MBSE 2025

FROM UNSTRUCTURED DATA TO INTELLIGENT MODELS

AEROMLM: ACCELERATING MBSE IN AEROSPACE

TRY THE DEMO

Maria-Theodora Folina
Department of Applied Informatics, University of Macedonia, Ics22059@uom.edu.gr

1 Introduction

No public structured datasets for aerospace requirements exist to support MBSE

- Aerospace engineers face thousands of unstructured requirements in lengthy PDF documents.
- Entity + relationship extraction needed for **SYSML/OPENMBEE**
- Addresses the event's objective of lowering adoption barriers through an automated migration path for existing data.

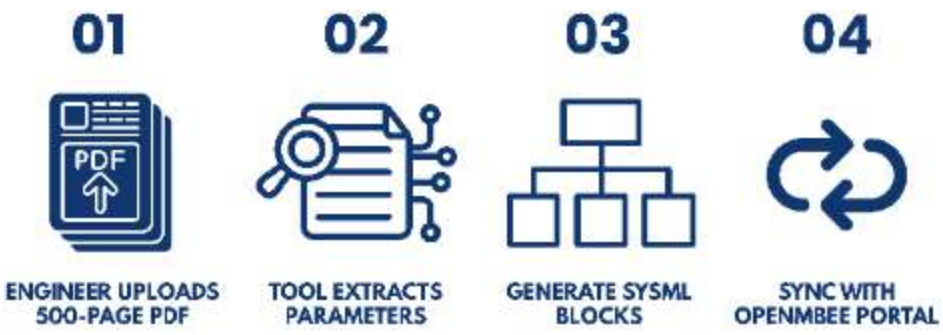
2 Solution

PURE + SCIBERT + AEROMLM → STRUCTURED OUTPUT

- Extend PURE (Public Requirements) dataset with custom aerospace-labeled data.
- Perform domain-adaptive pretraining (DAPT) of SciBERT on aerospace text (MLM) → AeroMLM.
- Fine-tune AeroMLM for NER and relation extraction in MBSE applications.

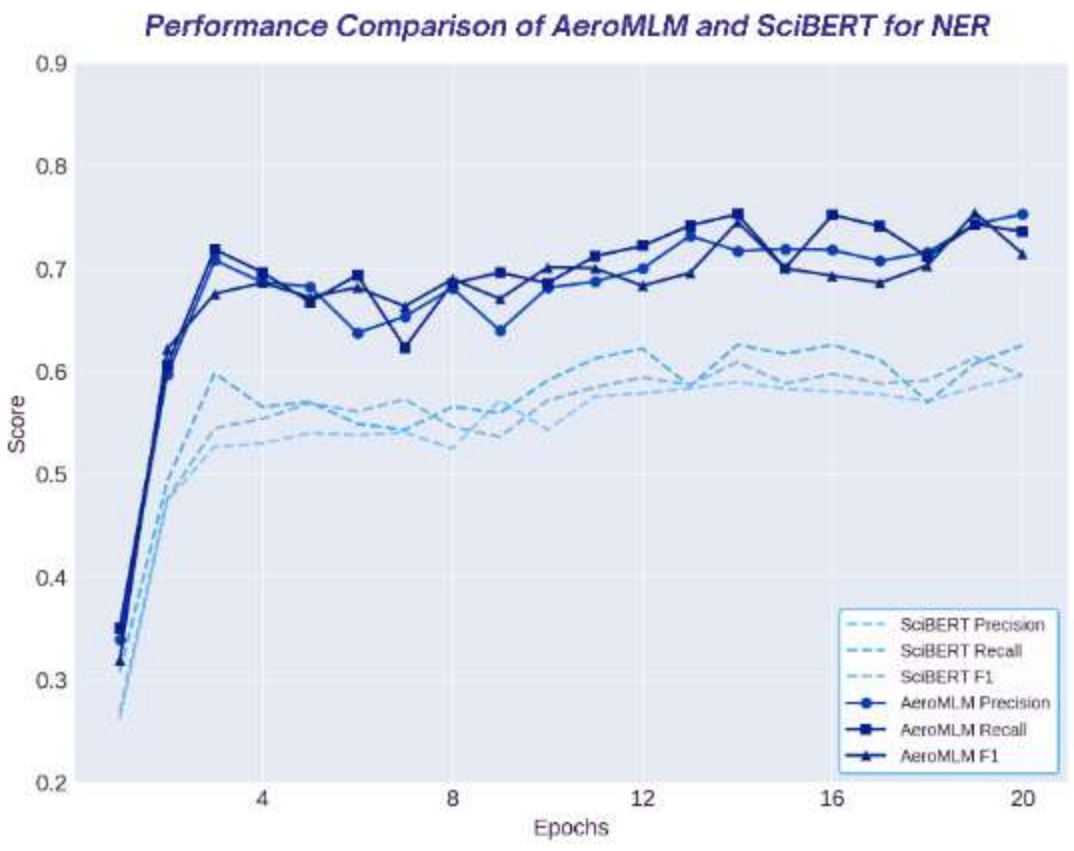
3 Methodology

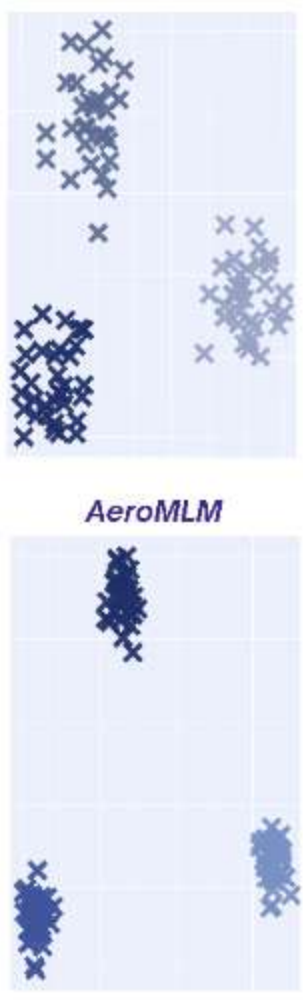
A schema is applied to PURE with an aerospace seed set. SciBERT undergoes domain-adaptive pretraining via MLM, then is fine-tuned for NER to map results.



4 Results

- 10× faster requirement parsing vs. manual effort.
- Weeks of effort reduced to days → accelerated MBSE adoption.
- SysML models generated automatically.
- Teams share a "single source of truth" in OpenMBEE.





Visualization of requirement embeddings

With applications in areas such as:



Conclusion

AeroMLM helps "bridge the gap" between traditional text-heavy processes and digital MBSE practices, driving faster adoption in aerospace engineering.

Related literature
Cai et al., "MBSE 2.0: Toward More integrated, Comprehensive, and Intelligent MBSE," Systems, 2025
Riesener et al., "Application of NLP for Requirement Management in MBSE," INCOSE IS 2021