10^{th} INTERNATIONAL SYSTEMS & CONCURRENT ENGINEERING FOR SPACE APPLICATIONS CONFERENCE (SECESA 2022)

5 - 7 OCTOBER 2022

A System Engineering Recommendation System based on Language Similarity Analysis: an Application to Space Systems Conceptual Design

Paul Darm(1), Audrey Berquand(2), Luis Mansilla(2), Annalisa Riccardi(1)

(1)***Affiliation***
16 Richmond St
Glasgow G1 1XQ
United Kingdom
Email: {paul.darm, annalisa.riccardi}@strath.ac.uk

(2)***Affiliation***
Keplerlaan 1
2201 AZ Noordwijk
Netherlands
Email: {audrey.berquand, luis.mansilla}@esa.int

Abstract

In Model-Based System Engineering (MBSE), the creation of complex engineering systems is facilitated by a standardised engineering data model and model version control, both of which generate valuable data after each conducted study. However, there are currently few to no approaches, reusing the information and knowledge from previous engineering studies. In this work, we present a new recommendation system, based on a widely adopted engineering data model, defined in the ECSS-E-TM-10-25A technical memorandum. This engineering data model is used by the European Space Agency (ESA), associated partners, as well as in other engineering domains. An engineering model (EM) is a hierarchical decomposition of an engineering system, providing information about the overall system, design options but also about low-level components. The novel recommendation system leverages a Knowledge Graph (KG) as a unified framework for storing multiple EMs. State-of-the-art semantic similarity Natural Language Processing (NLP) techniques are then used to define similarity between higher level information, so called metadata, associated with each EM. Textual information, such as the "Mission Objectives" of each study, are encoded with a neural language model into a vector representation, which allows to calculate a similarity metric between them, and then compare past-mission metadata with proposed metadata of a new study. In addition, a similarity between lower-level engineering components in the KG is described through the Jaccard metric, which compares components based by the set of parameters that each of them are associated with. By firstly clustering similar engineering designs through their associated metadata and then identifying analogous components in each cluster, the algorithm is able to recommend engineering components for new studies. In the results, the functionality of the approach is demonstrated as a pilot study for spacecraft conceptual design.

1 Introduction

As defined in [1],Model-Based System Engineering (MBSE) is a holistic system engineering approach, centered on the system model, the "*sole source of truth*". MBSE is quickly becoming a preferred design approach as it allows to maintain the consistency and manage the complexity of the system development process. Several organisations, across the manufacturing and production fields have initiated or completed the transition from a document-centric to a model-based approach [2].

The data model defined in the ECSS-E-TM-10-25A technical memorandum [3] contributes to the MBSE approach by defining guidelines for model-based data exchange for the early stages of engineering design. Furthermore, it contains a decomposition of a system, down to subsystem and equipment level, with a defined lists of parameters and disciplines. The ECSS, standing for European Cooperation for Space Standardization, is an initiative launched by the European Space Agency (ESA) in 1993 to define a coherent and single set of standards for all European space activities [4]. The ECSS-E-TM-10-25A is however not specific to the design of space systems and has also been applied within the maritime and defense fields.

While the ECSS-E-TM-10-25A provides the decomposition for a system, the comparison with past models is not yet feasible. Past-mission analysis is however a key initial step, often performed manually by engineers, to kick-start their design, and better evaluate initial parameters. It is therefore of interest to explore methods to enable analysis of past missions to provide insight for future studies. Multiple approaches have been developed in recent years by the National Aeronautics and Space Administration (NASA) to facilitate the design process of new studies. This includes a system that based on established rules-of-thumb, historical missions, and through multitude of metadata inputs by the user can pick from a catalogue adequate components or adapt mission parameters [5, 6]. Similarly, several algorithms are already used to predict the cost of a mission and verify that it will not exceed the allocated budget [7]. Access to previous data models can thus accelerate and enhance the design of new systems.

However, the complexity of space mission design and engineering systems in general with multiple design options create a high dimensional search space. With the amount of different design options, choosing the most important ones to accurately compare and match heritage information quickly becomes a complex task. An approach to enhance knowledge reuse is the one described in [8]. The authors propose to implement a conversational recommendation system, which would allow the user to have a more guided interaction to search for helpful design inputs, but would require a scheme of logical design combinations. Our approach instead focuses on using natural language to describe relevant design choices. In this paper, we present a recommendation system for the ECSS-E-TM-10-25A data model. The similarity analysis is performed at two levels, first comparing metadata describing the main engineering concept characteristics, and then comparing the content of the Engineering Models (EMs), which defines in more detail the architecture of the past studies. Metadata information in the domain of spacecraft engineering can be for example the "Mission Objectives", "Propulsion Type", or "Orbit type" associated to each study. The metadata and the EMs are firstly migrated to a common KG based on previous work presented in [9]. The recommendation system is then deployed on this KG. With the approximate metadata of a new study, the recommendation compares these new inputs to the ones contained in the KG and identifies past similar architectures. The recommendation system then suggests similar equipment and architecture elements, identified in the cluster of past study, for the new study.

To summarise, the main contributions of the paper are the following:

1. Define a similarity metric among a set of metadata that represents the high level engineering product features and adheres to the ECSS-E-TM-10-25A data model.
2. Describe a recommendation system that by leveraging on the aforementioned similarity metric and the Jaccard similarity measure, provides system engineering recommendations (overall system and components).
3. Demonstrate the functionality on the conceptual design of a space mission, which could be extended to any engineering systems that adheres to the ECSS-E-TM-10-25A data model.

Section 2 introduces in more details the conceptual design process of space missions, the case study of the methodology here proposed. Section 3 introduces the structure of the EMs adhering to the ECSS-E-TM-10-25A data model. Section 4 defines the mathematical problem of the textual metadata comparison, and introduces the Jaccard index for the EM's components comparison. Section 5 details the methodology followed to develop the recommendation system. Finally, the results are presented and discussed in Section 6.

2 Application

The conceptual design phase is the first stage of a spacecraft's life cycle, where its initial technical, programmatic and economic feasibility are established. The Concurrent Design Facility (CDF) team at ESA applies the concurrent engineering approach to optimise the feasibility studies of space missions. The keys elements of this design approach are (i) the concurrency of the design process, (ii) the cross-functionality of the team, (iii) a facility with adequate hardware and software infrastructure, and (iv) a design model [10]. The experts work simultaneously on the system design in a shared facility thus emphasising team communication. The system design is stored in an EM, and is used as source of truth by the experts. In the past 20 years, the CDF has performed over 250 studies.

To help kick-start a new study, system engineers usually look at past similar missions' reports to estimate initial parameters and foresee potential design drivers. However, the amount of accumulated data slows down this essential heritage analysis step. This task is often performed manually by the experts. Valuable design information is contained in EMs but those are not easily reused, queried, let alone compared. With an increasing number of space missions designed and the limitation of the manual approach, a recommendation system for this specific application is here proposed.

3 Data

The EM stores in a semi-structured way the information related to a single mission, its equipment and parameters. Figure 1 displays a simplified view of the EM's structure, a formal version is found in the UML data model of the ECSS-E-TM-

10-25A [3]. Each EM contains several design iterations of the system. Within each iteration, multiple design options are studied in parallel to facilitate trade-offs decisions. For instance, some design option might have a propulsion system and another option not. Each design option has a product tree organising the equipment and parameters per subsystems, for instance, a battery and its capacity would be listed under the power subsystem. Part of the parameters can be imported from a reference data library while others are directly defined by experts during the study. A group of parameters corresponding to a set of metadata is linked to the EM, the product tree root. The objects defined in the ECSS-E-TM-10-25A data model, for instance, *iteration* or *option* are not domain-specific and could therefore be applied to any engineering domain. This flexibility however comes at the cost of semantics. The information contained in each EM and respective associated metadata information are stored in a common KG. This is necessary as currently each EM is a standalone model. The migration of multiple EMs into a common database and conversion of the UML data model from ECSS-E-TM-10-25A is described in more detail in [9].

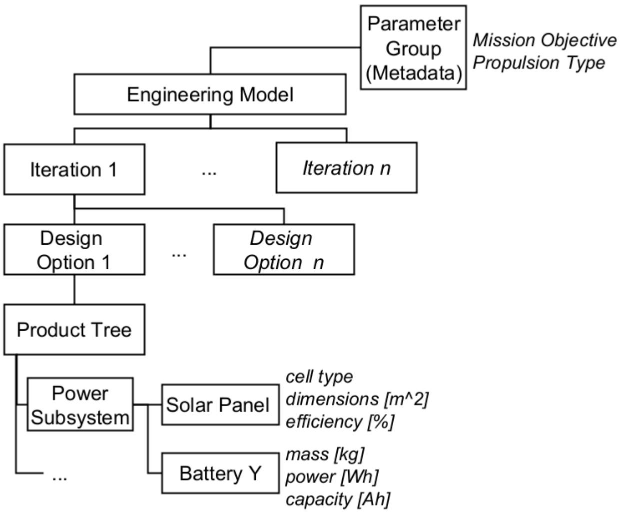


Figure 1: Simplified Schema of an Engineering Model

4 Problem Definition

4.1 Metadata Similarity

Being n the list of metadata associated to the conceptual design of an engineering system, one can define a set of metadata M as:

$$M = \{m_1, ..., m_n\} \in R^n$$

Let's consider two metadata sets P and Q, with p_i and q_i being their respective i-th single metadata entries. To calculate similarity between the two metadata sets and also the single metadata entries, the information they contain needs to be converted into a computer-interpretable format. More formally this can be represented as:

$$p_i \sim q_i \leftrightarrow s(x_{p_i}, x_{q_i}) > \varepsilon_1$$

, where $s(x_{p_i}, x_{q_i})$ is a function which defines a similarity metric between two vector representations of two single metadata entries. The representations x_{p_i} and x_{q_i} can be computed with a neural language model, where the input is the textual information of the single metadata entries and the output a vector or so called "embedding" of the textual information. This can be formulated as:

$$f(y_i, \theta_f) = x_i \in R^d$$

$$\forall y_i = p_1, q_1 .., q_n, p_n \in p_i \cup q_i$$

where θ_f represents the particular set of model parameters of the pre-trained language model f, which is discussed in more detail in the next section.

The semantic similarity between two embedding vectors can then be calculated following standard practice e.g. with the cosine similarity metric [11]:

$$s(x_{p_i}, x_{q_i}) = \frac{x_{p_i} \cdot x_{q_i}}{\|x_{p_i}\| \|x_{q_i}\|} \tag{1}$$

This allows to define a similarity criteria between a single metadata entry as:

$$S(p_i, q_i) = s(x_{p_i}, x_{q_i}) \quad \text{with} \quad p_i, q_i \in P, Q \tag{2}$$

Analogously, one can define the similarity between two sets of metadata by calculating the average over the single metadata similarities:

$$S_w(P, Q) = \frac{\sum_{i=1}^{n} S_(p_i, q_i)}{n} \tag{3}$$

4.2 Recommendation System

Let's consider a KG containing various EMs with m components $C = \{c_1, ..., c_m\}$ and k associated parameters $P = \{p_1, ..., p_k\}$ each. Components and parameters are connected by a set of relations R, e.g. $r = \{c_m, p_k\}$ means that component c_m contains parameter p_k. Given a component c_m there exist a set of parameters $N(c_m)$ that are associated with c_m in the following way:

$$N(c_m) = \{p_k \in P : r\{c_m, p_k\} \in R\} \tag{4}$$

This means that given two components c_a and c_b one can extract two sets of parameters $N_a = N(c_a)$ and $N_b = N(c_b)$, containing the parameters for c_a or c_b respectively. One can hence calculate a similarity between two components by calculating the Jaccard similarity between their two associated sets of parameters [12]:

$$J(c_a, c_b) = \frac{|N_a \cap N_b|}{|N_a \cup N_b|} \tag{5}$$

The Jaccard similarity quantifies the ratio of parameters associated with both components to the total amount of parameters of both components. Now, let P be the set of metadata for a new design system, the list of recommended components R^* is then defined as:

$$R^* = \{(c_a, c_b) \,|\, J(c_a, c_b) > \varepsilon_2, \forall c_a \in C_k, \forall c_b \in C_m \forall k, m \in D\} \tag{6}$$

where $D = \{J \,|\, S(q_j, p) > \varepsilon_1\}, \forall j = 1, .., n$ represents a cluster of J engineering systems that have single metadata entries q_j similar to p over n EMs contained in the KG. ε_1 and ε_2 are constant parameters that represent the similarity threshold set for metadata and components comparison respectively, determining if a pair of components it deemed to be similar as well as if a engineering system is similar to the new mission.

5 Methodology

For encoding textual metadata information, at sentence level, a neural language model from the family Sentence-Transformers is used [13]. These models create dense semantically meaningful vector representations taking as input the raw natural language sentences. They are trained so that the representations of two semantically related sentences have a high cosine similarity, while the cosine similarity between semantically distanced sentences is minimised. Sentence Transformers can lead to significantly improved embedding representations, if compared to existing word and sentence level models, and are the current state of the art as shown in [13]. As described in Section 4, the dense representations are used to compare different sets of past mission metadata. Based on the similarity of past-mission metadata to the metadata of a new mission, similar missions are identified with a threshold value. This simple threshold method could also be expanded by using a clustering algorithm e.g. k-nearest neighbours algorithm. If at least two similar past missions are identified, they can then be queried for for similar components defined in their EMs in order to identify patterns and then infer from these patterns possible recommendations for a new mission. The most simplest pattern would be to see what components has the highest number of similarity relations to other components and rank the components for recommendation accordingly. The complete methodology for the recommendation process described is presented in Algorithm 1.

6 Results

To demonstrate the recommendation system, a simple scenario was chosen where heritage information about three past missions is used to infer recommendations for a fourth new study. In this simplified scenario, the sole metadata parameters, which were considered, were *"Mission Objective"* and *"Propulsion Type"*. The *"Mission objective"* briefly describes

Algorithm 1 Component Recommendation System

Require: $x_1 \dots x_k \in X_k \leftarrow f(q_k, \theta_f)$ ▷ Embeddings X_k are precomputed for single metadata q
procedure SIMILAR MISSIONS(X_k, ε_2, p)
  for x_I in X_q, $I = 1 \dots q$ **do** ▷ Calculate pairwise similarities
    for x_J in X_q, $J = I \dots q$ **do**
      Compute similarity $S(x_I, x_J)$
      Add $S(x_I, x_J)$ to set of metadata similarities σ_{Meta}
    end for
  end for
  Calculate $\mathscr{X}$ clusters based on σ_{Meta}
  Metadata of new design mission $p \in D, D \in \mathscr{X}$
  return D
end procedure
procedure SIMILAR COMPONENTS(D, C, ε_1)
  for all pairs $mission_d, mission_e \subset D$ **do**
    for all c_d in C_d **do** ▷ Components $mission_d$
      for all c_e in C_e **do** ▷ Components $mission_e$
        Calculate Jaccard similarity $J(c_d, c_e)$
        if $J(c_d, c_e) > \varepsilon_1$ **then**
          Add c_d, c_e to set of similar components ω
        end if
      end for
    end for
  end for
  Define $D_c \leftarrow$ empty dictionary ▷ Get frequency of components
  for component c in ω **do**
    if component $\notin keys(D_c)$ **then**
      add component to $keys(D_c)$
      initialize $D_c(component) = 1$
    else $D_c(component) \leftarrow D_c(component) + 1$
    end if
  end for
  return D_c
end procedure

the main reason why a particular spacecraft or satellite was designed, for instance, to observe all boreal forests. The parameter *"Propulsion Type"* describes if and what kind of propulsion system is used for changing the spacecraft's orbit. The respective parameters for the three example missions are shown in Table 1.

For the new mission, the following metadata parameters are known and defined:

- *Mission objective:* autonomously seek out space debris with an short-range doppler radar, capture, and then de-orbit debris
- *Propulsion Type:* Gridded iodide ion thruster

The first step in the recommendation system is to compare the metadata information of the new study to the one of the past missions. As described in Section 4, textual metadata information is embedded into a vector representation with `mpnet-base`, a deep neural language model from the SentenceBERT family (https://huggingface.co/sentence-transformers/all-mpnet-base-v2) Subsequently, the pairwise cosine similarity between each pair of metadata information is calculated. In a first analysis, the "*Mission objective*" metadata of the new mission is compared to the ones of the heritage missions. The results of the cosine similarity comparison are shown in Table 2. A value close to 1 means a high similarity. The similarities between the different missions ranges between 0.2 and 0.7. The highest similarity is between *Mission 1* and *Mission 2* and the new mission with *Mission 2*. Multidimensional scaling is used for a better visualisation in two dimensions to allow for a better interpretation. Multidimensional scaling iteratively tries to minimise the square difference (distance represented by the similarity measure) between the representation in the original higher dimensional space and the smaller two-dimensional space.In Figure 2a, the similarity results are visualised for the parameter *"Mission objective"*. While *Mission 1*, *Mission 2*, and the *"new" Mission* all are relatively close to each other in the projected space, *Mission 3* is separated from this cluster. Considering the actual mission objectives of each mission, this is unsurprising

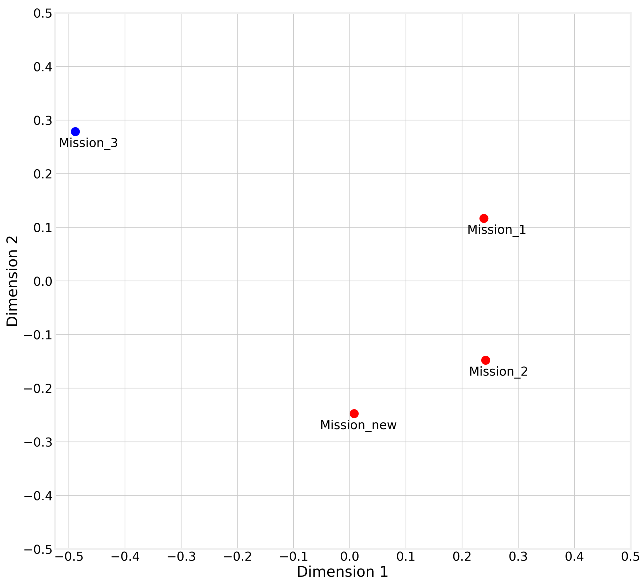
Table 1: **Spacecrafts metadata of three demonstrative heritage missions**

Models	Mission objective	PropType
Mission 1	launch a CubeSat into LEO as a space-based passive bistatic radar technology demonstrator where a signal processing algorithm to detect space debris will be tested	no propulsion module
Mission 2	study the (micro) space debris environment in LEO to complement the models by an active short-range radar"	RIT Xenon thruster
Mission 3	provide a close and regular monitoring of vegetation on Earth's surface with a super-spectral camera	Hall effect on Xenon thruster

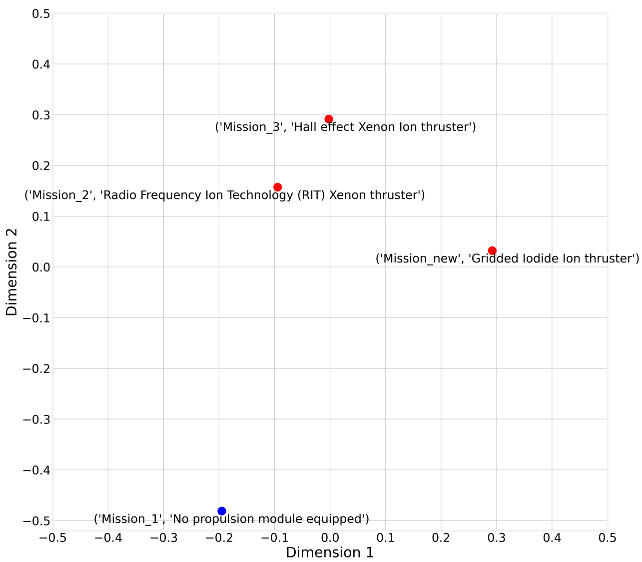
Table 2: **Pairwise similarities for metadata parameter "Mission objective"**

	Mission 1	Mission 2	Mission 3	New mission
Mission 1	1.00	**0.69**	0.23	0.62
Mission 2	**0.69**	1.00	0.20	**0.70**
Mission 3	0.23	0.20	1.00	0.25
New mission	0.62	**0.70**	0.25	1.00

as *Mission 3* describes an Earth observation mission for vegetation, while the other missions share as a goal using radar technology for detecting space debris.



(a) **MDS of "Mission objective" metadata**



(b) MDS of "Propulsion" metadata

Figure 2: **Multidimensional scaling (MDS) of pairwise similarities of "Propulsion" metadata**

Based on the semantic similarity of their *mission objectives*, past *Mission 1* and *Mission 2* are queried for shared similar components to recommend for the new study. An exemplary list of results is presented in Table 3. Similar components contained in *Mission 1* and *Mission 2* are firstly generic ones, such as *battery* or *solar cell*, which would occur in most studies. But secondly there are also specific ones such as *radar receiver* and *radar antenna*, which are expected in missions using radar technology.

After the comparison of the new mission based on similarities between their *mission objectives*, the next step is to consider the similarities between the "Propulsion Type" metadata parameter. Here, a different distribution of similarities is expected as *Mission 1*, with no propulsion module selected, is significantly different from the rest. Analogously to the first comparison, the cosine similarities are calculated, presented in Table 4, and subsequently visualised with multidimensional scaling in Figure 2b.

Mission 1 is as expected relatively separated from the concentration of other missions. This once more aligns with the actual information provided, since Mission 1 does not contain a type of propulsion while the other two heritage missions

Table 3: **Example components sharing high Jaccard similarity, showing the component type and relevant parameters for their identification for metadata ”Mission objective”**

Component type	Set of components	Relevant parameters
Cubesat battery	(”NanoAvionics EPS”, ”Battery”)	battery capacity, battery cell type
Cubesat solar cells	(”DHV-CS-10 Cubesat Solar Panel”, ”ISIS-PACE Small satellite solar panel”)	Solar cell type, solar array type
Sun sensor	(”SUN LENS Bison 64”, ”NSS Fine Sun Sensor”)	field of view,sun sensor bias
Radar receiver	(”XM1110_1103886”, ”ZOE-M8G-0”)	Sensitivity, modulation or protocol
Radar antenna	(”PulseLARSEN W3227”,”ISIS-GAPA-DSH-0001”)	Antenna Type, frequency Group

Table 4: **Pairwise similarities for metadata parameter “Propulsion Type”**

	Mission 1	Mission 2	Mission 3	Mission new
Mission 1	1.00	0.30	0.26	0.29
Mission 2	0.30	1.00	**0.78**	**0.61**
Mission 3	0.26	**0.78**	1.00	0.59
Mission new	0.29	**0.61**	0.59	1.00

have ion thrusters as their main propulsion system. Optimally, a value close to zero for the similarity of Mission 1 to the respective others would be expected, as having no propulsion system is semantically the opposite to having one. This could be improved by fine-tuning the used language model for the application in the technical domain, so it better grasps semantic nuances. The performance is adequate enough to differentiate between having a propulsion system and not having one.

The analysis of similar components is repeated for the similarity of “Propulsion Types” between the missions. The idea here is to find shared elements typically included in a propulsion system in the different EMs. An examplary list of components is presented in Table 5.

Both past missions interestingly include a type of cold gas thruster. This type of propulsion is not using electricity but relies on the expansion of the pressurised gas to create thrust. These thrusters are commonly used for small and precise manoeuvres. In this case, the available pressurised Xenon for the main engine was also used as propellant for this system, instead of installing a separate assembly with pipes and propellant tanks only for these type of manoeuvres. Although the new mission uses a different propellant (Iodine), this could also prove as a viable option. In contrast, novel electric propulsion systems are testing ways of not using a neutraliser for their setup. This shows that relaying just on heritage data cannot take into consideration new trends in design.

7 Conclusions

A new approach for providing recommendations for assisting in the design of engineering systems has been introduced in this work. EMs and associated metadata, described in natural language, build the basis of the approach. Firstly, similarity between the metadata information is established by calculating the cosine similarity of their representations, which are created by a neural language model from the *Sentence-Transformers*-family. Based on the similarity of the metadata, the respective associated EMs are clustered into similar missions, which in the second step of the analysis are investigated

Table 5: **Example components sharing high Jaccard similarity, showing the component type and relevant parameters for their identification for metadata ”Propulsion type”**

Component type	Set of components	Relevant parameters
Cold gas thruster	(”Xe Cold Gas Thruster/Valve Assy”, ”Cold Xenon Thruster”)	thrust, expansion ratio
Electric propulsion neutraliser assembly	(”EP neutraliser assembly”, ”Neutraliser assembly”)	neutralizer flow rate
Propellant tank	(”Propellant storage tank”, ”Xenon storage tank”)	Propellant, tank volume

for similar components. The link between components is established by using the Jaccard similarity metric between their respective associated parameters. The functionality of the method was demonstrated for space mission design. It was shown that by defining different characteristics semantic similar missions could be identified. Secondly, a list of example components could be generated for each respective characteristic as possible recommendations for a new mission. A possible downfall of this approach is that the representations created by the neural language model possibly do not completely capture the semantic meaning, as the chosen model is trained on generic language and not specialised for the technical domain. Additionally, the link between components is only possible if they share common parameters; for components such as *valves*, *pipes*, or even more specific components sometimes no characteristic parameters exist. This could lead to problems of ambiguity, when elements are predicted as similar, even if in reality they are not. Nonetheless, the approach can be easily expanded to other engineering domains, using the ECSS-E-TM-1025A data model or equivalent, provided relevant metadata parameters can be defined.

Acknowledgments

This study was fully funded by ESA under the OSIP Model-Based System Engineering Campaign – Studies Scheme, contract No.4000133311/20/NL/GLC. The authors would like to warmly thank Serge Valera (ESA) for providing expert knowledge on the ECSS requirements definition.

References

[1] Azad M. Madni and Michael Sievers. Model-based systems engineering: Motivation, current status, and research opportunities. *Systems Engineering*, 21(3):172–190, 5 2018.

[2] Aditya Akundi and Viviana Lopez. A Review on Application of Model Based Systems Engineering to Manufacturing and Production Engineering Systems. *Procedia Computer Science*, 185:101–108, 1 2021.

[3] ECSS Secretariat. ECSS-E-TM-10-25A Engineering design model data exchange CDF. Technical report, 2010.

[4] Kriedte W. ECSS - A Single Set of European Space Standards. In *Spacecraft Structures, Materials and Mechanical Testing*, pages 321–327, 1996.

[5] Kristina Hogstrom, Jonathan Murphy, Steven Zusack, Andrew Coffey, Chester Borden, Alan Didion, Damon Landau, Adam Nelessen, Macon Vining, and Robert Miller. From Cocktail Napkin to Concept Feasibility: Spacecraft Design in Early Formulation with TATER. In *2019 IEEE Aerospace Conference*, pages 1–18. IEEE, 3 2019.

[6] Alex Austin, Robert Miller, Jonathan Murphy, Mike Kolar, Jon Blossom, Tu-Anh Phan, Scott Doudrick, Kristina Hogstrom, and Anne Marinan. Rapid SmallSat Mission Formulation: Integrated and Concurrent Modeling in JPL's Team Xc. In *2020 IEEE Aerospace Conference*, pages 1–11. IEEE, 3 2020.

[7] Jairus Hihn, Matthew Ramirez, Anto Kolajian, Michael Saing, and Mason Takidin. Estimate Early and Estimate Often: How Your Concurrent Engineering Design Team Can Bootstrap Your Organizations Programmatic Capabilities. In *2020 IEEE Aerospace Conference*. IEEE, 2020.

[8] Francesca Mangili, Denis Broggini, Alessandro Antonucci, Marco Alberti, and Lorenzo Cimasoni. A Bayesian Approach to Conversational Recommendation Systems. In *AAAI 2020 Workshop on Interactive and Conversational Recommendation Systems (WICRS)*, 2020.

[9] Audrey Berquand and Annalisa Riccardi. From Engineering Models to Knowledge Graph: Delivering New Insights into Models. In *9th International Systems & Concurrent Engineering for Space Applications Conference (SECESA)*, Digital Edition, 2020.

[10] M Bandecchi, B Melton, B Gardini, and F Ongaro. The ESA/ESTEC Concurrent Design Facility. In *Proceedings of the 2nd European Systems Engineering Conference (EuSEC)*, pages 329–336, Munich, Germany, 2000.

[11] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. *Introduction to Information Retrieval*. Cambridge University Press, jul 2008.

[12] Paul Jaccard. THE DISTRIBUTION OF THE FLORA IN THE ALPINE ZONE.1. *New Phytol.*, 11(2):37–50, feb 1912.

[13] Nils Reimers and Iryna Gurevych. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. 8 2019.