Applied Research Document

Project: Role2Skill

1. Context

The project is conducted in the context of the online job market and focuses on analyzing job vacancies in the IT and technical domain. Job descriptions are commonly written as unstructured text and contain lists of required skills, which are difficult to interpret and compare.

The results of this research are intended to support students and junior job seekers by improving understanding of how job roles can be described using skills.

2. Problem Statement

Job vacancies often list many skills without a clear structure or priority. Because of this, it is difficult to understand which skills define a job role and how different roles relate to each other. The lack of structure limits meaningful comparison between job roles based on job descriptions alone.

3. Research Objective

The objective of this applied research is to explore how job roles can be represented using extracted skills in a structured way, so that roles can be compared more clearly.

4. Research Question

How can job roles be represented using extracted skills to enable meaningful comparison between roles?

Sub-questions:
- Which approach can be used to extract skills from job descriptions?
- How can extracted skills be normalized to reduce duplication?
- How can skill data be represented to compare job roles?

5. Research Approach and Methods

The research follows an applied approach and uses multiple methods. Literature and source review is used to understand common approaches to skill extraction and normalization. Data experimentation is used to test skill extraction and normalization on job description datasets. Comparative analysis is used to evaluate different ways of representing skill information for job roles.

6. Data and Operationalization

The research uses a dataset containing job titles and job descriptions. A skill is defined as a technical or soft competency mentioned in the text. Skill extraction is considered successful when relevant skills are correctly identified and represented consistently across job vacancies.

7. Results

The experiments show that extracting skills from job descriptions allows job roles to be represented as structured skill sets. Skill normalization reduces variation caused by synonyms and inconsistent naming. Using structured skill representations makes it easier to compare job roles based on their required skills.

8. Quality and Limitations

The results depend on the quality and coverage of the dataset used. Unstructured text and ambiguous skill names can affect extraction accuracy. To reduce this risk, the same processing steps are applied consistently across all job descriptions.

9. Conclusion

This applied research demonstrates that job roles can be represented through extracted and structured skill information. Such representations support clearer comparison between roles and reduce ambiguity in job descriptions.