Annotation Guidelines: Ideas from the Open Space Innovation Platform

The overall goal of the annotation process is to release a dataset of public ideas annotated with the type of information presented in the description of ideas.

An idea in the Open Space Innovation Platform is a free text description of a proposal to carry out a project or in some cases a phd thesis. The idea description can include: i) the motivation for the project describing the challenges or existing problems, ii) the main steps or activities of the project, iii) the expected benefits of the project or research, iv) the proposal itself, and v) the context.

We count on human annotators to identify **sentences** or **sequences of sentences** in ideas descriptions that can be classified as:

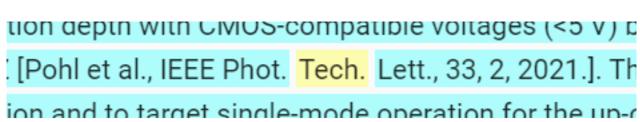
1. **General Problem/Challenge**: General Problem or challenge addressed in the idea
2. **Proposal/ Solution/objective**: Main proposal of the Idea. In this category we include the objective pursued in the idea, or the research goal.
3. **Benefits/Outcomes**: benefits and outcomes if the proposal is accepted or the project carried out.
4. **(Proposal/Solution) Steps/tasks/Elaboration**: Detailed description of the solution or a sequence of tasks to be carried out in the project
5. **Context**: General descriptive text including related facts

Note that the appearance of sentences of any of the previous types in an idea description is optional.

It is worth to note also that a few ideas are very badly described, sometimes making it impossible to understand the proposal. For such case we suggest discarding the idea and go to the next one.

Sentences with more than one label. Sometimes a sentence can mention both the problem/challenge and proposal/solution. In this case we mark the sentence as a proposal/solution. In the rest of cases, we assign the label corresponding to the first part of the sentence.

Bad sentence splitting. In some cases, the sentence splitting is not correct, and it identifies parts of a sentence as an independent sentence (see example below).



In those cases, we should annotate the subpart of the sentence with the same label as the whole correct sentence (in the example "Tech", should be annotated with the "Steps" label).

Annotation Tool

As annotation tool we use Label Studio. This tool allows you to easily select a sentence and label it with one of the predefined classes (see above). You will receive credentials to login in label studio, and a short tutorial on how to use the tool.

Annotation Process

You are assigned 60 ideas to annotate. Out of the 60 ideas, 30 of them are also annotated by another annotator B, and the other 30 by annotator C. When the annotation of the ideas has ended you will get together with annotator B to analyze the annotations on which you both disagree until you come to an agreement on the right class. Similarly, you will get together with annotator C to check on your disagreements.

Annotation Examples

Next, we provide 4 ideas that have been annotated using the 5 classes defined above and the following color scheme:

System analysis of a food processing facility for lunar habitats

Missions to the Moon come with a myriad of challenges and even though we may have the technology to bring humans to the Moon, we do not have the required food systems to sustain them. [Challenge] Ultimately, food is a limiting factor for future lunar habitats. [Challenge]

Due to resource and cost limitations and food acceptability challenges, future lunar habitats cannot rely solely on resupply missions from Earth. [Context] As a result, in situ food production becomes a necessity. [Context] Greenhouse modules for lunar habitats have come a long way, but there is still a research gap about the requirements for the interface between the greenhouse module and the habitat. [Context]

If the food production on the Moon is ever to become truly independent, lunar greenhouses must reach their full potential. [Context] For this to be achieved, a framework for a lunar food processing facility needs to be made, including food storage facilities and crew kitchens adapted for both early habitats and future settlements. [Solution]

As a foundation, any challenges within the provision of food for space missions must be clarified with a full system analysis of the space food system. [Context] This analysis will include a brief overview of the physiological and psychological challenges that food may counteract, as well as the prerequisites for nutritional qualities, followed by a systematic overview of the prospects and challenges of in situ food processing methods that are required for lunar food production. [Steps]

Processing methods and equipment need to be determined in regard to crops and other possible food supplies, e.g., future protein alternatives such as insects and cultured meat. [Steps] Not only will food processing extend shelf life, but also create new products, introduce new crop varieties, and increase the nutritional qualities of certain foods. [Benefits] Furthermore, a food processing facility may lead to better caretaking of side streams from the greenhouse or crew waste, which in turn can be refined into useful products that promote the circularity of the closed loop habitat. [Benefits]

Towards the understanding of the subsurface planetary oceans using ocean-induced magnetic field

Icy moons of the giant planets contain large subsurface oceans beneath external ice shells. [Context] The understanding of the subsurface oceans is important because the oceans transport heat and chemicals from the deep interior to the surface, influence the tectonic evolution of the ice crust and are possible candidates for hosting life. [Challenge] In this project, we propose to study the ocean-induced magnetic field (OIMF) which is generated by the circulation of the conductive seawater in presence of the ambient Jovian magnetic field via the process of electromagnetic (EM) induction. [Solution] In particular, our goal is to set limits of OIMF observability by space probes and assess the prospect of the EM methods in detailed investigation of icy moon interiors. [Solution]

The OIMF has played a crucial role in interpreting Galileo mission results and in detecting the subsurface oceans on Europa, Callisto and Ganymede (Zimmer et al., 2000; Khurana et al., 2002; Kivelson et al., 2002). [Context] However, it is likely that the potential of OIMF has not yet been fully explored. [Context] Moreover, two space missions to the satellites of Jupiter (NASA's Europa Clipper and ESA's JUICE) are expected to bring new data in early 2030s, providing unprecedentedly detailed information on the structure and temporal variations of the magnetic field in the vicinity of Europa and Ganymede. [Context]

The modelling of OIMF represents a multidisciplinary project which requires effective synergy between scientists from different geophysical branches. [Context] Our team is experienced in the modelling of OIMF (Šachl et al., 2019, 2022; Velímský et al., 2018, 2019) as well as the modelling of ocean circulation (Kvorka et al., 2022; Šachl et al., 2020) and ice shell deformation (Kvorka et al., 2018; Čadek et al., 2021) in icy moons. [Context] Thus, the readiness of our team for the proposed project is clearly high and we are capable of completing it successfully. [Context]

Enabling Real-Time Lunar Communication Services based on Bundle Protocol with Quality of Service Extensions

The near future is full of missions to the Moon that will need a robust communication system. [Context] This is where ESA's Moonlight project comes into play. [Context]

However, the question of how to establish real-time communications in spite of the unavoidable large delays within Near Space communications is still to be answered. [Context] Current data relays in space typically rely on a bent-pipe approach or circuit-switching. [Context] The Delay/Disruption Tolerant Networking (DTN) Bundle Protocol (BP) is an excellent starting point for this question to be answered, since it provides a packet-switching approach like it is used in terrestrial internet. [Challenge] It is desgined to provide end-to-end communications for highly stressed environments, that is, with large delays and intermittent connectivity. [Challenge] It does not however meet the expected reliability and maximum guaranteed bounds for latency and packet lost. [Challenge] Consequently, an extension to the BP is required, taking into account the necessary QoS and guarantees for the Moon-to-Earth connection to be stable and robust. [Solution]

Within this extension, scheduling of different segments (over multiple hops either pre-scheduled or on-demand) as well as the optimization of communication services must be taken into account. [Steps] For this to be efficient and achievable at a larger scale, pattern recognition and machine learning will have to be used. [Steps] Moreover, it is not only important for the real-time services to be developed, but also for them to successfully coexist with non-real-time services, such as store-and-forward scheduled data transmissions or messaging services. [Steps]

Furthermore, the creation of a scalable system must be a persistent goal. [Steps] Hence, a theoretical scalability analysis as well as a scalability simulation must be performed in order to ensure the future possibility of a larger number of communication services. [Steps]

AI-Support to generate conceptual designs for Concurrent Engineering studies

To improve performance, cost- and time-efficiency during the preliminary design phase of new space missions, ESA implemented the Concurrent Design Facility at ESTEC to conduct concurrent engineering (CE) studies. [Context] Human experts elaborate in a team effort a conceptual design - a fundamental understanding of a novel mission. [Context]
In today's era of New Space, cost- and time-efficiency is more important than ever. [Context] Furthermore, novel challenges like the vastly growing market of Commercial off-the-shelf components and unexplored mission ideas or business cases further exacerbate the design process, while known solutions are preferred over exploring new ideas - therefore decreasing the chance of potential innovations. [Challenge]

What if there was an AI assistant, that could support the process to drastically increase the scale of the explorable design space for the CE team and explore design options independent of "well-known solutions" to further improve the CE efficiency and level of innovation? [Solution]

Such an AI assistant would function as an extension to the CE design team, which would quantify the relevant mission requirements/constraints and use the generated conceptual system design for their own sessions within minutes. [Steps]
An AI, that gained its knowledge by applying well-established engineering calculation, just like the human engineers, would not only maximise the human-AI compatibility, but also make the underlying decision-making reasons adjustable and explainable. [Steps]
The proposed AI system would take a set of requirements and constraints as an input and forms a complete system design, based on a database of available components, therefore formalising a real-world design of a mission/system idea. [Steps]
The feasibility of such an AI system could already be proven, designing a CubeSat system utilising Deep Reinforcement Learning and real-world commercially available satellite components. [Steps]
Further research would expand the capabilities of such a design tool and implement other systems as well. [Steps]