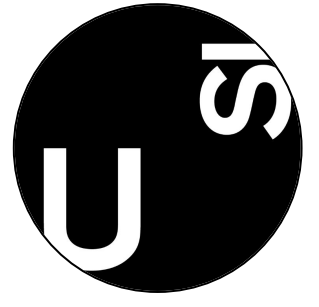




Paper

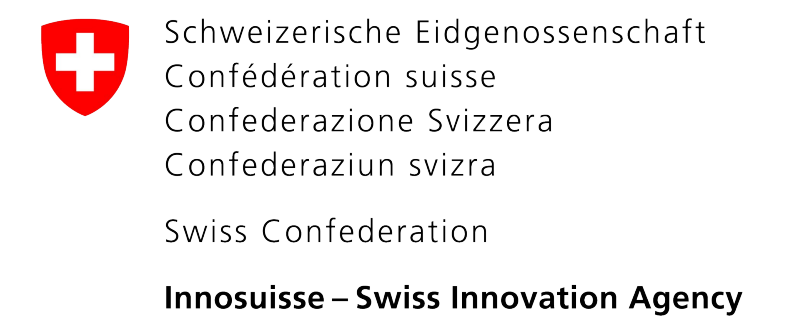


Github



Laura Vásquez-Rodríguez¹, Bertrand Audrin², Samuel Michel¹, Samuele Galli³,
Julneth Rogenhofer², Jacopo Negro Cusa³, Lonneke van der Plas^{4,1}
(1) Idiap Research Institute, CH, (2) EHL Hospitality Business School, HES-SO, University
of Applied Sciences and Arts Western Switzerland, CH, (3) Arca24.com SA, CH,
(4) Università della Svizzera italiana, CH

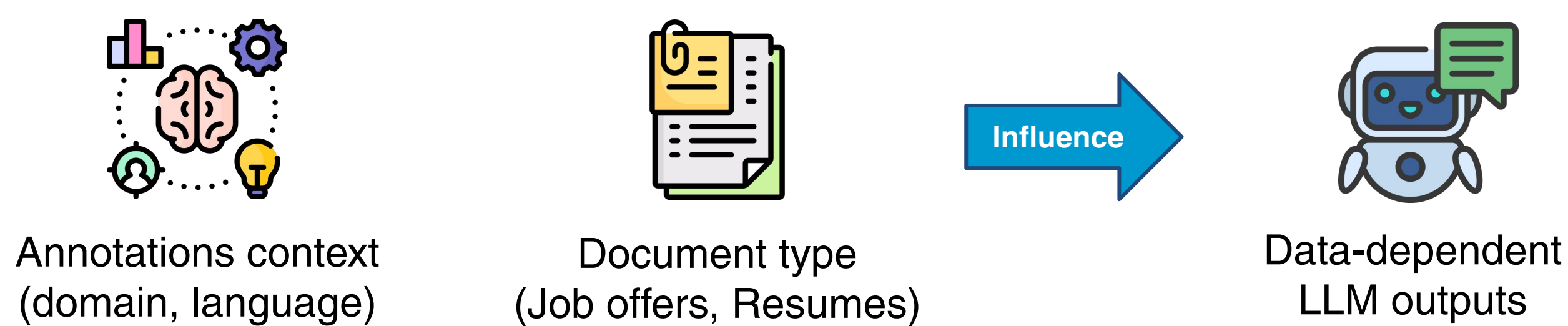
Innovation project
supported by



We annotated 800 datasets of **job offers** and **resumes** in four languages (**English**, **French**, **Italian**, and **Portuguese**) using **O*NET**- and **HEXACO**-based guidelines, resulting in **21** soft skill categories. These annotated datasets were then used to train **BERT**-based **multilingual** models for soft **skill extraction**.

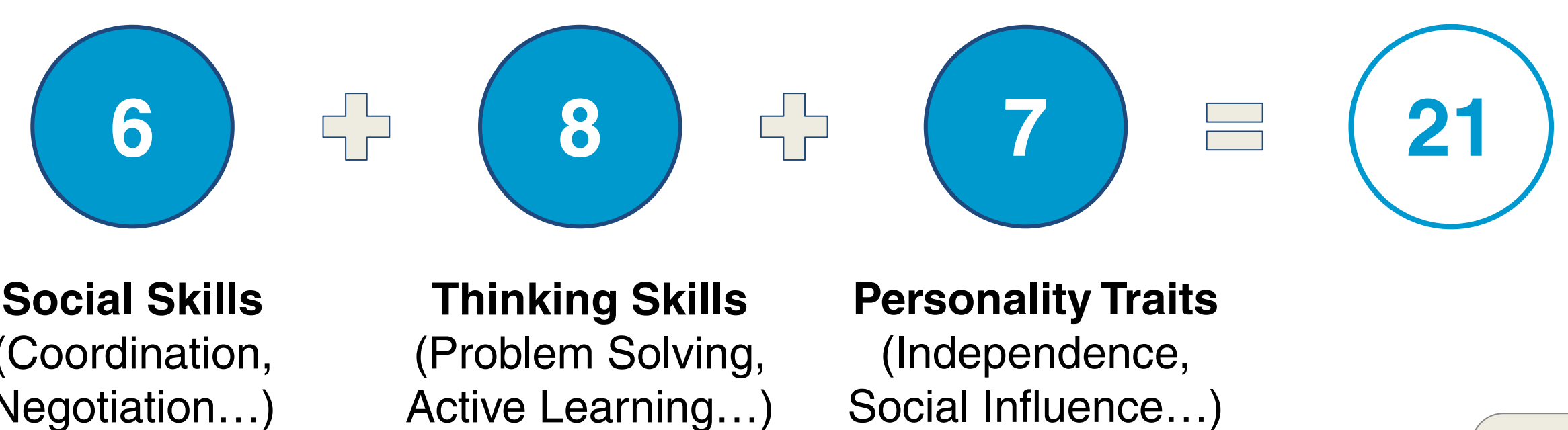
Summary

- **Applicant Tracking Systems** (ATS) often focused on hard skills, neglecting soft skills which are essential in candidate selection.
- Soft skills: **behavioral** and **social abilities** that people tend to develop through social interactions. These are challenging to identify and often **overlooked**.
- We **compare** soft skills, hard skills and occupations in shape and meaning highlighting their sources of **variability**.



Annotation Framework

- HR researchers created **annotation guidelines** based on O*NET taxonomy and HEXACO personality inventory.
- Soft skills were identified in **jobs** and **resumes** using the following criteria:



Contribution #3:
O*NET/HEXACO
based annotation
framework

Soft Skill Extraction in the Wild

- **10,000+** annotations (in-house and state-of-the-art) for soft skills
- Supervised system for token **classification** using a span-based approach.
 - independent (concept) → being action oriented (phrase, span)
- Use of **BERT**-based **multilingual** model (uncased) and fine-tuned it using our manually annotated skill datasets and state-of-the-art labels.
- Comparative evaluation of **soft** skills (left) vs **hard** skills (right)

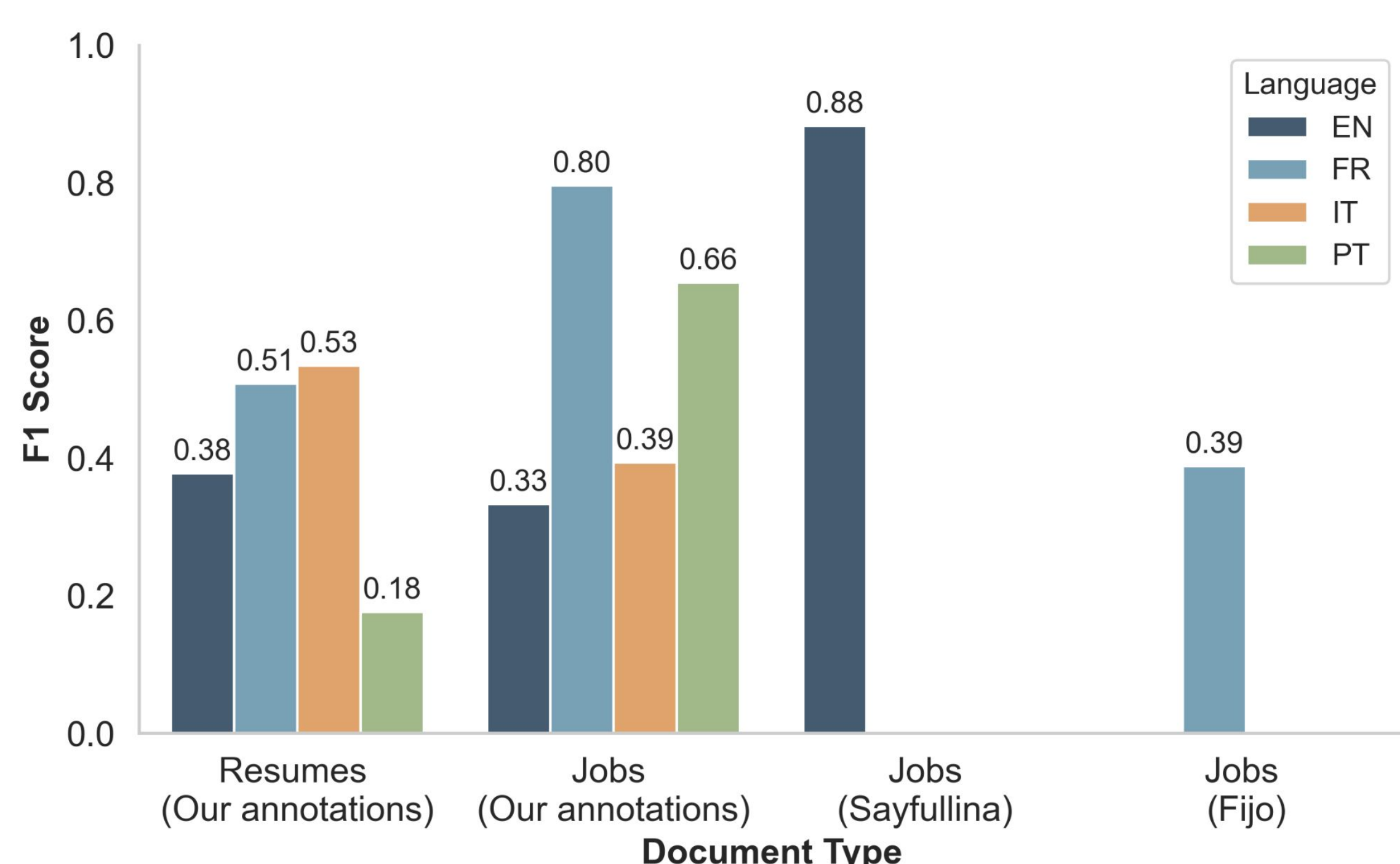
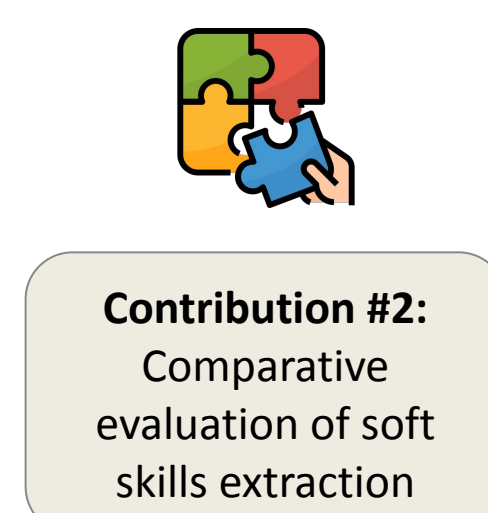


Table 1. Soft Skills



Contribution #2:
Comparative
evaluation of soft
skills extraction

Annotation and Detection Examples

- Annotation (Resume, Ours)**
- **curious** and **thoughtful** person with good inventive
 - **curious** and **thoughtful person** with **good inventive**
- Annotation (Jobs, Ours)**
- follow verbal and written instructions **be able to work quickly and concisely under pressure**
 - follow **verbal** and **written instructions** be able to **work quickly** and **concisely** under pressure
- Annotation (Jobs, Sayfullina):**
- be a national provider of **independent** sector complex healthcare
 - be a national provider of **independent** sector complex healthcare
(not a skill)

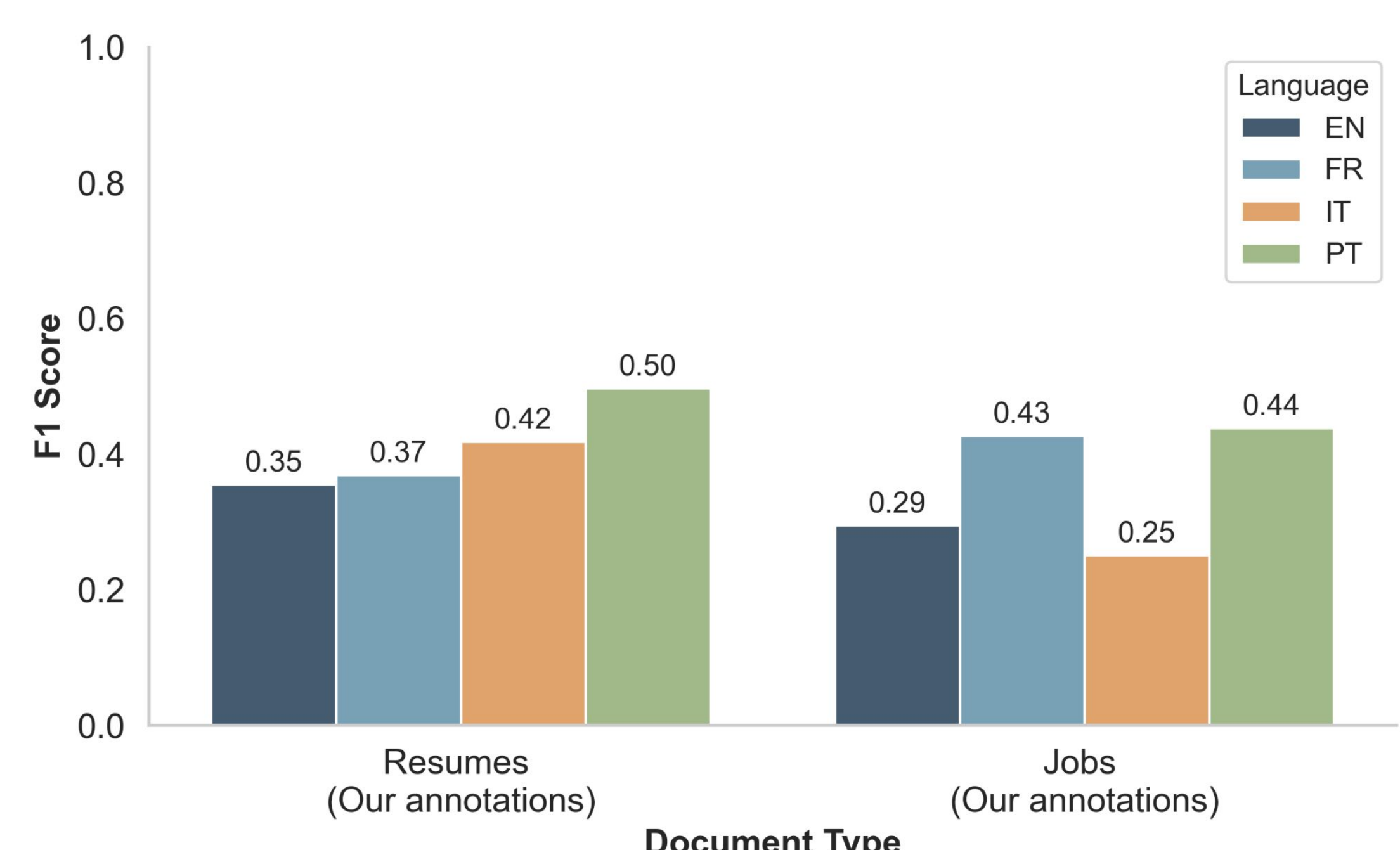


Table 2. Hard Skills

Semantic Analysis

- We analyze the semantic similarity between soft skills in different contexts using **SBERT** embeddings and the **t-SNE** algorithm.
- **What does it mean?** Similar data points are **grouped** in similar locations, clusters show **entities** that are potentially **equivalent** within the sample of documents.

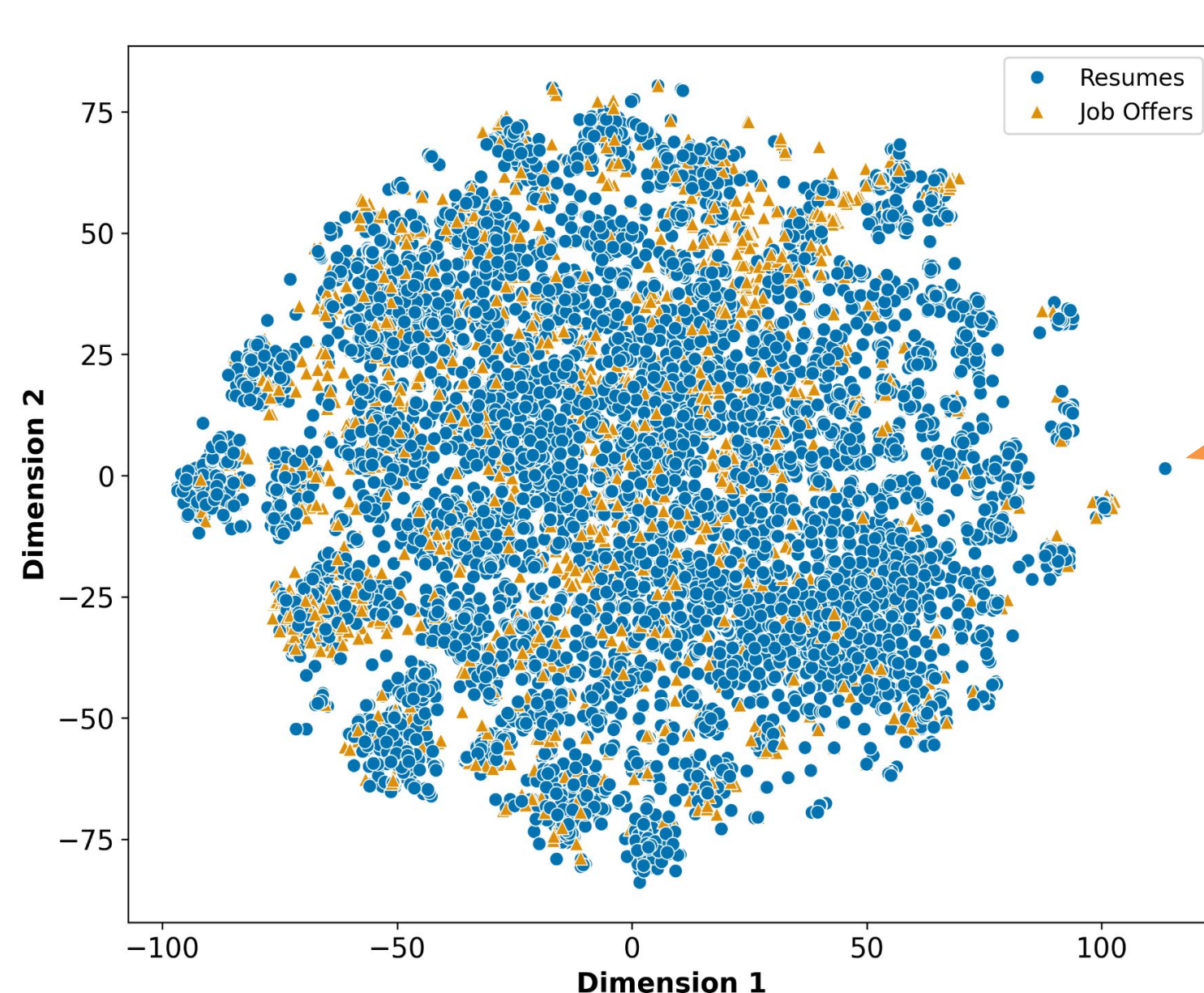


Figure 1. T-SNE Hard Skills

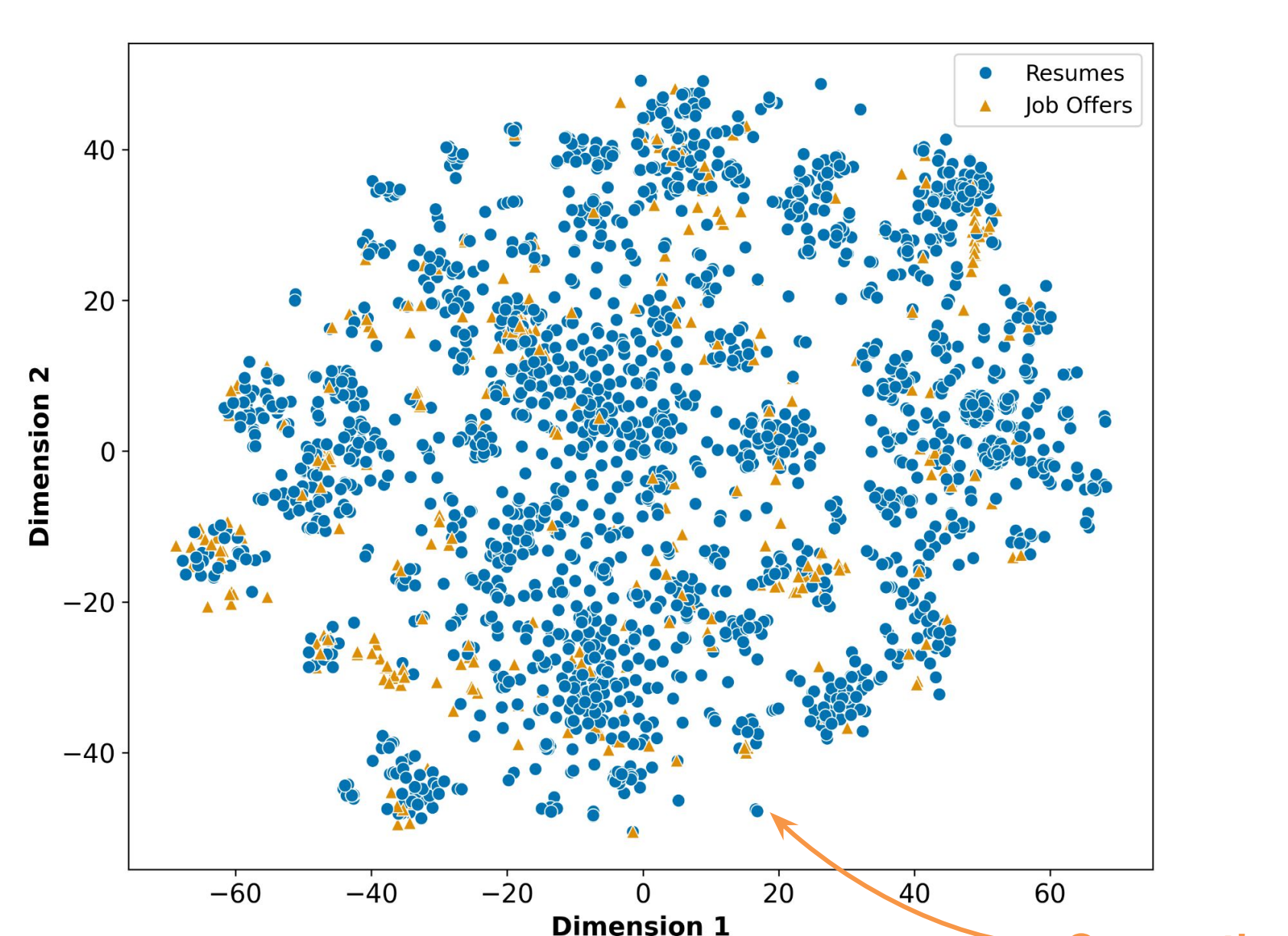


Figure 2. T-SNE Occupations Skills

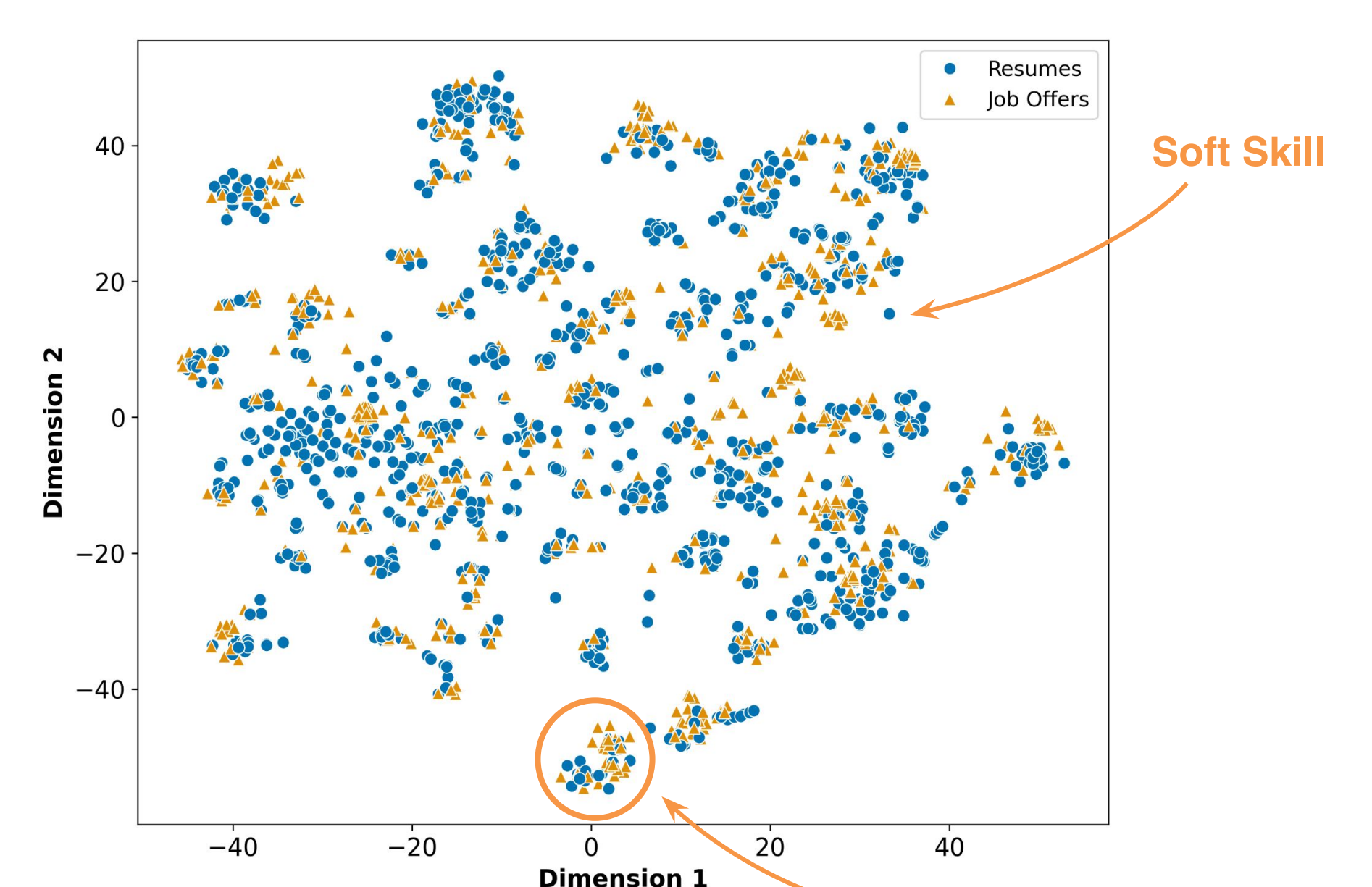


Figure 3. T-SNE Soft Skills



Contribution #1:
Variability
Analysis