**Capstone Project Concept Note and Implementation Plan**

1. Maka K. Armah

2. George W. Monibah  
3. Sekou S. Kamara

**Project Title: WorkWise “Where Talent Meets Opportunity”**

**1. Project Overview**

Youth unemployment in Liberia represents a national crisis that undermines economic growth, social stability, and the productive potential of the country’s young population. Despite thousands of graduates emerging annually from universities, vocational, and technical institutions, many are unable to secure meaningful employment. As a result, large numbers of young people are compelled to work in the informal economy, such as commercial motorcycling or pretty trading to sustain their livelihoods.

The WorkWise platform — “Where Talent Meets Opportunity” — proposes an innovative, AI-powered job matching solution that directly addresses this challenge. By leveraging natural language processing (NLP) and recommendation systems, the platform will match youth skills with employer needs, identify gaps, and recommend relevant training opportunities. This approach will reduce job search frictions, improve labor market efficiency, and contribute directly to Sustainable Development Goal 8 (Decent Work and Economic Growth) by promoting productive employment and reducing the proportion of youth not in employment, education, or training (NEET).

**2. Objectives**

The key objectives of the project are to:

* Bridge the gap between education and employment by connecting young jobseekers to relevant opportunities using AI-driven job matching.
* Identify and address skills mismatches by analyzing labor market data and recommending targeted upskilling and training programs.
* Support employers in finding qualified candidates more efficiently, reducing hiring time and recruitment costs.
* Enhance youth employability and inclusion by creating an accessible, locally adapted platform (via mobile app, SMS, or web) that serves both urban and rural populations.
* Contribute to national development goals by reducing youth unemployment, fostering entrepreneurship, and empowering young people to contribute productively to Liberia’s economy.

**3. Background**

Youth unemployment and underemployment remain among the **most pressing socio-economic challenges in Liberia**. Despite steady growth in the number of graduates from universities, vocational, and technical institutions, many young people cannot secure meaningful employment. According to the International Labour Organization (2022), youth unemployment globally is more than three times the adult rate, with **Sub-Saharan Africa bearing some of the highest levels of youth not in employment, education, or training (NEET)**.

In Liberia, this crisis is visible in the large number of graduates resorting to **survivalist livelihoods** such as motorcycle taxis, street trading, or casual manual labor simply to sustain their families. This mismatch between education outcomes and labor market demand not only wastes human potential but also perpetuates **cycles of poverty, inequality, and social instability**.

**Existing Solutions & Their Gaps**

While Liberia has seen attempts to address the problem through **generic job boards and donor-led employment initiatives**, these efforts remain insufficient.

* Many platforms are **poorly maintained, generic, or irrelevant** to Liberia’s job market.
* Employers continue to report **difficulty finding candidates with the right digital, technical, and soft skills**.
* Training programs exist, but they are **fragmented, limited in scope, and not integrated** with real-time labor market data.  
  As a result, youth face prolonged job searches, frustration, and lack of guidance on how to adapt their skills.

**Global Benchmarks**

Globally, there are promising initiatives that highlight the potential of **AI-assisted job matching**:

* **World Bank (Poland, 2023):** Machine learning-based job matching improved alignment between candidate skills and vacancies in public employment services.
* **Ajira Digital (Kenya):** Connected over one million youth to digital jobs while embedding AI-assisted training pathways.
* **Babajob (India):** Used mobile-based AI systems to link low-income workers with employers, proving viability in **low-resource settings similar to Liberia**.

These cases demonstrate that **AI reduces search friction, increases job match efficiency, and integrates skills development directly into employment pathways.**

**Why Machine Learning for Liberia**

A machine learning approach is especially valuable in Liberia’s labor market for three reasons:

**Skills Extraction & Normalization (NLP):**

* CVs and job descriptions in Liberia are often unstructured and non-standard.
* NLP models can translate informal skill descriptions (e.g., “computer fixing”) into formal categories (“IT Support Technician”), ensuring hidden talents are recognized.

**Personalized Matching & Training Recommendations:**

* A hybrid ML recommender system not only matches youth with jobs but also identifies skill gaps and **recommends targeted training programs**.
* This transforms the platform from a job board into a **career pathway ecosystem**.

**Scalability & Adaptability:**

* Unlike static systems, ML models **improve over time** through feedback loops (applications, interviews, hires, training completions).
* This makes the platform **resilient and adaptive** to Liberia’s evolving labor market.

**4. Methodology**

**1) Frame & KPIs**

1. **Objective:** Maximize quality job matches for Liberian youth and close skill gaps via targeted training.
2. **Metrics:**

* **Precision@K / nDCG:** measures how well top-K recommendations match job seeker needs.
* **Conversion Funnel:** application → interview → hire rate.
* **Time-to-Hire:** speed of successful placement.
* **Training Uptake:** % of youth enrolling in recommended training.
* **Fairness Metrics:** subgroup performance (urban vs rural, male vs female, educated vs vocational).

**2) Data Pipeline**

1. **Collection:** Multi-channel ingestion (web, mobile lite, CSV uploads from partners, scraped job boards).
2. **Processing:**

* **Cleaning & Normalization:** Map synonyms into a standardized ontology (e.g., “computer fixing” → “IT Support”).
* **Document Parsing:** PDFs and unstructured CVs/job descriptions parsed with **OCR + NLP**.

1. **Frameworks:** Python (Pandas, PySpark for scaling), spaCy (text cleaning, entity recognition).

**3) NLP & Search**

1. **CV/Job Parsing:**

* Named Entity Recognition (NER) for skills, education, experience, location.
* Sentence embeddings using **Sentence-BERT / Hugging Face Transformers** for semantic similarity.

1. **Hybrid Retrieval:**

* **BM25 (lexical search)** → captures keyword matches.
* **k-NN vector search (FAISS/Qdrant)** → finds semantically similar jobs even with typos or informal phrasing.
* **Result:** More robust matching for low-resource text in Liberia’s context.

**4) Matching Engine**

1. **Recommender System:**

* **Content-based filtering** → matches user profile attributes (skills, education) to job requirements.
* **Collaborative filtering (LightFM, matrix factorization)** → learns from user–job interaction history (e.g., similar youth hired for similar jobs).

1. **Re-Ranking:**

* Policy-driven constraints (local-first hiring, entry-level prioritization, diversity).
* **Explainable AI layer:** e.g., “Recommended because you have Excel + live within 5km.”

**5) Train, Evaluate, Ship**

1. **Offline Evaluation:** precision@K, coverage, fairness metrics across subgroups.
2. **Online Testing:** A/B tests on CTR (click-through rate), applications, interviews.
3. **Human-in-the-loop:** Employer accept/reject feedback + annotator corrections continuously improve model performance.
4. **MLOps:** Data validation with **Great Expectations**, model monitoring with **Evidently AI**, retraining triggered on drift detection.

**6) Ethics & Safety**

1. **Data Privacy:** PII minimization, encrypted storage, opt-in consent, deletion rights.
2. **Transparency:** Explanations for recommendations.
3. **Bias Mitigation:** Continuous monitoring of subgroup outcomes; fairness-aware re-ranking to ensure inclusivity.
4. **Accessibility:** Low-bandwidth options (SMS/USSD) to avoid digital exclusion.

**5. Architecture Design Diagram:**

This architecture is designed for **scalability, robustness, inclusivity, and continuous learning**, ensuring the system can adapt to Liberia’s dynamic labor market.

**System Architecture Overview**

The system is structured in **four layers** — Frontend, Backend, AI & Data Core, and Data Platform — connected via secure APIs and optimized for low-resource environments.

**1. Frontend (User-Facing Layer)**

* **Mobile App / Web Portal (React, Bootstrap, PWA-ready):** Provides youth and employers with intuitive, low-data usage interfaces.
* **SMS/USSD Gateway (Twilio, Africa’s Talking):** Ensures access for users with limited smartphones or internet connectivity.
* **Role:** Maximize accessibility and inclusivity; provide the primary entry point for job seekers and employers.

**2. Backend (API & Services Layer)**

* **API Gateway (FastAPI/Node.js):** Central hub connecting frontend to AI and data services; handles request routing.
* **Authentication & Authorization (JWT, OAuth2, Keycloak):** Secure identity management and role-based access.
* **Notification Service (SMS/Email/WhatsApp):** Delivers alerts, job recommendations, and training reminders.
* **Role:** Orchestrates secure, reliable communication between users and the AI engine.

**3. AI & Data Core**

* **NLP Pipeline (spaCy + Hugging Face Transformers):** Extracts structured skills/education/experience data from unstructured CVs and job postings.
* **Recommendation Engine (LightFM + Re-ranking Layer):** Matches youth to jobs and training opportunities; applies fairness-aware rules.
* **Feature Store (Feast/Hopsworks):** Central repository for features powering ML models; ensures consistency across training and inference.
* **Role:** The “intelligence” of WorkWise — learns patterns, generates insights, and drives smarter matches.

**4. Data Platform**

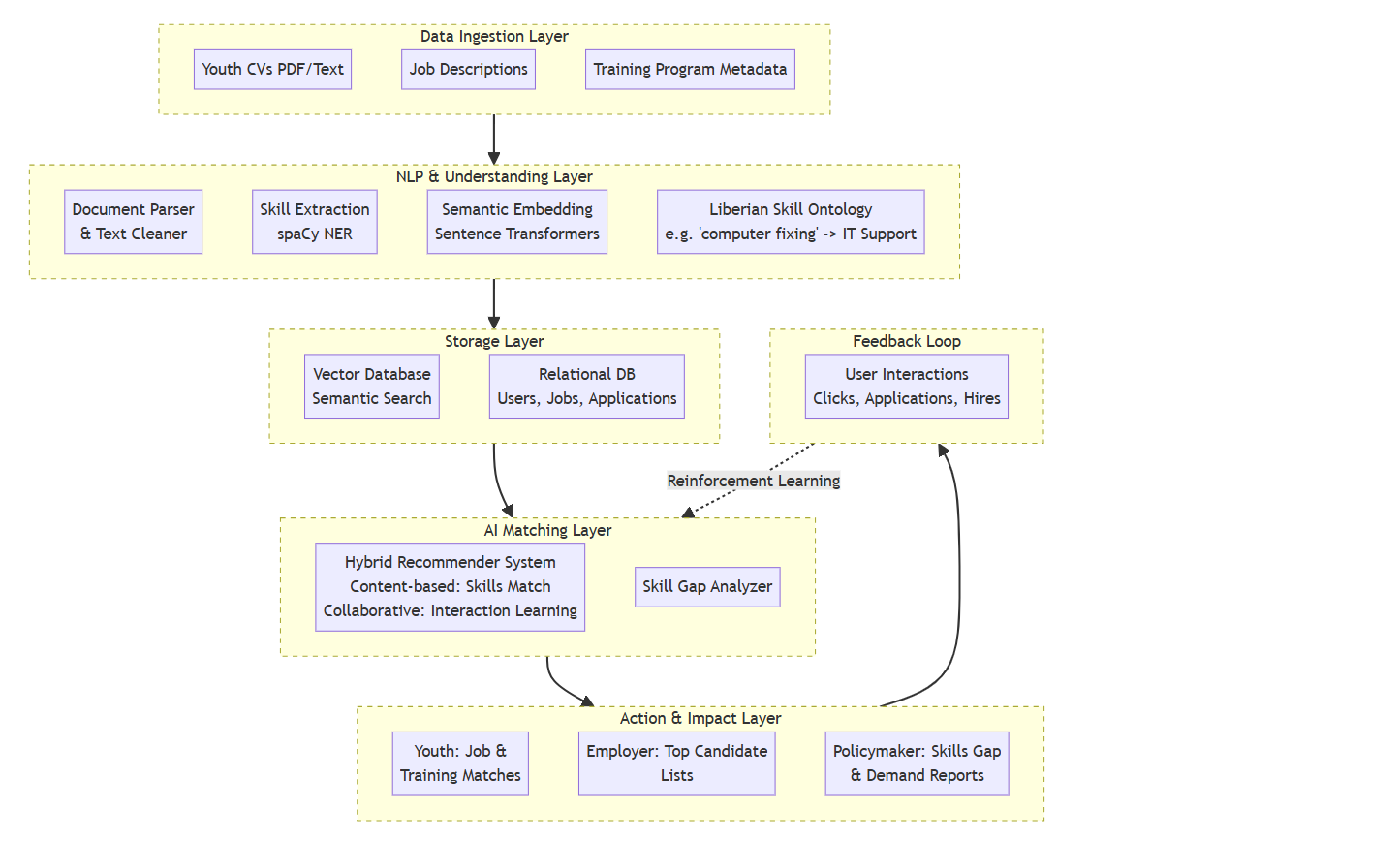
* **Vector Database (Qdrant/Pinecone):** Stores embeddings of CVs and jobs for semantic search and similarity matching.
* **Relational Database (PostgreSQL + PostGIS):** Stores structured user profiles, job listings, and geospatial data for location-aware matching.
* **Object Storage (S3-compatible, MinIO):** Secure storage for CVs, job descriptions, and training materials in PDF/text.
* **Role:** Foundation of data management — ensures structured, unstructured, and vectorized data flow seamlessly.

**5. Data Flows & Protocols**

* **HTTP/REST & GraphQL APIs:** Standard communication for frontend-backend interactions.
* **WebSockets:** Real-time notifications and status updates.
* **gRPC:** High-performance communication between AI services and databases.

**Design Diagram**

The following diagram illustrates the core data flow and AI processes:



**6. Data Sources**

The WorkWise platform will leverage multiple data sources to ensure accurate and inclusive job matching. Primary datasets include **youth-submitted CVs and profiles**, which provide information on skills, education, and work experience, and **job postings** collected from employers, government job boards, and NGO partners offering training opportunities. Additional data will come from **training program databases** to enable skill-gap analysis and personalized learning recommendations.

These datasets will be in mixed formats such as structured CSV files (profiles, job listings) and unstructured text or PDFs (CVs, job descriptions). Preprocessing will involve **data cleaning, normalization, and NLP-based parsing** to standardize informal or incomplete entries, such as mapping “computer fixing” to “IT Support Technician.” This approach ensures the system can effectively align youth capabilities with labor market demands while also highlighting training pathways to close identified skill gaps.

**7. Literature Review WorkWise:**

Youth unemployment remains a global crisis. The 2022 Global Employment Trends for Youth Report noted that the Covid pandemic led to a 34 million decline in global youth employment between 2019 and 2020, with an additional 7 million entering the potential labor force but unable to find jobs. By 2021, the global youth unemployment rate stood at 15.6%, over three times higher than adults, leaving 75 million young people unemployed and 732 million out of the labor force (International Labour Organization [ILO], 2022). Crucially, unemployment rates were similar for youth in and out of education, underscoring systemic barriers across all groups.

In low- and middle-income countries, including Liberia, recovery has lagged behind wealthier nations. For Liberia, youth unemployment is not just an economic statistic but a national crisis that threatens social stability, undermines growth, and wastes the potential of it youthful population. Research also shows that low levels of happiness and health among youth can predict poor future outcomes (Blanchflower, 2009). A visible symptom in Liberia is the large number of graduates from universities and vocational institutes forced into survivalist work such as motorcycle transport or small-scale trading—due to the lack of meaningful employment opportunities. This reflects a disconnect between education systems and labor market needs, leaving thousands of skilled graduates without pathways to employment.

Current solutions, such as generic job boards, are insufficient: they are poorly maintained, managed and lack relevance for Liberia’s labor market, and fail to address skills development gaps. As a result, youth face prolonged job searches, employers struggle to find qualified candidates, and cycles of poverty goes on and on.

This research proposes WorkWise, an AI-powered job matching platform designed to bridge the gap between talent and opportunity by:

* Reducing search frictions in Liberia’s job market.
* Providing data-driven matches between youth skills and employer needs.
* Recommending upskilling opportunities to close labor-market gaps.

The central questions—How can technology help reduce job search frictions in Liberia?

To what extent can data-driven systems improve youth employment outcomes? —are critical for shaping policies, aligning education with labor market demands, and unlocking Liberia’s demographic dividend.

A review of existing literature on AI-based employment solutions ensures that this research builds on proven foundations while tailoring innovations to Liberia’s unique context.

**Implementation Plan**

**1. Technology Stack**

**Modern, Open-Source, & Cloud-Native**

|  |  |  |
| --- | --- | --- |
| Layer | Technology | Justification |
| **Frontend** | - Mobile App / Website: HTML, CSS, Bootstrap, React - Web Portal: HTML, CSS, Bootstrap, React | Component – based, large ecosystem performant |
| |  |  | | --- | --- | | Backend/API |  | | |  |  | | --- | --- | | API Gateway: Node.js, Python - Auth Service: JWT, OAuth - Notification Service: SMS/Email |  | | |  | | --- | | FastAPI for excellent async support & auto-docs. Node.js for real-time features. | |  | |
| |  |  | | --- | --- | | AI/ML Core |  | | |  |  | | --- | --- | | - NLP Pipeline: spaCy, Transformers - Recommendation Engine: LightFM - Feature Store: Hopsworks, Feast |  | | |  | | --- | | LightFM is perfect for hybrid recommendation. spaCy is fast and accurate for NLP. | |  | |
| |  |  | | --- | --- | | Data Storage |  | | |  |  | | --- | --- | | - Vector Database: Qdrant, Pinecone - Relational Database: PostgreSQL with PostGIS - Object Storage: S3-compatible (MinIO) |  | | |  | | --- | | PG for relational data, Qdrant for vector search, MinIO for cost-effective file storage. | |  | |
| **Comms** | |  |  | | --- | --- | | Twilio API (SMS), Socket.io (Real-time) |  | | |  | | --- | | Twilio is the gold standard for programmable SMS. | |  | |
| **Monitoring** | |  | | --- | | Evidently AI, Great Expectations | | **Critical:** Monitor for data and model drift from the very beginning. |

**2. Timeline**

**A 5-Week Agile Execution Plan**

* This timeline is aggressive but achievable by focusing on a **Learning MVP** (Minimum Viable Product), not a perfect one.

**Theme for Week 1-2: Foundation & Data Intelligence**

1. **Goal:** Build the data pipelines and prove the AI can "understand" Liberian CVs and jobs.
2. **Key Activities:**
   * Finalize the **Liberian Skill Ontology** (a curated list of skills mapped to common local phrases).
   * Develop the **NLP Pipeline**: Ingest sample CVs (PDFs/text) and job posts, extract skills and entities using spaCy, and generate embeddings.
   * Set up **PostgreSQL** schemas for users, jobs, and applications.
   * **Success Metric:** The system can process a CV and output a structured list of normalized skills.

**Theme for Week 3: The Matching Brain**

1. **Goal:** Build and train the core recommendation engine.
2. **Key Activities:**

* Develop the **Hybrid Recommender** using LightFM. Use content features (extracted skills) and prepare for collaborative filtering (will be sparse initially).
* Implement **Semantic Search** using the vector database for "find jobs like this" functionality.
* Build the **Skill Gap Analysis** module (compare youth skills to job skills, recommend missing ones).
* **Success Metric:** Given a youth and a job, the engine can generate a match score and list missing skills.

**Theme for Week 4: Integrate & Learn**

1. **Goal:** Connect the AI brain to a simple interface and start the feedback loop.
2. **Key Activities:**

* Develop a **simple web interface** (using - Mobile App / Website: HTML, CSS, Bootstrap, React  
  - Web Portal: HTML, CSS, Bootstrap, React) for youth to upload CVs and see matches.
* Build core **APIs** (e.g., /upload\_cv, /get\_matches, /get\_skill\_gap).
* **Run a mini-pilot** with 10-20 known youth and real job posts.
* **Manually validate** matches and collect feedback.
* **Success Metric:** A youth can upload their CV and receive a list of relevant job matches with explanations.

**Theme for Week 5: Scale & Strategize**

1. **Goal:** Document, refine, and plan for the future.
2. **Key Activities:**
   * **Refine models** based on pilot feedback.
   * **Document** the entire architecture, code, and lessons learned.
   * Develop a **Sustainability & Scaling Roadmap**: Outline Phase 2 (employer portal, SMS integration, advanced MLOps).
   * Prepare a **demonstration** for stakeholders.
   * **Success Metric:** A working prototype, a performance report, and a convincing roadmap for future investment and development.

**3. Milestones**

The following milestones will guide the development and implementation of the **WorkWise Platform** over a **4–5 week pilot timeline**. Each milestone represents a critical step in moving from concept to a functional AI-assisted job matching system.

**Milestone 1 — Week 1: Project Setup & Data Preparation**

* Finalize project scope, objectives, and measurable success criteria.
* Collect initial datasets (youth CVs, job postings, training program data).
* Preprocess and clean data (NLP-ready CVs, standardized job requirements).

**Milestone 2 — Week 2: Model Development (NLP + Skills Extraction)**

* Implement **NLP pipeline** for parsing CVs and job descriptions.
* Develop **skills ontology** to normalize informal terms (e.g., “computer fixing” → “IT Support”).
* Train initial ML models for **job–candidate matching**.

**Milestone 3 — Week 3: Recommendation Engine & Prototype**

* Build **hybrid recommendation system** (content-based + collaborative filtering).
* Integrate **training recommendations** for skill-gap closure.
* Connect backend models to a lightweight **prototype interface** (web/mobile)

.

**Milestone 4 — Week 4: Testing, Feedback & Refinement**

* Conduct **pilot testing** with a small group of youth and employers.
* Evaluate accuracy of matches, system usability, and user satisfaction.
* Refine algorithms and features based on **real-world feedback loops**.

**Milestone 5 Week 5: Final Deployment & Reporting**

* Deploy a **stable prototype version**.
* Document **lessons learned**, performance metrics, and user adoption insights.
* Develop a **sustainability and scaling roadmap** (government, NGO, and private sector partnerships).

**4. Challenges and Mitigation Strategies**

1. **Limited Timeframe (4–5 weeks)**
   * Challenge: Building NLP + recommender within a month is ambitious.
   * **Mitigation**: Focus on a **Minimum Viable Product (MVP)**: basic CV parsing, job matching, and training recommendations. Advanced features (bias detection, USSD/SMS integration) deferred to Phase 2.
2. **Data Quality & Availability**
   * Challenge: Youth CVs and job postings may be incomplete or inconsistent.
   * Mitigation: Begin with a **small curated dataset**; apply manual validation and synthetic data generation for testing.
3. **Technical Complexity**
   * Challenge: NLP + recommendation integration is technically intensive.
   * Mitigation: Use **pre-trained NLP models (spaCy, Hugging Face)** and open-source recommendation libraries (**LightFM, Surprise**) to accelerate development.
4. **Connectivity & Access Constraints**
   * Challenge: Many Liberian youth face **low internet access**.
   * Mitigation: For the pilot, focus on **mobile-first lightweight web apps**. Plan SMS/USSD for Phase 2 expansion.
5. **User Adoption & Trust**
   * Challenge: Employers and jobseekers may hesitate to use a new system.
   * Mitigation: Engage **trusted early adopters** (anchor employers + youth groups). Provide training and highlight **explainability of matches**.
6. **Sustainability Beyond Pilot**
   * Challenge: Risk of platform being abandoned post-pilot.
   * Mitigation: Capture lessons learned, demonstrate measurable impact, and build a **partnership roadmap** for scaling.

**5. Ethical Considerations for WorkWise**

**1. Data Privacy and Protection**

WorkWise will handle sensitive personal data such as CVs, employment history, and skill profiles. Ethical responsibility demands:

* Compliance with **data protection standards** (encryption at rest and in transit).
* **Minimal data collection** (only what is necessary for matching).
* **Anonymization** for analytics to protect individual identities.
* Clear **informed consent** mechanisms so youth understand how their data will be used, shared, and stored.

**2. Bias and Fairness in AI**

AI-driven systems risk amplifying existing inequalities (gender, education, location, or socioeconomic status). To address this:

* Conduct **regular fairness audits** on model outputs.
* Apply **bias mitigation techniques** (balanced training datasets, fairness-aware algorithms).
* Ensure subgroup monitoring (urban vs rural, male vs female, educated vs vocational) to guarantee **equitable opportunities**.

**3. Transparency and Explainability**

Trust depends on clarity.

* Users should **understand recommendations** through **explainable AI (XAI)**: e.g., “You were recommended because your Excel skills matched employer needs.”
* Employers should see **why candidates are matched**, avoiding black-box decisions.
* This transparency increases adoption, accountability, and trust.

**4. Accessibility and the Digital Divide**

Liberia faces stark **urban–rural internet disparities**. To ensure inclusivity:

* Prioritize **mobile-first lightweight design**.
* Provide **SMS/USSD access** for low-connectivity areas.
* Deploy **community kiosks or NGO partners** for assisted access in rural areas.

**5. Informed Consent and Autonomy**

* Users opt in voluntarily and have **control over their data**.
* Profiles can be updated, deactivated, or deleted at any time.
* Employers must follow ethical recruitment standards (no exploitative practices).

**6. Accountability and Human Oversight**

* If the AI produces harmful or misleading recommendations, **the development team bears responsibility** for correction.
* **Human oversight** complements AI outputs, ensuring critical hiring decisions are not left entirely to algorithms.
* Establish a **grievance mechanism** for users to report errors or abuses.

**7. Sustainability and Long-Term Impact**

* A short pilot must not create **false hope**. WorkWise commits to building **continuity pathways** by partnering with government, NGOs, and private sector.
* Ethical responsibility extends beyond job matching — the system must **deliver long-term employability and trust**.

**8. Labor Rights and Protection from Exploitation**

* Only employers who meet **basic labor standards** should be onboarded.
* Job postings must be **verified** to prevent scams or unsafe work.
* The platform includes **reporting tools** so jobseekers can flag exploitative practices.

**9. Equity in Skills Development Recommendations**

* Training recommendations should not bias toward expensive, inaccessible programs.
* WorkWise will **prioritize free, subsidized, or NGO-sponsored training** to ensure affordability.
* Transparent partnerships with training providers will guarantee **fair access for low-income youth**.

**6. References**

1. International Labour Organization. (2022). Global Employment Trends for Youth 2022: Investing in transforming futures for young people. ILO. https://www.ilo.org
2. World Bank. (2021). AI for Job Matching in Developing Economies: Pilot results from Kenya. World Bank Group. [https://www.worldbank.org](https://www.worldbank.org?utm_source=chatgpt.com)
3. World Bank. (2023). Leveraging Machine Learning for Job Matching in Public Employment Services: Evidence from Poland. World Bank Group.
4. Ajira Digital Program. (2022). Annual Impact Report. Ministry of ICT, Innovation and Youth Affairs, Government of Kenya. https://ajiradigital.go.ke
5. Babajob. (2016). Connecting India’s Informal Workers to Better Jobs. Babajob.com (archived).
6. IEEE. (2022). Skill-based Employment Matching Using Machine Learning. Proceedings of the IEEE International Conference on Data Science and Advanced Analytics (DSAA).
7. Babcock, L., & Engler, M. (2021). Reducing labor market frictions with AI-powered platforms: Evidence from emerging markets. Journal of Development Economics, 152, 102697.
8. United Nations. (2015). Transforming our world: The 2030 Agenda for Sustainable Development. United Nations. [https://sdgs.un.org](https://sdgs.un.org?utm_source=chatgpt.com)