ARTIFICIAL INTELLIGENCE (AI) AND NATURAL LANGUAGE PROCESSING (NLP) TO SUPPORT SPACE ENGINEERING ACTIVITIES

Gérald Garcia [1], Gaëtan Pruvost[2], Serge Valera [3], Luis Mansilla[3], Audrey Berquand [3]

[1] Thales Alenia Space, Cannes (FR), firstname.lastname@thalesaleniaspace.com
[2] Thales SIX / Theresis, Palaiseau (FR), firstname.lastname@thalesgroup.com
[3] European Space Agency (ESA), Noordwijk (NL), firstname.lastname@esa.int

LINK WITH THE CONFERENCE OBJECTIVES

This paper relate to the following MBSE 2022 objectives :

- S-2 : Current and future R&D programmes with focus on MBSE, digital spacecraft, big data, machine learning and artificial intelligence
- T-3: Showcase novel approaches to MBSE that have the potential to significantly improve the state-of-practice in space mission design and development

ABSTRACT

This work has been performed in the frame of an OSIP ESA study (ESA Contract Nr. 4000133471/20/NL/GLC/kk).

Model based approaches have proved to be efficient in supporting engineering activities, models replacing traditional document-based approaches. Nevertheless even in most advanced deployments, a lot of engineering artifacts are textual either because the return on investment of introducing models is too low on this particular case or because the information is more efficiently expressed in natural language even if consistency and correctness issues appears.

Huge progress has been made recently in AI-based Natural language processing (NLP), mainly driven by chat bots and vocal home assistant usages. The proposed idea consists in spinning in these technologies into space engineering process, studying how natural language processing can help the space engineer in daily activities.

Many engineering domains can take advantages of these technologies. The most evident one are the requirement management domains as most of the requirements are textual and are generally not formally modeled even if they have a certain level of structure and rules.

Using NLP technology semantic information can be extracted from textual requirements, which may have several advantages that will be evaluated: find related requirements to a given one, enhance search in requirement database, check consistency of traceability, smart comparison of specification contents, identification of suspicious requirements, identification of overlapping requirements, ...

Natural language processing is also promising in linking related engineering textual artifacts (like design reports, justification reports, …) even with models contents.
The top-level goals for this deployment as stated in the proposal are to:

- Capturing use-cases relevant for AI and NLP technics.
- Define an overall framework for implementing solutions
- Implement search capability with business knowledge
- Implement search in structured data
- Explore similarity in a particular set of artefacts
- Explore graph based representations and merging with NLP

After a selection of use-cases, 5 use cases has been demonstrated in this study :

- “Smart searching” taking into account space domain : improving the capabilities of search engines by injecting a knowledge linked to our domain (for example Ariane 6 is a launcher).
- Searching extended to non-textual content : providing an unified search engine based on textual query that is capable to search both in textual content (requirements, documents, …) but also in model artefacts (functions, components, mass budgets, …)
- Similarities into a particular set of artefacts : improve search by relying on semantic distance between two artefacts using models trained on space domain corpus.

- Similarities into heterogeneous set of artefact : implementation of a trace assistant capable to propose to the user the most probable trace links between two set of artefacts (for example two specification for satisfaction links, between functions and related requirements, …)
- Integrating structured data (models) and unstructured data (text) into a single representation to improve textual similarities by model information.

These use-case has been implemented as proof of concept into this study. For each use-case an investigation phase has been done where multiple implementation option has been studied and the retained principle has been integrated into a micro-service single search architecture (with an unified web interface) capable to support almost all use-cases and sourcing its data (the engineering artefacts) into a knowledge graph structured by a ontology very similar to Osmose one.
This study has validated the interest and the feasibility of the different use-cases. The proposed use-cases for implementation has increasing level of difficulty with regards to current state of the art. Some use cases are very close to operational usage (search engine), some use-case has very encouraging results (for example the transfer learning on top of a Space BERT model to perform trace link recommendations) that has to be confirmed in other data set. Some experiment (in particular the fusion between information coming from a model and textual content, ie merging NLP transformers approach with graph ML) has to be investigated further and are still at very low TRL.

1. Overall solution architecture

Cross-cutting all the use-cases implementation, we setup a dedicated infrastructure called EDA and capable to support the implementation of each use-case while delivering convincing demonstration from operational data.

This architecture is depicted here bellow and composed of a set of micro-services (enabling the replacement of each individual bricks as the state of the art is evolving).

It is composed of :
- An indexing database (based on OpenSearch)
- A set of NLP algorithms exposed as REST services that are connected to the database
- A end-user web interface
- A command line interface to control the whole architecture (ingestion, test, processing, …)

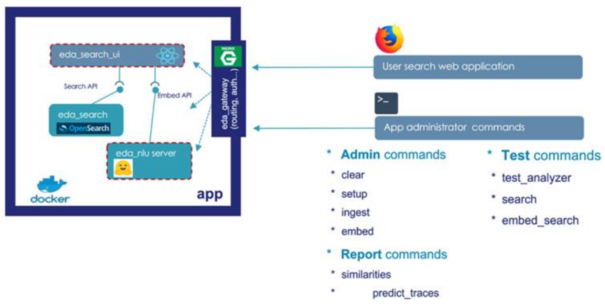

Figure 1 : Overall architecture

In order to feed the data, we rely on the Exago Thales Alenia Space infrastructure that is maintaining a knowledge graph composed of engineering data extracted from the different engineering authoring tools. This knowledge graph is structured by an OSMoSE (see [1]) like ontology. Exago permit to this demonstrator to have access to all cross-linked engineering data without the need to develop tool specific bridges. This coupling between EDA and Exago is based on standard interfaces and other knowledge graph could be plugged.

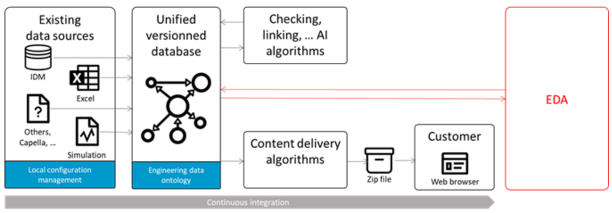

Figure 2 : Integration between EDA and Exago to feed data to EDA.

2. Example of applications

Here are two examples coming from the demonstrated use-cases.

The first one is demonstrating the capability of EDA to detect similar engineering elements. Engineering elements can be of any kind known by the knowledge graph. It can be requirements, Capella elements, budgets elements (mass, power, …) and any other type of artefact.

In order to perform semantic search and similarity, all the knowledge graph elements are verbalised taking into account their context in the graph and then transformed to vectors using space domain dedicated transformer model [2]. Then similar elements propositions are integrated into the digital review web site proposed by Exago.

This permit to the reviewer to find related elements to the current reviewed one, even if they are not formally linked in the knowledge graph. This is demonstrated in the screenshot here after (when reviewer look at mechanical alignment requirement for gyro, the similar requirements for the other equipment are proposed).

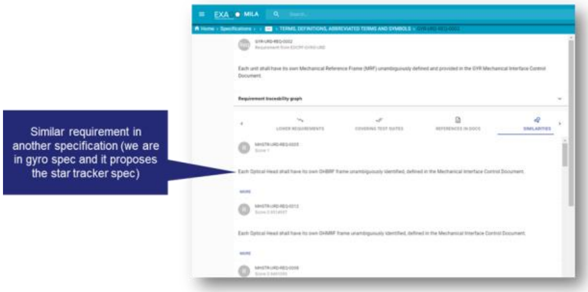

Figure 3 : Similarities find by EDA embedded into the digital review tool proposed by Exago

The second application is demonstrate the capability to implements a trace assistant which permit to propose automatically trace link between two specifications, a specification and a test plan or a specification and a Capella model.

In this application we have performed transfer learning using an existing transformer model and train specialising it to classify requirement trace links. We trained with an operational space project data stored in the knowledge graph from a DOORS project.

We evaluate the results of this training and it gives very encouraging results (Precision of 0.86 and recall of 0.75) which validate the feasibility of the approach and open the door to operational application permitting to save a lot of hours of engineering work and improving the quality of the tracing between engineering artefacts

The demonstrator is a web based interface where the user select to “containers” (for example specifications) to trace and the algorithm is proposing pairs of requirements to be trace from most probable one to less probable one. The engineers is then accepting or not the proposal having always the final word on the decision.

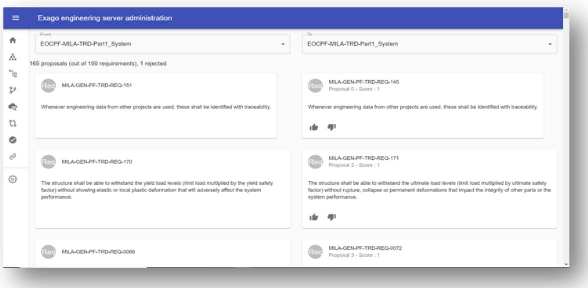

Figure 4 : Screenshoot of the trace assistant

3. Conclusions

This study is a path finder to further usage of artificial intelligence as an effective support to the operational engineering work. The proposed use-cases are useful all along the file cycle of a space program and as the potential to provide improvement of the engineering work all along this life cycle.

These study has validated the interest and the feasibility of the different use-cases and confirmed that the topic of artificial intelligence based engineering assistant is relevant not only for medium or long term application but also for immediate applicability. This is possible thanks to the effort put on this subject by many teams (academics, developers, …) all around the world and by the fact that this is accessible (mainly open source even for big trained models). Most of the time this work is performed on domains far away from the space domain (social networks, chat bot assistants, …) but we have demonstrated that these technology bricks can be reused and combined to serve the space domain. We can spin in all this material and this research to improve right now the engineering work efficiency on space projects.

The domain is evolving at ultrafast pace and things that are difficult today will certainly be solved simply by dedicated libraries or algorithm just tomorrow. This permit to be very optimistic on the operational deployment but is also a risk for investments on deployment as the obsolescence of the technology is also very quick and have no possible comparison with space project time frame.

This open the door to operational deployments but also to more academic work to come.

REFERENCES

[1] OSMoSE initiative, https://mb4se.esa.int/OSMOSE_Main.html

[2] SpaceTransformers: Language Modeling for Space Systems. Berquand, Audrey and Darm, Paul and Riccardi, Annalisa. IEEE Access 2021, vol 9.