

# Using Natural Language Processing for Agent Evaluation

Evan Wells

**Abstract- This project aims to automate the Agent Evaluation step in Role-Based Collaboration (RBC) by leveraging Natural Language Processing (NLP) to interpret qualitative agent data—such as resumes, reviews, surveys, and written feedback. By converting unstructured or semi-structured text into a structured qualification matrix (Q matrix), the system enables more accurate and scalable assignment decisions. This process replaces manual evaluation with a data-driven approach, improving both efficiency and reliability in collaborative environments.**

## I. Problem statement/Objectives

Assigning agents to roles in collaborative systems often relies on a member of administration evaluating each agent individually. However, much of the useful agent-related information exists in unstructured text (e.g., feedback reports, resumes). The problem lies in designing a system that can automatically interpret this information and use it to generate a Q matrix, which can be used for optimal assignments using the GRA model.

The objectives of this project are to simplify agent evaluation through automated analysis of unstructured or semi structured data and to use this analysis to produce a qualification matrix from the data. Lastly, we want to smoothly integrate this process with the GRA algorithm.

## II. Methodology

The system will be implemented in Python. Agent data (resumes, reviews, etc.) will be preprocessed using NLP libraries. These extracted attributes will be mapped to role requirements to build a numerical Q matrix, representing the agent's suitability for each role. This matrix will then serve as input to the Group Role Assignment algorithm, implemented using PuLP for linear optimization.

## III. Libraries/APIs

- PuLP, for GRA and GMRA algorithm - <https://coin-or.github.io/pulp/>
- NumPy, for computing – <https://numpy.org/>
- Pandas, for data analysis - <https://pandas.pydata.org/>
- spaCy, for NLP – <https://spacy.io/>  
OR  
Transformers, for NLP – <https://huggingface.co/docs/transformers/en/index>
- PyTorch, for deep learning – <https://pytorch.org/>  
OR  
TensorFlow, for deep learning – <https://www.tensorflow.org/>

## IV. Expected Contributions

This project will contribute a practical framework that brings NLP-based agent evaluation into the RBC workflow. It bridges the gap between qualitative human data and quantitative optimization methods.

The system can be adapted for various domains—corporate team-building, project allocation, or even educational groupings—where structured evaluation data may not be available.

## V. Desired Timeline

The timeline can be approximated as follows:

Week 1-2	Research into NLP, feature extraction and creating sample data.
Week 3-4	Prototyping the approach and proof-of-concept
Week 5-6	Prototyping the approach and proof-of-concept (2)
Week 7-8	Creating a minimalist UI
Week 9-10	Implementing the prototype into UI
Week 11-12	Implementing the prototype into UI (2)
Week 13-14	Debug, fix issues and improve design
Week 15-16	Debug, fix issues and improve design (2)

## VI. Feasibility

The success of this project heavily relies on the quality and quantity of available training data. Given the nature of NLP, real-world, diverse data sources—such as forms, PDFs, and Word documents—will yield the best results. Although this will be my first experience working with NLP, I am prepared to take on the challenge and confident in my ability to learn and implement the required techniques.

## VII. Datasets Required

- Resume Dataset - <https://www.kaggle.com/search?q=resume+in%3Adatasets>
- Employee Review - <https://www.kaggle.com/datasets/fiodarryzhykau/employee-review>
- Employee Engagement - <https://www.kaggle.com/datasets/gun4hire/employee-engagement-dataset>
- Sentiment Analysis - <https://www.kaggle.com/datasets/jp797498e/twitter-entity-sentiment-analysis>  
or <https://www.kaggle.com/datasets/kazanov/sentiment140>
- Employee Survey - <https://www.kaggle.com/datasets/mexwell/employee-performance-and-productivity-data>
- Skill and job matching - <https://www.kaggle.com/datasets/ravindrasinghrana/job-description-dataset>
- Job-Skills - <https://www.kaggle.com/datasets/ravindrasinghrana/job-description-dataset>

## VIII. Similar\Related Research

Although no existing research directly mirrors the proposed project's goal of fully automating agent evaluation using NLP methods, several studies incorporate foundational components relevant to this work. Notably, research in aspect-based sentiment analysis has demonstrated how qualitative textual data—such as reviews or feedback—can be broken down

into analyzable components to derive meaningful judgments. Similarly, systems have been developed to automatically identify and normalize skills from resumes.

The proposed project would generalize and integrate these approaches, using NLP not only to assess sentiment or extract skills, but to produce a quantitative qualification matrix (Q matrix) suitable for optimization algorithms such as GRA. In doing so, it aims to bridge existing research silos into a unified, automated evaluation framework.

#### **Similar Research:**

Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. SemEval-2014 Task 4: Aspect Based Sentiment Analysis. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 27–35, Dublin, Ireland. Association for Computational Linguistics.

Zhao, M., Javed, F., Jacob, F., & McNair, M. (2015). SKILL: A System for Skill Identification and Normalization. *Proceedings of the AAAI Conference on Artificial Intelligence*, 29(2), 4012-4017. <https://doi.org/10.1609/aaai.v29i2.19064>