BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY

Excel@FIT 2026
http://excel.fit.vutbr.cz

Application for recommending participation in public tenders

Tony Pham*

Abstract

Public tenders are often published on heterogeneous web portals and accompanied by extensive documentation in formats such as PDF, DOCX, XLSX, and HTML. For supplier companies, deciding whether to participate is a time-consuming process that requires manual analysis of requirements, deadlines, contractual conditions, and tender relevance. The aim of this thesis is to design and implement a web application that supports this process by automatically collecting, processing, and analysing tender documentation.

The proposed solution uses web scraping, document processing, natural language processing, and large language models. It retrieves tender data from different portals, extracts text from attached documents, identifies key information, classifies the tender domain, and generates a structured summary with a recommendation for participation.

The result is a functional prototype tested on real tender samples and verified in cooperation with GORDIC. The application reduces manual work, improves orientation in tender documentation, and supports more consistent and informed decision-making.

*xphamt00@stud.fit.vutbr.cz, *Faculty of Information Technology, Brno University of Technology*

1. Introduction

Supplier companies often have to decide quickly whether a public tender is worth pursuing. This decision is difficult because tender documentation is published across different procurement portals and stored in various formats such as PDF, DOCX, XLSX, and HTML. Important information about technical requirements, deadlines, qualification criteria, integrations, and contractual conditions must usually be checked manually. Similar challenges in public procurement and public administration are discussed in AI-based decision support systems [1] and information extraction approaches [2].

The goal of this work is to support the initial evaluation of public tenders. The proposed web application automates the entire process—from obtaining tender data to analysis and recommendation. The overall system architecture is shown in Figure 1, while the processing workflow is illustrated in Figure 2. The user interacts with the system through a simple interface for submitting a tender URL (Figure 3).

2. Tender Acquisition and Document Processing

The first part of the solution focuses on obtaining tender data from heterogeneous sources. The application accepts a tender URL via an API endpoint and selects a suitable adapter for the source portal, as shown in Figure 1. The implemented prototype supports several procurement systems, including NEN, E-ZAK, TenderArena, PROEBIZ, and Vhodné uveřejnění.

After acquiring the data, the system downloads available documents and processes them into a unified text representation. As illustrated in Figure 1, this includes extracting text from formats such as PDF, DOCX, XLSX, TXT, and HTML, followed by normalization and splitting into smaller chunks. This preprocessing step is necessary for handling unstructured and heterogeneous documents, which remains a key challenge in modern information extraction systems [3].

3. Language Model Analysis

The core analytical part uses large language models through Azure OpenAI, as illustrated in the workflow in Figure 2. The system performs domain classifica-

tion and extracts structured information from the processed documents.

The extracted information includes the subject of the tender, key requirements, deadlines, technical conditions, integrations, and security aspects. The use of large language models for tender analysis is supported by recent research demonstrating their effectiveness in extracting relevant information from procurement documents [4]. The results are stored as structured outputs (e.g., JSON files), as shown in Figure 2. Identified and missing areas of the tender documentation are presented to the user in a structured way (Figure 4).

4. Scoring and Decision Support

For selected domains, such as electronic records management systems, the application computes a fit score based on the extracted information. This scoring process is part of the overall workflow shown in Figure 2.

The score consists of multiple components, including section coverage, concept match, module match, capability match, and integration and security fit. The final result is presented to the user as a clear recommendation with supporting metrics, as illustrated in Figure 5.

This approach follows the concept of AI-based decision support systems, where the system assists the user but does not replace expert judgement [1].

5. Conclusions

This work presents a functional prototype of a web application for automated public tender analysis. The solution integrates data acquisition, document processing, AI-based analysis, and scoring into a unified architecture (Figure 1) and workflow (Figure 2).

The main contribution is a practical decision-support tool that reduces manual effort and improves the efficiency of evaluating public tenders. The system presents results through a user-friendly interface (Figure 3) and structured outputs (Figures 4 and 5), enabling faster and more informed decisions.

Future work may focus on extending support for additional procurement platforms, improving processing of scanned documents, and evaluating the system on larger datasets..

Acknowledgements

I would like to thank my supervisor Ing. Jiří Hynek for his guidance and support. I would also like to thank my girlfriend, my family, my rabbit, and my dog for their support.

References

[1] Lucia Siciliani, Vittorio Taccardi, Pierpaolo Basile, Michele Di Ciano, and Pasquale Lops. Ai-based decision support system for public procurement. *Information Systems*, 119:102284, 2023.

[2] Lucia Siciliani, Elena Ghizzota, Pierpaolo Basile, and Pasquale Lops. Oie4pa: Open information extraction for the public administration. *Journal of Intelligent Information Systems*, 62(1):273–294, 2024.

[3] Derong Xu, Wei Chen, Wenwei Peng, Chao Zhang, Tong Xu, Xiangyu Zhao, and Enhong Chen. Large language models for generative information extraction: A survey. *Frontiers of Computer Science*, 18(6):186357, 2024.

[4] V. P. Mäkinen. Cost-effective tender documents analysis with large language models, 2025.