# Space Futures Initiative Research Agenda

Madeleine Chang[1], Jack Davies[2], Michael Dello-Iacovo[3], Carson Ezell[4], Fin Moorhouse[5], Yannick Mühlhäuser[6], George Profitiliotis[7], and Olaf Willner[8]

[1]The School of Civil  Environmental Engineering, Cornell University, 527 College Ave., Ithaca, NY 14853
[2]Centre for Global Challenges, Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, The Netherlands
[3]Sentience Institute
[4]Harvard University, 56 Linnaean St., Cambridge, MA 02138, USA
[5]Future of Humanity Institute, University of Oxford, Trajan House, Mill St, Oxford OX2 0DJ, United Kingdom
[6]Institute of Physics, University of Tuebingen, Auf der Morgenstelle 14, D–72076 Tuebingen, Germany
[7]Department of Humanities, Social Sciences, and Law, School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Zografou Campus, 15780, Athens, Greece
[8]Bowers College of Computing and Information Science, Cornell University, 402 Gates Hall, Cornell University, Ithaca, NY 14853

September 11, 2022

# Contents

## 1. Space and Artificial Intelligence

*Introduction*

Space settlement depends on artificial intelligence (AI) because of the physical limitations of humans in the outer space domain. The intense physical characteristics of the space environment, including high energy radiation and low gravity, pose challenges to biological creatures. These challenges are not faced by AI systems. Hence, robotics and AI will be relied upon to perform many space activities to support settlement. These activities may include agriculture, mining, manufacturing, construction, and more.

Transformative artificial intelligence (TAI) is expected to significantly speed up scientific and technological progress, which are among the bottlenecks to further space development. TAI is generally defined as an AI system that significantly outperforms humans' cognitive abilities in all domains. Many space activities require technologies which do not yet exist or still require significant development, including in-space manufacturing, in-situ space resource utilisation, space-based energy, nuclear propulsion systems, terraforming, and more. TAI may be able to generate engineering solutions for these activities more quickly and effectively, or identify new scientific breakthroughs which allow for further space activities. Our trajectory in space may also be permanently affected by technologies dependent on TAI, including self-replicating probes and autonomous, self-expanding industry.

TAI systems could also handle enormously complex logistical processes necessary for autonomous space coordination, operations, and settlement. TAI systems may also be able to establish efficient operations on newly explored celestial bodies more quickly by coordinating the production of equipment and infrastructure based on local materials and conditions.

Furthermore, TAI may help to develop new governance mechanisms for space which outperform those we can generate alone. This may allow for a completely new space governance mechanism that replaces existing institutions. On the whole, the prospect of TAI introduces deep uncertainty into our thinking about humanity's future in space.

The ethics of using TAI systems may differ on space from on earth. Earth's inhabited regions are naturally safe for humans already, while space and celestial bodies present adverse conditions to human life, requiring some intervention. Furthermore, if TAI systems allow humans to develop self-sustainable colonies in space, the chance of a non-TAI related existential catastrophe may significantly decrease. However, once TAI is utilised for increasingly complex industrial and coordination processes in space, accidental or unforeseen disruptions can have catastrophic or existential consequences on an astronomical scale.

**A. Transformative Artificial Intelligence (TAI)**

- How will our space activities and governance systems prior to the deployment of a TAI influence long-term space outcomes?
- How will the geographic considerations of the space domain affect our ability to regulate the usage of TAI, given the implications of space geography for surveillance and rogue actors?
- What regulations should be placed on TAI usage or applications in the space domain?
- What scenarios can we expect if there are multiple TAI systems competing for space resources and space development?
- How can TAI be most effectively, ethically, and safely leveraged and regulated to support long-term space activities?
- Are there possible scenarios where TAI systems are developed in space, and what are the implications?
- To what extent could space settlement, or the desire for space settlement, accelerate TAI development timelines?
- How would TAI modify or accelerate timelines for space settlement and other space activities?
- How would TAI systems with different capabilities cause space development to proceed differently?
- How plausible is the rapid settlement scenario using self-replicating probes described by Armstrong and Sandberg (2013), on what timescales, and what would be the implications and risks?
- How plausible is self-sustaining, self-expanding industry? How would we expect this to look in practice, and what are the risks?

**B. Existential Risks**

- How do the ethics/risks of TAI usage in space compare to those on earth?
- How does the outer space domain impact TAI-related existential risks?
- To what extent do our concerns about malicious usage of TAI change when humanity becomes an interstellar species?
- How do technological innovations that may follow TAI—such as digital minds and self-replicating probes—impact existential risks, including non-extinction x-risks?

## 2. Space Technologies

*Introduction*

Our timelines and capacities for space development are dependent on many technological uncertainties. Some technologies, such as nuclear propulsion systems, might be possible within the next couple decades. Others, such as large rotating enclosures (i.e. O'Neill Cylinders) to support civilizations, may be further away. We have more uncertainty surrounding the technologies in very long term futures, many of which may not even be yet hypothesised. However, by exploring the fundamental laws of physics, we can gain insights into the outermost limitations on what might be technologically possible to accomplish.

Space manufacturing is likely to become the most cost-effective way to produce goods for settlers beyond earth. Furthermore, the microgravity environment permits the production of new products which cannot be produced under earth's gravity. For example, 3D-printed organs can be made in space, whereas the tissues would collapse during printing on earth. Such products may be manufactured for space-based or earth-based consumers. Manufacturing facilities can also theoretically replicate themselves using resources in space (i.e. self-expanding industry), increasing the rate of production without expenditures from earth.

Prolonged habitation beyond the Earth's biosphere requires a means to provide the necessary resources and conditions for life (i.e. water, air, food, favourable environmental conditions, proper waste management). Current life-support system technologies are unable to provide full loop closure by purely physicochemical means, so biological compartments are needed. The creation of a robust totally closed ecological life-support system is needed to ensure long-term survival of humanity beyond Earth.

Equally critical to the development of closed ecological life-support systems is the development of advanced energy production systems. Moving into deep space will require research into technologies for energy generation, storage, and distribution that will move beyond the currently available options of radiothermal generators, solar panels, and fuel cells. Candidate technologies include nuclear fission, fusion, and space-based solar power.

*Questions*

### A. Travel and Propulsion Systems

- Which propulsion systems are most cost-effective and feasible for space exploration to the outer solar system and beyond for human and robotic missions?
- What are the timelines of development for various propulsion systems? What are the main scientific and technological bottlenecks they face?
- How concerned should we be about the dual use nature of propulsion systems such as nuclear fusion? How should this impact the extent to which controls

are introduced on access to such propulsion systems?

- What further standards can we develop for various propulsion systems? How effective are existing standards with applications to propulsion systems such as the Safety Framework for Nuclear Power Source Applications in Outer Space?

**B. Settlements**

- Can we expect off-Earth settlements to become self-sustaining (economically, technologically, etc.)? Under what timelines?
- What are the biggest obstacles to self-sustainability of settlements on other celestial bodies?
- Which proposals and possible strategies for space settlements are most likely to come into existence (e.g. O'Neill cylinders, terraformed celestial bodies, etc.)?
- How do O'Neill cylinders compare in safety and expected suffering to other proposals for large-scale space settlement, such as terraforming celestial bodies?
- How would governance mechanisms work for megastructures containing large populations? What are the pros and cons of self-governance compared to international cooperation regimes?
- How can 'freedom engineering', or deliberately engineered solutions to maximise liberty in outer space, help mitigate malicious governance and totalitarianism in outer space?

**C. Biological**

- To what extent do we expect biological settlers to be involved in further space settlement relative to technological, autonomous systems?
- How will closed ecological life-support systems impact how, where, and when humans settle in space?
- How might the ethical considerations of human genetic modification be perceived differently in space than an earth, given the adversarial conditions?
- What are the limitations on space travel over long distances for humans, and how do we expect these limitations to impact our space trajectory?

**D. Energy and Manufacturing**

- In what ways can space-based materials and manufacturing in microgravity environments transform our economic, global health, and other systems?
- What regulatory regimes and protocols should be established for off-earth inventions and manufacturing facilities in space?

- What will be the terrestrial and space-based impact of advanced energy production and space-based manufacturing systems, including economic and environmental effects?
- How do energy levels available in space compare to the computational requirements for high energy demand projects, such as sentient simulations?
- How will projects with energy requirements that can only be met in outer space impact power dynamics?

---

## 3. Space Conflict

*Introduction*

There are currently few restrictions on the militarization of outer space beyond an agreement to not place weapons of mass destruction (WMD) in outer space via the Outer Space Treaty (OST). Many experts believe that great powers such as the United States and China are currently in a 'space race' to develop advanced technologies to defend their interests in space and have the ability to carry out counter space missions against adversaries. Placing limitations on space militarization is inhibited by political tensions between the West, China, and Russia, as well as the vast differences in perspectives on spacefaring between the Global North and Global South.

Anti-satellite (ASAT) tests, which can produce significant quantities of space debris, have been carried out by India and Russia in the past several years, and they have previously been conducted by China and the United States. Non-kinetic tactics such as signal jamming are also a threat to space assets. Almost all space assets can also have dual-use purposes, making restrictions difficult to create. While on-orbit servicing or rendezvous and proximity operations (RPOs) can be used to repair satellites, they can also be used to interfere with adversarial space assets. Asteroid deflection technologies for planetary defence could be used against adversaries, and there may be applications of terraforming to offensive operations against adversarial territories.

Furthermore, technologies such as higher resolution satellites can enhance surveillance on earth or celestial bodies, improving warfare capabilities in other domains. In the longrun, we can consider how the weaponization of space will evolve, which space technologies may be weaponized, and whether there are any rules and norms that we can establish to limit or prevent the weaponization of space altogether.

*Questions*

### A. Militarization

- Should any militarization of space be allowed? To what extent is it possible to prevent space militarization?
- How could a ban on the militarization of space altogether go poorly (e.g. non-compliant great powers or rogue actors develop an asymmetric advantage)? How can we mitigate these risks?
- How should we enforce a ban or limitations on the militarization of space?
- Should outer space militarization rules differ for various regions of space, such as in-orbit compared to other celestial bodies?
- What should be considered a space 'weapon' for the purpose of legal regulations? What should we do about edge cases such as earth-to-space weapons?

- Given the dual use nature of most space objects, should our definitions of space weapons be based on original purpose or intent?
- What other technologies used for spacefaring may have a high risk of being used for dual use purposes?
- How might planetary defence technologies, such as asteroid deflection, be used for dual use purposes? How does their development impact safety?
- Who should be responsible for ownership of planetary defence technologies? Should there be restrictions on their development and ownership?
- What factors would influence the frequency of first strikes in the space domain, both in the near-term and long-term future?
- How do we expect space weapons to evolve into the long-term future? How does the nature of warfare change with different levels of technological maturity?
- What should we expect an interplanetary, interstellar or intergalactic war to look like? What does the offence-defence balance look like?

**B. Deterrence**

- To what extent should there be widespread surveillance of the outer space domain?
- How could surveillance of outer space lead to asymmetric power advantages, and how can we mitigate the associated risks?
- What should outer space surveillance systems look like, and who should be responsible or allowed to maintain them?
- How does widespread surveillance in space affect the chance of conflict?
- What measures can we take to ensure that dual use technologies are not used for military purposes?
- What learnings from nuclear deterrence can we take into space weapons in general? What policies could have been put in place before nuclear weapons were a reality that would have positively impacted the future? Can we apply that to space weapons?
- How does deterrence apply over long time and distance scales in space? To what extent are actions attributable and threats credible?
- How does the principle of mutually assured destruction apply to use of weapons of mass destruction in the space domain?

---

## 4. Space Governance and Norms

*Introduction*

The Outer Space Treaty (1967) and subsequent treaties from the 1970s form the basis of modern international space law. Since then, the United Nations has passed several non-binding resolutions and principles, but the regulation of space lags significantly behind modern space technologies and developments. For example, there are only vague rules on space resource utilisation and space settlement (i.e. property rights). Competing interests at the United Nations have resulted in slow progress, but bilateral agreements have also played a role in shaping space law, particularly the United States-led Artemis Accords. Hence, our challenges in outer space norm-setting not only involve deciding upon the ideal principles to govern space usage, but also effective and inclusive institutional frameworks for addressing emerging and future space concerns.

Currently, binding resolutions at the United Nations apply to all space activities by parties who are signatories. Non-binding standards and guidelines outlined by the UN are also adopted by many nations, but some nations may not adopt these standards because of resource constraints or a desire to have a more friendly regulatory environment. However, technical standards are increasingly developed by industry-led groups (e.g. CONFERS) or intergovernmental organisations outside the UN (e.g. IADC). Maintaining many sources of standards and governing authority results in quicker rules and more technical expertise in the policy domain, but it can lead to fragmentation of governing authority and rules.

Having a high-level governing authority such as the UN can make rules more coherent at the expense of speed of implementing new rules, specialised technical knowledge, and a lack of experimentation. As humans settle further into space, new political entities may be discussed and established. This leads to questions such as whether there should be an authority to govern all of humanity's sphere of influence in space, which would have the ability to delegate authority to space-based political entities, or whether such entities would be entirely independent.

*Questions*

**A. Enforcement**

- How can we create scalable verification mechanisms for rules and norms as civilization expands across space and time?
- How can we ensure that objects launched into space are registered, including military objects?
- What further transparency and confidence building measures (TCBMs) could we create to increase trust between space actors?
- Should registration be extended beyond launches alone to include other space activities such as resource extraction, in-orbit servicing, and

rendezvous and proximity operations (RPOs)? Launches from other celestial bodies? Who should maintain such a system?

- If a register of space activities is developed, how can we be certain all actors are complying with the requirements?
- How can national governments effectively supervise private in-space activities? Is this system scalable?
- Should there be a unified international database to track all space objects, and if so, who should be responsible for maintaining the database?
- How can we create scalable space situational awareness (SSA) and space traffic management (STM) systems, and who should be responsible for them?

### B. Norms and Values

- What norms and values are most important to promote among space actors to mitigate long-term risks?
- To what extent are binding rules for outer space more impactful than non-binding norms or shifts in values?
- How effective are technical standards at promoting safety and mitigating risks in outer space relative to shifts in values and norms?
- To what extent should laws and expectations in space be based on clear definitions rather than expected norms and behaviours, as promoted by the United Kingdom with its UN resolution on reducing space threats?
- How much divergence in norms and values can we expect once humanity spreads over interplanetary and interstellar distances? What are the potential risks of such divergence, and how can we mitigate them?
- How might principles of values handshakes and acausal trade be applicable to space governance, including pre-commitments to actions or extraterrestrial contact scenarios?

### C. Distribution of Benefits

- What measures can we take to ensure that all of humanity benefits from space exploration and development?
- To what extent should we consider antitrust measures, capping profits, or distributing revenues to prevent extreme wealth imbalances within the space economy?
- How does Article I of the Outer Space Treaty relate to the distribution of benefits related to space resources?

- What are the legal implications of a state or company that develops a monopoly on exploitation of a particular mineral or celestial body?
- How may the geographical structure of outer space contribute to the formation of monopolies over critical resources, or the lack thereof?

### D. Governance Mechanisms

- How much authority over space governance should be maintained by high-level actors, and how much autonomy should be delegated to local actors and settlements?
- To what extent can there be high-level governing frameworks across long distances in space (i.e. interstellar)?
- To what extent can we apply polycentric governance to outer space? What concrete measures could be taken to promote polycentric governance?
- How worried should we be about a regulatory 'race to the bottom,' where countries compete to be the most favourable places to register private space companies and spacecraft (similar to 'flags of convenience')?
- What role should the UN, individual nations and non-state actors (NGOs, NewSpace, etc.) have in space governance authorities and decision-making?
- To what extent could there be communication and coordination between various interplanetary and interstellar human or post-human settlements?
- Should there be governance authorities in outer space that provide large-scale shared infrastructure to all actors (e.g. public goods)? How feasible is this, and what would the implications be for coordination?
- How effective could interplanetary or interstellar space institutions be for conflict resolution or responding to disasters (i.e. asteroid strike or supernova)?

---

## 5. Space Sustainability

*Introduction*

Low Earth Orbit (LEO) describes an orbit between about 160 km and 1000 km above earth's surface. In comparison, objects in geosynchronous orbit (GEO), which maintain a constant position relative to earth's surface, orbit at 35,786 km. Most space activities currently occur within LEO, and there are a limited number of orbital slots available within GEO. Hence, the governance of both regions is relevant to international and national space policy.

LEO is becoming increasingly congested because of increasing space activities, growing space debris, and the proliferation of small satellite constellations. Such congestion creates a risk of runaway Kessler syndrome, where collisions of space debris occur in positive feedback loops, rendering LEO unsafe and unusable. Furthermore, increased usage of LEO can cause radio frequency interference for GEO satellites. Satellite mega constellations can also interfere with astronomical observation on earth's surface.

A sustainable governance regime for space debris should be supported by effective and adaptive technical standards. There are guidelines for satellites to be removed from LEO and GEO at the end of their life, but proper practices are not always followed. Furthermore, best practices are evolving over time based on learning. Methods such as active debris removal (ADR) can also be used to reduce space debris, but there are dual-use concerns associated with using the same technology for interference with adversarial satellites. Hence, clear communication protocols and the recording of activities is essential.

In the long-run, space resource utilisation may also lead to unsustainable practices associated with space resources. Concerns about resource depletion in the long-term future may parallel similar concerns that exist on earth today. We may be able to apply our best practices with common pool resource (CPR) management to the space domain to ensure sustainable use of space resources.

*Questions*

### A. Orbital Debris

- What would be the implications and risks of various levels of debris in low-Earth orbit? What risks does our current trajectory suggest we will face?
- How expensive and feasible are efforts to reduce and clean up space debris likely to be? Which methods and technologies are most promising?
- If orbital debris limits our space access in the near future, what would be the implications for our trajectory in space? Would there be any long-term implications for governance frameworks, cooperation, and differential technological development?

- How likely is Kessler syndrome? How catastrophic would Kessler syndrome be for civilization in both the near-term and the long-term?

**B. Space Resources**

- To what extent can we expect the depletion of solar system resources, and over what timescales?
- Should we establish monitoring and governance regimes for space resource systems? What would such systems look like?
- Which space resources are most likely to be depleted? What are the implications for settlements in the solar system?
- How can we apply common pool resource (CPR) governance strategies to space resource systems? To what extent would they be effective at promoting sustainability?
- How will the accessibility and possible uses of space resources evolve over time given technological development, and what are the implications for the scarcity and value of various resources?
- How should the environmental benefits and costs of off-Earth mining be accounted for?

---

## 6. Space Resources and Property Rights

*Introduction*

Definitions and the specification of protections and rights are essential to ensuring space activities can be conducted safely, and that those who conduct them can receive the benefits in a manner aligned with international law. The legality of extracting outer space resources, such as water ice or metals, has been a contested subject at the international level, in large part because of the non-appropriation clause of the Outer Space Treaty. However, most experts agree that space resource extraction is legal under the Outer Space Treaty, and several national governments—including the United States, Luxembourg, the United Arab Emirates, and Japan—have passed laws to permit the extraction of space resources for scientific or commercial use.

The non-appropriation clause is even more evidently opposed to the creation of permanent property rights in space, raising questions about how actors can conduct commercial activities and other space activities free from interference. A popular proposal, which is included in the United States-led Artemis Accords, suggests creating temporary safety zones around space activities to prevent harmful interference; protection from harmful interference is granted in Article IX of the Outer Space Treaty. Alternative systems to safety zones for creating a resemblance of property rights could also be considered.

Other resources are particularly scarce and valuable in outer space, including both orbital slots and space resources. Governance mechanisms will need to be established for orbital slots and radio frequency bands on celestial bodies—the International Telecommunication Union (ITU) does this for earth's orbit. Certain regions of celestial bodies, such as the peaks of eternal light on the Moon, may be particularly valuable to both scientific and commercial activities. The Lagrange points are locations in space where objects are stable because of being gravitationally bound by two larger objects, and these locations may become scarce and highly sought after for space activities.

Space resource extraction brings up questions of how and to what extent benefits should be distributed to all of humanity. This may include the sharing of revenues, or it could refer to mandated capacity building efforts such as technology transfer and knowledge sharing. Space resources could also be operated as a utility—humanity's common endowment. In addition, space resources should be protected for future generations, including benefits from their use.

A failure to create effective systems for allocation resources and property could lead to dangerous power asymmetries in outer space, or inhibit coordinated space development. Both of these outcomes may negatively impact long-term space development. Hence, as we consider governance regimes for space resources and property, we should consider how effectively the regimes scale to longer futures and minimise risks of adverse outcomes.

*Questions*

**A. Property Ownership**

- How should safety zones, mentioned in the Artemis Accords, be defined and enforced?
- Are there any alternatives to safety zones that we should consider for protection of space activities in the near future?
- What do ownership and property rights mean in outer space? Does mining/extracting resources imply property ownership or appropriation? Or does the term simply apply to settlement?
- How sustainable is a lack of private property in outer space in the long-term future?
- What will ownership and property rights look like for space settlements on celestial bodies in the near future and long-term future?
- What political mechanisms could we consider for allocating property rights on celestial bodies? How could such political mechanisms be used to limit totalitarian rule?
- How may property rights differ across the celestial environment (e.g. planets, asteroids and comets, orbits, star systems, interstellar space)?
- What are the necessary conditions for an effective outer space property rights regime?
- Should rules for property rights be rigid and established before exploration, or adaptive and changing with learning?
- What are the implications of ownership of large-scale projects, such as Dyson spheres, for power dynamics?

**B. Zoning and Development**

- How could zoning be used as a mechanism to support decentralisation of power and mitigation of totalitarian outcome risks?
- How should the zoning process work for space settlements? Should master planning occur, or should zoning be adaptive as settlements evolve?
- Who should be responsible for creating and enforcing zoning rules on celestial bodies?
- How should competing interests be considered via zoning, such as science (ex: radio quiet zones) and commerce (ex: mining on the lunar far side)?
- What effects of various celestial environments (i.e. low gravity means airborne dust travels further) are necessary to consider when creating

zoning laws? How should we update zoning laws based on new learning about the environment?

- What are the implications of a lack of zoning considerations before space settlement, including due to a lack of coordination between multiple actors (e.g. U.S. and China)?
- Should there be zoning regulations in outer space outside of celestial bodies (i.e. interplanetary or interstellar space)? What would the zoning and zoning process look like?
- How should we govern usage of Lagrange points and travel corridors?

**C. Resource Allocation**

- How could monopolies lead to catastrophic risks associated with power asymmetries? To what extent should there be antitrust laws and other measures to prevent monopolies in outer space?
- How should we balance the interests of promoting space exploration, limiting first-come first-serve rules, preventing dangerous power asymmetries, and protecting future generations when creating regimes for resource allocation?
- How should the right to use valuable locations on the surface of celestial bodies, such as the Lunar south pole, be decided?
- How should radio frequencies and orbital slots be allocated on celestial bodies in an equitable manner?
- What economic mechanisms (e.g. auction or sharing) would be best for allocating space in low-Earth orbit, or orbits around other celestial bodies when these orbits become much more valuable and crowded?
- What timelines should we expect for asteroid mining? From where should we expect demand for space resources (i.e. in space or on earth)? How will this impact the economy?
- What could the medium- and long-run economics of space look like? How will the value and scarcity of goods change over time in the space economy, and how will space economies influence power dynamics?
- How should access to critical resources needed for survival be allocated to mitigate risks of civilizational collapse or totalitarianism in outer space?
- How should benefits from space resource extraction be distributed to benefit all of humanity and future generations?

---

## 7. Existential and Catastrophic Risks

*Introduction*

One argument in favour of early space settlement is to avoid placing all of our 'eggs in one basket', reducing existential risks for humanity if a catastrophic event were to occur on earth. However, for the upcoming decades and centuries, a space settlement would likely still be highly dependent on earth for resource needs. It also is unclear how the cost-effectiveness of space settlements as a protection against existential risks compares to other interventions like isolated refuges. Nevertheless, if large scale space settlements are built, it seems likely that strategic planning and governance could steer this process in a direction that increases their contribution to x-risk reduction.

Poor space governance and irresponsible space activities can also increase the chance of existential risk. For example, the need for robotic exploration of space, and possibly genetic engineering to allow humans to survive in the adverse conditions of space, may also speed up TAI timelines or lead to irresponsible deployment, raising the probability of an TAI-induced existential catastrophe. Larger space settlements also increase the possibility of very destructive space-based warfare and prolonged arms races that would increase existential risk .

There are also existential risks associated with extraterrestrial intelligence (ETI). Better understanding the nature and frequency of ETI can adjust our strategies in contact scenarios and prioritisation of thinking about ETI contact. Humanity has also conducted past missions related to messaging extraterrestrial intelligence (METI), which may be an existential risk factor.

Forward or backward biological contamination could also have catastrophic implications, both for humanity or natural astrobiological life. Other origins of life that we discover in outer space have value, and thus it is important that we take measures for their protection and consider the implications of our space activities.

Other non-human sentient beings may also exist in outer space, including animals that we bring (wild or farmed) or digital minds that we create. Hence, we should consider the suffering risks (s-risks) that can arise in long-term space futures and the actions we can take to mitigate them.

If space development goes poorly, there is also a risk of becoming locked into a suboptimal state where humanity fails to reach its potential and large-scale suffering, domination and warfare remain the norm. Measures that we can take to avoid locking in negative norms, values, or power structures are worthy of consideration.

For existential risks related to TAI, see section Space and Artificial Intelligence.

*Questions*

**A. Distributed Risk**

- How effective is space settlement as a strategy for mitigating existential risks?
- Is settling Mars a realistic strategy for mitigating existential risk this century? How cost effective would this strategy be relative to other options (i.e. terraforming earth or colonising a different celestial body)?
- How long should we expect between landing the first humans on Mars, establishing the first settlement, and establishing a self-sustaining settlement?
- How could the settlement process be steered towards an outcome that provides maximum protection against existential risk?
- Which existential catastrophes might space settlement this century protect against? Which would they not help with?

**B. SETI and METI**

- How can we further investigate the problems of the "Great Filter" and the Fermi Paradox effectively?
- How should we reform our response strategy to extraterrestrial contact scenarios? How can we respond in a way that minimises the probability of a negative outcome?
- Who should be responsible for coordinating communication with ETI?
- How should we think about the potential contact scenarios with ETI? What are the implications for existential risks?
- How can we empirically analyse unidentified aerial phenomena (UAP), and what are their implications for SETI?
- How can we improve our global system of improving, screening, and responding to technosignature sightings?
- What are the potential implications of a mission such as Breakthrough Starshot or searches for technosignatures? How should we update our bounds on ETI probabilities based on various degrees of inconclusive findings?
- What information hazards may we create by revealing information about humanity to a hostile ETI?
- How can we use our knowledge of advanced artificial intelligence systems to better understand the nature and intent of possible ETI we discover?

- How meaningful are various models for thinking about the quantity and spreading of extraterrestrial civilizations?
- What are the benefits and risks associated with messaging extraterrestrial intelligence (METI)?
- How can we ensure sufficient oversight on decisions about METI? If we do broadcast messages, what information, if any, is safe to include?
- How might an ETI respond to METI signals, and how can we interpret the response? How might an ETI distort, filter, or falsify information that it sends us about its existence?

**C. Biological Contamination**

- What risks could be posed to astrobiological life or human settlements on other celestial bodies due to the introduction of terrestrial microbial life (i.e. forward contamination)?
- What risks could be posed to the Earth's biosphere due to the introduction of extraterrestrial microbial life to the terrestrial environment (i.e. backward contamination)?
- How likely are negative outcomes from biological contamination, what can we expect them to look like, and what are the extreme risks associated with cascading failures?
- How does the increase in spacefaring driven by commercial actors influence risks of biological contamination?
- How can we update standards relating to preventing contamination in light of the new space actors and activities?

**D. Suffering Risks**

- To what extent should we be concerned about the existence and suffering of non-human sentient beings in outer space?
- How could we mitigate future digital minds suffering in outer space?
- What measures could we take to prevent animal suffering in outer space, including both wild animals and farmed animals?
- How might conflict or moral warfare between civilizations with differing values lead to s-risks?
- To what extent are there s-risks associated with terraforming (via wild animal suffering) or directed panspermia? Can we conduct such activities in a manner that mitigates s-risks?
- What are other possible s-risks or related scenarios in outer space that we should take measures to avoid?

**E. Values Lock-In**

- How can we ensure positive norms and values are spread to other regions of the galaxy as expansion occurs?
- How can we mitigate negative value drift in the far future?
- How concerned should we be about value lock-in in the near future, particularly as a result of transformative technologies such as TAI?
- What are the end states for space expansion that we can expect to be stable over the very long term? What are likely states and paths that will not be stable?
- How effective are pre-commitments to certain norms and values before expansion occurs for positively influencing the far future?

---

## 8. Meta

*Introduction*

Much has been written about long-term space futures and governance in science fiction novels, and increasingly academic literature. However, the impact of this literature on policymaking and technological development is under researched. Furthermore, it is still unclear the extent to which ongoing space activities and governance can influence far futures. Understanding the pathways through which we can improve long-term outcomes and mitigate associated risks would be highly beneficial.

Space activities and governance also have few precedents of similar nature and weight because of the unique geography of the space domain in terms of scale and habitability, as well as the potential to lock-in certain long-term outcomes given the advanced technologies necessary for space activities. A consideration of various methodologies, including the application of potential precedents to space governance, could improve research in this emerging field. Furthermore, the outer space domain has implications for a wide variety of disciplines. The long-term impacts of space activities on agriculture, medicine, energy, commodities, and many other fields requires knowledge of experts in these areas. An application of their methodologies to thinking about space futures could help shape our understanding of space futures.

In order to understand the path dependencies of space futures on current activities and propose improvements, one also needs to have a robust understanding of the current space governance framework and the goals and interests of the relevant actors in the space domain. Such an understanding can lead to an identification of tractable policy solutions and improvements, opportunities to improve norms and coordination, and the most significant risks given the current state of space governance. The Center for Space Governance is developing a mapping of the space governance landscape, which we intend to use to inform our research.

*Questions*

**A. Impact**

- To what extent is space governance important, tractable, and neglected?
- What is the theory of change behind space governance? How impactful is influencing governance institutions compared to influencing norms and values of space actors?
- How important is space governance relative to other cause areas, and how contingent is the importance of space governance on reducing other risks (i.e. misaligned AI)?
- What is the current state of literature in space governance outside of long-termism? What are the key similarities and differences in approaches and values?

- To what extent is long-term space governance path dependent on modern space policy and governance? How does near-term space policy and coordination affect long-term outcomes?
- What are the path dependencies that should we consider for how space development will turn out? What are the implications of our existing cultural norms and technological progress on long-term space expansion?
- What factors influence how space actors and policymakers address space development and governance? What is the role of science fiction, academic literature, and civil society?
- What are the knowledge pathways through which space-related scientific advancements lead to commercial space activities, and how can we shape them to mitigate risks?
- How can actors outside the traditional space governance framework impact space policy, norms, and values?
- How does the rate of space development affect the outcome? Could space settlement and expansion be sped up or slowed down to minimise the chance of a bad outcome?

**B. Methodologies**

- To what extent can we make accurate statements about future space development with high uncertainty about the space environment and our future technological progress? How robust are different methods to those problems?
- Can we use decision-making models such as decision-making under deep uncertainty (DMDU) approaches to mitigate risks and uncertainties in outer space?
- How accurately can we quantify certain probabilities and risks involved in space development, such as Kessler Syndrome?
- What can we learn from historical treaties when creating future treaties for outer space (e.g. Antarctic Treaty, Seabed Arms Control Treaty, Partial Nuclear Test Ban Treaty, etc.)?
- What can we learn from historical examples about how settling beyond Earth could play out politically and economically (e.g. colonisation, Gold Rushes)?
- To what extent should we focus on avoiding the worst-case scenarios by considering the maxipok rule when thinking about space futures, relative to maximising probabilities of best-case scenario outcomes?
- How can we use the vulnerable world hypothesis to improve our thinking about existential risk management efforts in outer space?

**C. Current Landscape**

- Who are the key stakeholders in the space domain? How will their relative importance change over time?
- How exactly is space policy made in the US/Russia/EU/China? What are the relevant state agencies, companies, industry associations, think tanks and interest groups?
- How will presently non-spacefaring states or emerging spacefaring states impact space geopolitics in the future?
- What are the key differences in preferences between key stakeholders in the space domain?
- To what extent does the current framework for space governance favour the preferences of particular actors and neglect the preferences of others?
- To what extent will existing space governance frameworks have an impact on shaping future space development, including the Outer Space Treaty or Artemis Accords?
- What are the implications of the Outer Space Treaty for space resource extraction and other emerging and future space activities?
- How will existing relationships and levels of coordination between actors in the space domain impact future trajectories?
- To what extent do competitive pressures currently exist between competing states and corporations in the space domain, and how does this influence space development?
- How effective are the procedures and approaches of the Conference on Disarmament on military-related diplomacy in outer space?
- How does current great power conflict (e.g. U.S., China, Russia) influence long-term outcomes in the space domain?