# Explicit-Implicit Skill Extraction

**Presented by: Yonatan Elman, Michael Kovalchuk, Roni Fadlon**






# Motivating Use Case

**Background-** Resumes are free text; skill information is scattered, explicit and implicit.

**Importance-** Accurate skill mapping is key for reliable candidate matching.

**Challenge-** Implicit skills don't appear by name and vary in phrasing.

**Today-** Manual review and keyword-based tools miss many implicit skills.

# Project task description

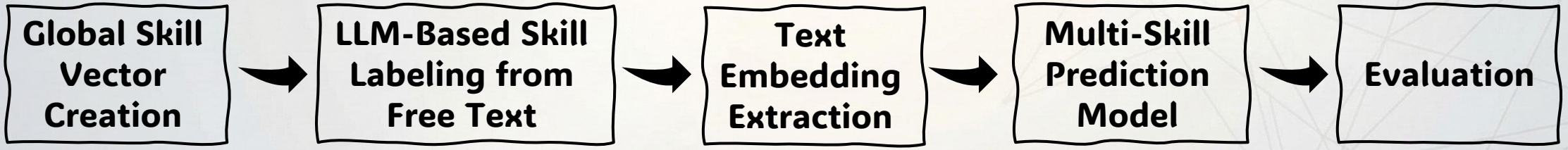
Given a resume job-role text chunk, classify each skill in the global skill vector as None/Implicit/ Explicit.

Input: Free-text job-role description.

- ✓ Identifies implicit skills, not only explicit mentions.
- ✓ Uses a continuous 0–0.5–1 evidence scale.

Output: Evidence vector (0 / 0.5 / 1) aligned with the global skill vector.

# Models and methods




**Core Models:**

- Main Model: Fine-tuned Transformer (BERT / RoBERTa)
- Baseline 1: Keyword Matching
- Baseline 2: Zero-Shot LLM
- Refinement: Hyperparameter Tuning

**Adjustments & Improvements:**

- Fine-tuning on synthetic labeled data
- Handling rare skills (class balancing)

# Data specification and generatio

Each example: a free-text job-role description & a label vector over the global skill list (0 / 0.5 / 1), with synthetic Train / Val / Test splits.

An LLM assigns each skill one of: Explicit / Implicit / None → 1 / 0.5 / 0.

A large synthetic dataset of LLM-generated job-role descriptions aligned with the global skill vector.

Sample skills + evidence levels, prompt an LLM to generate a matching job-role description and store each example as:(job-role text, skill-label vector).

# Metrics and KPIs

**Result Measurement**
Compare model predictions (0 / 0.5 / 1) to synthetic LLM-generated ground-truth labels.

**Model Quality**
Evaluate on validation set using **Accuracy** and **Macro-F1**, with special focus on Implicit skills.

**Metrics & Protocol**
Precision / Recall / F1 per skill + Confusion Matrix for None / Implicit / Explicit.