A Semantic Approach for Generating Graphical Representation from Aircraft Maintenance Text

Thi-Bich-Ngoc Hoang[a], Ba-Huy Tran[b] and Marzieh Mozafari[c]
Capgemini Engineering, France
{firstname.lastname}@capgemini.com

Keywords: Maintenance Task, Process Ontology, Graphical Illustration, Technical Language Processing, Information Extraction, Aircraft Maintenance Manual.

Abstract: Industrial maintenance is a strategic business function. Over the past twenty years, the role of maintenance in companies has become increasingly important both technologically and economically. However, maintenance service has not taken into account the frequent change in maintenance knowledge, the users' perspective (the training, the origins, or the cultures), and the users' support documents preferences. In this article, we propose preliminary results of an approach to make industrial maintenance universal. We use natural language processing techniques to extract core information from maintenance text and then construct a knowledge base to store all relevant information about maintenance processes, domain information, and corresponding graphics. Finally, we generate a graphical representation of input text to help better understand the procedure, thus increasing the user experience and the performance of maintenance operations in terms of reducing time and cost. This approach is first applied to aircraft maintenance and can be applied to maintenance in other industry domains as well.

1 INTRODUCTION

Industrial maintenance is a strategic business function. It can be defined as all the troubleshooting and repair actions, adjustment, overhaul, control and verification of material or even immaterial equipment. Over the past twenty years, the role of maintenance in companies has become increasingly important, both technologically and economically. Whether it is industrial maintenance expenditure or dedicated staff, the maintenance sector shows a significant increase on all points. In France, annual expenditure on maintenance is around 18 billion euros and requires 70,000 jobs[1].

In addition, the needs of user maintenance players evolve over time and cannot be satisfied by the services currently provided by computer maintenance support systems on the market. Indeed, these services are based on the knowledge initially formalized but which is not systematically updated. Thus the services offered after a few years are no longer in line with current knowledge. We must take into account the dynamic aspect of knowledge, to meet the needs of users and improve the performance of help software offering these services.

Different activities in industrial maintenance generate a vast volume of written data in the form of reports, historic records, plans, and schedules. Many text processing tasks out of these textual data can be effectively automated using Natural Language Processing (NLP). Extracting practical information from maintenance documentations presents a unique challenge in the domain of information extraction (IE), because these instructional texts include multiple steps with specific objects which should be performed sequentially. In addition, this process highly depends on the quality of the raw data and the way it is processed with NLP such as pre-processing, Tokenization, Part-of-Speech tagging (POS), Name Entity Recognition (NER), etc.

The aim of this study is to leverage NLP along with knowledge bases to improve the performance of maintenance documents analysis, simplify aircraft maintenance processes, and insure semantic interoperability. To that end, we extract core information in aircraft maintenance documents including description, warnings, cautions, notes, actions, related com-

[a] https://orcid.org/0000-0002-8693-9195
[b] https://orcid.org/0000-0002-2578-9138
[c] https://orcid.org/0000-0002-1384-7548
[1] https://metgroupe.fr/les-chiffres-cles-maintenance/

Hoang, T., Tran, B. and Mozafari, M.
A Semantic Approach for Generating Graphical Representation from Aircraft Maintenance Text.
DOI: 10.5220/0011543000003335
In *Proceedings of the 14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2022) - Volume 3: KMIS*, pages 169-176
ISBN: 978-989-758-614-9; ISSN: 2184-3228

ponents, and information related to each component using NLP techniques. Then we develop an ontology to describe the maintenance procedure and domain, and also generate illustrations using the graphic representation of resources or contexts in our knowledge base to better guide the user. To the best of our knowledge, this is the first attempt in the aeronautic maintenance domain, and this approach can be applied to other industrial domains as well.

The rest of the paper is organized as follows. Section 2 presents a literature review of the information extraction from text and semantic representation of maintenance procedure. The NLP techniques for extracting the information from maintenance documents and the methodology for generating ontology and graphical representation of maintenance procedures are described in Section 3. Finally, Section 4 draws some conclusions and offers a view of possible future work.

2 RELATED WORK

2.1 Information Extraction from Text

Information extraction from text has attracted a number of studies in recent years. However, most of the information extraction systems have been developed for domains such as medical (Xu et al., 2017a; Ge et al., 2020; Lopes et al., 2019), biomedical (Nayel et al., 2019; Gao et al., 2021; Yang et al., 2022), and others (Beltagy et al., 2019; Wadden et al., 2020; Aly et al., 2021). Extracting domain specific information from maintenance documents has received limited attention (Dixit et al., 2021; Sharp et al., 2017).

NLP is an approach to extract information from text written by humans. Approaches applied in recent NLP systems are grouped into two categories: Rule-Based Methods (RBM) and Learning-Based Methods (LBM). RBM refers to the modeling where the relationships and patterns in data are defined by the human while in LBM, these relationships and patterns are figured out and trended out by the machine. RBM is interpretable and suitable for rapid development and domain transfer, and requires pre-defined vocabularies (Adnan and Akbar, 2019; Valenzuela-Escárcega et al., 2015; Patel and Tanwani, 2019). In general, the performance of LBM is better in terms of precision and recall but appropriate feature selection is important. In addition, generating training data is a time-consuming task.

Xu et al. (Xu et al., 2017b) recognized medical concepts and terminology such as diseases, drugs, treatments, or procedures from unstructured medical text. They used the bidirectional Long-Short Term Memory (biLSTM) and conditional random fields to identify medical named entity relies on character-based word representations learned from the supervised corpus. Using another approach, Beltagy et al. (Beltagy et al., 2019) introduced SciBERT which is based on BERT (Devlin et al., 2018) but retrained on a large scientific paper corpus. The authors showed that their model is effective on NLP tasks such as sequence tagging, sentence classification and dependency parsing.

In the domain of maintenance, Dixit et al. (Dixit et al., 2021) proposed a method to extract entities of interest from maintenance records based on their extends on an existing domain dictionary. This dictionary includes lists of components, positions, observations, and actions and is used to identify corresponding elements in an input sentence. Their approach got preliminary result, which mostly depends on the quality of the domain dictionary. Sharp et al. (Sharp et al., 2017) introduced a proof-of-concept pipeline combining machine learning and natural language processing techniques to cluster and tag maintenance data. They achieved the accuracy around 70% when categorizing and labelling a free form maintenance log entry from a set of known labels.

The above approaches neither target to our objectives nor are available online for public use. Thus, in our work, we only consider NLP tools that are efficient and available for use as follows.

Stanford CoreNLP (Singh et al., 2013): allows users to perform a variety of NLP tasks, such as part-of-speech tagging, tokenization, or named entity recognition. The advantages of this tool are the scalability and optimization for speed, making this tool relevant for processing large amounts of data, and performing complex operations. Spacy (Honnibal and Montani, 2017) : Spacy offers components for uses an object-oriented approach to NLP handling and supports pre-trained statistical neural network models and word vectors. Comparing to NLTK and other libraries, Spacy well interfaces with all major deep learning frameworks and performs faster; however, it lacks flexibility and does not support many languages. Natural Language Toolkit - NLTK (Honnibal and Montani, 2017) supports common tasks in NLP by offering a model trained on a wide range of corpora and lexical resources. It processes and presents all data in form of strings and does not support object-oriented. One of the disadvantages of NLTK is that it requires significant resources and time when performing on massive amount of data (Al Omran and Treude, 2017).

As showed in previous research (Honnibal and

Montani, 2017; Al Omran and Treude, 2017), when evaluating these NLP tools in various data collections such as Java API documents, Stack Overflow, and Github Readme files, Spacy makes the best performance for NLP tasks, especially for POS tagging. In addition, Spacy supports transfer learning, which can be used to import knowledge from annotated examples into the pipeline to improve its efficiency. Thus, we choose Spacy POS tagging to identify actions, related objects, warnings, guides, and other information in aircraft maintenance text.

2.2 Semantic Representation of Maintenance Procedures

To develop our generic ontology helping capture all necessary knowledge from which illustrations can be generated, we examine three types of ontologies.

Ontologies for Procedure: A maintenance text can be viewed as a set of maintenance procedures, that are our main objects of study. We consider a maintenance procedure as a particular procedure for doing something involving one or more steps or operations. In the literature, there exist many ontologies describing such a procedure. For business process modeling, the Business Process Model and Notation (BPMN) is a widely used standard. The specifications gave rise to the construction of ontologies, such as the BPMN Ontology (Rospocher et al., 2014) or BPMN 2.0 Ontology (Natschläger, 2011). For industrial procedure modeling, Karray et al. (Karray et al., 2012) has introduced an ontology for industrial maintenance, and Chungoora et al. (Chungoora et al., 2013) introduced an ontology for the manufacturing process. Besides, there are several works inspired by the Process Specification Language (PSL), a framework to describe the structure of process executions such as that of (Grüninger, 2009). Ontologies for procedures can be built based on ISO specifications, as presented in (Fraga et al., 2018). Except for the last one, these models are complex and only the fragment of them that deals with process description is related to our study. Furthermore, the proposed models (and their submodels) aren't adapted to our needs. One solution is to reuse and extend the models as introduced in (Annane et al., 2019) or (Tarbouriech et al., 2021).

Domain Ontologies: We next examine the domain ontologies representing the knowledge of the domain where the maintenance is applied on. In the aeronautical field, to our best knowledge, there's only an ontology, called Aircraft Ontology (Ast et al., 2014), which is reused in (Stefanidis et al., 2020). However, the ontology is not well-structured and does not cover all aircraft components we need.

Ontologies for Graphics: To represent knowledge about graphics, an ontology is required. As one of the earlier works, Niknam et al. (Niknam and Kemke, 2011) presented an ontology for basic computer graphics and geometric shapes. There are afterward several studies on iconic ontologies as introduced in (Kuicheu et al., 2012; Ma and Cahier, 2014; Lamy and Soualmia, 2017). The work of (Lamy and Soualmia, 2017) is the most relevant to our context as the authors propose an iconic ontology that plays the main role and is linked with the domain ontology through a mapping ontology.

2.3 Generating Graphics from Knowledge Base

In the literature, a common method of generating graphics from information, specifically from a knowledge base, is to rely on SVG graphics which is an XML-based markup language for describing vector graphics based on two dimensions. This method has been used in several works such as the automatic generation of maps (Ipfelkofer et al., 2006), simulation of 2D models (Lehtonen and Karhela, 2006), generation of medical icons (Lamy et al., 2008) or generation of traditional medicine recipes (Kouame et al., 2020). Although several works have been proposed in this area, no application or source code has been publicly available. Very few studies have been invested on graph generation methodology. (Kouame et al., 2020) drew pictograms using Inkscape, and automatically generated icons and recipes as SVG image files from ontology knowledge, using Python scripts.

3 METHODOLOGY

In this section, we present our proposed semantic-centric approach to simplify maintenance processes by graphical representation. We first use NLP techniques to extract core information from maintenance documents, and then use these pieces of information to populate our ontology. We also use some external resources to improve our knowledge base and integrate images corresponding to each aircraft component. Finally, we illustrate a graphical representation of maintenance processes using information from our built knowledge base.

Our approach can be divided into three steps presented in Figure 1.

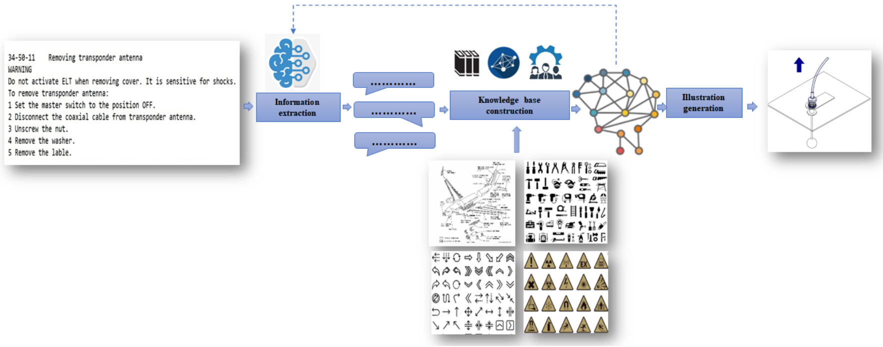

Figure 1: The workflow of our approach to generate graphical representation from maintenance text.

3.1 Extracting Core Information from Aircraft Maintenance Text

We conducted experiments and evaluated our model on the Aircraft Maintenance Manual (AMM) for WT9 Dynamic LSA issued by Aerospool on 22 May 2017[2], and it is freely available and usable for all applications.

3.1.1 Tools and Techniques

In this work, we target to an NLP tool which is flexible enough to be tailored to suit our demands and requirements for analyzing the aircraft maintenance text. In addition, the tool should be high performance and effective. Therefore, we choose Spacy (Honnibal and Montani, 2017) which is showed as the most effective NLP tools when evaluated on several data collections compared to the other current open-source tools (Al Omran and Treude, 2017; Bird et al., 2009). When applying Spacy to analyse aircraft maintenance, we tailor the model to fulfil our requirements such as retraining the model and applying rules to re-tokenize text complement to the Spacy tokenization.

3.1.2 Extracting Core Information from Text

Our objective is to extract core pieces of information from aircraft maintenance texts. Each maintenance, that we call process, is assumed to be structured in a format as illustrated in Figure 2.

Each process includes several pieces of information such as identification number, name, description, warnings, cautions, notes, and steps of instruction. We identify and extract these pieces of information as follows:

- The process identification and name are included in the first line of each process. The identification is the first part of the line which includes numbers connected to each other by '-' while the name is the rest part of the first line.

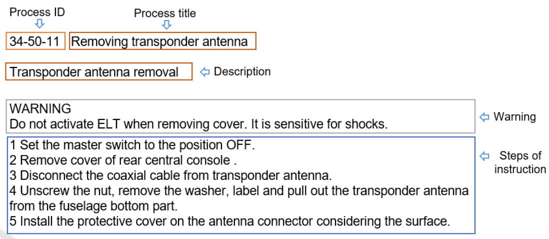

Figure 2: The process of 'Removing transponder antenna'.

- The warnings, cautions, and notes of each process are recognized by the paragraph right after corresponding key words such as Warnings, Cautions, and Notes. If there is no keyword specified, we consider paragraphs which contain words in a list ('*must*', '*never*', '*do not allow*', '*ensure*', '*guarantee*', '*should*', '*have to*', '*need*') as the warning of the process.
- The description of the process is the paragraph which appears after warnings, cautions, notes (if any) and contains words to guide how to perform process name.
- The steps of instruction are the rest of the process text. Each step starts by an ordinary number and contain at least one sentence. Each sentence is considered as one task. Each task can have one or several subtasks (clauses) which include at least one action and one related component. Each task/subtask points to its previous task/subtask.

Our main objective is extracting core information from steps. For each step in a maintenance process, we identify actions, related components, and information related to each component such as the status, direction, and position. In addition, we also recognize corresponding warnings and cautions of each instruction step.

All these elements are recognized by Spacy with

[2]https://www.aerospool.sk/downloads/RTC/AS-AMM-01-000_I1_R1_20180202.pdf

our own customization and adaptation. We base on the POS, head text, and children text of each token in an instruction step to determine the classification (action, component...) of each token or group of token.

Actions: An action is the main verb of the sentence/clause which is recognized by the NLP tool. We hypothesize that the subtask A is the 'previous subtask' of the subtask B if the main verb in the subtask A is the head text of the main verb in the subtask B. As shown in the analysis of the step 4 of the process *34-50-11* illustrated in Figure 3, the action of the subtask_1 is *'unscrew'* while the action of subtask_2 is *'remove'*. The head text of *'remove'* is *'unscrew'* thus the subtask_1 is the 'previous_subtask' of the subtask_2.

Components: Components related to a certain action in a subtask is the nouns of which their head text is that action. These nouns are recognized by Spacy with our customization and adaptation. In Figure 3, the component related to the action '*unscrew*' is '*the nut*' while '*the washer*' and '*labels*' are components related to '*remove*'.

Information Related to a Certain Component: We identify information related to a certain component such as the status, the position, the direction, and other quantity. These elements are recognized as adverb or adverb positions by Spacy and then they are checked with our predefined lists to determine the corresponding information related to components. Our predefined list of words describing the position of a component is *('among', 'around', 'behind', 'beneath', 'between', 'by', 'in', 'into', 'inside', 'near', 'next to', 'on', 'over', 'across', 'below', 'above', 'against', 'under', 'beside', 'in front of', 'through', 'underneath')* while the list of words describing the status of a component is ('*ON*', '*OFF*') and the list of words describing the direction of a component is ('*to*', '*from*'). The number of each object is stored in the '*quantity*'. In the step 4 of the process 34-50-11, the '*transponder*' is pulled out '*from*' (the direction) '*the fuselage bottom part*'.

Warning and Guide: We identify warnings in a subtask by checking if the subtask contains gerund verb in our predefined list. The warning phrase will start from the gerund verb to the end of the sentence/clause. Guides is specified in a similar way to warnings considering gerund verbs that are not in the gerund warning verb list. In the step 4 of the process 34-50-11 (Figure 2, the phrase *'considering the surface'* is specified as a general 'guide' for this step.

The analysis of the step 4 of the process *34-50-11* (*'Unscrew the nut; remove the washer, label; and pull out the transponder antenna from the fuselage bottom part.'*) is presented in the Figure 3.

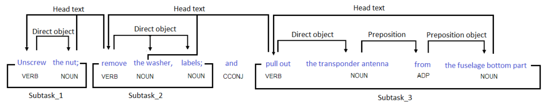

Figure 3: The analysis of the task 4 in the '34-50-11 Removing transponder antenna' process.

While using Spacy to identify actions (verbs), components (nouns), and other information, we find that this tool works well on technical documents, specifically on aircraft maintenance manual. However, there still exists a number of false positive (FP) and false negative (FN). We deal with this issue by customizing the Spacy model. We first create a list of aircraft components which is constructed and verified by experts in this domain. This list is then used to retrain the Spacy to recognize components that the original Spacy can not. This list could also be used to verified components recognized by Spacy to eliminate wrong recognition. In addition, we retokenize and apply some matched rules to improve the Spacy performance.

3.1.3 Storing the Core Information Extracted from Maintenance Text

```
{
"id_process":"wt9-34-50-11",
"id_step":wt9-4-S4,
"tasks":[
   {
      "id":"wt9-T4",
      "subtasks":[
       ...
       {
          "id":"wt9-T4-3",
          "action":["pull out"],
          "objects":[
             {
                "id":"5",
                "name":"transponder antenna",
                "direction":"from",
                "position":"",
                "status":"",
                "quantity":""
             },
             {
                "id":"10",
                "name":"fuselage bottom part",
                "direction":"",
                "position":"",
                "status":"",
                "quantity":""
             }
          ],
          "warning":"",
          "guide":"",
          "previoussubtask":"wt9-T4-2"
       }
      ],
      "previoustask":"3"
   }
 ]
}
```

Listing 1: An excerpt of the output in json given by the analysis of task 4 of the process 35-50-11.

After analyzing the maintenance text, we store the output in JSON structure as shown in the Code 1. Each step has an ID, a name, a list of tasks, and point to a previous task. The previous task of a task can be itself if the current task is the first one of the process. For each task, we store an Id, a name, a list of subtasks, and a previous task ID. In the substask, pieces of information included are the identification, the list of actions, the related objects, the warning, and the guide. The ID, name, direction, position, status, and quantity are identified for all objects related to a certain action. The JSON output of the analysis for the step 4, process 34-50-11 is illustrated by the Code 1.

3.2 Building Knowledge Base

3.2.1 Ontology Development

We try to propose a more generic ontology as possible to apply our approach to the aeronautic maintenance domain. As maintenance documents of this domain aren't always publicly available, we developed our ontology based on an XA41/XA42 AMM, a WT9 Dynamic LSA AMM, and some internal Airbus AMMs. Figure 4 depicts our ontology composed of two parts:

1. **Maintenance Procedure Description:** To describe maintenance procedures, we consider the following concepts:
   - Process, task, and subtask: The *Process* class represents the general maintenance procedure that is composed of successive tasks. A *Task* can contain in turn subtasks (*Subtask*) and can refer to another process.
   - Resources: *Resource* is held by an *Enterprise*. It's categorized as a *Device* or a *Tool* used to perform a manual operation (an *Act*) on a *Component* or a part of a component (*Component Part*).
   - Context: This abstract class represents additional information about a process or task. It could be a warning, a preliminary condition for action, a state to switch to, or the precise position of the component.
2. **Graphical Information:** To represent information about graphical objects, we use the *Graphic* class that is specialized by *Resource Graphic* and *Context Graphic*. *hasGraphic* is introduced as an annotation property so that individual of *Context* and concepts representing a resource can be linked a *Graphic*.

3.2.2 Knowledge Base Construction

We integrate several resources to enrich our knowledge base as follows:

Aircraft Maintenance Procedures: These pieces of information were extracted using the NLP techniques presented in Section 3.1. Extracted concepts could be enriched by many ways: using dictionary, such as WordNet[3] (for common nouns) or VerbNet[4] (for actions); using open data, such as DBPedia[5] or Wikidata[6] (for named entity); or by domain experts. This part is under investigation. So far we applied WordNet and VerbNet while the other resources will be left for future work.

Aircraft Components: Aircraft components are presented by concepts belonging to the domain ontology and are validated by experts to enrich the vocabulary.

Tools and Devices: We have populated a preliminary dataset of popular tools and devices used for maintenance tasks.

Graphics: Aircraft components are drawn by specialists or given by enterprises. Information of interval objects inside the graphics and the corresponding aircraft components are imported. In addition, we imported icons for the context (warnings, cautions or notes) and tools.

Figure 5 represents an excerpt of the knowledge graph describing the task 4 of the process, using the developed ontology. A number of classes have been populated beforehand, for example, action verbs like *Install* and *Pull*, that have *Act* as parent class; or aircraft components like *WT9 Antenna* and *WT9 Fuselage*, that specialize the *Component* class.

3.3 Generating Illustration

The generating process is inspired from (Kouame et al., 2020). As presented, a resource or a context can have a graphic representation based on that we can generate the corresponding illustration for a maintenance task. Regarding Figure 6, the first image (Init) shows our origin SVG image (wt9_uc.svg), inside each component has a proper ID corresponding to the one of the knowledge graph. Thanks to the identifier and ordering, components can be removed (or hidden) or even animated. Furthermore, an arrow can be also added to represent the direction of the maintenance action (currently only up and down), as demonstrated by the next image. In this manner, the whole maintenance process can be demonstrated by a sequence of illustrations, each of which corresponds to a particular task.

[3]https://wordnet.princeton.edu/
[4]https://verbs.colorado.edu/verbnet/
[5]https://www.dbpedia.org/
[6]https://www.wikidata.org/

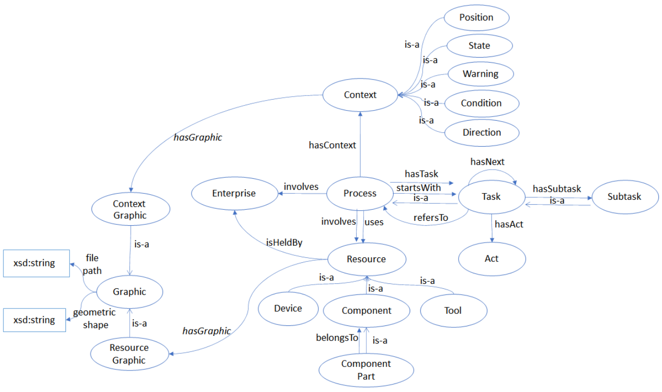

Figure 4: An ontology representing maintenance procedures.

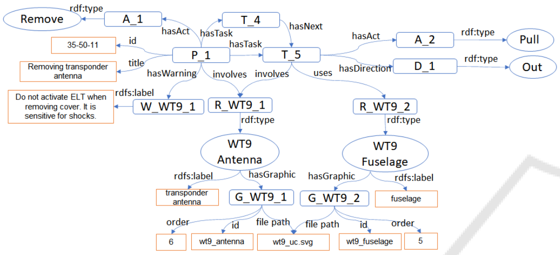

Figure 5: An excerpt of the knowledge graph describing a task of the maintenance process.

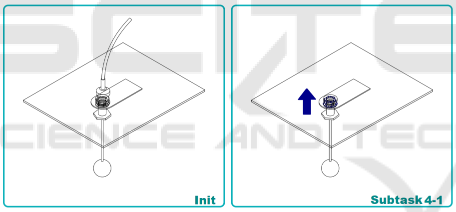

Figure 6: Graphical representation of Subtask 4-1 in the process 34-50-11.

4 CONCLUSION

In this paper, we introduce a method to simplify aircraft maintenance processes combining NLP techniques, knowledge base, and graphical representation. We first extract actions, components and related information in processes and then use these results as an input to populate our ontology. We also use external resources to improve our knowledge base. In addition, we integrate images corresponding to aircraft components into the ontology and use them to generate a graphical representation for each task in the processes.

In the future, we have a plan to build a ground truth on a big dataset and use it to evaluate our NLP method. In addition, we will improve our knowledge base by apply more external resources and by having validation from domain experts.We would also want to construct a collection of images corresponding to all aircraft components. This will help us effectively generate graphical representation of tasks in maintenance process.

We suppose that the approach of simplifying maintenance processes we built have a broad range of applications in several industrial domains such as car maintenance, ship maintenance or mechanic machine maintenance.

REFERENCES

Adnan, K. and Akbar, R. (2019). An analytical study of information extraction from unstructured and multidimensional big data. *Journal of Big Data*, 6(1).

Al Omran, F. N. A. and Treude, C. (2017). Choosing an nlp library for analyzing software documentation: a systematic literature review and a series of experiments. In *2017 IEEE/ACM 14th international conference on mining software repositories (MSR)*. IEEE.

Aly, R., Guo, Z., Schlichtkrull, M., Thorne, J., Vlachos, A., Christodoulopoulos, C., Cocarascu, O., and Mittal, A. (2021). Feverous: Fact extraction and verification over unstructured and structured information. *arXiv preprint arXiv:2106.05707*.

Annane, A., Aussenac-Gilles, N., and Kamel, M. (2019). Bbo: Bpmn 2.0 based ontology for business process representation. In *20th European Conference on Knowledge Management (ECKM 2019)*.

Ast, M., Glas, M., Roehm, T., and Luftfahrt, V. (2014). *Creating an ontology for aircraft design*. Deutsche Gesellschaft für Luft-und Raumfahrt-Lilienthal-Oberth eV.

Beltagy, I., Lo, K., and Cohan, A. (2019). Scibert: A pretrained language model for scientific text. *arXiv preprint arXiv:1903.10676*.

Bird, S., Klein, E., and Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.".

Chungoora, N., Young, R. I., Gunendran, G., Palmer, C., Usman, Z., Anjum, N. A., Cutting-Decelle, A.-F., Harding, J. A., and Case, K. (2013). A model-driven ontology approach for manufacturing system interoperability and knowledge sharing. *Computers in industry*.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Dixit, S., Mulwad, V., and Saxena, A. (2021). Extracting semantics from maintenance records. *arXiv preprint arXiv:2108.05454*.

Fraga, A. L., Vegetti, M., and Leone, H. P. (2018). Semantic interoperability among industrial product data standards using an ontology network. In *ICEIS (2)*.

Gao, S., Kotevska, O., Sorokine, A., and Christian, J. B. (2021). A pre-training and self-training approach for biomedical named entity recognition. *PloS one*.

Ge, S., Wu, F., Wu, C., Qi, T., Huang, Y., and Xie, X. (2020). Fedner: Privacy-preserving medical named entity recognition with federated learning. *arXiv preprint arXiv:2003.09288*.

Grüninger, M. (2009). Using the psl ontology. In *Handbook on Ontologies*. Springer.

Honnibal, M. and Montani, I. (2017). Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing. *Unpublished software application. https://spacy. io*.

Ipfelkofer, F., Lorenz, B., and Ohlbach, H. J. (2006). Ontology driven visualisation of maps with svg-an example for semantic programming. In *Tenth International Conference on Information Visualisation (IV'06)*. IEEE.

Karray, M. H., Chebel-Morello, B., and Zerhouni, N. (2012). A formal ontology for industrial maintenance. *Applied ontology*.

Kouame, A., Brou, K. M., Lo, M., and Lamy, J.-B. (2020). Visual representation of african traditional medicine recipes using icons and a formal ontology, ontomedtrad. In *MIE*.

Kuicheu, N. C., Wang, N., Tchuissang, G. N. F., Siewe, F., and Xu, D. (2012). Description logic based icons semantics: An ontology for icons. In *2012 IEEE 11th International Conference on Signal Processing*. IEEE.

Lamy, J.-B., Duclos, C., Bar-Hen, A., Ouvrard, P., and Venot, A. (2008). An iconic language for the graphical representation of medical concepts. *BMC medical informatics and decision making*.

Lamy, J.-B. and Soualmia, L. F. (2017). Formalization of the semantics of iconic languages: An ontology-based method and four semantic-powered applications. *Knowledge-Based Systems*.

Lehtonen, T. and Karhela, T. (2006). Ontology approach for building and visualising process simulation models using 2d vector graphics. In *SIMS Proceedings of the 47th Conference on Simulation and Modeling*. Finnish Society of Automation, SIMS-Scandinavian Simulation Society.

Lopes, F., Teixeira, C., and Oliveira, H. G. (2019). Contributions to clinical named entity recognition in portuguese. In *Proceedings of the 18th BioNLP Workshop and Shared Task*.

Ma, X. and Cahier, J.-P. (2014). Graphically structured icons for knowledge tagging. *Journal of information science*.

Natschläger, C. (2011). Towards a bpmn 2.0 ontology. In *International Workshop on Business Process Modeling Notation*. Springer.

Nayel, H. A., Shashirekha, H., Shindo, H., and Matsumoto, Y. (2019). Improving multi-word entity recognition for biomedical texts. *arXiv preprint arXiv:1908.05691*.

Niknam, M. and Kemke, C. (2011). Modeling shapes and graphics concepts in an ontology. In *SHAPES*.

Patel, R. and Tanwani, S. (2019). Application of machine learning techniques in clinical information extraction. In *Smart Techniques for a Smarter Planet*. Springer.

Rospocher, M., Ghidini, C., and Serafini, L. (2014). An ontology for the business process modelling notation. In *FOIS*.

Sharp, M., Sexton, T., and Brundage, M. P. (2017). Toward semi-autonomous information. In *IFIP International Conference on Advances in Production Management Systems*. Springer.

Singh, J., Joshi, N., and Mathur, I. (2013). Development of marathi part of speech tagger using statistical approach. In *2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE.

Stefanidis, D., Christodoulou, C., Symeonidis, M., Pallis, G., Dikaiakos, M., Pouis, L., Orphanou, K., Lampathaki, F., and Alexandrou, D. (2020). The icarus ontology: A general aviation ontology developed using a multi-layer approach. In *Proceedings of the 10th International Conference on Web Intelligence, Mining and Semantics*.

Tarbouriech, C., Bernard, D., Vieu, L., Barton, A., and Éthier, J.-F. (2021). Flight procedures description using semantic roles. In *CEUR Workshop Proceedings*.

Valenzuela-Escárcega, M. A., Hahn-Powell, G., Surdeanu, M., and Hicks, T. (2015). A domain-independent rule-based framework for event extraction. In *Proceedings of ACL-IJCNLP 2015 System Demonstrations*.

Wadden, D., Lin, S., Lo, K., Wang, L. L., van Zuylen, M., Cohan, A., and Hajishirzi, H. (2020). Fact or fiction: Verifying scientific claims. *arXiv preprint arXiv:2004.14974*.

Xu, K., Zhou, Z., Hao, T., and Liu, W. (2017a). A bidirectional lstm and conditional random fields approach to medical named entity recognition. In *International Conference on Advanced Intelligent Systems and Informatics*. Springer.

Xu, K., Zhou, Z., Hao, T., and Liu, W. (2017b). A bidirectional lstm and conditional random fields approach to medical named entity recognition. In *International Conference on Advanced Intelligent Systems and Informatics*. Springer.

Yang, T., He, Y., and Yang, N. (2022). Named entity recognition of medical text based on the deep neural network. *Journal of Healthcare Engineering*, 2022.