المؤتمر الدولي الأول التعليم الإلكتروني
وتكنولوجيا المعلومات
The First International Conference on E-learning and Information Technology
19-20 نوفمبر2025

# AI-Driven University Quality Assessment across Libya Using Digital Data Sources

[1]Mustafa Alkhunni,[2] Kameelah Alnmri, [3]Asmaeil Balg
[1]Training & Development Department, Mellitah Oil & Gas B.V., Tripoli, Libya
[2] Computer Department, Faculty of Education, Al-Ajelat, University of Zawia, Al-Ajelat, Libya
[3]Computer Department, Faculty of Education, Al-Ajelat, University of Zawia, ,Al-Ajelat, Libya
Emails: [1]mustafa.alkhunni@mellitahog.ly; [2]k.alnmri@zu.edu.ly; [3]a.belg@zu.edu.ly

**Abstract**:
this paper is prepared to build an end-to-end unsupervised analyzer using Python. This reproducible model on hand is developed based on a human-centered, data-driven assessment of Libyan universities that combines structural indicators as a data sample (student counts, number of faculties, teaching staff, research, and quality scores) with a real-time sentiment signal derived from public online sources in Arabic and English. After cleaning and standardizing all features, the institution-level sentiment is estimated and join it with structural metrics. An unsupervised clustering pipeline using (k-means algorithm) reveals three distinct institutional profiles. Enhanced visualizations (Figs. 8–10) make the findings accessible to decision-makers [see appendix I. Results show substantial variation in both structural performance and public perception and illustrate how perception data complements traditional indicators. The approach offers a scalable framework that can be continuously updated for monitoring higher-education quality in Libya [See the Appendix I for the entire code].

**Index Terms**: sentiment analysis, clustering, visualization, higher education, Libya, reproducibility

**الملخص:**

تهدف هذه الدراسة إلى تطوير نموذج تحليلي غير خاضع للإشراف باستخدام لغة بايثون، يعتمد على نهج إنساني متمحور حول البيانات لتقييم الجامعات الليبية. يجمع النموذج بين مؤشرات هيكلية كمجموعة بيانات (أعداد الطلاب، عدد الكليات، أعضاء هيئة التدريس، البحث العلمي، ومؤشرات الجودة) وبين إشارات آنية للمشاعر العامة المستخلصة من المصادر الإلكترونية المفتوحة باللغتين العربية والإنجليزية. بعد إجراء عمليات التنظيف والتوحيد لجميع السمات، يتم تقدير مستوى المشاعر على مستوى المؤسسة وربطه بالمقاييس الهيكلية. يكشف خط التجميع غير الخاضع للإشراف باستخدام خوارزمية K–meansعن ثلاثة أنماط مميزة للمؤسسات. وتتيح التصورات البيانية المحسّنة (الأشكال 8–10) عرض النتائج بشكل يسهل على صانعي القرار استيعابها [انظر الملحق الأول]. أظهرت النتائج وجود تباين كبير في الأداء الهيكلي والإدراك العام،

مما يوضح كيف يمكن لبيانات الإدراك أن تكمل المؤشرات التقليدية. يقدم هذا النهج إطارًا قابلاً للتوسع يمكن تحديثه بشكل مستمر، لمراقبة جودة التعليم العالي في ليبيا [انظر الملحق الأول للاطلاع على البرمجة الكاملة للموديل ]

**الكلمات المفتاحية:** تحليل المشاعر، التجميع، التصور البياني، التعليم العالي، ليبيا، قابلية إعادة الإنتاج

## I. Introduction

Quality assurance in higher education has traditionally relied on periodic peer review and rubric-based evaluations. Global perspectives underscore the need for continuous, evidence-based quality monitoring [1]. The availability of digital traces—student feedback, social media, and web metrics—now enables near real-time analysis. In Libya, the university system has faced sustained stress and structural challenges [2]. Rankings and public metrics also shape institutional behavior and reputation, influencing policy and stakeholder expectations [3]. This paper contributes the following: (i) a curated multi-source indicator set and perception signals for Libyan universities; (ii) an unsupervised pipeline that groups institutions by shared characteristics; and (iii) visual summaries that support interpretation and policy dialogue.

## II. Data and Setting

We compile publicly available indicators: faculty and student counts, proxies for research activity, and web visibility. For the University of Zawia, we updated the student count to 32,147; public ranking context places Zawia approximately 9,743 out of 24,511 worldwide and 419 out of 1,534 in Africa. Numerical features are standardized (z-score) before clustering. In terms of sentiment, public posts that mention Libyan universities are collected from social platforms and forums in Arabic and English. Messages are de-duplicated, language-detected, and normalized (diacritics and elongation removal; URL, hashtag, and emoji handling). Each post receives a polarity label {positive, neutral, negative}[11] and an optional continuous sentiment score in the range $[-1, 1]$. Institution-level indicators—class shares, mean sentiment, and weekly volume were aggregated so that perception signals can be compared alongside structural metrics. In parallel, widely used ranking methodologies provide comparable, public indicators that complement our structural dataset [5]– [8].

## III. Methodology

### III-A. Pipeline Overview

The workflow (Fig. 1) comprises five stages: data collection, feature preparation, sentiment estimation, unsupervised clustering with validation, and visualization and interpretation. All analyses operate on standardized features to ensure comparability across scales.


**Figure 1.** End-to-end pipeline for the proposed analyzer

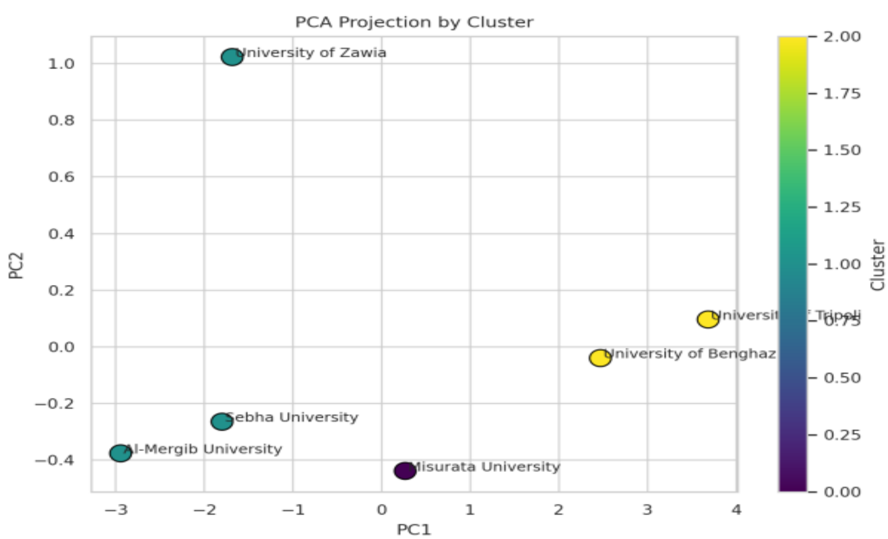


**Figure 2.** PCA Projection by cluster

### III-B. Data Preparation

Cleaning and alignment: structural sources are de-duplicated and aligned on canonical institution names (e.g., "University of Benghazi").

Type handling: numeric fields (Student Count, Teaching_ Staff_ Count, Research Score, quality Score) are coerced to numeric; rows with essential missing values are removed [appendix I]-[appendix II].

Standardization: to remove scale effects, numeric features are z-scored: $x' = (x - \mu) / \sigma$.

### III-C. Feature Engineering and Sentiment Estimation

Structural features: Student Count, Number of Faculties, Teaching Staff Count, Research Score, and Quality Score [See the appendix I] and [appendix II].

Sentiment feature (real-time): public posts are collected from selected platforms and forums (Arabic/English).

Pre-processing: language detection; Arabic normalization (diacritics removal, elongation squashing); URL/emoji/hashtag handling; de-duplication and bot/spam filtering.

Modeling: a polarity classifier (supervised transformer or lexicon fallback) yields a label $\in$ {positive, neutral, negative} and an optional score $\in [-1, 1]$ [11].

Aggregation: for each university and period (daily/weekly), we compute class shares (% positive/neutral/negative) and a mean sentiment score. The main modeling feature is Sentiment Score (or % positive if scores are unavailable) [11].

### III-D. Clustering and Model Selection

Algorithm: we adopt k-means for its simplicity and transparency in stakeholder discussions; models run on the standardized feature matrix.

Choosing k: we examine a small range of k (e.g., 2–6) and select the operating point using the elbow criterion (inertia vs. k) and the silhouette score to balance cohesion and separation.

### III-E. Validation and Robustness Checks

Internal metrics: we report the silhouette score as the primary indicator (higher is better). Where space allows, also the Calinski – Harabasz and Davies–Bouldin standards used to compute indices [9]-[10].

### III-F. Visualization and Interpretation

Dimensionality reduction: PCA provides an interpretable 2-D projection; t-SNE is used sparingly for local structure. These are for illustration (Figs. 2–3).

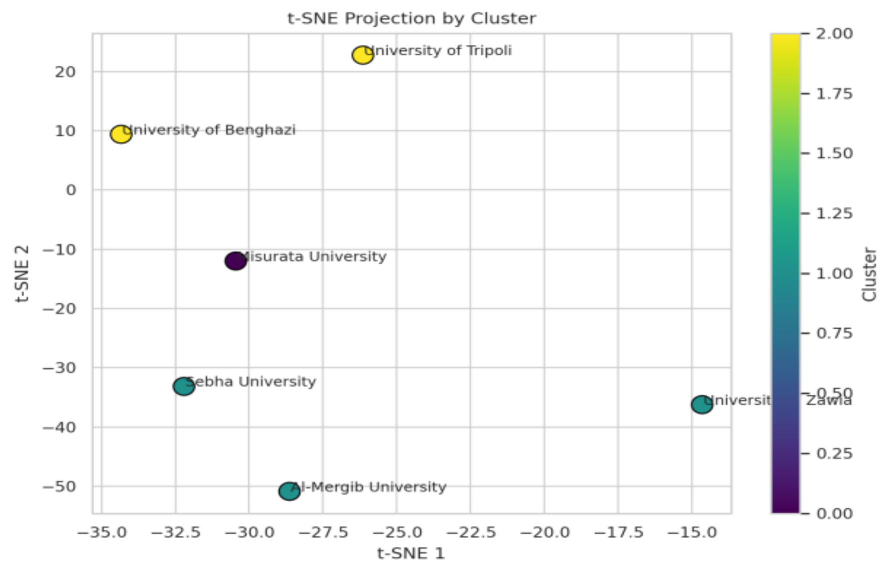


**Figure 3.** t-SNE projection (illustrative on small N)

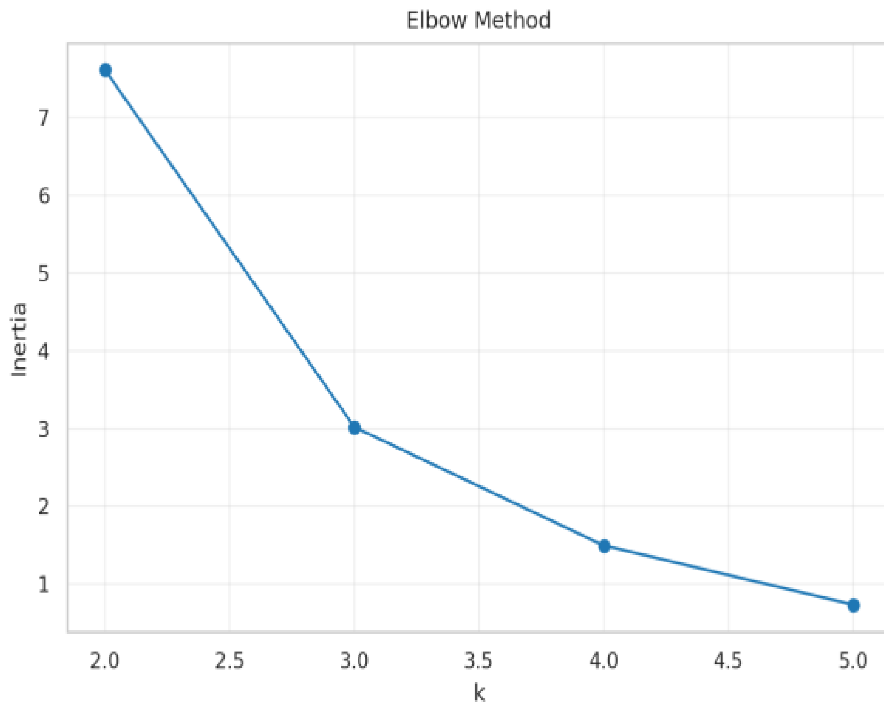


**Figure 4**. Elbow method: (inertia vs. k).

**III-G. Real-Time Sentiment Ingestion and Scoring (Operational Detail)**

Flow: streaming connector → rate-limited collection → language normalization → polarity classification → university-level aggregation.

**III-H. Ethics, Privacy, and Data Governance**

Public based content is used only which is permitted with no restriction related to data privacy, minimize personally identifiable information, honor content removal requests, and filter sensitive topics [12].

**III-I. Reproducibility**

We set explicit random seeds (e.g., random state = 42), document library versions, and provide a minimal script that regenerates figures.

## IV. Results

On the standardized feature set, k = 3 yields a silhouette score of 0.39. Figures 4–7 present model-selection diagnostics and feature profiles; Figures 8–10 provide dashboard-style visualizations for readability and stakeholder communication.

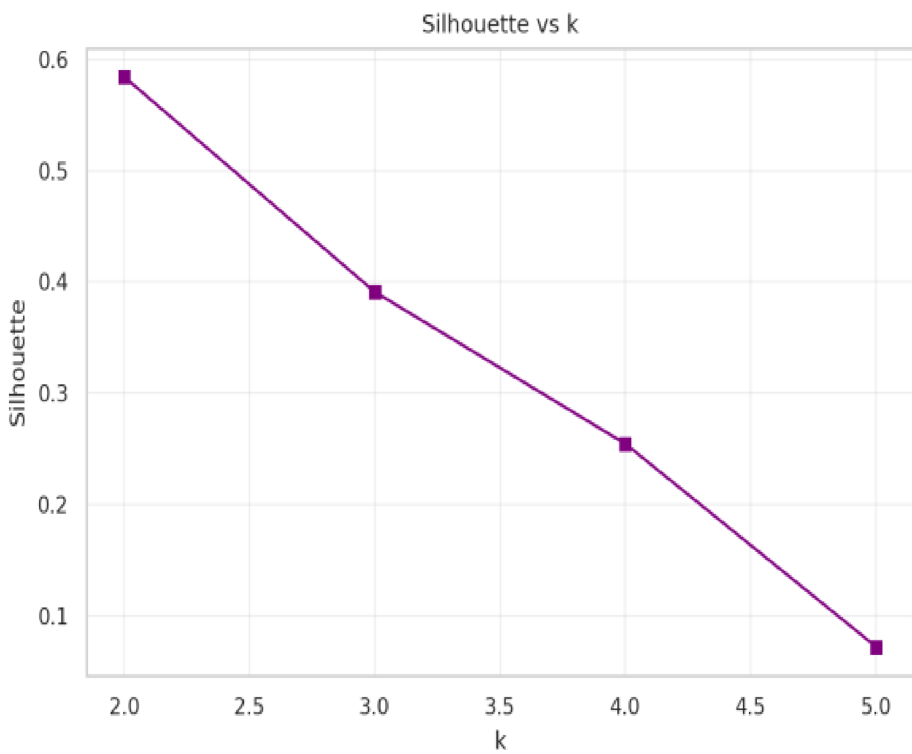


**Figure 5.** Silhouette score across k.

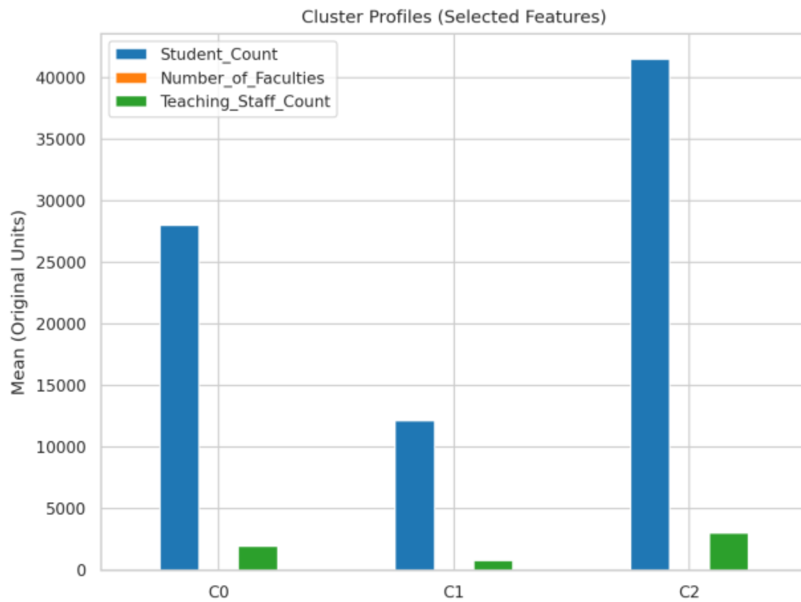


**Figure 6.** Cluster profiles (mean of selected features).

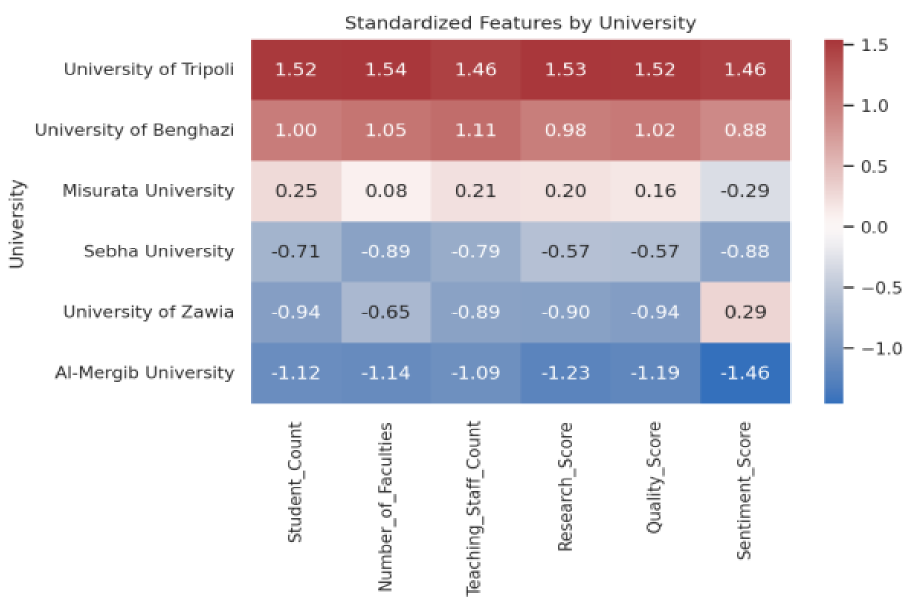


**Figure 7.** Standardized feature heat map by university.

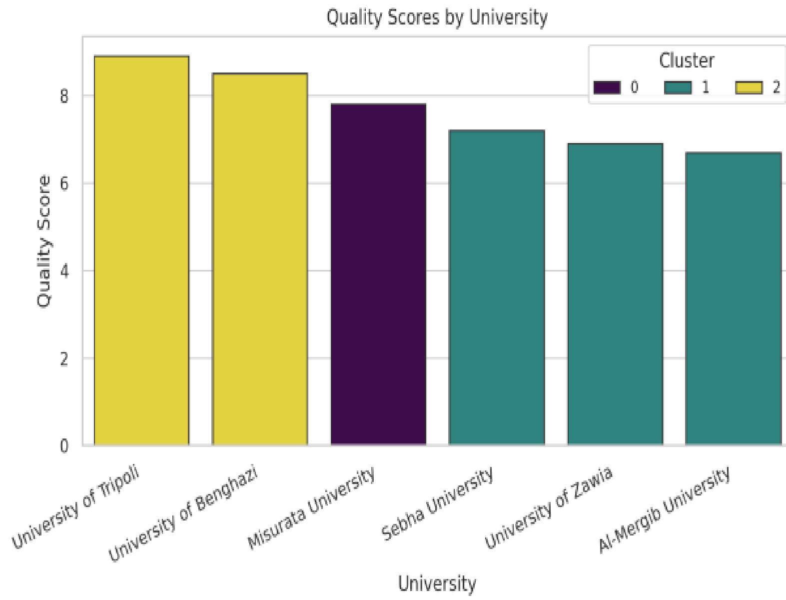


**Figure 8.** Quality score distribution by university (colored by cluster).

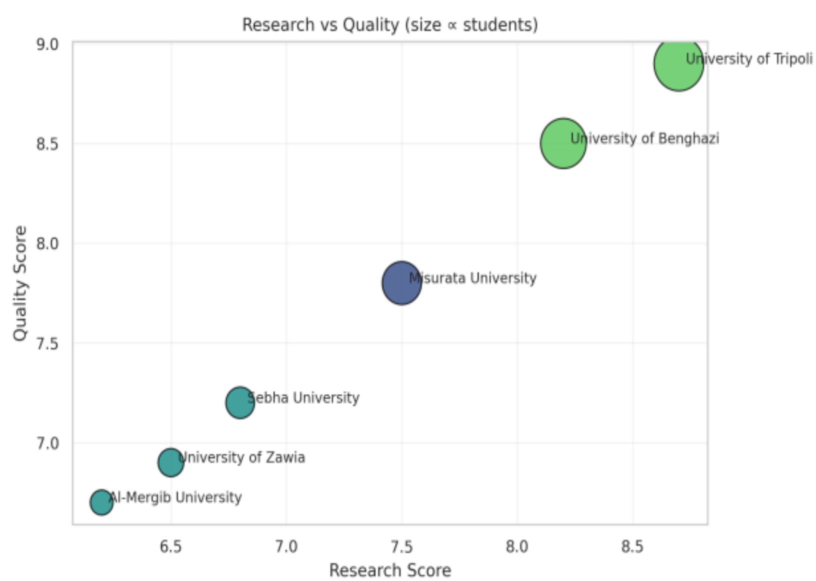


**Figure 9.** Research vs. quality with bubble size proportional to student count; color shows cluster.

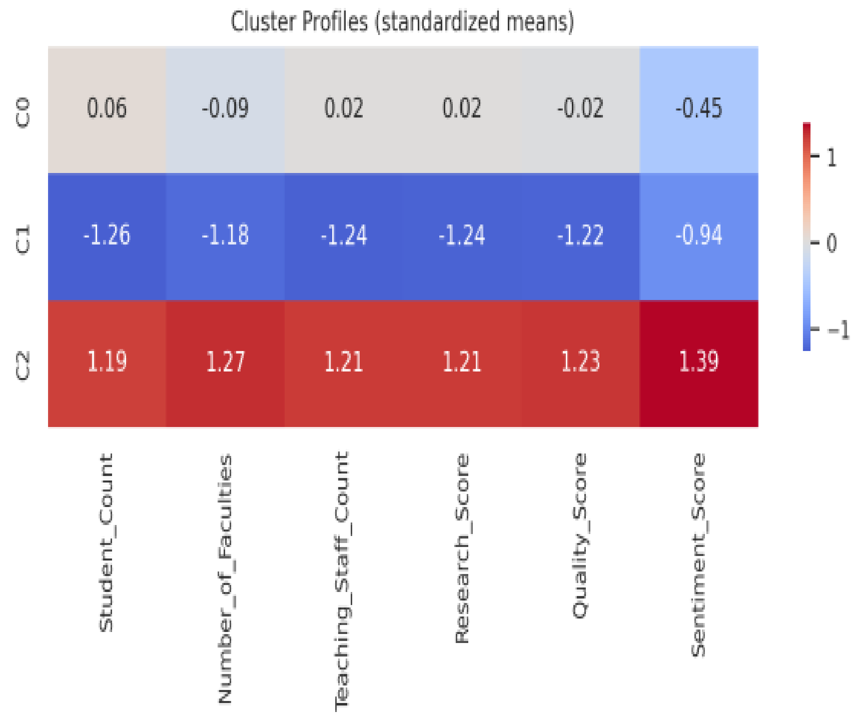


**Figure 10.** Cluster profile heat map (Z-scored feature means).

**Discussion and Limitations**

Profiles can inform differentiated strategies: invest in research capacity at comprehensive universities; strengthen teaching quality and community engagement at regional institutions; and provide foundational infrastructure at smaller universities. These considerations align with broader analyses of higher-education systems in fragile contexts [4] and documented stresses in Libyan higher education [2]. In addition, the reputational dynamics introduced by rankings should be interpreted with care when comparing institutions [3]. Limitations include the small illustrative dataset and reliance on public data; sentiment signals should be expanded and triangulated with administrative and bibliometric sources.

## VI. Conclusion and Future Work

AI-driven profiling of Libyan universities is feasible with publicly available signals and yields interpretable clusters that complement traditional evaluations. Next steps include expanding data coverage to include much more indicators such as student feedback and student insights, facilities, institutional accreditation and etc., adding sentiment visualizations per institution and deploying database web-based model in order to increase the functionality of an interactive data analysis. This developed profile can be upgraded to be real time tool for Libyan higher education assessment.

## VII. Acknowledgment

The authors thank the Training & Development Department at Mellitah Oil & Gas B.V. and the Faculty of Education at Al Ajaylat, University of Zawia, for their support and feedback. Any remaining errors are the authors' own.

## VIII. References

[1] P. G. Altbach, Global Perspectives on Higher Education. Johns Hopkins Univ. Press, 2016.
[2] H. T. S. Benamer and A. Bredan, "Higher education in Libya: System under stress," Libyan Journal of Medicine, vol. 9, no. 1, 2014.
[3] E. Hazelkorn, Rankings and the Reshaping of Higher Education: The Battle for World-Class Excellence. Palgrave Macmillan, 2015.
[4] World Bank, The Changing Nature of Work: Higher Education in Fragile Contexts. World Bank, 2020.
[5] uniRank, "World University Rankings Methodology," accessed Aug. 21, 2025.
[6] Webometrics, "Ranking Methodology," accessed Aug. 21, 2025.
[7] QS Arab Region Rankings, "Methodology," accessed Aug. 21, 2025.
[8] SCImago Institutions Rankings, "Methodology," accessed Aug. 21, 2025.
[9] D. L. Davies and D. W. Bouldin, "A cluster separation measure," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-1, no. 2, pp. 224–227, 1979.
[10] Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. Communications in Statistics, 3(1), 1–27.
[11] Liu, B. (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, 5(1), 1–167.
[12] IEEE Digital Privacy, "Ethical Issues Related to Data Privacy and Security: Why We Must Balance Ethical and Legal Requirements in the Connected World," 2024.
[13] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," Journal of Computational and Applied Mathematics, vol. 20, pp. 53–65, 1987.
[14] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed. Springer, 2009.