Research Proposal

Title:  
Job Market Analysis and Matching platform using Artificial Intelligence: A Case Study of South Sudan

# Introduction

South Sudan faces a growing challenge of unemployment and labour market fragmentation. Despite the presence of international organizations, government agencies, and private sector actors, there is no centralized and intelligent system to connect job seekers to opportunities. This research proposes the development of an AI-powered job market analysis and matching platform to bridge the gap between job supply and demand using natural language processing (natural language processing) and machine learning techniques.

# Problem Statement

The labour market in South Sudan lacks a structured mechanism to:

* Aggregate job postings from different sources.
* Analyse labour demand trends.
* Match candidates to jobs based on skills, education, and experience.
* Identify skills gaps and recommend training areas.  
  This fragmentation limits access to employment information, especially for underrepresented groups and emerging professionals.

# Research Contribution

This research contributes to addressing the dual challenges of labour market fragmentation and skill mismatch in South Sudan by:

Developing a comprehensive job market analysis framework that uses artificial intelligence techniques to extract, classify, and analyse job data from multiple sources. This will identify high-demand job roles, emerging employment sectors, and frequently requested skills across regions.

Creating a job-candidate matching system that uses intelligent text analysis to align job descriptions with candidate profiles based on qualifications, experience, and skillsets.

Generating actionable insights to support employment policy, training initiatives, and digital employment platforms. These insights will inform government agencies, development organizations, and training institutions about current labour trends and skills gaps.

# Research Objectives

To collect and analyse job postings and candidate profiles using AI techniques.

To design a natural language processing-based system for extracting and classifying job and skill-related information.

To develop a recommendation engine that matches job seekers with relevant opportunities.

To provide insights into labour market trends and skill gaps through interactive dashboards.

# Research Questions

1. What are the most in-demand job roles and skills across different sectors and regions in South Sudan?  
2. How can artificial intelligence techniques be used to automatically extract and classify information from job postings and candidate profiles?  
3. To what extent do candidate qualifications and skills match current labour market demands?  
4. How effective is a job matching system based on artificial intelligence in recommending suitable job opportunities to job seekers?  
5. What insights can be drawn from labour market data to guide training programs and employment policies?

## Methodology

## Data Collection

1. Scrape job postings from NGO portals, UN sites, LinkedIn, and local job boards.
2. Collect anonymized CVs and profiles from sample candidates.
3. Utilize public datasets like ESCO or O\*NET for skill taxonomy.

## Data Pre-processing

1. Clean and normalize text data.
2. Extract relevant fields (title, location, skills, qualifications) using NLP.

## Model Development

1. Apply Named Entity Recognition (NER) and semantic similarity models (e.g., BERT) to extract and match data.
2. Use TF-IDF and embeddings for text vectorization.
3. Implement clustering for job market segmentation.
4. Build a recommendation engine using content-based filtering.

## Visualization and Reporting

Develop dashboards using Stream lit or Power BI to visualize trends, mismatches, and recommendations.

# Expected Outcomes

1. A functional prototype of an AI-based job matching system.
2. Analytical insights into employment trends and skills demand.
3. Recommendations for job seekers and policymakers.

# Significance of the Study

This research aligns with national priorities for economic development and employment promotion. The outcome can inform workforce planning, improve transparency in the job market, and support digital transformation efforts in South Sudan.

# Timeline

| Phase | Activity | Duration |
| --- | --- | --- |
| Phase 1 | Data collection & cleaning | 2 weeks |
| Phase 2 | NLP model development | 2 weeks |
| Phase 3 | Recommendation system | 1 week |
| Phase 4 | Dashboard and report | 1 week |
| Phase 5 | Review & finalization | 1 week |

# Tools and Technologies

Python (NLTK, spaCy, scikit-learn, Transformers)

Streamlit / Power BI

BeautifulSoup / Selenium (for scraping)

Jupyter Notebooks / Google Colab