

*Internship Report*

Development Of Ontologies And Knowledge Graphs For The Manufacturing Domain

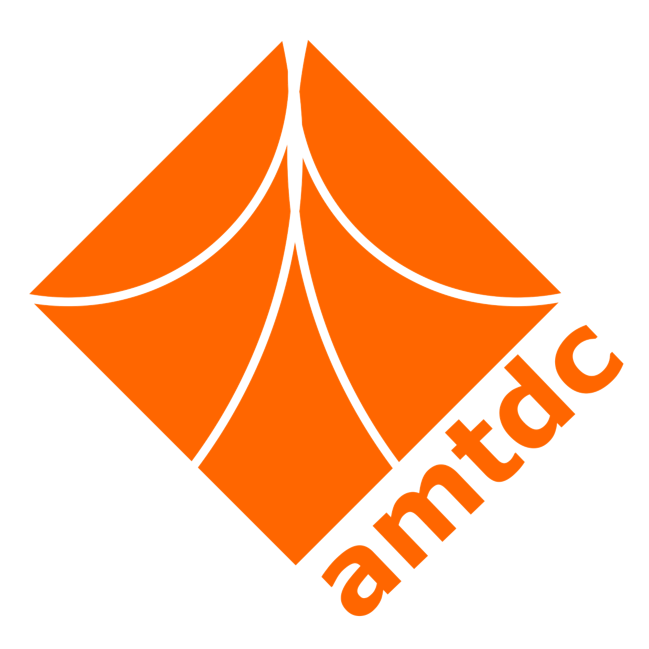
Submitted by

Shriram G

Year II Semester IV

Anna University Regional Campus Coimbatore

as a part of the internship conducted by



Sponsored by



Ministry of Heavy Industries

Government of India

**April 2022 - July 2022**

**TABLE OF CONTENTS**

[**ACKNOWLEDGEMENTS**](#_8usxl8e75s94) **3**

[**A**](#_ugftai8owsx6)**BSTRACT 4**

[**SUMMARY**](#_7xbqcgc7r38c) **OF INTERNSHIP ACTIVITIES 4**

**INTRODUCTION 5**

**WORK DONE 6**

[**SUMMARY**](#_pui4zj45chjf) **7**

[**REFERENCES**](#_kqfp3eq3dd6b) **8**

[**APPENDIX**](#_brxn2wy46svh) **9**

# **ACKNOWLEDGEMENTS**

I would like to express my deepest gratitude to Dr. Srikanth Ramasubramaniam and Dr. Vairamuthu Rajagopal for their exceptional guidance and mentorship throughout my internship. Their expertise, unwavering support, and insightful suggestions have been invaluable in shaping the success of this project. Their dedication and commitment to excellence have truly inspired me, and I am immensely grateful for the opportunity to work under their supervision.

I would also like to extend my heartfelt appreciation to Rahul Singhal and Shruti S, my esteemed teammates, for their unwavering commitment and collaborative spirit. Their valuable contributions, creative ideas, and willingness to go the extra mile have played a significant role in achieving the objectives of this project. The synergy and camaraderie within our team have made this internship experience both enjoyable and enriching.

Lastly, I would like to express my gratitude to the entire organization and all its members for providing me with this remarkable opportunity. The stimulating work environment, encouragement, and resources made available to me have contributed significantly to my professional growth.

# **ABSTRACT**

The manufacturing industry generates vast amounts of data across various domains, making it challenging to effectively manage, organize, and utilize this information. Ontologies and knowledge graphs have emerged as powerful tools for knowledge representation, capturing relationships between concepts, and facilitating intelligent information retrieval. By leveraging natural language processing techniques, these knowledge structures can be enriched and populated with relevant data extracted from textual sources.

The primary motivation behind this project was to enhance knowledge representation within the manufacturing domain, enabling researchers, practitioners, and decision-makers to navigate and explore information effectively. By constructing a knowledge graph that captures domain-specific concepts, relationships, and attributes, it becomes possible to uncover valuable insights, support decision-making processes, and enable efficient knowledge retrieval. Meanwhile NLP techniques excel in the function of advanced information retrieval of information corpora made available in the database of the user.

A user-friendly prototype was developed and presented to provide the semblance for a final working proof-of-concept model that allowed for efficient testing and dynamic update of requirements as and when necessary to provide an intuitive and interactive experience.

This report presents the outcomes of the internship project aimed at developing ontologies and knowledge graphs for the manufacturing domain using natural language processing (NLP) techniques. The project employed Neo4j as the graph database, Spacy and SpanBERT for NLP analysis, and Gradio for an interactive interface. The goal was to utilize NLP models to extract relevant information from manufacturing text data, construct ontologies and create a knowledge graph that captures domain-specific concepts and relationships. The project aimed to improve knowledge representation and facilitate efficient exploration of manufacturing-related information. The report provides an overview of the project's methodology, implementation, functionality, results. The findings demonstrate the successful integration of NLP techniques, showcasing the effectiveness of this approach in constructing a robust NLP techniques for the manufacturing domain. The report concludes by discussing the implications of the project and potential avenues for future enhancements, highlighting its contribution to knowledge representation and extraction in manufacturing using NLP methodologies.

# **INTRODUCTION**:

* Brief introduction about the problem / challenge
* Objectives of the work
* Scope of your work

# 

# Work / activities carried out:

1. Methodology adopted and workflow
2. Elaborate on various activities within the workflow
3. Results of your internship problem

# Summary/conclusion

The internship project has successfully achieved its objectives in the development of ontologies and knowledge graphs for the manufacturing domain using NLP techniques. The integration of Spacy and SpanBERT enabled us to extract relevant information from textual data, enriching the knowledge graph with domain-specific concepts and relationships. The Neo4j graph database served as a robust infrastructure for storing and querying the developed knowledge graph efficiently. Furthermore, the implementation of the Gradio interface provided a user-friendly and interactive platform for users to explore and query the knowledge graph.

The project's outcomes contribute to the advancement of knowledge representation and extraction in the manufacturing domain, offering researchers and practitioners a valuable resource for accessing and analyzing manufacturing-related information. The developed ontologies and knowledge graphs provide a structured and comprehensive view of the domain, facilitating informed decision-making and supporting various applications such as process optimization, quality control, and supply chain management.

Although the project has achieved its primary goals, there are potential areas for further enhancement. Future work may involve expanding the ontology to incorporate additional domain-specific concepts, integrating more advanced NLP models, and integrating data from diverse sources to further enrich the knowledge graph. Furthermore, user feedback and usability studies can help refine the Gradio interface and improve the overall user experience.

# **REFERENCES**

1. Mastering spaCy : An End To End Practical Guide by Duygu Altinok

2. Natural Language Processing With Tensorflow by Thushan Ganegedara

3. Speech and Language Processing by Daniel Jurafsky and James Martin

4. Getting Started With Google BERT by Sudarshan Ravichandhiran

5. Learning Neo4j by Jerome Baton, Rik Van Bruggen

6. Machine Learning With Pytorch and Scikit-Learn by Sebastian Rachka

7. Deep Learning by Francois Chollet

8. Introduction To Information Retrieval by Christopher Manning and Heinrich Schutze

# **APPENDIX**

\*\*\*\*\*\*\*\*\*