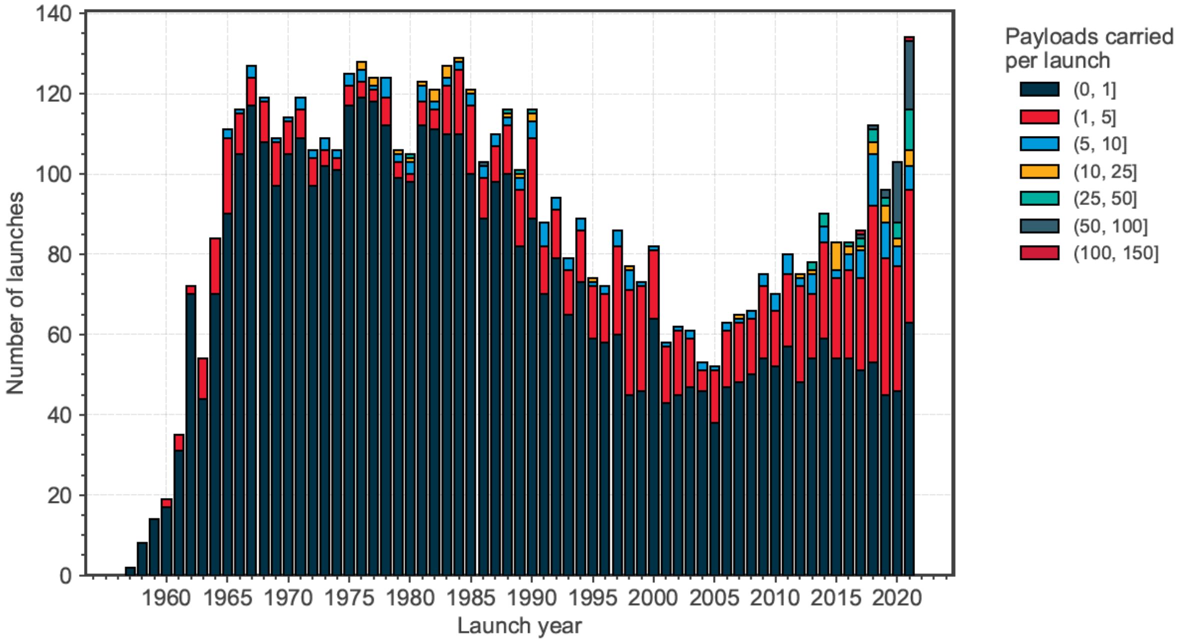


Figure 12: Development of number of launches, marked with number of payloads carried per launch. Adapted from (ESA, 2022).

Figure 13 and 14 describe the lifetime distribution of spacecraft which are not part of a constellation and which are part of a constellation, respectively. Non-constellation spacecraft which are often scientific, military or communication spacecraft which designed for a longer life span due to their specific task, single-spacecraft nature (one main idea of constellations is of course that erroneous spacecraft can be replaced by other spacecraft in the constellation – quantity over quality). This is clearly visible in Figure 13 which depicts the lifetime distribution for non-constellation spacecraft. The main part of the spacecraft in this sector is characterized by a lifespan of 2 to 15 years. This stands in stark contrast to the lifespan distribution for constellation spacecraft, depicted in Figure 14. Here, the vast majority has a lifespan of up to 2 years only. Here, clearly the abovementioned concept of quantity over quality applies.

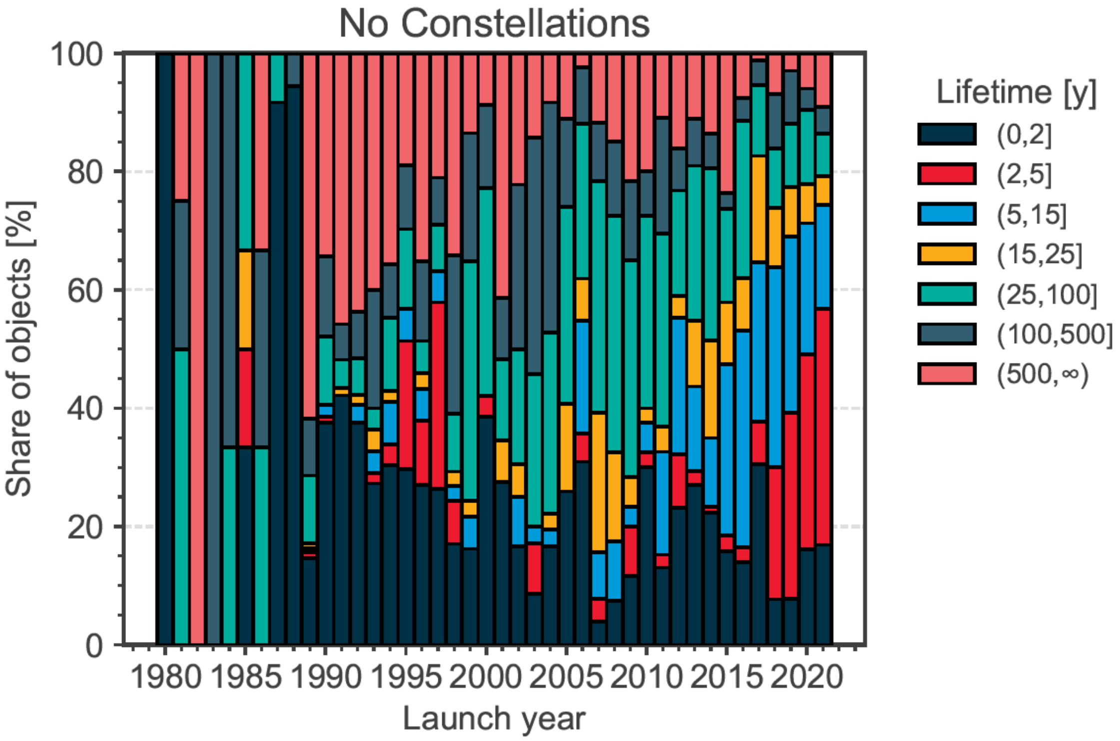


Figure 13: Lifetime distribution for non-constellation spacecraft. Adapted from (ESA, 2022).

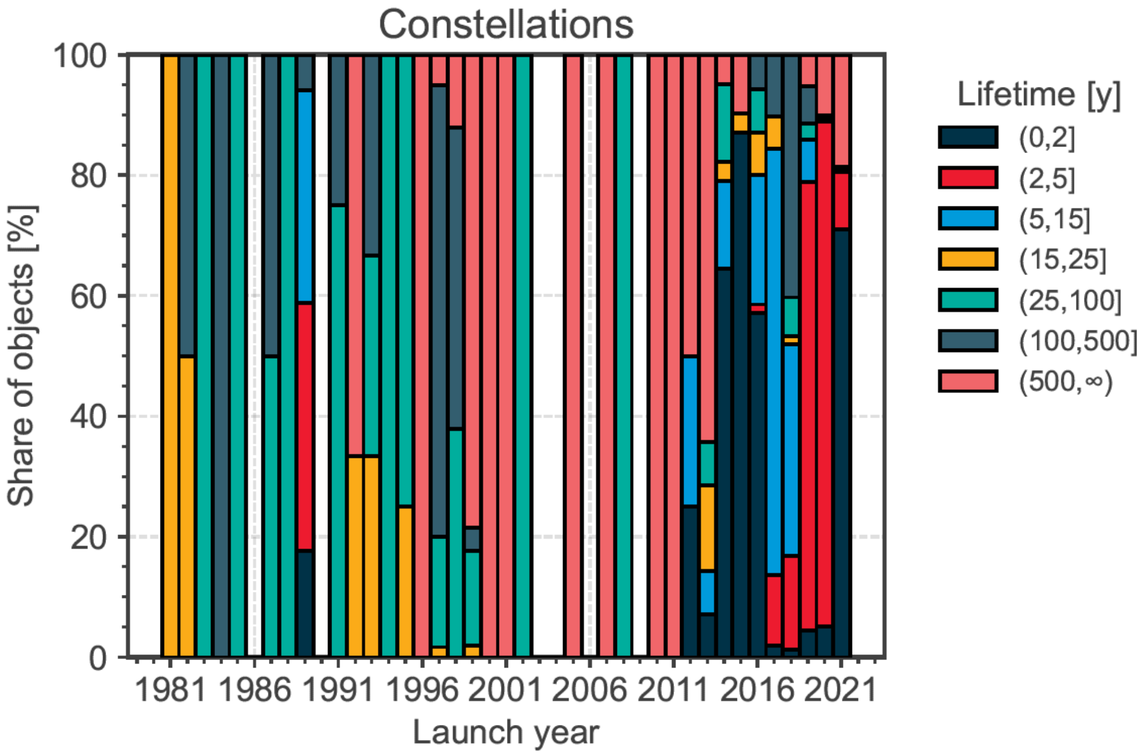


Figure 14: Lifetime distribution for constellation spacecraft. Adapted from (ESA, 2022).

Figure 15 shows the temporal change of the non-constellation vs. constellation ratio of launched payloads. Constellations first started to make an appearance in orbit just before 2000 and stand now for roughly 50% of all spacecraft in orbit. This will be of particular relevance for the later discussion concerning collision avoidance between non-constellation and constellation spacecraft.

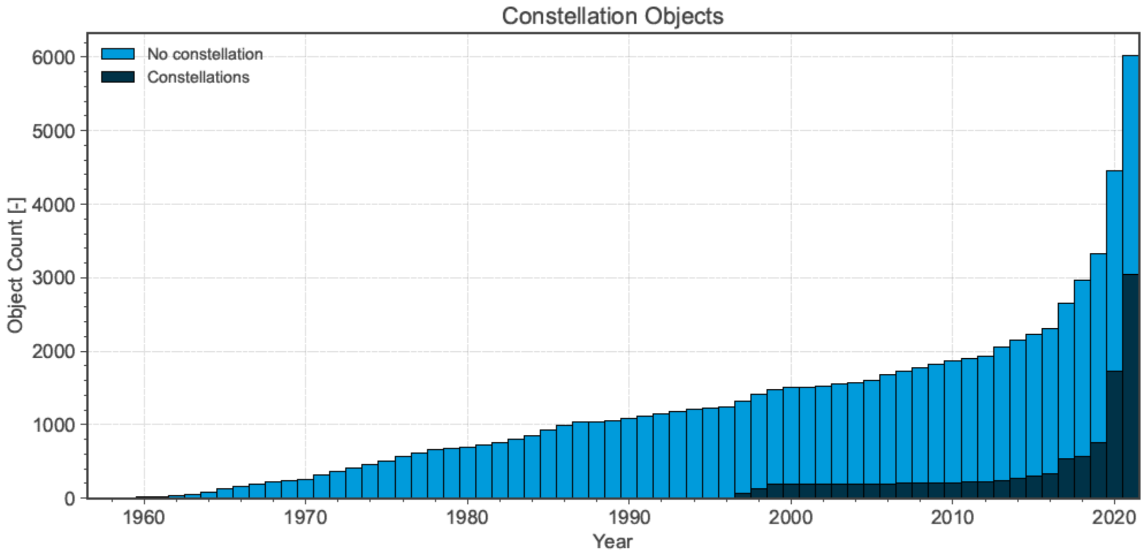


Figure 15: Development of non-constellation vs. constellation spacecraft in LEO. Adapted from (ESA, 2022).

Figure 16 gives on overview of the distribution of both catalogued objects and operational payloads vs. their respective orbital shells. Two distinct peaks can clearly be seen: around an altitude of approximately 550 kilometres and around 1200 kilometres. They correspond to the constellations of Starlink and its competitor Oneweb, respectively. In addition to the Starlink constellation, there is a significant amount of catalogued objects that Oneweb needs to transit in order to reach its designed orbit, something which later will discussed, in particular in the light of lacking traffic rules.

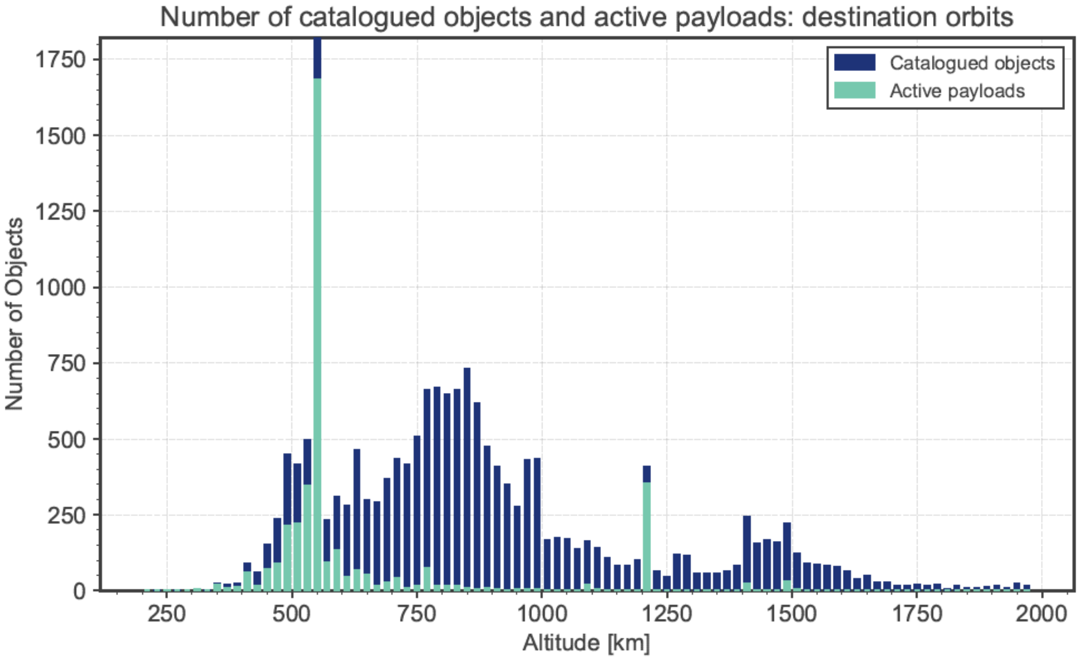


Figure 16: Number of active payloads and catalogued objects vs. altitude. Adapted from (ESA, 2022).

Another interesting question is the origin of registered objects. Figure 17 shows the distribution of space debris objects using the denotations introduced earlier. It is possible to make out three specific observations: (1) Until the year 2000, rocket fragmentations (RF) played a great role. (2) However, two distinct peaks for the years 2007 and 2009 resembling mainly payload fragmentations (PF) are clearly visible. As mentioned above, these peaks are related to the fragmentation of the Fengyun-1 spacecraft (as a result of anti-spacecraft rocket test) and the accidental collision between the Iridium-33 and Cosmos-2251 satellites. (3) A significant increase in payloads after approximately the year 2016 can be observed, i.e., operational spacecraft or calibration objects.

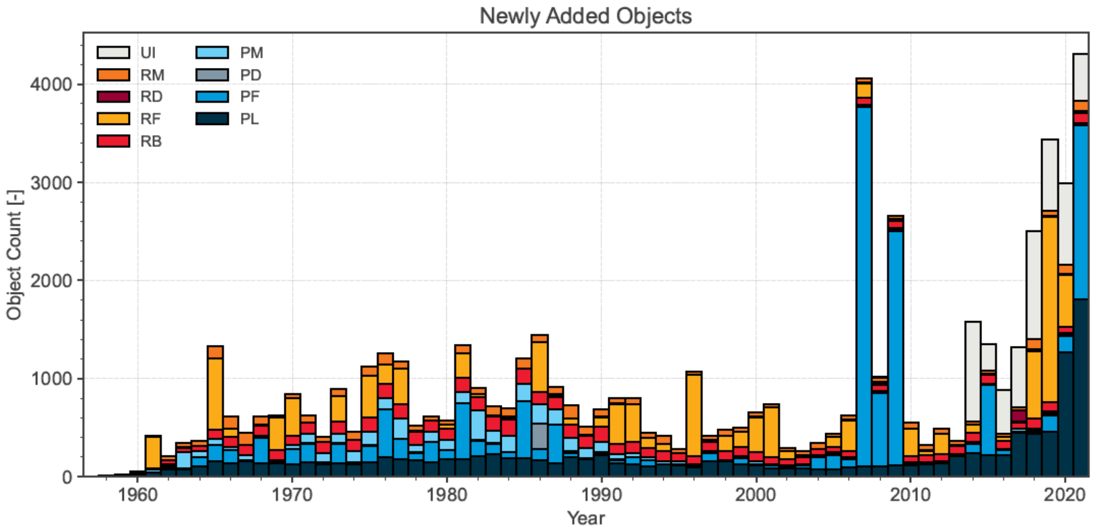


Figure 17: Distribution of newly added space debris by type. Note the significant contribution of PF (payload fragmentation) during the time period 2007-2009 which is associated with the fragmentation of Fengyun-1, Cosmos-2251 and Iridium-33. Adapted from (ESA, 2022).

Figure 18 gives an spatial overview of where which types of space debris and other objects have been inserted. By the far, most of the objects have been registered in LEO.

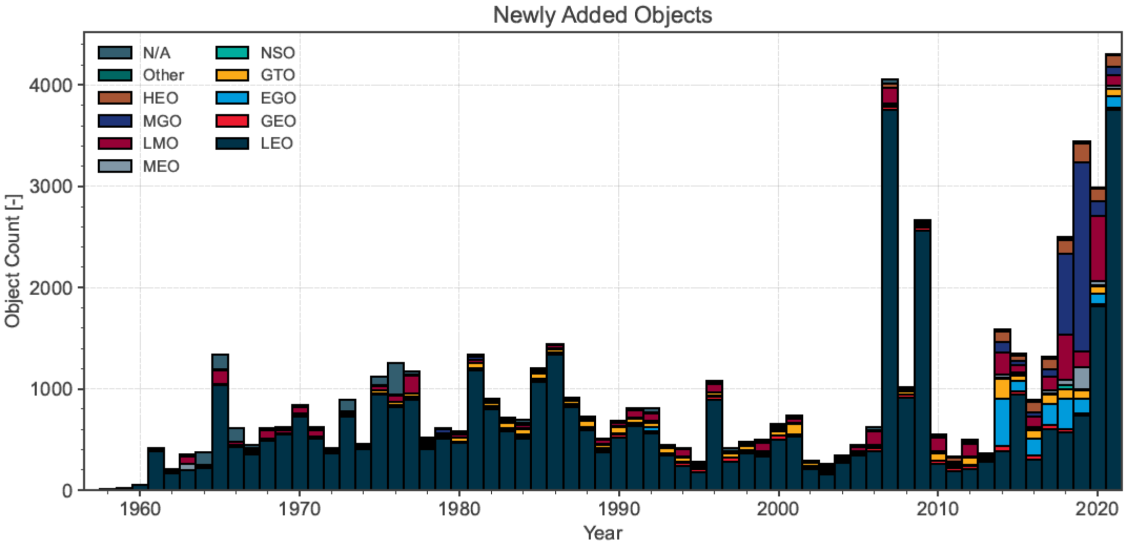


Figure 18: Distribution of newly added space debris by orbital shell. Note the significant contribution to LEO during the time period 2007-2009 which is associated with the fragmentation of Fengyun-1, Cosmos-2251 and Iridium-33. Adapted from (ESA, 2022).

3.1.3 Predicted future development

Until now, only past and present observations concerning space debris have been reported. In the light of the generation of mega-constellations such as Starlink, it is of great interest to study predicted future development. Figure 19 depicts the development of the number of objects in in LEO with a size larger than 10 centimetres for two scenarios. (1) The blue line indicates the predicted development if no further launches were to be undertaken, but present space debris generation rates (by means of fragmentations and explosions in the light of present disposal rates). Although in this scenario no launch activity will take place, the number of objects of this size will double from the 2025 to 2225. (2) The red line resembles to continuation of present activities. In this case, the number of objects will increase by a factor of 8, approximately.

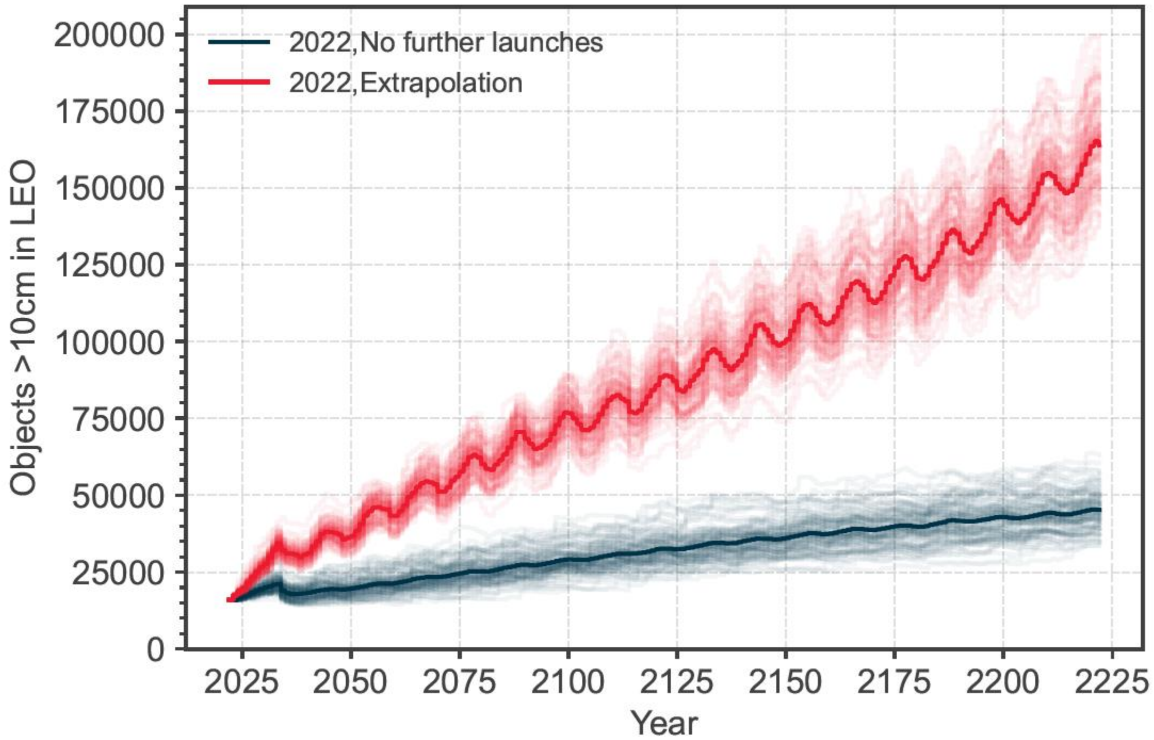


Figure 19: Predicted development of the number of objects larger than 10 cm in protected LEO over time, assuming no further launches (dark line) and an extrapolation of current behaviour concerning launch traffic, generation of space debris through explosions and fragmentation, and disposal rates. Adapted from (ESA, 2022).

The development presented in Figure 19 does of course describe the numbers for all orbits, but for the topic of this thesis LEO is of particular interest, since Starlink resides in this orbital shell. Figure 20 depicts the predicted number of catastrophic collisions in the protected LEO shell for the same two scenarios as given in Figure 19. (1) Assuming no further launches, one would still except a rise from zero catastrophic collisions at present to over 100 in 2225. (2) Assuming a continuation of today's trend, one would expect almost 900 catastrophic collisions in 200 years. In this case, one is approaching the Kessler syndrome rendering the protected LEO shell useless.

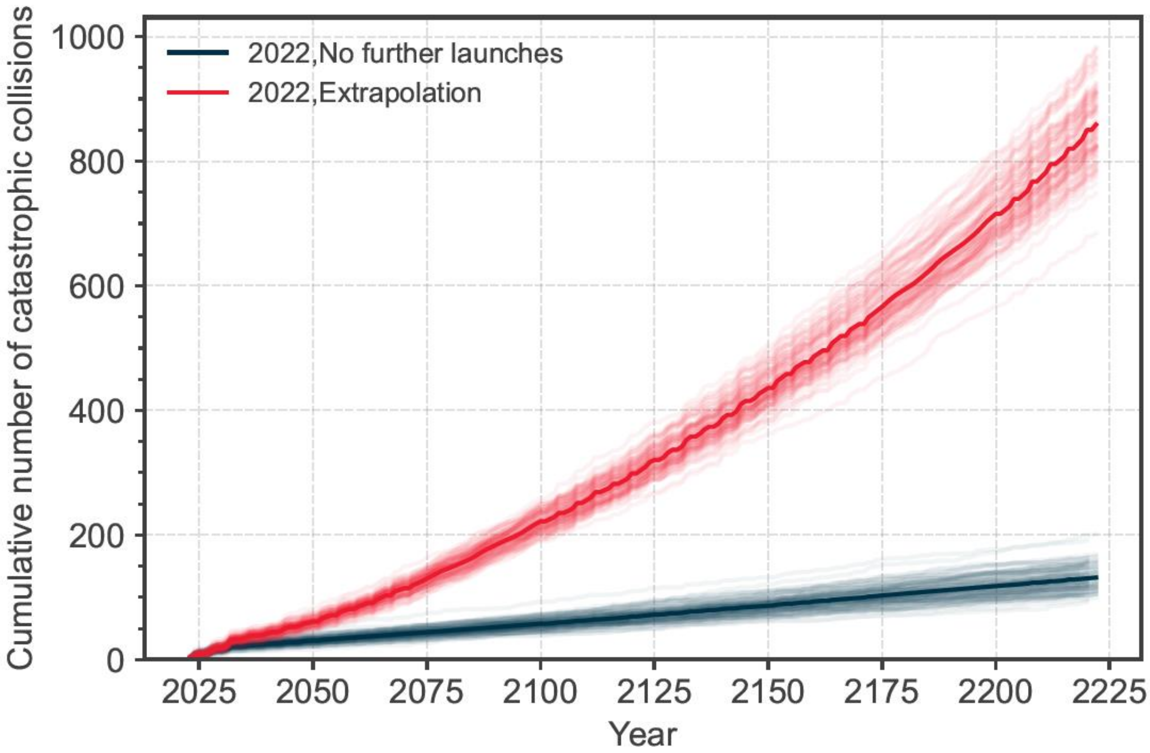


Figure 20: Predicted development of the number of cumulative collisions in protected LEO over time, assuming no further launches (dark line) and an extrapolation of current behaviour concerning launch traffic, generation of space debris through explosions and fragmentation, and disposal rates. Adapted from (ESA, 2022).

3.1.4. Mitigation of space debris

Since space debris can lead to the damage or destruction of spacecraft – and as a worst-case scenario by means of the Kessler-syndrome rendering access to certain orbital shells impossible, the majority of space debris mitigation objectives can be summarized as follows (ESA, 2022):

Limitation of space debris generated by normal operations: Both payloads and rocket parts should not lead to the release of space debris during nominal operations and should be designed in such a way. If space debris is created anyway, it should be limited as much as possible.

Decreasing the risk for in-orbit fragmentations: Break-up during operational phases should be avoided or minimized as much as possible. Intentional fragmentations should not be performed and if undertaken anyway, performed at low altitude such as the debris deorbits naturally.

Introduction of after-end-of-life disposal: One has declared two protected regions, namely Low Earth Orbit (LEO) and Geostationary Orbit (GEO) which should be void of space debris. Both payloads and rocket parts that have ended their operational phase should either actively be deorbited or moved to other lower regions in order to reduce their lifetime.

Minimizing the risk for in-orbit collisions: Tracking of active parts and space debris and established avoidance mechanisms.

3.2. Risk, risk analysis and risk management

3.2.1 Risk

Risk is a central concept in this thesis and it is essential to provide an understanding on how risk is defined, "measured" and dealt with. Although risk is an every-day concept, there is no universal definition. Equally essential is the concept of risk management which describes how an identified risk (through risk analysis) is managed.

In order to present a context for risk, risk analysis and risk management, Figure 21 depicts a general picture of a risk management process according to ISO Standard No. 31000:2018 (International Organization for Standardization, 2018).

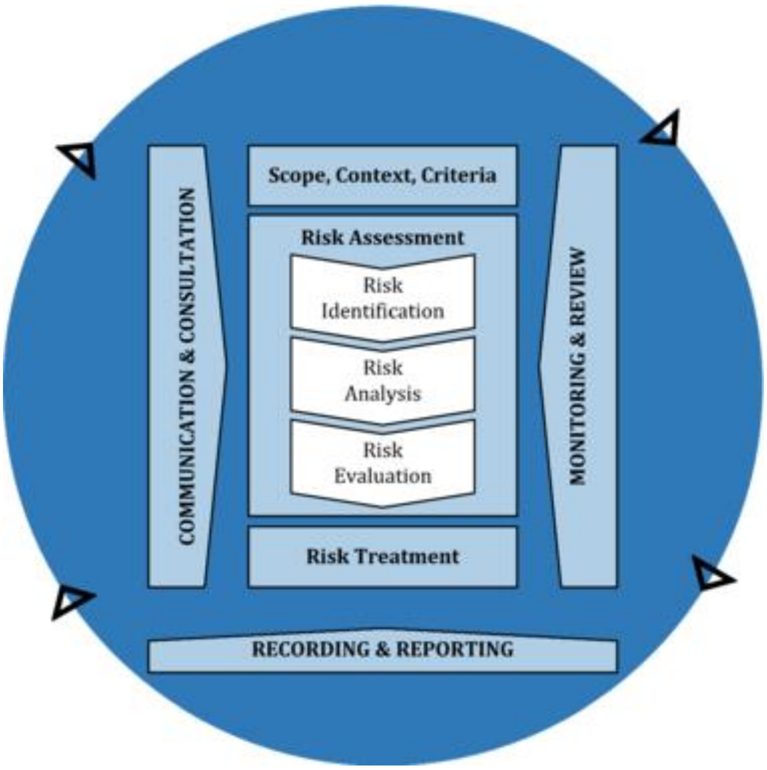


Figure 21: Risk management process. Adapted from (International Organization for Standardization, 2018).

Given that space debris is considered a risk for space activities, it is obvious that one wants to manage that risk, hence risk management. A risk management process based on ISO 31000:2018 is presented in Figure 21 and is divided into three main elements:

1. Planning,
2. Risk assessment,
3. Risk treatment.

After the context of the risk management process is established, the risk assessment consisting of risk analysis and risk evaluation is executed. This is followed by the risk treatment and later monitoring. Communication and consultation is ongoing in all subprocesses.

A risk R is often associated with an incident A and a consequence C. Simply speaking, one does not know if the incident A will occur and what its consequence will be. That implies that an uncertainty U is associated with A and C. The probability P describes how likely it is that the incident A with the specific consequence C will occur. P is based on the background knowledge K given a certain quality (strength) of this knowledge SK (Aven, Røed, & Wiencke, 2008).

Given these parameters, risk can be described by using a variety of different approaches. One possibility is to describe the risk R by a combination of C and P: (C, P). The probability P is a number and therefore will the risk be quantified (Aven, 2015). Another and related description of risk is the combination of C and U: (C, U). The uncertainty U is quantified by the possibility P in this context.

This quantification of risk does not provide an estimate of the quality of the assessment, since it does not provide any information on the quality of background knowledge which was used in order to obtain values for the possibilities or the risks that have been studied in the risk analysis. Thus, it is important to evaluate the background knowledge in order to assess the strength of the risk analysis (Aven, Røed, & Wiencke, 2008). In a worst case scenario, the knowledge a risk analysis is based upon is flawed and can result in an erroneous risk analysis. The faultiness can include (1) the wrong incidents A being investigated, (2) a false level of "resolution" of the analysis (too detailed or too superficial) or (3) obtained wrong numbers for possibilities. Simply speaking, one could claim that one always will receive an answer when performing a risk analysis but is important that the analysis is based on sufficient background knowledge in order to provide a CORRECT answer (Aven, 2015). If the knowledge K which the risk analysis is based upon and the related strength of this knowledge SK is included, then risk can be described by (C, P, SK, K).

The strength of the background knowledge SK can be assessed by answering a couple of questions:

- Are the prerequisites sufficient for the analysis?
- Is the data- and modelling foundation sufficient for the analysis?
- Do expert evaluations agree?
- Is there a good understanding of the included incidents? (Aven, Røed, & Wiencke, 2008)

Equally important is the knowledge level of the persons working with the risk analysis and the required actions. Do these persons have a sufficient level of knowledge? Aven, Røed, & Wiencke (2008) mention a few criteria that are connected to a strong knowledge level:

- The prerequisites that have been done are considered very reasonable.
- A wealth of reliable and relevant data/information is available.
- There is broad agreement between the experts.
- The involved phenomena are well understood, and the included models are known for providing predictions with the required accuracy.

A factor that is associated with uncertainty or background knowledge is the phenomenon of "black swans" (Aven, 2015). These are surprising incidents – surprising concerning to the knowledge which the risk analysis is based upon. These incidents are not included in the analysis and can have potentially great consequences.

3.2.2 Risk analysis

The risk analysis is one of the vital parts of a risk management process, mostly because it provides the input for the potential actions to be taken. It is of great importance to choose an appropriate level of detail for the risk analysis: An analysis which is too detailed may loose the overview of the larger picture. In addition, a breakdown into many "small" incidents which by themselves represent an acceptable risk level may blur the fact that their collective impact might still lead to an unacceptable risk at a larger level. In contrast, a risk analysis which is performed too broad may overlook certain smaller incidents which it is necessary to find actions against (Aven, Røed, & Wiencke, 2008).

3.2.3 Risk management

With the risk analysis in place, the ultimate goal is to manage the risk. Simply speaking, there are five possible ways to manage risk:

- Remove,
- Reduce,
- Optimise,
- Transfer,
- Keep the risk. (Aven, 2015)

A risk management process is an iterative process meaning that one constantly revisits the assumptions that have been made in the light of new knowledge and experience from handling the risk in question. The reason for this iterative approach is the possibility to adjust and improve risk management actions.

Risk management means decision-making under uncertainty (Aven, Røed, & Wiencke, 2008). This is emphasised by situations with high risk and great uncertainties leading to difficulties predicting the consequences of decisions being made. Aven (2015) presented a model which describes decision-making under uncertainty where the most important actions in this context are review and assessment by the management of the company. Especially the latter will be of great importance when discussing the potential impact of the utilization of emerging technologies where the human control factor is potentially excluded.

3.3. Legal framework for operations in space

Risk analysis and management are often related to legal issues. However, in contrast to, e.g., maritime travel, space is not as thoroughly regulated in the legal sense. Some treaties, legal principles and guidelines exist governing space traffic (UNCOPUOS, 2017):

United Nations treaties:

1. The “Outer 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 (1967),
2. The “Rescue Agreement”: Agreement on the Rescue of Astronauts, the Return of Astronauts and the Return of Objects Launched into Outer Space¸ entered into force 1968,
3. The “Liability Convention”: Convention on International Liability for Damage Caused by Space Objects, entered into force 1972,
4. The “Registration Convention”: Convention on Registration of Objects Launched into Outer Space, entered into force 1976.
5. The “Moon Agreement”: Agreement Governing the Activities of States on the Moon and Other Celestial Bodies, entered into force 1984 (UNCOPUOS, 2017).

These treaties are legally binding instruments. However, they are outdated and not likely to be updated.

In addition, several principles have been adopted by General Assembly of the United Nations:

1. The “Declaration of Legal Principles”: Declaration of Legal Principles Governing the Activities of States in the Exploration and Uses of Outer Space, 1963,
2. The “Broadcasting Principles”: The Principles Governing the Use by States of Artificial Earth Satellites for International Direct Television Broadcasting, 1982,
3. The “Remote Sensing Principles”: The Principles Relating to Remote Sensing of the Earth from Outer Space, 1986,
4. The “Nuclear Power Sources Principles”: The Principles Relevant to the Use of Nuclear Power Sources in Outer Space, 1992,
5. The “Benefits Declaration”: The Declaration on International Cooperation in the Exploration and Use of Outer Space for the Benefit and in the Interest of All States,

Taking into Particular Account the Needs of Developing Countries, 1996 (UNCOPUOS, 2017).

The most relevant documents are non-legally binding instruments such as guidelines and best practices. The two most common are the (1) *IADC Space Debris Mitigation Guidelines* (IADC, 2002) and (2) *Guidelines for the long-term sustainability of outer space activities* (UNCOPUOS, 2021). Many nations attempt to deal with the non-legally binding nature of these guidelines by incorporating them into national space law where they in their turn become legally binding. However, only a few countries have introduced national space laws, thus limiting this approach.

Further initiatives include ratings and industrial guidelines. Firstly, the *Space Sustainability Rating* has been introduced in 2022 (Space Sustainability Rating, 2023). The rating assesses the sustainability of a space mission and returns a score very much alike a rating for, e.g., a refrigerator. Amongst other collision avoidance, trackability, data sharing etc are rated. Although this is neither a legally binding instrument, it provides an incentive to spacecraft operators to follow more sustainable procedures. Secondly, the space industry itself is becoming more and more involved in policy making which resulted, e.g., in the *Space Industry Debris Mitigation Recommendations* (Foust, 2023). The novelty of these recommendations is that they were incorporated by the space industry itself, not regulatory bodies. That being said, it is interesting to note that the content of the industrial recommendations is rather similar to regulatory guidelines.

3.4. Legal and regulatory considerations related to emerging technologies such as Artificial Intelligence (AI)

When addressing new technologies and their potential impact on operations and thus resulting risk assessments, “it is not so much the technology itself, but a particular product or service making use of the technology, that poses a risk.” (European Commission, 2019, p. 11). This explains the reason for looking at liability. Risk includes consequence and a possible consequence is being liable for a damage. This in its turn can mean monetary risk and reputation.

Also, important stakeholders have acknowledged the possible challenges arising from the introduction of new technologies: “As with many disruptive innovations, AI presents risks and challenges that could affect its adoption, and therefore our business. AI algorithms may be flawed. Datasets may be insufficient or contain biased information. Inappropriate or controversial data practices by Microsoft or others could impair the acceptance of AI solutions. These deficiencies could undermine the decisions, predictions, or analysis AI applications produce, subjecting us to competitive harm, legal liability, and brand or reputational harm.“ (UNITED STATES SECURITIES AND EXCHANGE COMISSION, 2018)

In a report from the Expert Group on Liability and New Technologies of the European Commission (European Commission, 2019) the issue of “Liability for Artificial Intelligence and Other Emerging Digital Technologies” is discussed. It should be noted that this report focuses on actions taking place on Earth. Nevertheless, the report is relevant to this thesis, since collision avoidance actions may be decided upon on Earth, e.g., by a company offering a collision avoidance service, or in space, e.g., by a satellite utilizing onboard monitoring and computing facilities. The report establishes the fact that basic protection for liability claims are given in the member states of the European Union – even if a claim is failed for compensation following an incident involving new technologies. The following key findings of the report relevant for this thesis are listed below:

- “A person using a technology which has a certain degree of autonomy should not be less accountable for ensuring harm than if said harm had been used by a human auxiliary”. (European Commission, 2019, p. 3)
- “Manufacturers of products or digital content incorporating emerging digital technology should be liable for damage caused by defects in their products, even if the defect was

caused by changes made to the product under the producer's control after it had been placed on the market". (European Commission, 2019, p. 3)

- "For situations exposing third parties to an increased risk of harm, compulsory liability insurance could give victims better access to compensation and protect potential tortfeasors against the risk of liability." (European Commission, 2019, p. 4)
- "Where a particular technology increases the difficulties of proving the existence of an element of liability beyond what can be reasonably expected, victims should be entitled to facilitation of proof." (European Commission, 2019, p. 4)
- "Emerging digital technologies should come with logging features, where appropriate in the circumstances, and failure to log, or to provide reasonable access to logged data, should result in a reversal of the burden of proof in order not be to the detriment of the victim." (European Commission, 2019, p. 4)
- "It is not necessary to give devices or autonomous systems a legal personality, as the harm these may cause can and should be attributable to existing persons or bodies." (European Commission, 2019, p. 4)
- "While existing rules on liability offer solutions with regard to the risks created by emerging digital technologies, the outcomes may not always seem appropriate, given the failure to achieve:

  > (a) a fair and efficient allocation of loss, in particular because it could not be attributed to those …. whose objectionable behaviour caused the damage; or
  > … who were in control of the risk that materialised … or who were cheapest cost avoiders or cheapest takers of insurance." (European Commission, 2019, p. 5)

- "It is therefore necessary to consider adaptations and amendments to existing liability regimes, bearing in mind that, given the diversity of emerging digital technologies and the correspondingly diverse range of risks these may pose, it is impossible to come up with a single solution suitable for the entire spectrum of risks." (European Commission, 2019, p. 5)
- "Comparable risks should be addressed by similar liability regimes, existing differences among these should ideally be eliminated. This should also determine which losses are recoverable to what extent." (European Commission, 2019, p. 5)
- "Strict liability is an appropriate response to the risks posed by emerging digital technologies, if, for example, they are operated in non-private environments and may typically cause significant harm. Strict liability should lie with the person who is in control

of the risk connected with the operation of emerging digital technologies and who benefits from their operation (operator). If there are two or more operators, in particular (a) the person primarily deciding on and benefitting from the use of the relevant technology (frontend operator) and (b) the person continuously defining the features of the relevant technology and providing essential and ongoing backend support (backend operator), strict liability should lie with the one who has more control over the risks of the operation." (European Commission, 2019, p. 6)

- "If harm is caused by autonomous technology used in a way functionally equivalent to the employment of human auxiliaries, the operator's liability for making use of the technology should correspond to the otherwise existing vicarious liability regime of a principal for such auxiliaries." (European Commission, 2019, p. 7)
- "The more frequent or severe potential harm resulting from emerging digital technology, and the less likely the operator is able to indemnify victims individually, the more suitable mandatory liability insurance for such risks may be." (European Commission, 2019, p. 8)

One can elaborate on the problems arising from proofing causality in connection with potential liability issues: "However, the less evident the sequence of events was that led to the victim's loss, the more complex the interplay of various factors that either jointly or separately contributed to the damage, the more crucial links in the chain of events are within the defendant's control, the more difficult it will be for the victim to succeed in establishing causation without alleviating their burden of proof." (European Commission, 2019, p. 20). The more new technologies are involved in this process, the harder it will be to prove their responsibilities in possible incidents. Ultimately, "It is even harder if the algorithm suspected of causing harm has been developed or modified by some AI system fuelled by machine learning and deep learning techniques, on the basis of multiple external data collected since the start of its operation. Even without changes to the original software design, the embedded criteria steering the collection and analysis of data and the decision-making process may not be readily explicable and often require costly analysis by experts." (European Commission, 2019, p. 20). If an AI thus modifies an algorithm, this modification can change towards the next iteration, since AIs themselves can be updated or their operational procedures will change due to data or input that the AI itself will collect and implement into its system: "Not only may such data be flawed in itself, but the processing of otherwise correct data may also be imperfect. The latter may be due to original defects in designing the handling of data, or the consequence of distortions of the system's self learning abilities due to the bulk of data collected, whose randomness may lead the AI system in question to misperceive and miscategorise subsequent input.” (European Commission, 2019, p. 21). Problems of causation can be considered analogue to problems in risk management, in particular when addressing potential causes leading to an incident and the respective barriers which can be implemented to decrease probability of the incident happening or the consequences if the incident happens anyway: “Even if something is proven to have triggered the harm (for example, because an autonomous car collided with a tree), the real reason for it is not always equally evident. The car may have been poorly designed (be it its hardware, pre-installed software, or both), but it may also have either misread correct, or received incorrect, data, or a software update done by the original producer or by some third party may have been flawed, or the user may have failed to install an update which would have prevented the collision, to give just a few examples, not to mention a combination of multiple such factors.” (European Commission, 2019, p. 22).

In terms of risk management, the role of human control is significant: “Emerging digital technologies make it difficult to apply fault-based liability rules, due to the lack of well established models of proper functioning of these technologies and the possibility of their developing as a result of learning without direct human control.” (European Commission, 2019, p. 23). In particular, the last part of the citation is of great interest, since AI can develop themselves without human control. In addition: “Legal requirements have to be distinguished from industry standards (or practices) not yet recognised by the lawmaker.“ (European Commission, 2019, p. 23).

The same report also discusses the matter of complexity, opacity, openness, autonomy, predictability, data-driveness and vulnerability. *Complexity* is of importance in terms of risk management of a complex system. New technologies are characterized by various parts that interact, including digital parts such as AI. It can be difficult to determine the cause of a fault in such a complex architecture or even if different architectures interact. Thus, the complexity of interaction between different architectures is potentially problematic, but also the complexity within the algorithms involved itself. *Opacity* stands for the possibility of being able to understand or comprehend the processes taking place within the new technology. This is being worsened by the fact that already difficult to understand technologies becoming even more difficult to understand due to their ability to develop themselves independently. Often, one relates to new technologies as black boxes. This is somewhat related to *openness* which refers to the fact that emerging technologies are not finished when they are implemented, but they will continuously be improved and thus changed. In order to facilitate this continuous development, new technologies need to be open for input from other systems or data sources. This of course influences risk management decisions and will be discussed later. *Autonomy* resembles the most challenging issue in relation to risk management, namely that new technologies operate entirely on their own with no - or little – human control. “They are themselves capable of altering the initial algorithms due to self-learning capabilities that process external data collected during the operation. The choice of such data and the degree of impact it has on the outcome is constantly adjusted by the evolving algorithms themselves.“ (European Commission, 2019, p. 33). *Predictability* concerns the difficulty of anticipating a precise reaction of the new technology, in particular when taking into account that new technology are supposed to develop themselves and thus providing new types of responses that may differ from the pre-defined reaction. A new reaction to a given accident means of course an altered risk management. *Data-driveness* describes the dependency of new technologies on external information, data that is either provided as an input from the outside or build-in sensors. These data may “be flawed or missing altogether, be it due to communication errors or problems of the external data source, due to flaws of the internal sensors or the built-in algorithms designed to analyse, verify and process such data.” (European Commission, 2019, p. 33). The last point refers to cybersecurity, namely *vulnerability*: a constant interaction of the new technology with the outside through updates or simply the possibility of providing input with outside data.

Whereas AI is relatively new in space operations, it has been employed in aviation extensively (Kashyap, 2019): The author of this study describes a procedure for AI’s in aviation consisting of four steps: retrieve, reuse, revise and retain. The retrieve step is looking for old incidents that are similar to the new case. Here, different old cases might provide the information needed to evaluate a response to the new case. However: “This adjustment process requires space information and is exceptionally unpredictable.“ (Kashyap, 2019). It is furthermore suggested that the solution is evaluated: “The assessment procedure is regularly performed by numerical instruments. This procedure additionally requires area information, as does the repair process. The outcome is a ‘tried/repaired case’, or an affirmed answer for the issue. The learning perspective is actualized by including data about the affirmed answer for the case-base.“ (Kashyap, 2019).

Abovementioned material takes base in the challenges that are faced when utilizing AI’s on Earth. However, this is further complicated by lack of regulations in space in addition to the existence of a variety of guidelines, national space laws etc. (Pagallo, Bassi, & Durante, 2023). The authors state clearly that “the unique features of AI technologies … will require the adoption of new legal standards”. (Pagallo, Bassi, & Durante, 2023). Since legal standards often are related to liability issues which in their turn are related to risk management, this raises the questions if the introduction of AI technologies making their own decisions also should imply the adoption of new risk management standards (such as is done, e.g, for complex systems). “The increasing use of AI systems that augment or replace analysis and decision-making by humans, making sense of huge streams of data or defining and modifying decision-making rules autonomously, is a sound example of how third-order technology works. The effects of such autonomy, from a legal viewpoint, can be nonetheless controversial.” (Pagallo, Bassi, & Durante, 2023). This last statement clearly shows the challenges related to decision-making of AI systems – a problem which propagates further into risk management processes. An often-cited solution to this problem is “to determine and enforce a “meaningful human control” (MHC) over the entire technological cycle and system functioning. … There is no such alternative between human control over autonomous technologies on the one hand and, on the other, the further development of autonomous systems that take crucial decisions by themselves. MHC is the subject of academic courses on the risks and threats brought forth by possible losses of control of AI systems, so as to keep them in check.” (Pagallo, Bassi, & Durante, 2023). The authors elaborate further claiming that “the aim should be to attain a fair balance between human control and AI autonomy. In the field of space technologies, it is even obvious that the autonomy of AI systems should be strengthened, for example, for the management of complex constellations that should reduce the workload of ground operators or for the guidance, navigation, and control of rovers that should remove human scheduling errors and find their way across unspecified fields, navigating around obstacles either on the Moon or on Mars. The balance that shall be struck between human control and AI autonomy can be measured by the degree of “social acceptability” concerning the risks inherent in the automation process, as well as the level of social and political cohesion that regards the values and principles that are at stake with the development of increasingly autonomous technologies …” (Pagallo, Bassi, & Durante, 2023). Note that these authors explicitly mention the risk connected to the automation process itself something which will be discussed later in this thesis. The authors go even further and claim that “the law should reinforce current strict liability doctrines for the risks posed by third-order technologies.” (Pagallo, Bassi, & Durante, 2023).

The topic of human control or feedback is indeed a widely discussed issue: " …as more and more is generated by unsupervised learning and the AI itself, the demand for more comprehensive and sophisticated validation of results will increase. This will require a new set of tasks that will be based on manual work—human user feedback will be essential to the language models and the generative AI algorithms they feed. There is nothing quite as dangerous in the workplace as overconfidence based on "unconscious incompetence." Without "human in the loop," AI at its most powerful looks like a case of "unconscious incompetence" writ large. Machines simply don't know what they don't know." (Omar, 2023). This relates to the definition of risk mentioned in an earlier chapter where uncertainty and background information are essential components when quantifying or describing risk. The author suggest that there are several ways of involving humans into the AI architecture, e.g., reinforcement learning with human feedback (RLHF). That point is here that RLHF is a special form of reinforcement learning (RL) but in addition combined with human response with the goal of providing improvement of the AI operation. It is mentioned that the RLHF approach can provide an alternative way of feedback when normal RL operations are not sufficient for complex operations, such as collision avoidance with thousands of spacecraft. If RLHF is utilized, it is of great importance that the correct way of providing human feedback is chosen, such as rankings, ratings or quality markers. Despite the fact the RLHF is already employed in a variety of areas such chatbots, financial trading or autonomous vehicles, "as generative AI is applied across other use cases, this will only multiply the need for RLHF, particularly for more mission-critical applications where accuracy is essential for both productivity as well as safety." (Omar, 2023).

The importance of quality-managing the work of AI has also been mentioned by other authors: "The retraining phase is optional and should take place in case the classifier does not identify correctly the distances. For the retraining phase, another number of images should be collected for each space dimension and the to be labelled as it is described in the training phase. Then, a new classifier more accurate is created." (Makris & Aivaliotis, 2022).

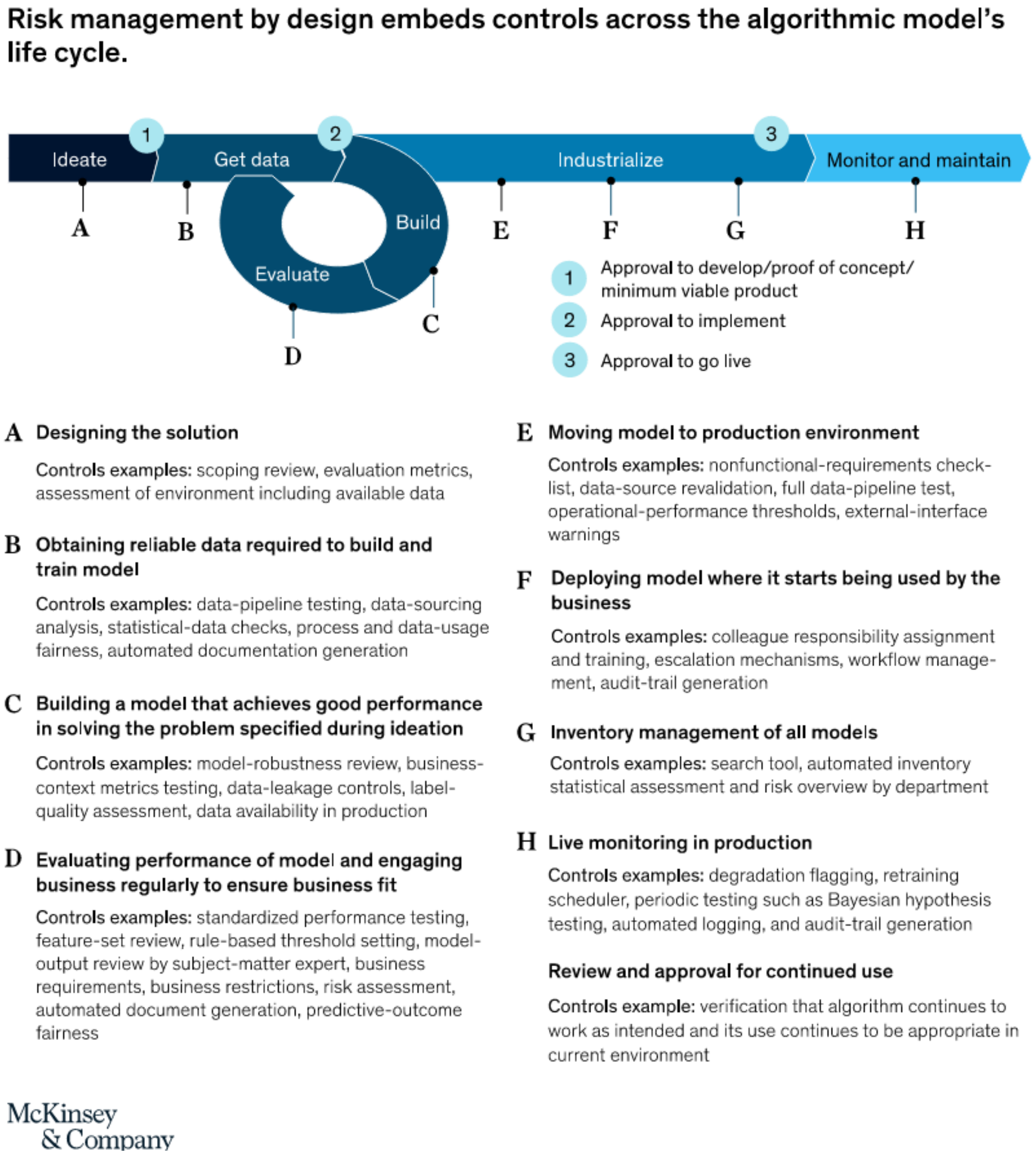
Until now, focus has been drawn on the implications and risks new technologies bring with them. It has been indicated that AI has been used for autonomous operations, thus making decisions. Hence, AI can assist in risk management processes, but at the same time the inclusion poses a threat to traditional risk management approaches: "In this complex and fast-moving environment, traditional approaches to risk management may not be the answer … Risk management cannot be an afterthought or addressed only by model-validation functions such as those that currently exist in financial services. Companies need to build risk management directly into their AI initiatives, so that oversight is constant and concurrent with internal development and external provisioning of AI across the enterprise." (Baquero, Burkhard, Govindarajan, & Wallace, 2020). The authors note that there are three factors which explain why it is challenging to deal with risks related to AI:

- New risks and novel responsibilities,
- Problems with identifying the level of AI's inclusion in a company's structure,
- Development of own AI risk management procedure.

The authors explain why "traditional model risk management (MRM) is insufficient" (Baquero, Burkhard, Govindarajan, & Wallace, 2020):

- MRM assumes static models, i.e., between assessments the assumptions and boundary conditions for the risk assessment are not changing. AI however, will constantly change.
- MRM workflows are often prone to length review processes before a new development can be implemented. AI development happens at a much faster rate and cannot be aligned with MRM work patterns.
- MRM often related to classical risk types which do not include the new risk types arising from the introduction of AI.
- AI are often included in larger systems (preferably also AI). Thus, these combined systems become more complex and more difficult to risk-assess, also in terms of third-party risks (vendor).

The authors suggest that the inclusion of risk management into the AI development as a solution: "To tackle these challenges without constraining AI innovation and disrupting the agile ways of working that enable it, we believe companies need to adopt a new approach to risk management: derisking AI by design. Risk management by design allows developers and their business stakeholders to build AI models that are consistent with the company's values and risk appetite. Tools such as model interpretability, bias detection, and performance monitoring are built in so that oversight is constant and concurrent with AI development activities and consistent across the enterprise. In this approach, standards, testing, and controls are embedded into various stages of the analytics model's life cycle, from development to deployment and use …" (Baquero, Burkhard, Govindarajan, & Wallace, 2020). Figure 22 describes the schematics of the suggested approach.



Risk management by design embeds controls across the algorithmic model's life cycle.

A Designing the solution

Controls examples: scoping review, evaluation metrics, assessment of environment including available data

B Obtaining reliable data required to build and train model

Controls examples: data-pipeline testing, data-sourcing analysis, statistical-data checks, process and data-usage fairness, automated documentation generation

C Building a model that achieves good performance in solving the problem specified during ideation

Controls examples: model-robustness review, business-context metrics testing, data-leakage controls, label-quality assessment, data availability in production

D Evaluating performance of model and engaging business regularly to ensure business fit

Controls examples: standardized performance testing, feature-set review, rule-based threshold setting, model-output review by subject-matter expert, business requirements, business restrictions, risk assessment, automated document generation, predictive-outcome fairness

E Moving model to production environment

Controls examples: nonfunctional-requirements checklist, data-source revalidation, full data-pipeline test, operational-performance thresholds, external-interface warnings

F Deploying model where it starts being used by the business

Controls examples: colleague responsibility assignment and training, escalation mechanisms, workflow management, audit-trail generation

G Inventory management of all models

Controls examples: search tool, automated inventory statistical assessment and risk overview by department

H Live monitoring in production

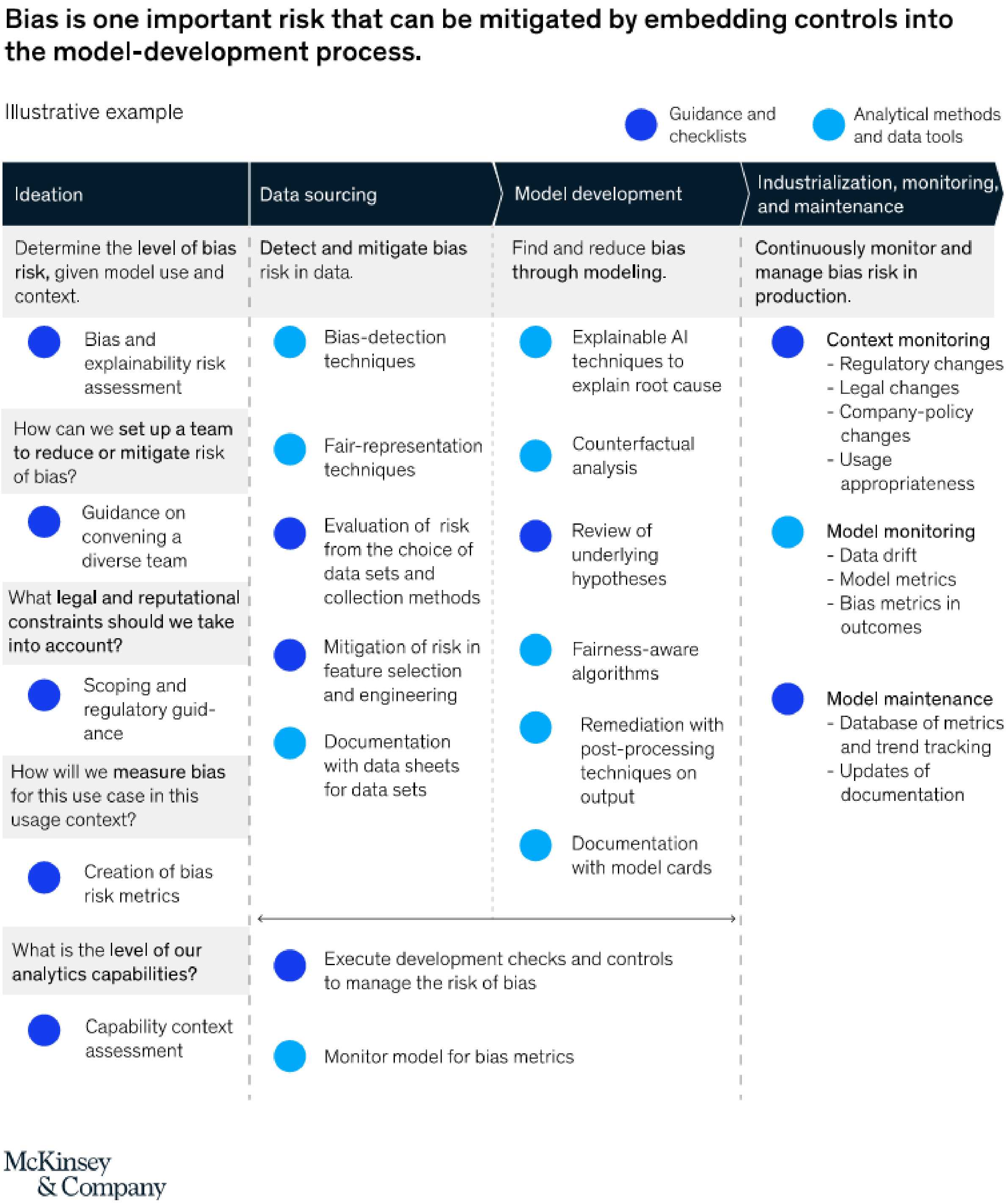
Controls examples: degradation flagging, retraining scheduler, periodic testing such as Bayesian hypothesis testing, automated logging, and audit-trail generation

Review and approval for continued use

Controls example: verification that algorithm continues to work as intended and its use continues to be appropriate in current environment

Figure 22. A suggested model for risk management by design in order cope with risks inherent to emerging technologies. Adapted from (Baquero, Burkhard, Govindarajan, & Wallace, 2020).

The authors point out that if companies acquire AI products from vendors it is of outmost importance that control groups, business groups and vendors interact early in the process in order to identify risks and their related risk management processes. To continue that thought it is mentioned that – in order to reduce costs – a desired approach is to include "risk identification and assessment, together with associated control requirements, directly into the development and procurement cycles. This approach also speeds up pre-implementation checks, since the majority of risks have already been accounted for and mitigated." (Baquero, Burkhard, Govindarajan, & Wallace, 2020). This last statement is of great interest for this thesis and will be discussed later. The authors provide an example of one of the risks related to AI, namely biasing in data input and methodology. In order to handle this type of risk, one can implement a variety of checks in the analytics-development process, see Figure 23.



Bias is one important risk that can be mitigated by embedding controls into the model-development process.

Illustrative example

● Guidance and checklists
🔵 Analytical methods and data tools

Ideation

Determine the level of bias risk, given model use and context.

- ● Bias and explainability risk assessment

How can we **set up a team to reduce or mitigate** risk of bias?

- ● Guidance on convening a diverse team

What **legal and reputational constraints should we take into account?**

- ● Scoping and regulatory guidance

How will we **measure bias** for this use case in this usage context?

- ● Creation of bias risk metrics

What is the level of our **analytics capabilities?**

- ● Capability context assessment

Data sourcing

Detect and mitigate bias risk in data.

- 🔵 Bias-detection techniques
- 🔵 Fair-representation techniques
- ● Evaluation of risk from the choice of data sets and collection methods
- ● Mitigation of risk in feature selection and engineering
- 🔵 Documentation with data sheets for data sets

Model development

Find and reduce **bias through modeling.**

- 🔵 Explainable AI techniques to explain root cause
- 🔵 Counterfactual analysis
- ● Review of underlying hypotheses
- 🔵 Fairness-aware algorithms
- 🔵 Remediation with post-processing techniques on output
- 🔵 Documentation with model cards

- ● Execute development checks and controls to manage the risk of bias
- 🔵 Monitor model for bias metrics

Industrialization, monitoring, and maintenance

Continuously monitor and manage bias risk in production.

- ● **Context monitoring**
  - Regulatory changes
  - Legal changes
  - Company-policy changes
  - Usage appropriateness
- 🔵 **Model monitoring**
  - Data drift
  - Model metrics
  - Bias metrics in outcomes
- ● **Model maintenance**
  - Database of metrics and trend tracking
  - Updates of documentation

McKinsey & Company

Figure 23. Suggested model in order to mitigate risk associated with biases. Adapted from (Baquero, Burkhard, Govindarajan, & Wallace, 2020).

The authors provide four steps in order to mitigate the risk of bias in an AI:

- Ideation: This resembles the early phase of a risk analysis. The purpose of the AI is analysed and its associated risks catalogued: "An early understanding of the risks of the use case will help define the appropriate requirements around the data and methodologies. All the stakeholders ask, "What could go wrong?" and use their answers to create appropriate controls at the design phase." (Baquero, Burkhard, Govindarajan, & Wallace, 2020)
- Data sourcing: An essential of AI's is the input of data. As mentioned earlier in this thesis, this resembles a significant risk. One needs to define which data sets that cannot be used and what kind of biases the input of old data represents which can propagate within the AI system.
- Model development: Since space industry represents are commercial market, only a few methods are present in the AI market – something which is decided by a few companies only and thus biases AI development. Interestingly, the authors claim that "one black-box methods will not be allowed in high-risk use cases" (Baquero, Burkhard, Govindarajan, & Wallace, 2020) – something which will be discussed later concerning the black-box approach that Starlink utilizes in terms of collision avoidance.
- Monitoring and maintenance: Here, it is defined how the performance of the AI is monitored – something which will be influenced by the risk associated with the AI and how often the AI is updated or used. If self-learning is included, the monitoring needs to be executed on a more automatic basis.

Keeping these four steps in mind, it is stated that the fast-evolving nature of AI's "make large-scale and rapid deployment incredibly difficult for traditional risk managers to support. To adjust, they will need to integrate their review and approvals into agile or sprint-based development approaches, relying more on developer testing and input from analytics teams, so they can focus on review rather than taking responsibility for the majority of testing and quality control. Additionally, they will need to reduce one-off "static" exercises and build in the capability to monitor AI on a dynamic, ongoing basis and support iterative development processes. But monitoring AI risk cannot fall solely on risk managers. Different teams affected by analytics risk need to coordinate oversight to ensure end-to-end coverage without overlap, support agile ways of working, and reduce the time from analytics concept to value …" (Baquero, Burkhard, Govindarajan, & Wallace, 2020). The key-takeaways here are the multi-team approach for risk-management and the necessary monitoring of AI's in order to utilize and support the AI's further development.

In order for this multi-team risk-management approach to work properly, it essential that the different teams expand their knowledge and horizons so that they can complement each other. This is often regarded as a rather challenging approach which is characterised by the fact that the different groups "do not speak the same language". In particular, risk managers need to obtain sufficient knowledge in data science, and, on the opposite, data scientists need to expand their knowledge in risk management processes. This of course requires a departure from the classical silo-thinking and brings with it a necessary organizational change. Most importantly, it is required to monitor and interpret the quality of the AI's operations.

4. Methodology

The topic of this thesis represents a great challenge, since it resides in an area with great legal ambiguities combined with an extremely fast technological development which puts question marks on established risk management procedures. Although representing such a challenge, it is still obvious from the discussion in the next chapter that this is a relevant topic which needs attention.

In order to pave the way for the following discussion, the author considered it of importance to provide the theoretical background in the previous chapter. It is essential to understand WHY space debris is such a great problem for humanity and HOW the nature of the problem has changed over time – mainly with the introduction of new space actors and in particular mega-constellations, introducing thousands of spacecraft into orbit. The most common tool to deal with the increase of objects in orbit is collision avoidance which represents a risk management tool. Before the new space era, conjunctions were rare and, thus, collision avoidance represented a classical risk management procedure as described in the previous chapter. Legal challenges were not of great importance, since incidents were extremely rare and, thus, it was not needed to address possible updates of the legal framework. With the introduction of the new space era, the number of spacecraft and their operators increased significantly. This highlighted the legal challenges and put stronger demands on collision avoidance. Classical manual collision avoidance procedures will soon not be possible anymore and already now a number of operators relies on new technologies such as AI to cope with the increase in number of collision warnings and the amount of data these are based on. Thus, it was important to focus on the legal background of "classical space operations", but also on the legal aspect of AI operations and their implications for risk management operations. Ultimately, it is shown that classical risk management procedures are not sufficient to cope with AI operations, basically since AI make decisions on their own and because of the non-static problem situation. The latter includes the constant change in orbits of the spacecraft in question and the constant change of the AI itself. This highly dynamic environment requires a new approach in terms of risk management, also addressing the risk the usage of AI itself represents. Thus, the legal background for space operations and classical risk management understanding in combination with considerations on legal aspects related to emerging technologies were presented in the previous chapter. Finally, suggestions on how risk management procedures can be adapted to the challenges provided by the use of AI were presented. This approach was chosen in order to provide the basis for the following discussion. As mentioned, the topic covers a rather wide range from space operations, legal aspects, risk management etc. However, the author feels confident that the theoretical background has been covered sufficiently in order to provide enough material for the discussion. Covering such a wide range of material requires a certain selection, but the well-documented nature of space operations, collision avoidance and legal aspects proved it feasible to confine the description of the theoretical aspects to an acceptable level.

It was originally planned to perform surveys among spacecraft operators and service providers for collision avoidance. However, the space industry exhibits a huge variety of transparency and opacity concerning the operations – something which will be demonstrated in the following chapter. Such diverging approaches to data and information sharing among the possible survey targets would inherently have led to a strong bias in the results. Therefore, it has been decided to make use of freely available data and information – either in digital or paper form. When performing such a “literature search” it is of course of great importance to shed light on the problem from all possible angles. This is itself represents a form of a bias, because all information available will not per definition represent all possible explanations equally, since marketing, economic and political forces can enhance a certain side of the problem. In addition, one needs to confine the literature search to an amount which is manageable within the scope of this thesis.

5. Empiri

Starlink describes itself as “the world’s first and largest satellite constellation using a low Earth orbit to deliver broadband internet capable of supporting streaming, online gaming, video calls and more. Leveraging advanced satellites and user hardware coupled with our deep experience with both spacecraft and on-orbit operations, Starlink delivers high-speed, low-latency internet to users all over the world.” (Starlink, 2022a).

Starlink inserts its spacecraft into extremely low orbit in order to further mitigate risks. Initially, the spacecraft are deployed into altitudes of less than 350 kilometres and later lifted to the operational altitude of circa 550 kilometres. The insertion into extremely low orbit enables it for Starlink to easily deorbit spacecraft that do not meet operational standards. However, Starlink itself acknowledges that insertion and flying at these extremely low altitudes is challenging and complex and provides itself an example for this: the loss of 38 spacecraft due to increased atmospheric drag caused by a geomagnetic storm. Starlink does not provide information on space weather related effects are covered by their autonomous collision avoidance system as described below, but states: “Despite such challenges, SpaceX firmly believes that a low insertion altitude is key for ensuring responsible space operations.” (Starlink, 2022b). The explanation for this is a decreased time which is required for space debris to deorbit, e.g., to be naturally removed from orbit. Starlink describes this as “self-cleaning orbits” characterized by the fact the space debris deorbits due to atmospheric drag within approximately 5 years, see Figure 24.

This complies with the newly introduced requirement by the FCC that all spacecraft should be deorbited latest 5 years after their end of life (FCC, 2022). Starlink claims that other spacecraft operators inserting their satellites into higher orbits take into account increased risk of generation of space debris, since it will take significantly longer for space debris to naturally deorbit at higher altitudes. Starlink supports its statement by depicting space debris as a function of altitude, as shown in Figure 25. It is clearly visible that higher orbital altitudes suffer already from a higher density of space debris which can be assumed to increase further due to a longer time that it is required for natural deorbiting.

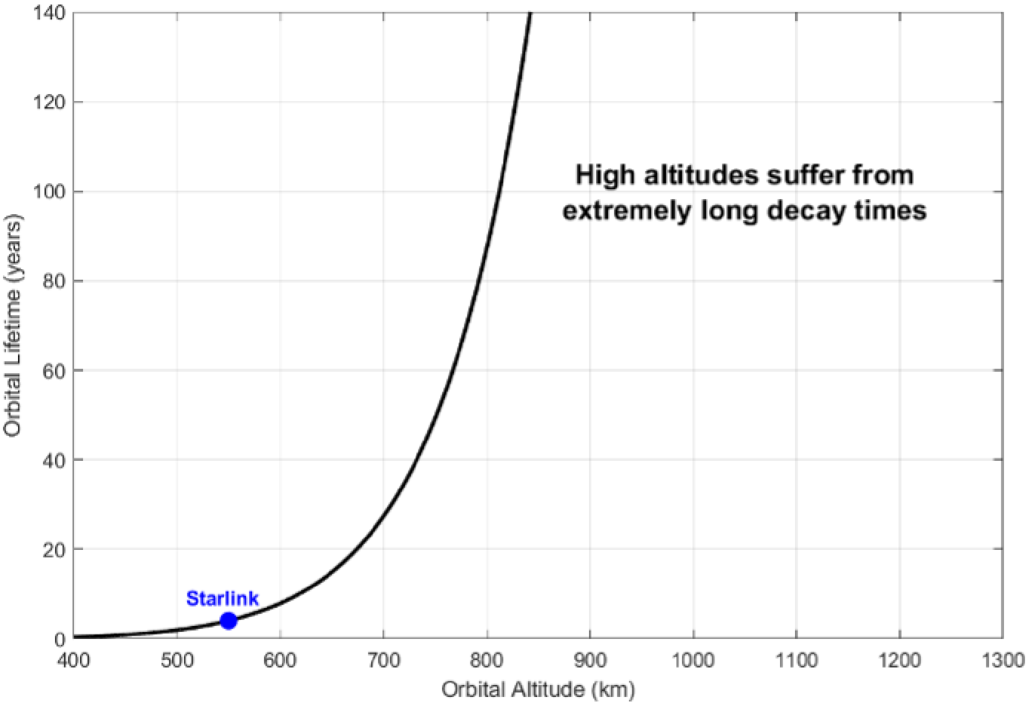


Figure 24: Orbital lifetime for a Starlink spacecraft. Adapted from (Starlink, 2022b).

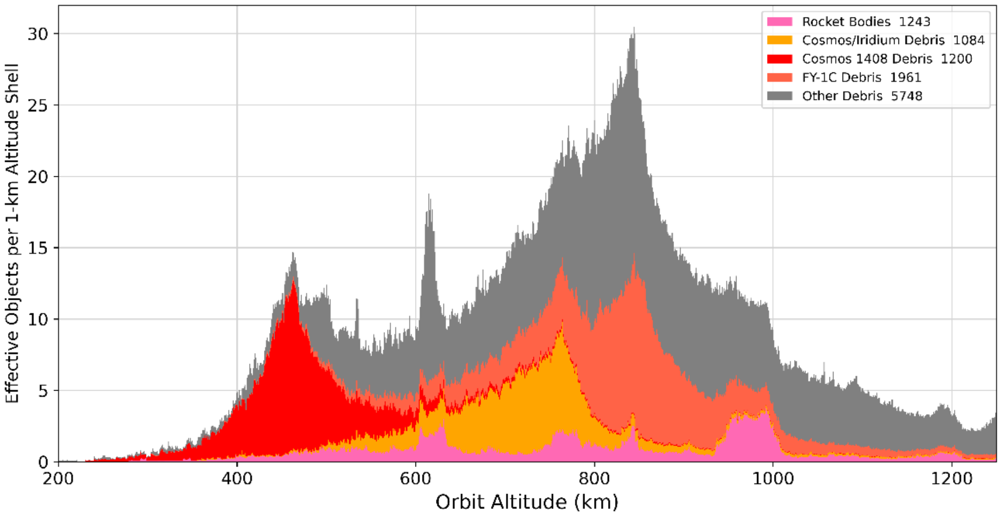


Figure 25: The number of debris objects per 1 kilometre shell vs. orbit altitude. Adapted from (Starlink, 2022b).

Starlink also takes a proactive role concerning transparency and data sharing. It shares future position and velocity prediction data as well as covariance data (statistical uncertainty of the predictions on space-track.org. Starlink envisions that access to space-track.org should be made available without login requirements in order to make it possible for everyone to access the provided data. Starlink claims also that it has been the first and only operator to provide routine system health reports to the FCC on a voluntary basis. These reports include amongst other the number of manoeuvres undertaken by the Starlink constellation. Figure 26 gives an example of the numbers divided into different causes. It is clearly visible that the major part of manoeuvres is due to space debris and other satellites. However, it is also worth noting that a small, but significant, number of avoidances manoeuvres is due to possible collisions with spacecraft from the same constellation.

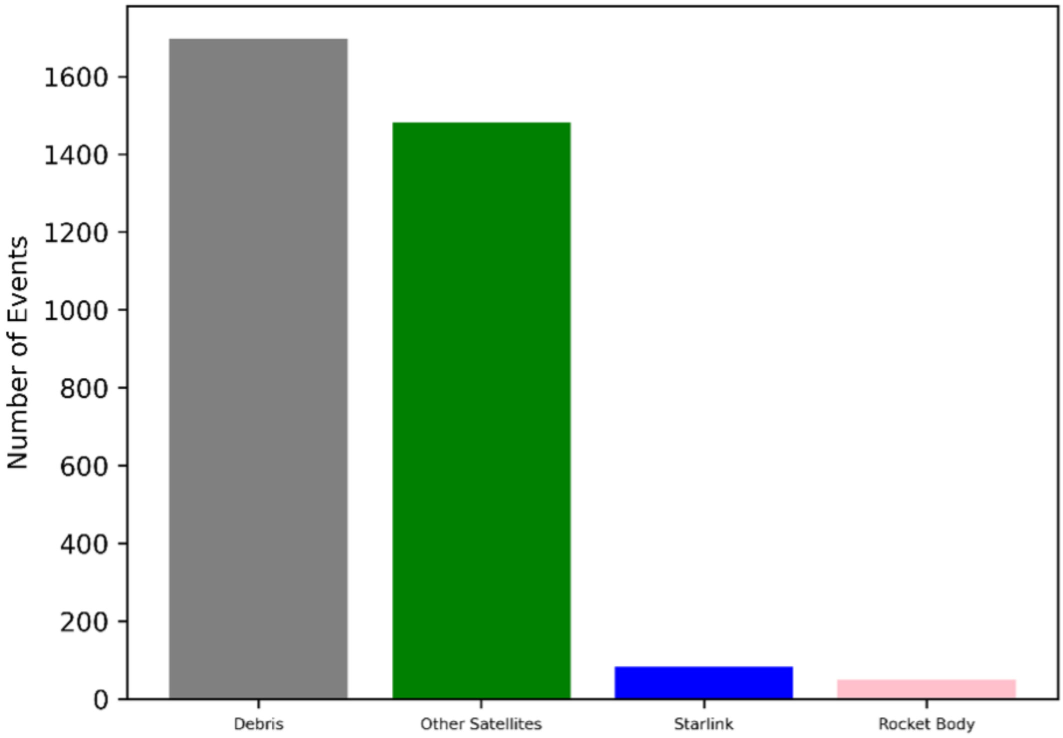


Figure 26: Number of Starlink manoeuvres for a six-month period (July-December 2021). Adapted from (Starlink, 2022b).

Starlink focuses also on collision avoidance and describes several steps that are utilized. Starlink actively shares information such as location data (based on onboard GPS data), prediction data and information on planned manoeuvres with the U.S. Space Force and other operators working with collision avoidance screening, such as LeoLabs. Future ephemerides provided by Starlink are uploaded to space-track.org three times per day where they are in their turn accessed by LeoLabs and in cooperation with the U.S. Space Force evaluated against the trajectories of other spacecraft and known space debris. Possible conjunctions are reported back Starlink and other operators as CDMs (Conjunction Data Messages) containing information such as satellite vectors, uncertainties on position and manoeuvrability. Subsequently, these messages are upload to the Starlink spacecraft that require this information. Due to the complexity of the operations and the number of spacecraft involved, Starlink has installed an onboard, autonomous collision avoidance system on each spacecraft. The spacecraft in question will plan an avoidance manoeuvre if there is a probability for a collision greater than 1/100000. It is worth noting that Starlink - according to its own statement – requires a reaction if the probability for a collision is 10 times higher than the industry standard (which is 1/10000). The implications of that will be discussed later. Also, the autonomous collision avoidance system assesses the risk of increasing the risk for further collisions which are cased by a potential manoeuvre – without providing information on how this is obtained. In order to further minimize the collision probability, Starlink spacecraft can autonomously reduce their cross-section for a potential impact, claiming the probability for a collision is therefore reduced by a factor of 4-10, see Figure 27.

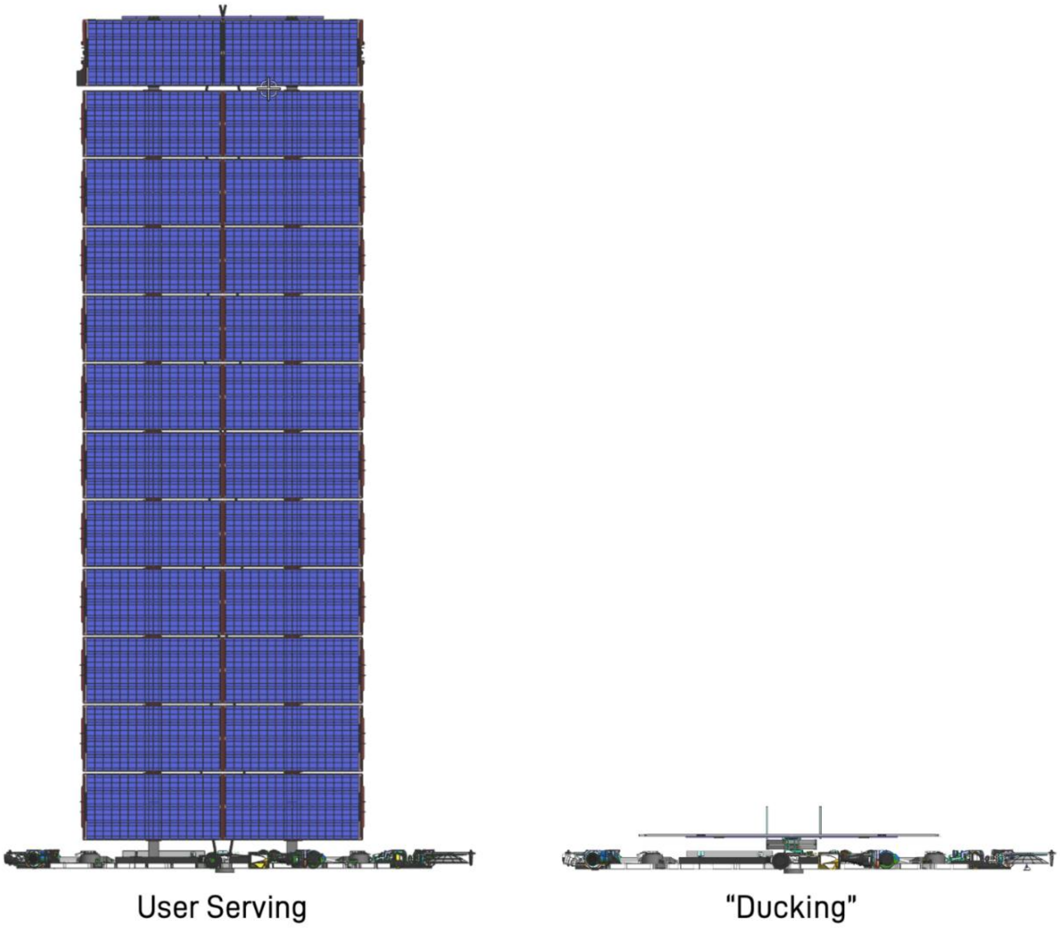


Figure 27: Autonomous reduction of the cross-section by a Starlink spacecraft. Adapted from (Starlink, 2022b).

In terms of legal concerns, Starlink accepts responsibility for all manoeuvres related to conjunction events. However, in contrast to the autonomous collision avoidance described above, Starlink also – somewhat confusingly – states that Starlink coordinates with the operator of another spacecraft that is in operation and can manoeuvre in case of high-probability conjunctions. It is not mentioned where the threshold for high-probability is placed. Starlink proclaims that their own spacecraft can receive commands not to manoeuvre in case the other operator takes responsibility for the manoeuvre (Starlink, 2022b).

In order to assess its quality, Starlink has arranged for their collision avoidance system to be reviewed by NASA's CARA (Conjunction Assessment and Risk Analysis) program under a Space Act Agreement (SAA) with NASA. As a result, Starlink claims that NASA relies on Starlinks collision avoidance system to prevent collisions with NASA spacecraft (Starlink, 2022b).

The abovementioned statement is however in great contrast with the concerns about the Starlink system which were raised in a letter by the National Aeronautics and Space Administration (NASA) which was sent to the Federal Communications Commission (FCC) in 2022 (Spaceref, 2022). These concerns include the potential impact on Low-Earth-Orbit (LEO) caused by the insertion of thousands of spacecraft, greater risk for collisions etc. As described above, Starlink utilizes an autonomous collisions avoidance system and this is what NASA is concerned about in particular. NASA acknowledges that Starlink might be able – by means of an autonomous system – to reduce the risk for collisions for any single spacecraft with larger objects to zero. However, NASA raises doubt concerning the claim that the collision risk also could be reduced to zero for constellations of thousands of satellites. This is further worsened by the potential autonomous interaction of different constellations consisting of thousands of spacecraft. The lack of traffic rules in orbit complicate this issue even further: "… the concern remains that other vendors proposing large constellations would also use auto-maneuvering capability within altitude ranges occupied by Starlink, thereby requiring multiple autonomous constellations to maneuver out of each other's way without clearly defined rules of the road for such interactions." (Spaceref, 2022). NASA claims that " … the assumption of zero risk from a system-level standpoint lacks statistical substantiation" and "… recommends SpaceX commission a risk analysis that addresses the efficacy of autonomous-vs.-autonomous constellation conjunction assessments and mitigation actions to provide confidence that the situation could be sufficiently addressed. If the recommended analysis is conducted, NASA requests the opportunity to review the results to help ensure there will be minimal risk to NASA and other assets.“ (Spaceref, 2022).

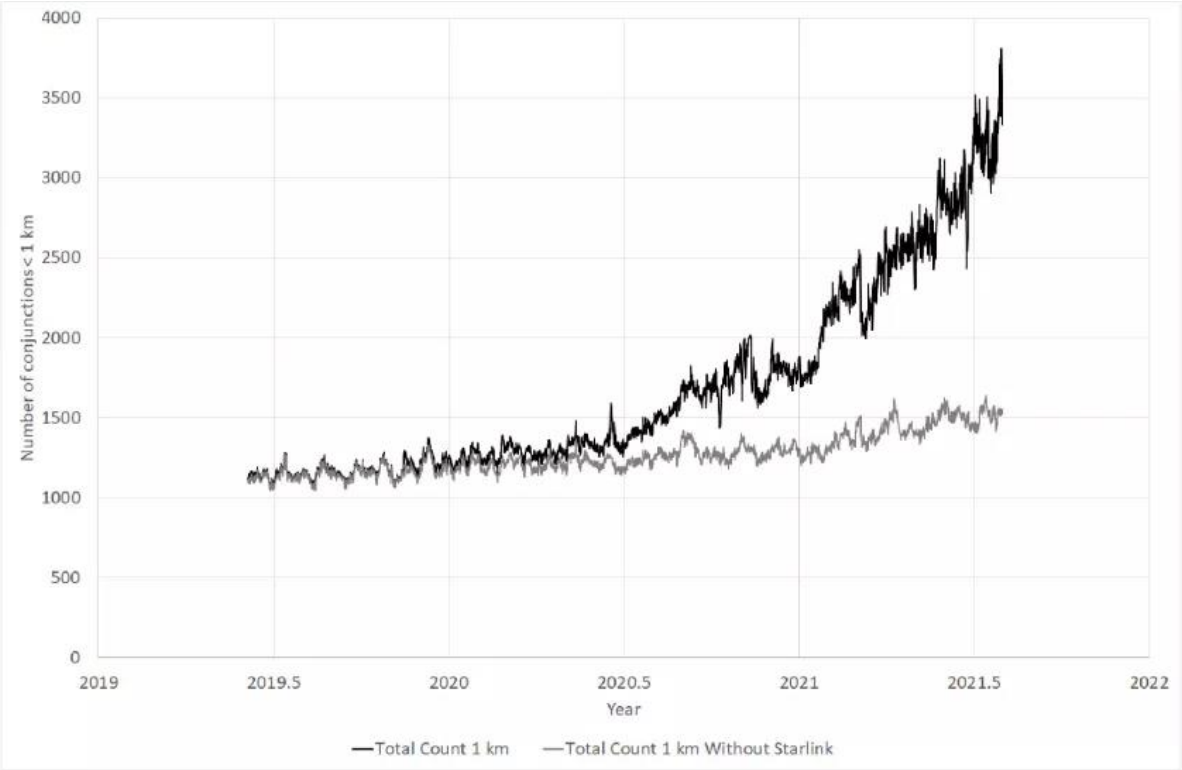


Figure 28: The number of close encounters between spacecraft. The black curve depicts the total count and the grey curve depicts the encounters excluding Starlink spacecraft. Adapted from (Pultarova, 2021).

It is of course important to evaluate statements provided by a commercial entity on its own homepage. Therefore, statements concerning Starlink’s operations made by other entities are listed below.

According to Pultarova (2021) Starlink spacecraft are at the moment responsible for over 50% of all conjunctions in orbit with a potential increase up to 90% in the near feature. Every week, Starlink spacecraft are involved in about 1600 conjunctions (with a distance smaller than 1 kilometre) with other spacecraft. The author refers to Hugh Lewis from the Astronautics Research Group at the University of Southampton (United Kingdom). Lewis utilizes the Socrates (Satellite Orbital Conjunction Reports Assessing Threatening Encounters in Space) database which is managed by Celestrack in order to evaluate future collisions risks. Lewis states the

Starlink is involved in about 50% of all conjunctions that are visible in the Socrates database.

Figure 28 visualized not only the significant increase of encounters over time, but also the fact that Starlink spacecraft stand for a significant part of these conjunctions. As a prediction, given the fact that only 1700 of the planned 12000 Starlink spacecraft are launched, Starlink will be involved in 90% of all close conjunctions once the full operational constellation number is reached. Pultarova (2021) also refers to Siemak Hesar of Kayhan Space that is working with the development of a commercial autonomous space traffic management system. Hesar claims that an operator which is responsible for 50 spacecraft will need to handle up to 300 conjunction alerts per week. He assumes that up to ten alerts are escalated leading to collision avoidance manoeuvrers that need to be performed. The U.S. Space Surveillance Network provides the basis for Kayhan Space estimates. The network tracks 30000 satellites in different phases of operation (live to defunct) and space debris down to a size of 10 centimetres providing a significant unknown population of space debris (smaller than 10 centimetres). The database will experience an increase ten times which partly can been related to the growth of larger constellations of spacecraft such as Starlink. In addition, means of tracking space debris are expected to improve meaning that also smaller pieces of debris can be discovered and tracked. According to Hesar: "This problem is really getting out of control, …The processes that are currently in place are very manual, not scalable, and there is not enough information sharing between parties that might be affected if a collision happens." (Pultarova, 2021). Hesar points to a number of potential threats:

- With a significantly increasing number of conjunctions, the probability of a spacecraft operator making an erroneous decision also increases.
- Collision avoidance manoeuvrers also include an economic factor: they are related to fuel (which is extremely precious in space), time and effort. Thus, operators might be pressured to re-evaluate a decision purely made on a statistical basis for collision.
- The uncertainty related to positions can amount to the order of 100 metres for active spacecraft or 1 kilometre or more for a piece of debris. Thus, the sphere of uncertainty can obtain a diameter of a few kilometres.

Hesar concludes: "In a situation when you are receiving alerts on a daily basis, you can't maneuver for everything … The maneuvers use propellant, the satellite cannot provide service.

So there must be some threshold. But that means you are accepting a certain amount of risk. The problem is that at some point, you are likely to make a wrong decision." (Pultarova, 2021). Furthermore, Pultarova (2021) refers to Lewis which shares his concerns arising from the monopoly that Starlink possesses in the mega-constellation arena. Lewis is worried if Starlink has the required knowledge as a spacecraft operator, since they only have been a launch provider before. In addition, Lewis emphasizes that an autonomous collision avoidance system, such as Starlink utilizes, not only has positive sides, but can carry with it further complications: an automatic collisions avoidance will lead to an orbital change which in its turn alters predicted ephemeris existing in the system which future collision avoidance operations rely upon. This render subsequent collision avoidance operations more complicated, according to Lewis: “Starlink doesn't publicize all the maneuvers that they're making, but it is believed that they are making a lot of small corrections and adjustments all the time," Lewis said (Pultarova, 2021). "But that causes problems for everybody else because no one knows where the satellite is going to be and what it is going to do in the next few days." (Pultarova, 2021).

TheVerge (2021) reports on a close conjunction between a Starlink spacecraft and a satellite operated by its competitor OneWeb on March 30, 2021. OneWeb operates its spacecraft at higher altitudes than Starlink, meaning that OneWebs satellites need to move through the orbital shell that Starlink populates. The U.S. Space Force 18th Space Control Squadron provided a red alert warning of a potential collision with a probability of 1.3 percent with a closest approach of 190 feet. Given the lack of legal requirements for spacecraft operators to react to collision warnings, the OneWeb operators connected to Starlink operators in order to coordinate manoeuvres. According to (TheVerge, 2021), Starlink turned off its artificial intelligence powered avoidance system – the reason being unknown. This is in line with concerns within the space industry relating to Starlink’s autonomous collision avoidance system. The main concern is simply that operator on the opposite site has no way of knowing what the reaction of Starlink’s system will be: “Coordination is the issue … It is not sufficient to say ‘I’ve got an automated system,’ because the other guy may not have, and won’t understand what yours is trying to do.” (TheVerge, 2021).

However, TheVerge (2021) sheds light on further nuances of the same issue. A positive part of the incident is the contact between the operators of both companies was established and the issue resolved “given the constraints of global best practice” in the absence of any regulations. The lack of any regulations is made even worse by the speed of Starlink’s expansion. Although

Starlink has increased its transparency concerning the position of its spacecraft and predictions of future orbits, “its automated system for avoiding collisions is a closed book where openness and coordination are needed the most” and “What is the point of having it if you have to turn it off when there’s going to be a potential collision?” (TheVerge, 2021). Also, it is emphasized that it is unclear who the responsible part is in the case of an actual collision. The legal problem is worsened by the fact that Starlink populates lower orbital shells, meaning that spacecraft operators, such as Oneweb, which operate at higher altitudes need to traverse Starlink’s operational area on a regular basis, although with uncertain legal requirements: “OneWeb and others will have to transit through Starlink to reach their destinations, so SpaceX needs to ensure now that other satellite operators can do that safely.” (TheVerge, 2021). However, the absence of an established legal framework and intra-operator procedures still looms over future operations: “The scary situation is when one of the operators is not communicative, and then it’s just crossing your fingers.” (TheVerge, 2021).

The complexity of the inter-operator communication is described by Brodkin (2021): Starlink accuses OneWeb of “spreading a false story claiming that the companies’ satellites nearly crashed into each other”. Starlink claimed that OneWeb asked for the autonomous collision avoidance system to be turned off – not Starlink turning it off on its own. According to Starlink’s version, it offered to perform a voluntary manoeuvre, but that during the discussion with OneWeb it was agreed upon to wait for another CDM (conjunction data message). During another contact between the operators of both parties, Starlink claims to have suggested to wait for another CDM, whereas OneWeb – requiring more time for the planning of a manoeuvre than Starlink – insisted on performing its own manoeuvre. Both spacecraft are reported to have passed each other by approximately 1000 metres.

Another close call where Starlink was involved was the near-collision with the Aeolus spacecraft in 2019. Mack (2019) describes how the European Space Agency (ESA) was forced to move the Aeolus spacecraft in order to avoid a collision with a Starlink satellite. The final conclusion was that Starlink did not see the email informing them to move their spacecraft in order to avoid the collision. ESA moved their own satellite about ½ an orbit before the potential collision. The U.S. Air Force informed Starlink about an increase in probability of collision between the two spacecraft, exceeding the industrial threshold of 1 in 10000. Starlink admitted that "A bug in our on-call paging system prevented the Starlink operator from seeing the follow on correspondence on this probability increase … Had the Starlink operator seen the correspondence, we would have coordinated with ESA to determine best approach with their continuing with their maneuver or our performing a maneuver." (Mack, 2019). Another interesting comment by ESA is the following: "It is very rare to perform collision avoidance maneuvers with active satellites. The vast majority of ESA avoidance maneuvers are the result of dead satellites or fragments from previous collisions." (Mack, 2019).

Further information on this incident is provided by O'Callaghan (2019). According to ESA, Starlink refused to move out of Aelous' way. The ESA spacecraft altered its orbit and then returned to its operational orbit after clearing the potential collision. "According to Holger Krag, head of the Space Debris Office at ESA, the risk of collision between the two satellites was 1 in 1,000 – ten times higher than the threshold that requires a collision avoidance maneuver. However, despite Aeolus occupying this region of space nine months before Starlink 44, SpaceX declined to move their satellite after the two were alerted to the impact risk by the U.S. military, who monitor space traffic.“ (O'Callaghan, 2019). ESA decided to move its spacecraft since the probability was approximately 1 in 1000. ESA also acknowledged that there are no traffic rules and that Starlink did not act erroneously, but that in the light of an increased number of spacecraft in orbit legal regulations and traffic management need to be implemented. (O'Callaghan, 2019) states that despite Starlink's autonomous collision avoidance system being supposedly capable of avoiding collisions, the system seems not have been used during this incident. ESA also raises concerns that more manoeuvres relating to active satellites and in particular spacecraft of mega-constellations will be required in the future: "We see it as part of our changing environment … We want to raise awareness in this sense, that there's quite a bit of work that needs to be done on how to make sure that these type of operations will run smoothly in the future." (O'Callaghan, 2019).

O'Callaghan (2019) raises more issues related to mega-constellations. Firstly, the tracking of mega-constellations is not feasible with today's tracking technology and new approaches need to employed in order to pave the way for further growth. Secondly, Starlink's black-box approach concerning its own autonomous collision avoidance system raises concerns.

Holger Krag (Head of Space Debris Office at ESA) is referred to with the following words: "My concern is how often will we have such events in the future? These are just two satellites. Now they will add several thousand, and they will also be disposed and end up at various altitudes. And there's no rule or law on how to react, it's all goodwill. What I want is an organized way of doing space traffic. It must be clear when you have such a situation who has to react. And of course automating the system. It cannot be when we have 10,000 satellites in space that there are operators writing the email what to do. This is not how I imagine modern spaceflight." (O'Callaghan, 2019).

Starlink has commented on this issue with the following statement: "Our Starlink team last exchanged an email with the Aeolus operations team on August 28, when the probability of collision was only in the 2.2e-5 range (or 1 in 50k), well below the 1e-4 (or 1 in 10k) industry standard threshold and 75 times lower than the final estimate. At that point, both SpaceX and ESA determined a maneuver was not necessary. Then, the U.S. Air Force's updates showed the probability increased to 1.69e-3 (or more than 1 in 10k) but a bug in our on-call paging system prevented the Starlink operator from seeing the follow on correspondence on this probability increase – SpaceX is still investigating the issue and will implement corrective actions. However, had the Starlink operator seen the correspondence, we would have coordinated with ESA to determine best approach with their continuing with their maneuver or our performing a maneuver." (O'Callaghan, 2019)

Another incident involving Starlink led to a response from the U.S. Department of Transportation's Federal Aviation Administration (FAA). Starlink did not provide trajectory data for a launch executed on August 19, 2022. The purpose of these data is evaluate the collisions risk with existing objects in orbit around Earth. The fine was set to USD 175000 (FAA, 2023).

Foust (2021) reports on an agreement which has been signed between NASA and Starlink. The main point of the agreement is that "NASA will rely on the autonomous maneuvering system on SpaceX's Starlink satellites to avoid any close approaches between that fleet of spacecraft and NASA spacecraft, like the International Space Station." (Foust, NASA and SpaceX sign agreement on spaceflight safety, 2021). Furthermore "the agreement … is intended to formalize both parties' strong interest in the sharing of information to maintain and improve space safety. … With commercial companies launching more and more satellites, it's critical we increase communications, exchange data, and establish best practices to ensure we all maintain a safe space environment" (Foust, 2021). In clear speech, this means that NASA spacecraft will remain on its path whereas Starlink satellite have to move in order to avoid any potential collision.

Concerning the International Space Station, Starlink is required to avoid any conjunctions within a box of 50 by 50 by 4 kilometres centred on the station.

Foust (2021) finished the report with the following statement which is of great interest for this thesis: “Every operator I’ve talked to has a tremendous amount of automation in their systems … Effectively, these vehicles are on ‘auto-drive’ and the operators themselves may not know when a vehicle chooses to maneuver. … When these vehicles are doing autonomous maneuvers, small maneuvers, they are making the assumptions that went into avoiding them in the first place invalid … If you put two large constellations near one another, all these get greatly magnified.”

It seems logical that the introduction of so-called mega-constellations (although the term mega-constellation is misleading, since these constellations are planned to consist of thousands of spacecraft and not millions) will increase the probability of collisions. It is interesting to note that addressing collisions means collisions between spacecraft of the same constellation or between the constellation’s spacecraft and other objects. Increasing the number of spacecraft in orbit is also associated with an economical factor: more objects in space will require more manoeuvrers – something which is rather expensive since all fuel needs to be carried by the spacecraft. A general problem when addressing collision risk associated with mega-constellations is their lack of transparency concerning the operational procedures of the respective collision avoidance system (or simply: which way will the other one turn?) or underlying assumptions such if the spacecraft operator adheres to the space debris mitigation guidelines or not. Arroyo-Parejo, Sanchez-Ortiz, & Dominguez-Gonzalez (2021) performed a study in order to evaluate the collision risk related to the Oneweb and Starlink missions. The author emphasize that a variety of assumptions had to be made (refering to the lack of transparency from the spacecraft operators), something which directly relates to the quality of background knowledge and ultimately the probability which is calculated: “Moreover, additional guesses such as spacers or fairing have been assumed in an attempt to obtain more conservative data.” (Arroyo-Parejo, Sanchez-Ortiz, & Dominguez-Gonzalez, 2021). The authors describe the probability of a collision is given by $P_{i=n} = \frac{N^n}{n!} e^{-N}$, where $N = F\ A_C\ T$ is the mean number of collisions - which gives $N_\alpha = F\ A_C\ T\ \alpha$, where α is the number of spacecraft in the constellation (Arroyo-

Parejo, Sanchez-Ortiz, & Dominguez-Gonzalez, 2021). Ultimately, the authors obtain the probability for collision to occur in a constellation consisting of α spacecraft in the same orbit: $P_{i \geq \alpha} = 1 - e^{-N_{\alpha}}$.

Arroyo-Parejo, Sanchez-Ortiz, & Dominguez-Gonzalez (2021) use the derived equations and conclude that moving the spacecraft within the constellations leads to an increase for the collision risk by a factor of 10%. Oneweb had originally planned to deploy 48000 spacecraft and the authors find that the potential deployment of such a vast number of spacecraft would represent an increase of the probability of a collision by 20%: “For the Oneweb case it has been concluded that the effect of adding 48000 second generation satellites would be hazardous in terms of collision probability.” (Arroyo-Parejo, Sanchez-Ortiz, & Dominguez-Gonzalez, 2021) The work of Arroyo-Parejo, Sanchez-Ortiz, & Dominguez-Gonzalez (2021) where collision risks were calculcated can be put into context by the work of McKnight et al. (2022) which descibes statistical collision risks in LEO by accumulating the conjunction data messages (CDM) – the latter which describes the probability of collision (PC). The usual threshold for PC for which action is taken is in the range of 1/10000 to 1/100000. The threshold is different for different spacecraft operators and depends on risk acceptance and of course the quality of the PC. The abovementiond threshold triggers then a further evaulation of the conjunction and usually for a PC value between 1/10000 and 1/1000 a collision avoidance maneuver is initiated. The PC is dependent on several factors such as miss distance, the size of the objects (refering to the ability of the Starlink spacecraft to reduce the solar panel area) and the uncertainty of the position and speed of the objects. Modern spacecraft include a Global Positiong System (GPS) receiver providing accuracies in the order of a few meters, where is groundbased monitoring returns accuracies of the order of 20 meters (McKnight, et al., 2022).

In terms of identifying the statistically most concerning objects which should be chosen for Active Debris Removal (ADR) mission, meaning spacecraft that will be launched in order to remove these high-risk objects, McKnight et al. (2022) acknowledge the fact that the PC can be obtained both empiracally and statistically. The selection process for the most concerning objects is depicted in Figure 29. It should be pointed out that McKnight et al. (2022) describe risk as being the product of consequence and probability – neglecting other parameters such as quality of knowledge. The authors compared statistical risk with accumulated CDM PC and found a large conversion between these two on a general basis. They present the short time frame for the statistical approach (six months) as a likely explanation for this divergence. This is confirmed by the fact that the statistical risk converges with accumulated CDM PC for older debris such as rocket bodies for which a longer monitoring period could be acchieved.

McKnight et al. (2022) also point out several caveats, in particular related to data input into the collision avoidance model, something which is of great importance for this thesis: Firstly, the quality of the PC values can assessed by means of checking if the covariance in the analysis is larger than the miss distance between the two objects. Simply speaking, one can always get a result from the collision analysis, but one needs to make certain that the result gives physical meaning – something which is of great importance when employing AI's in collision avoidance and the suggest human monitoring of the quality of the AI's work. The authors find that the main part of such cases is related to operational payloads, i.e., payloads with maneuver capabilites. This makes perfect sense, since collision risk estimates are made on the basis that both objects maintain constant speed and orbit. However, in the case of a maneuvering payloads the quality of the risk assessment is severely degraded. Secondly, the objects that are likely to collide can inhabit different orbits meaning different velocities. Since the relative velocity is a main input into the collision risk assessment and since in the authors study a constant relative velocity was assumed, this will likely lead to a divergence between between risk assessment types. Lastly, McKnight et al. (2022) point out that all discussions connected to collision risks need to take into account that they only include potential collisions between known objects. Only objects down to a size of roughly 10 cm are catalogued, leaving a huge number of small objects which do not find their way into the risk assessment: "As a result, there is no ability for operational satellites to avoid collisions with these objects likely to produce mission-terminating effects upon impact." (McKnight, et al., 2022).

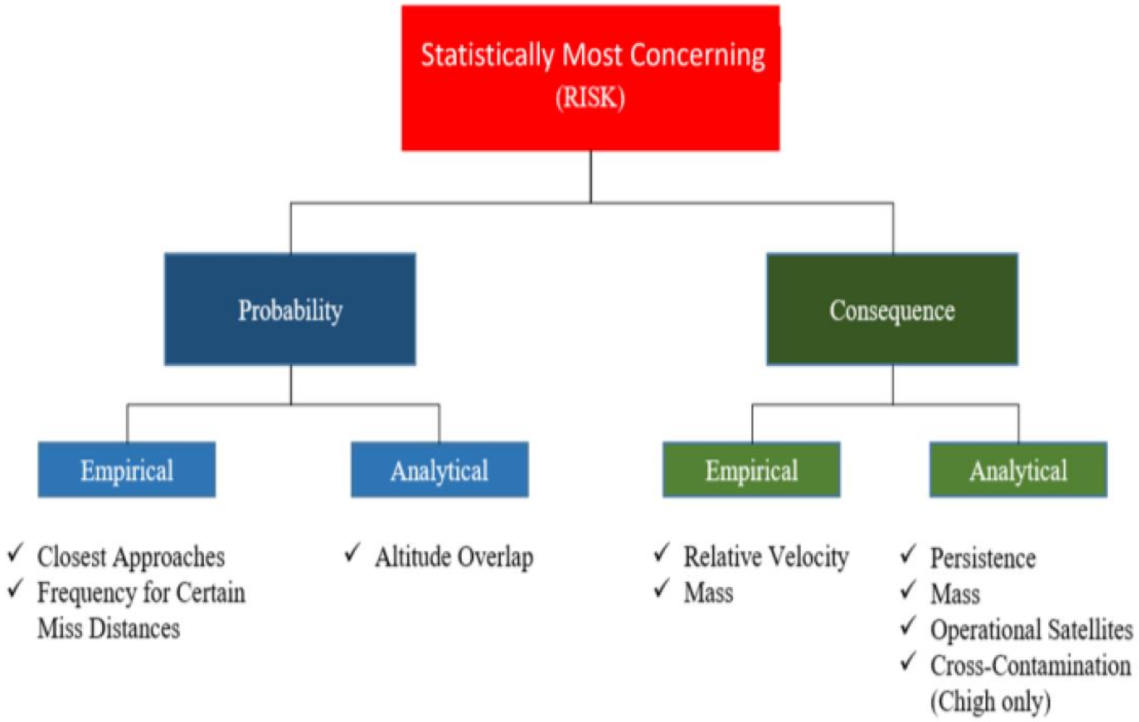


Figure 29. Risk assessment for statistically most concerning objects. Adapted from (McKnight, et al., 2022).

The European Space Agency (ESA) is at the forefront of space debris mitigation with several initiatives. One example is CREAM (Collision Risk Estimation and Automated Mitigation) which attempts to address several issues of collision risk estimation with automated reactions (and the related challenges described above). The main point of CREAM is to automate collision avoidance handling, since the workload related to collision avoidance has risen significantly and will increase even further with the anticipated launches of ten thousand of spacecraft the next years. At the moment, one can expect hundreds of conjunction warnings for a LEO spacecraft every week resulting in about two warnings leading to an actual manoeuvre. It is obvious, that such a setup requires around the clock standby resembling a significant economic impact: in 2014, 14 million Euros were used to react to false collision alerts and thus manoeuvring spacecraft. Here, CREAM is anticipated to step in for human experts by means of machine learning with goal of a “continuous collision avoidance process in replacement of the classical impulsive maneuvering” (Virgili, Flohrer, Krag, Merz, & Lemmens, 2019). Note that in the study of McKnight et al. (2022) objects with manoeuvre capabilities caused problems estimating collision probabilities as constant manoeuvring will require a higher frequency in terms of monitoring of objects. This means that one is moving from a semi-static system (in terms of risk assessment) to a system that is continuously changing. CREAM is planned to “conduct safe and efficient collision avoidance maneuvers without human intervention” (Virgili, Flohrer, Krag, Merz, & Lemmens, 2019). In order to achieve that, the authors state that one needs to consider additional criteria in addition to the calculated collision risk calling for new development steps including, among others the development of risk evolution, inter-company coordination (with special focus on mega-constellation), establishing acceptable delays between the reception of orbital elements and subsequent conjunction updates; an overall consensus on thresholds.

The main idea for CREAM is to use the existing database at ESA consisting of over 1 million close approaches in the past in order to establish parameters for unsupervised identification of collision alerts. CREAM also addresses a possible next step, i.e., on-board decision making. This would mean that the spacecraft itself makes the decision to move or not to move, instead of an operator or AI on Earth. However, this requires that the data which is necessary in order to make the decision is uploaded to the spacecraft– which is similar to the commands sent to spacecraft in case a decision is made on Earth. CREAM also addresses the issue of false alerts which is mainly caused by uncertainties of the orbital elements which form the basis of collision avoidance alerts. Only better tracking of orbital objects, mainly by onboard GPS, will lead to an increase of orbital uncertainty and thus enables later decision making.

Rebhan, Einecke, Losing, Limmer, & Schmitt (2020) reported on their experience from machine learning on collision avoidance datasets and described it as “very difficult”. The problem is that there is a huge dataset on “near misses” from the past – as used for CREAM – but there is very little data on actual collisions. This last point is extremely interesting for this thesis: how should an AI learn to avoid collisions if there are very little examples? Could there be a statistical basis for all these “near misses” mainly due to orbital uncertainty? The authors suggest further collection of data and possibilities to “predict expert decision” – in a way a human monitoring.

SpaceWatch Global (2022) reports that Astroscale has been chosen to perform a study on CREAM. It is mentioned that also inter-satellite links and onboard processing will be studied in order provide suggestions for commanding structures to facilitate late collision avoidance maneuvers.

Astroscale (2022) gives more details on the projecct and states that “CREAM is designed to generate technologies for automated systems that can determine the likelihood of orbital collisions with greater accuracy, reduce the number of false alerts by allowing reduced time between maneuver decisions and close approaches, and finally to optimize maneuver plans that are uploaded to satellites”. Whereas it was earlier stated that CREAM would fully automatically, this approach seems to be somewhat nuanced: “In the frame of CREAM we aim to develop the technology for automated collision avoidance to reduce the manual interventions needed, … investigating robust decision criteria and maneuver designs, improved means for coordination among operators of spacecraft, as well as ways to guarantee late access to spacecraft thus enabling later involvement of human operators and analysts and reducing false alarms.” (Astroscale, 2022). Note that a certain human involvement is thus still anticipated in CREAM.

6. Discussion

Before engaging in the discussion, the problem is shortly outlined again: With increasing numbers of spacecraft launched in the space, the number of debris also increases. Space debris resembles a danger for functioning satellites: they can be damaged and in the worst case destroyed. There is an ultimate fear that the number of space debris will reach such high levels that spacecraft operations will be made impossible. In order to avoid collisions between spacecraft and debris or several spacecraft, the orbital elements of all objects (both defunct and functional) are monitored and used to forecast collisions between them. In case a certain threshold is met, collision avoidance manoeuvres are executed.

In the early space age with a few spacecraft, potential collisions were rare and the number of objects to be monitored few. Collision avoidance operations were performed manually by the operators. With the new space era, spacecraft numbers and thus possible collisions increase significantly – mostly by the introduction of so-called mega-constellations consisting of many thousands spacecraft. Thus, manual collision avoidance handling became more challenging reaching a limit where more automated operations were necessary. For the latter, AI is often employed.

This change from manual to automated operations utilizing AI is challenging for established risk management procedures and regulations concerning liability. Manual operations are covered by classical risk management actions under uncertainty with the main points being evaluation of the proposed measures to minimize the risk by the management before implementation and an iterative process where the actual risk management process is revisited after evaluating its impact and underlying assumptions. Here, one can consider the framework in which the risk management process is undertaken to be semi-static: the underlying assumptions will not change significantly during the course of the reevaluation. In contrast, the framework in which the automated collision avoidance programs are performed is highly dynamic, since collision avoidance manoeuvres are performed constantly changing the underlying assumptions subsequently all the time. In addition, this is enhanced by the fact the AI is supposed to constantly learn and alter itself. Hence, not only the framework in which the risk management process is utilized is changing but the risk management process itself as well. Furthermore, the human reevaluation as such is being removed. This is context for this thesis.

Firstly, the issue of quality assertion of the risk management process is discussed. Referring to Aven (2015) and Aven, Røed, & Wiencke (2008), the incident A is a collision between two objects in orbit around the Earth. Several combinations are possible: between to active spacecraft, a spacecraft and space debris or two pieces of space debris. Active spacecraft can also be divided in manouverable and non-manouverable satellites. The possible consequence C of a collision (the main focus here in on a collision of an active spacecraft with another object) include damage on the spacecraft without operational impact or with operational impact (such as degradation of solar panels which deliver power to the spacecraft) over total operational impact (the spacecraft is not destroyed, but a vital part is damaged such as the spacecraft is not operating anymore) to complete destruction of the spacecraft with the creation of thousands of pieces of space debris. In order to put collision avoidance into context, one can describe it as one of the probability reducing barriers. Other probability reducing barriers are space traffic management (something which is not in place, yet), orbital monitoring of objects in space, data exchange and following the different guidelines and initiatives mentioned above. Consequence reducing barriers are spacecraft design (providing, e.g., redundancy of essential operational parts, choosing material that generates little debris if hit, spacecraft that can manoeuvre etc) or future technologies such as Active Debris Removal or In Orbit Servicing. Aven (2015) and Aven, Røed, & Wiencke (2008) also use the background knowledge K with a certain quality of this knowledge SK in order to describe risk. Here, the background knowledge describes the orbits of all spacecraft and space debris in orbit and their temporal development. However, only space debris down to a size of 10 cm is monitored and thus included in collision avoidance warnings. Hence, the quality of the background knowledge is very weak for objects smaller than 10 cm in size and strongly depending on the way the orbital elements are obtained (either by groundbased radars or in-orbit measurements) leaving different levels of quality for the probability of the collision warning (McKnight, et al., 2022).

Aven (2015) describes the goal of risk management to remove, reduce, optimize, transfer or keep the risk. All five ways are theoretical approaches for potential collisions. Removing the risk would mean removing space debris (which will be performed in a few years time) and removing the operators own spacecraft (which would mathematically reduce the risk to zero, but would not make any economic sense). Reducing the risk resembles a collision avoidance manoeuvre moving the spacecraft out of the path of the debris. Optimising the risk is the scope of collision avoidance handling by new technologies such as AI. Transferring the risk is a not practical here, which keeping the risk obviously is: if the risk is acceptable, then the operator can keep it. In order to manage the risk, a risk management process is performed (International Organization for Standardization, 2018). The main point is that the process is of an iterative nature. Furthermore, risk management under uncertainty requires an evaluation of the proposed measures by the management (Aven, 2015). As mentioned above, this makes sense for a semi-static environment where the assumptions hold within the reevaluation cycle.

With the significant increase of collision warnings which could not be handled anymore by human operators alone, automated solutions such as AI with self-learning capabilities were introduced with the aim of continuous orbital adjustments in contrast to incremental orbit changes as before. Now an automated system is supposed to make the decision to move a spacecraft (or not) instead of a human operator. Human operations are of course associated with a certain risk, something which risk management is supposed to handle – amongst others. The use of an AI also includes a risk: the actual software can be erroneous as well as the data input. In addition, the AI is supposed to be self-learning – all without human control (UNITED STATES SECURITIES AND EXCHANGE COMISSION, 2018). These risks can then lead to liability issues which has led to the report from the expert group of the European Commission (European Commission, 2019). Main points from this report are: (1) The manufacturers of an AI are responsible (liable) for damage which results from the AI's operations. This is interesting, since it then would be in the interest of the manufacturer to perform a quality assessment throughout the products lifetime. This is strengthened by the request of the expert group to include logging features in each AI in order to evaluate liability. This could provide an incentive for the manufacturer to use this feature for human quality assessment of the AI's performance. The liability issue and connected risk is further complicated by the fact failures can be caused by the manufacturer or the AI or the operator making use of an AI. Here, the question for a risk manager is on how to define risk for these three different parts. The risk manager might not even have control over the other parties, e.g., the risk manager for an operator cannot require the manufacturer of an AI or the AI itself to perform certain actions, thus nullifying the risk management process. However, the report states that "liability should lie with the one who has more control over the risks of the operation." (European Commission, 2019, p. 6). Thus, liability can be used in order to define risk management processes. However, collision avoidance operations become more and more complex, leading to difficulties assessing liability for any damage because of poor understanding of causality which in its turn renders risk management processes unvaluable (European Commission, 2019, p. 20). Risk analysis relies also an "old data": when studying an incident, risk managers can learn from it, e.g., understand causality, which then can play into a future risk analysis which forms the basis for risk management procedures. If risk managers do not (fully) understand what or which line of events caused an incident, it is difficult to establish probability and consequence reducing barriers with a great chance of placing them wrongly (both in location and in time): “Even without changes to the original software design, the embedded criteria steering the col-lection and analysis of data and the decision-making process may not be readily explicable and often require costly analysis by experts.” (European Commission, 2019, p. 20). The problem is then multiplied by the fact that the AI changes itself relying on the input of data. Both actions can be erroneous, since the update of the AI is self-initiated without human control and the input data can also be erroneous if not quality-checked: “Not only may such data be flawed in itself, but the processing of otherwise correct data may also be imperfect. The latter may be due to original defects in designing the handling of data, or the consequence of distortions of the system’s self learning abilities due to the bulk of data collected, whose randomness may lead the AI system in question to misperceive and miscategorise subsequent input.” (European Commission, 2019, p. 21). In context, this would mean that an autonomous, self-learning system will adapt its operations from collision avoidance actions it has performed, from data input into its system (orbit data from different sources) and so on. The positive aspect of this operation namely that the system can handle many more operations than a human operator could thus be side-lined by the potential erroneous actions or changes to operations. Since traffic management rules are missing in orbit, several scenarios could emerge: (1) an encounter between an operator utilizing autonomous avoidance and an operator handling this manually, (2) an encounter between a spacecraft that can change its orbit and a spacecraft which cannot, (3) an encounter between spacecraft operators both utilizing AI’s but with different levels of human involvement etc. Starlink (2022b) shows that little less than 50% of the collision avoidance operations for Starlink are related to debris and little less than 50% are related to non-Starlink spacecraft. Only a minor part are related to pure Starlink conjunctions.

A problem with these several different scenarios is the different levels of transparency of the used system, e.g., Starlink is not providing information on if space weather related issues are covered by their AI system. It would be surprising if space weather is not included, since space weather plays a significant role at the low orbits which Starlink is using for their insertion. A big question mark can be raised after Starlink lost 38 spacecraft due to a space weather related issue during spacecraft insertion (Starlink, 2022b). In addition, Starlink is not revealing information on how its collision avoidance system actually works – despite the need for transparency and data sharing. This is in strong contrast to the report from the expert group of the European Commission. Furthermore, despite the expectation of Starlinks system begin autonomous there are indications that the system has been turned off with respect to potential collisions (Brodkin, 2021; Mack, 2019; O'Callaghan, 2019; TheVerge, 2021).

This is further complicated by the fact that different operators have different risk accept thresholds, different reaction times, different latency times (meaning a temporal limit of a reaction before a possible encounter: one operator would longer than another operator before a reaction has to be performed), different policies concerning data sharing or transparency of their operations etc.

The issue of lacking human involvement with AI was also highlighted by the report of the European Commission, stating that the possibility of self-learning without human control can imply difficulties to use existing liability rules which translates then into risk assessment (European Commission, 2019, p. 23). The use of AI in space operations is relatively new, but the same technical solutions have been employed in other areas for a longer period. There the issue of quality control has already been addressed, something which also should be performed for space operations. Kashyap (2019) reported on AI in aviation where the third step in the operational procedure is “revise”. However, the author describes the revision process as “unpredictable” and stated the need for evaluation. Thus, in other areas than in space operations, a certain form of quality check is already envisioned for AI operations which indicates that the same should be performed for space operations.

To give a concreate example: NASA raised concerns on Starlinks AI collision avoidance tool – for addressing collision risk for constellations of thousands of spacecraft: “… the concern remains that other vendors proposing large constellations would also use auto-maneuvering capability within altitude ranges occupied by Starlink, thereby requiring multiple autonomous constellations to maneuver out of each other’s way without clearly defined rules of the road for such interactions.” (Spaceref, 2022). Furthermore, NASA suggests that a risk analysis is performed with respect to autonomous-vs-autonomous collisions avoidance procedures. This is so to say the worst case scenario where two autonomous systems have to react to each other without any boundary conditions (such as traffic rules). The least desirable scenario would be a collision because both systems decided not to react or to move into the same direction. The request by NASA to perform a risk analysis is interesting, since a risk analysis forms the basis for a risk management process which a collision avoidance resembles. It remains a question if operators of mega-constellations actually have performed a risk analysis concerning the mentioned autonomous-vs-autonomous operations.

This shows that the introduction of emerging technologies, such AI, provides an operational advance in terms of faster operations, but brings with it challenges connected with risk management: Firstly, the decision making within the risk management process is shifted and secondly, the utilization of the AI itself represents a risk which needs to be addressed on the same line as human operators represent a risk. These challenges are augmented by lack of regulations in space which Pagallo, Bassi, & Durante (2023) point out. It is obvious from the discussion up to this point that – similar to aviation operations including AI utilizing a certain amount of human interaction to quality assess the AI's operation – human control should be introduced to quality check AI operations related to space: "to determine and enforce a "meaningful human control" (MHC) over the entire technological cycle and system functioning. … There is no such alternative between human control over autonomous technologies on the one hand and, on the other, the further development of autonomous systems that take crucial decisions by themselves. MHC is the subject of academic courses on the risks and threats brought forth by possible losses of control of AI systems, so as to keep them in check." (Pagallo, Bassi, & Durante, 2023). This is of course a sensitive step: Moving to far with providing human control over AI could be considered a step back, returning to human operations with strong help from AI. In case of space operations this is not desirable (and not feasible), because the amount of collision avoidance operations were the main reason for introducing AI. Moving to little would not provide the quality control which this discussion has established to be necessary. Pagallo, Bassi, & Durante (2023) point to that balance between human quality control and the AI's self-governing process. However, this is a rather abstract statement and in terms of risk management where risk often is described quantitatively a more detailed statement is desirable: "The balance that shall be struck between human control and AI autonomy can be measured by the degree of "social acceptability" concerning the risks inherent in the automation process, as well as the level of social and political cohesion that regards the values and principles that are at stake with the development of increasingly autonomous technologies …" (Pagallo, Bassi, & Durante, 2023).

This last citation requires further discussion, however any discussion concerning the "social acceptability" of the risk associated with AI operations is above the scope of this thesis, whereas the political approach fit well within the scope of this work. A certain level of work is done by the space industry to raise awareness within the society concerning the level of dependency of daily life on spacecraft operations. If this awareness can be anchored within society, a better understanding of space sustainability is obtained. Society will hopefully understand that AI operations are necessary (but also bringing with them an inherent risk) to cope with the sheer amount of collision avoidance operations. The political side of this discussion can be related to the handling of the legal aspects on international and national level. There is consensus on international level that space debris is a major threat and requires new technologies to handle it. However, the regulatory aspects of space operations have not kept pace with the technological development which have led to the introduction of a variety of non-legally binding instruments. Hence, the political cohesion is present, but work still remains to be done on regulating disruptive technologies without restraining their development.

Omar (2023) develops this line of thought further and states that validation of the models is important and human feedback required: "Machines simply don't what they don't know". This is an important in terms of discussing risk management for AI operations. Knowledge and its quality are two important components for describing risk (Aven, 2015) and without the possibility of assessing the background knowledge and – in particular – its quality, the risk management process becomes almost useless. The system can always certainly provide a quantitative description of risk in this case, but the quality of the numbers cannot be evaluated.

McKnight et al. (2022) noted that human control is essential in terms asserting physical meaningful results: if the covariance in the analysis is larger than the miss distance between the objects than any collision possibility does not make any physical sense. This is a perfect example that shows the importance of a quality check of an AI's results.

Foust (2021) visualised this issue explaining that many operators utilize autonomous systems – on "autodrive". These operators are often not aware of any potential movements of the spacecraft. This resembles two problems: (1) The operators do not have any control over the system and, especially, cannot verify if the movement was qualitatively in order. (2) By performing manoeuvers the background assumptions are rendered useless, since the orbits have been altered and new risk calculations need to be performed. Note that the suggested CREAM solution anticipates continuous orbit alterations – in contrast to Foust (2021) statement. McKnight et al. (2022) also emphasised that continuous movement of spacecraft impacted negatively on the estimation of collision probabilities. Also the initial suggestion of fully automatic operations for CREAM was later softened and modified in order to make possible the ”enabling later involvement of human operators and analysts…” (Astroscale, 2022).

This can be visualised by the following example: Imagine an AI which provides collision avoidance warnings for a spacecraft operator and decides on manoeuvring on its own. The input data into this system can be data from groundbased radars that track spacecraft and space debris and GPS-data from spacecraft. Assume furthermore, that the AI incorporates the data from the radars and GPS with certain error margins, but these error margins are exceed by the radar data due to the instrumental failure within the radar and by the GPS data due ionospheric disturbances. Thus the risk for collisions will increase significantly due to the larger error of the underlying data. The AI will underestimate the risk because of the assumption of unchanged error margins. A quality check by a human operator will visualise the changed background assumptions and providing this feedback into the AI can lead to a more realistic collision probability.

Omar (2023) provided a suggestion for the correct form of human feedback: RLHF (reinforcement learning with human feedback). The goal with RLHF is not only to provide feedback as a quality check, but also to aid the improvement of the AI. Omar (2023) points out that the correct form of human feedback needs to be chosen in forms of ratings, markers, rankings etc.

When focusing on risk management processes more, it is obvious that AI can resemble both positive and negative impacts on risk management. On the positive side, AI can aid risk management processes, but on the negative AI are often utilized in fast-developing areas where classical risk management procedures are not sufficient: “In this complex and fast-moving environment, traditional approaches to risk management may not be the answer …” (Baquero, Burkhard, Govindarajan, & Wallace, 2020). These authors also provide an explanation on why classical risk management models do not suffice: the assumption of semi-static models where the assumptions and boundary conditions are not changing and where workflows are connected to lengthy processes and their employment for classical risk types in rather simple systems. All this is challenged in the environments where AI’s are used. Firstly, orbital regimes are changing constantly and rapidly: new satellites are launched all the time and collision avoidance manoeuvres are subsequently performed all the time. Furthermore, a change from iterative to continuous orbital changes is foreseen. In addition, the AI's in operation are also developing. Hence, this resembles are highly dynamic system. Secondly, the AI's development is occurring at a high rate in a dynamic environment – in full contrast to traditional workflow where review processes are lengthy. Thirdly, space systems are inherently complex and bring with them new risk types, in particular the risk of employing AI's. Hence, it is obvious that classical risk management cannot address the challenges related to AI operations and new approaches are necessary. Baquero, Burkhard, Govindarajan, & Wallace (2020) confirm that such a new approach is of need, suggesting "derisking AI by design". As discussed above it is of essence to oversight the quality of the AI's operations without hindering the further development of the AI and without restricting the operations the AI was thought to perform from the beginning: "Tools such as model interpretability, bias detection, and performance monitoring are built in so that oversight is constant and concurrent with AI development activities and consistent across the enterprise. In this approach, standards, testing, and controls are embedded into various stages of the analytics model's life cycle, from development to deployment and use …" (Baquero, Burkhard, Govindarajan, & Wallace, 2020). Figure 22 describes this approach figuratively with two points to highlight: (1) the steps B-C-D-E resembling "get data" – "build" – "evaluate" – "industrialize" and (2) H resembling "review and approval for continued use". The first point reminds of the schematics for risk management under uncertainty proposed by Aven, Røed, & Wiencke (2008). Both schematics include a returning loop in order to reevaluate the performance of the operation. However, in this supposed risk management model for AI operations this loop appears earlier than compared with the model by Aven, Røed, & Wiencke (2008). The second point shows that the model for the AI contains a second point – not loop – to "review and approval for continued use".

Another problem when addressing quantification of collision risk and potential risks is orbital overcrowding and economical concerns. Operators have to weight collision probabilities against to ability to move the spacecraft. Orbits became more and more overcrowded, so there might be little space to move the spacecraft. In addition, moving a spacecraft costs fuel which costs money. Furthermore, fuel is very limited which puts further constraints. The question is how AI's will assess and weight such factors against a pure number which resembles a collision probability. Pultorova (2021) states: "So there must be threshold. But that means you are accepting a certain amount of risk." Here it is essential that human feedback is inserted into the decision making process in order to assess the acceptability of the risk in addition to the pure collision risk.

7. Conclusions

The topic of this thesis is “How do emerging technologies for collision avoidance in Earth’s orbit impact risk management processes?” The conclusion is that emerging technologies can aid risk management processes, but show an inherent risk attached to them. Not only is it necessary to address that risk, but also to develop new risk management processes where emerging technologies are employed.

References

Arroyo-Parejo, C., Sanchez-Ortiz, N., & Dominguez-Gonzalez, R. (2021). Effect of mega-constellations on collision risk in space. *Proceedings of the 8th European Conference on Space Debris* . Darmstadt.

Astroscale. (2022, March 10). *Astroscale Awarded 800K Euro ESA Contract to Boost Collision Avoidance Maneuver Capability on Congested Orbital Highways*. Retrieved from https://astroscale.com/astroscale-awarded-800k-euro-esa-contract-to-boost-collision-avoidance-maneuver-capability-on-congested-orbital-highways/

Aven, T. (2015). *Risikostyring* (2 ed.). Universitetsforlaget.

Aven, T., Røed, W., & Wiencke, H. S. (2008). *Risikoanalyse.* Universitetsforlaget.

Baquero, J. A., Burkhard, R., Govindarajan, A., & Wallace, T. (2020, August 13). *Derisking AI by design: How to build risk management into AI development*. Retrieved from https://www.mckinsey.com/capabilities/quantumblack/our-insights/derisking-ai-by-design-how-to-build-risk-management-into-ai-development

Beredskapsstyrelsen. (2022). *Nationalt Risikobilde.* Retrieved from https://www.brs.dk/da/redningsberedskab-myndighed/krisestyring2-og-beredskabsplanlagning/nationalt-risikobillede/

Brodkin, J. (2021, April 21). Retrieved from SpaceX denies claim that Starlink and OneWeb satellites almost collided: https://arstechnica.com/information-technology/2021/04/spacex-says-oneweb-spread-false-story-of-near-miss-satellite-collision/

ESA. (2022). *ESA's Annual Space Enivorenment Report.* ESA Space Debris Office. Retrieved from https://download.esa.int/esoc/downloads/BR-336_Space_Debris_WEB.pdf

European Commission. (2019). *Liability for Artificial Intelligence and other emerging technologies.* Retrieved from https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf

FCC. (2022, September 29). *FCC Adopts New '5-Year Rule' for Deorbiting Satellites*. Retrieved from https://www.fcc.gov/document/fcc-adopts-new-5-year-rule-deorbiting-satellites

Foust, J. (2021, March 19). *NASA and SpaceX sign agreement on spaceflight safety*. Retrieved from https://spacenews.com/nasa-and-spacex-sign-agreement-on-spaceflight-safety/

Foust, J. (2023, June 14). Retrieved from World Economic Forum offers new debris mitigation guidelines: https://spacenews.com/world-economic-forum-offers-new-debris-mitigation-guidelines/

FAA. (2023, February 17). Retrieved from FAA Proposes $175,000 Fine Against SpaceX for Not Submitting Required Pre-Launch Data: https://www.faa.gov/newsroom/faa-proposes-175000-fine-against-spacex-not-submitting-required-pre-launch-data

IADC. (2002). *Space Debris Mitigation Guidelines.* Retrieved from https://www.iadc-home.org/documents_public/file_down/id/5249

International Organization for Standardization. (2018). *Risk management — Guidelines (ISO Standard No. 31000-2018)*. Retrieved from https://www.iso.org/standard/65694.html

Kashyap, R. (2019). Artificial Intelligence Systems in Aviation. In *Cases on Modern Computer Systems in Aviation* (pp. 1-26). doi:10.4018/978-1-5225-7588-7.ch001

Kessler, D., & Cour-Palais, B. (1978). Collision frequency of artificial satellites: The creation of a debris belt. *Journal of Geophysical Research*, 2637-2646.

Mack, E. (2019, September 3). *ESA's near collision with SpaceX Starlink satellite blamed on a 'bug'*. Retrieved from https://www.cnet.com/science/esa-spacex-starlink-satellite-nearly-collides-with-european-aeolus-satellite/

Makris, S., & Aivaliotis, P. (2022). AI-based vision system for collision detection in HRC applications. *Procedia CIRP*, 156-161. doi:https://doi.org/10.1016/j.procir.2022.02.171

McKnight, D., Dale, E., Bhatia, R., Kunstadter, C., Stevenson, M., & Patel, M. (2022). A map of the statiscal collision risk in LEO. *73rd Internatonal Astronautical Congress.* Paris.

NASA. (2022, Mar). *Orbital Debris Quarterly News* . Retrieved from https://orbitaldebris.jsc.nasa.gov/quarterly-news/pdfs/ODQNv26i1.pdf

O'Callaghan, J. (2019, Sep 2). *SpaceX Declined To Move A Starlink Satellite At Risk Of Collision With A European Satellite*. Retrieved from https://www.forbes.com/sites/jonathanocallaghan/2019/09/02/spacex-refused-to-move-a-starlink-satellite-at-risk-of-collision-with-a-european-satellite/?sh=3356cab41f62

OECD. (2022, Sep 15). *Earth's Orbits at Risk - The Economics of Space Sustainability*. Retrieved from https://www.oecd.org/innovation/earth-s-orbits-at-risk-16543990-en.htm

Omar, M. (2023, 05 25). The Growing Need For Human Feedback With Generative AI And LLMs. Retrieved from https://www.forbes.com/sites/forbestechcouncil/2023/05/25/the-growing-need-for-human-feedback-with-generative-ai-and-llms/

Pagallo, U., Bassi, E., & Durante, M. (2023). he Normative Challenges of AI in Outer Space: Law, Ethics, and the Realignment of Terrestrial Standards. *Philos. Technol., 36*. doi:https://doi.org/10.1007/s13347-023-00626-7

Pultarova, T. (2021, August 18). *SpaceX Starlink satellites responsible for over half of close encounters in orbit, scientist says*. Retrieved from https://www.space.com/spacex-starlink-satellite-collision-alerts-on-the-rise

Rebhan, S., Einecke, N., Losing, V., Limmer, S., & Schmitt, S. (2020). Engineered Decision Tree for Judging Spacecraft Collision Risks. *8th European Conference on Space Debris.*

Space Sustainability Rating. (2023, June 17). *SPACE SUSTAINABILITY RATING*. Retrieved from https://spacesustainabilityrating.org/

Spaceref. (2022, February 14). *NASA Letter To FCC Regarding SpaceX Starlink Gen 2 System*. Retrieved from https://spaceref.com/status-report/nasa-letter-to-fcc-regarding-spacex-starlink-gen-2-system/

SpaceWatch Global. (2022, April 7). *Astroscale to lead ESA's collision avoidance study*. Retrieved from https://spacewatch.global/2022/04/astroscale-to-lead-esas-collision-avoidance-study/

Starlink. (2022a, October 08). *World's mostadvanced broadband satellite internet*. Retrieved from https://www.starlink.com/technology

Starlink. (2022b, October 08). Retrieved from Update: https://www.spacex.com/updates/#sustainability

TheVerge. (2021, April 9). *OneWeb, SpaceX satellites dodged a potential collision in orbit*. Retrieved from https://www.theverge.com/2021/4/9/22374262/oneweb-spacex-satellites-dodged-potential-collision-orbit-space-force

UNCOPUOS. (2017). *International Space Law: United Nations Instruments.* Retrieved from https://www.unoosa.org/res/oosadoc/data/documents/2017/stspace/stspace61rev_2_0_html/V1605998-ENGLISH.pdf

UNCOPUOS. (2021). *Guidelines for the long-term sustainability of outer space activities* . Retrieved from https://www.unoosa.org/documents/pdf/PromotingSpaceSustainability/Publication_Final_English_June2021.pdf

UNITED STATES SECURITIES AND EXCHANGE COMISSION. (2018). *ANNUAL REPORT PURSUANT TO SECTION 13 OR 15(d) OF THE SECURITIES EXCHANGE ACT OF 1934*. (https://www.sec.gov/Archives/edgar/data/789019/000156459018019062/msft-10k_20180630.htm, Producer) Retrieved June 2, 2023, from https://www.sec.gov/Archives/edgar/data/789019/000156459018019062/msft-10k_20180630.htm

Virgili, B., Flohrer, T., Krag, H., Merz, K., & Lemmens, S. (2019). CREAM - ESA's proposal for collision risk estimation and automated mitigation. *First International Orbital Debris Conference.*