

JARGES: Detecting and Decoding Jargon for Enterprise Search.



1. SUMMARY

Newcomers to an organisation often struggle with unfamiliar internal vocabulary, which can affect their ability to retrieve relevant information. Enterprise Search (ES) systems frequently underperform when queries contain jargon or terminology that is specific to the organisation. This paper introduces 'JARGES', a novel feature for i) detecting and ii) decoding jargon for ES. It is designed to enhance a ranking model combining Learning to Rank (LTR) and transformer-based synonym expansion.

2. BACKGROUND & MOTIVATION

Enterprise Search (ES) can be simply defined as finding the information needed from within an organisation [Bentley, 2011] or as a service that "enables employees to find all the information the company possesses without knowing where the information is stored" [White, 2015]. Jargon is enterprise-specific vocabulary that employees/members can understand. It encompasses words, phrases, expressions, and idioms that are not universally familiar or properly understood.

An organisation's ES service is critical for accessing internal knowledge. Users often struggle with relevance due to specialised organisational jargon, causing vocabulary mismatch that hinders effective searches, particularly for newcomers, leading to poor query results and reduced productivity.

Jargon is not necessarily a bad thing!

Although excessive use of jargon and terminology in organisations is often perceived as exclusionary, we use the term here in a positive context for conveying complex ideas, processes, or services among employees/members who share common knowledge of the enterprise. In this context, jargon and terminology facilitate efficient communication.

3. AIM and OBJECTIVES

To address this problem, we propose a ranking model that integrates Learning to Rank (LTR) and Language Modelling (LM). The LM component is called 'JARGES' (JARGOn for Enterprise Search), and is designed to detect and decode jargon in ES queries and corpora. This study aims to evaluate the effectiveness of this approach in improving search result rankings, particularly for content rich in organisational terminology.

To test this hypothesis, we perform a quantitative evaluation of the performance of an LTR ranking model with and without JARGES using the LTR-formatted ENTRP-SRCH dataset (2,544 human-annotated Q-D pairs). We subsequently perform a qualitative analysis of the decoded jargon terms via their contextual synonyms.

4. DETECTING JARGON

Comparative TF-IDF scoring and sparse matrix vectorisation can be used to distinguish word or phrase salience patterns between corpora [Belfathi et al., 2024].

"The University's Blackboard VLE can record results via the rubrics scoring scheme"

The JARGES algorithm flags the words "blackboard", "vle" and "rubrics" as jargon because they do not occur frequently in common vocabulary but have a high TF-IDF score in the organisation's corpus.

The words "university", "scoring" and "scheme" are classified as non-jargon (as the TF-IDF score divergence between common terms and corpus-specific terms is minimal).

9. REFERENCES

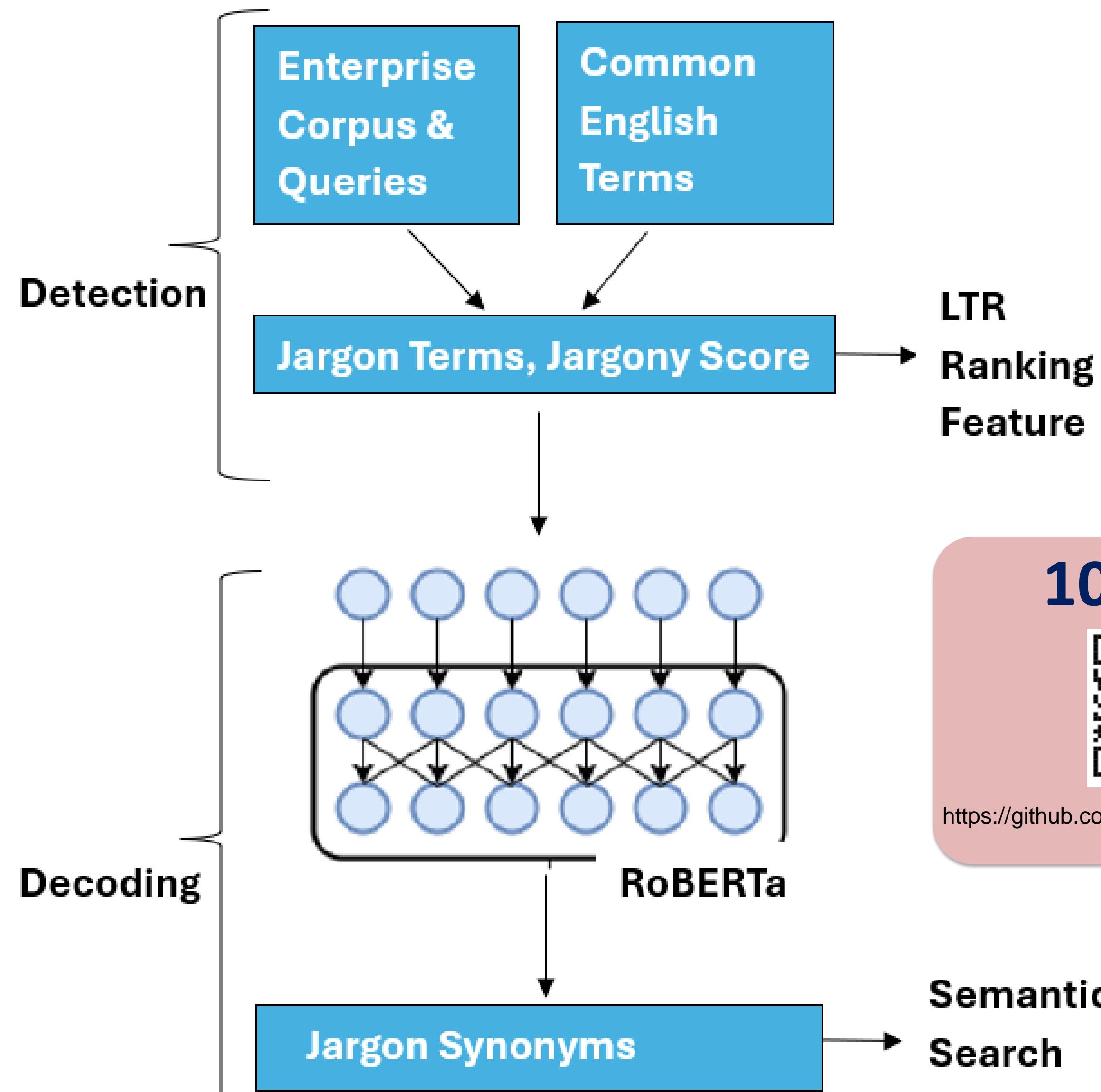
[Belfathi et al., 2024] Belfathi, A., Gallina, Y., Hernandez, N., Dufour, R., and Monceaux, L. (2024). Language Model Adaptation to Specialized Domains through Selective Masking based on Genre and Topical Characteristics. arXiv, 2402.12036.

[Bentley, 2011] Bentley, J. (2011). Mind the Enterprise Search Gap: Smartlogic Sponsor MindMetre Research Report.

[Daly, 2025] Daly, C. and Hederman, L. (2025). JARGES: Detecting and Decoding Jargon for Enterprise Search. In 17th International Conference on Knowledge Discovery and Information Retrieval (KDIR).

[White, 2015] White, M. (2015). Critical success factors for ES.

5. MODEL ARCHITECTURE



10. GitHub:



<https://github.com/ColinDaly75/ES-Ranking/>

6. DECODING & SYNONYMS

RoBERTa LM is adapted for synonym generation by fine-tuning it on the organisation's corpus. RoBERTa is used to predict plausible alternatives for the previously detected jargon terms. Synonyms can enhance semantic search by enabling the system to recognise and retrieve conceptually related terms beyond exact keyword matches.

7. EVALUATION / RANKING

Quantitative testing using the ENTRP-SRCH dataset failed to demonstrate a statistically significant increase in ranking performance, as evidenced by an nDCG@10 change of less than 0.001 (where $p > 0.05$). This result is disappointing, but not entirely unexpected, as the ENTRP-SRCH is small, with limited query diversity and a scarcity of jargon-rich Q-D pairs. In spite of this, the qualitative analysis of the generated synonyms revealed promising results for recall as a foundation for semantic search.

A/B test results for ranking models with the percentage change in nDCG score after implementation of the JARGES feature.

Feature	nDCG@10
Base LTR model	0.9646 ± 0.001
With JARGES	0.9639 ± 0.001
Percentage change	0.0007%

8. FUTURE WORK

Future work plans will address the limitations of our ENTRP-SRCH dataset, which centres on just twenty of the most frequently submitted queries. The use of click-through data in place of human judgements for Q-D pair annotation would facilitate larger and more diverse ES datasets that are better able to capture the nuances of enterprise specific terminology. Finally, it would be interesting to perform a longitudinal study to gauge the JARGES impact on semantic and exploratory search, based on query expansion and recall on a real-world ES system.