Knowledge Processing Studies. 2026, 18, 6(1): 12-30.

International Journal of Knowledge Processing Studies
(KPS)

Homepage: http://kps.artahub.ir/

ORIGINAL RESEARCH ARTICLE

Entrepreneurship Discourse in the Arabic Twitter Sphere: A Sentiment and Content Analysis

Mahboobeh Vahabi Abyaneh[1*], Amirmahdi Adib[2], Ali Mobini Dehkordi[3]

[1] *Assistant Professor, Faculty of Entrepreneurship, College of Management, University of Tehran. vahabi.m@ut.ac.ir, 0000-0002-2797-7669.*

[2] *Ph.D. Student, Faculty of Entrepreneurship, College of Management, University of Tehran. a.mahdi310@gmail.com, 0009-0005-0235-9089.*

[3] *Professor, Faculty of Entrepreneurship, College of Management, University of Tehran. mobini@ut.ac.ir, 0000-0003-0838-4945.*

ARTICLE INFO

Article History:
Received: 2025-12-05
Revised: 2026-01-12
Accepted: 2026-02-02
Published Online: 2026-03-01

Keywords:
Entrepreneurship, Sentiment Analysis, Arabic Twitter, Data Mining, Vision 2030.

Number of Reference:72
Number of Figures: 12
Number of Tables:1

DOI: ***10.22034/kps.2026.580547.1275***

ABSTRACT

This study investigates the structure and dynamics of entrepreneurial discourse on the Arabic Twitter (X) sphere within the Gulf Cooperation Council (GCC) states. It examines how entrepreneurship is socially constructed amidst the transition from a rentier to a knowledge-based economy, focusing on state narratives versus public sentiment. Adopting a computational social science approach, the study analyzed a corpus of 48,841 content units harvested between June and December 2025. To ensure statistical independence and prevent double-voting bias, the research employed a Confidence-Calibrated Ensemble architecture. This pipeline integrates fine-tuned models (MARBERT) with pseudo-independent Large Language Model configurations (GPT-4 in zero-shot and few-shot settings) using calibrated tie-breaking. Techniques included Sentiment Analysis, Named Entity Recognition (NER), and demographic profiling of 1,888 active users. Descriptive analysis of the collected corpus identifies Saudi Arabia as the primary discourse locus (accounting for approximately 47% of traffic). Semantic analysis reveals a significant discursive shift from economic to cultural themes, termed ‘Entrepreneurial Nationalism,’ highlighted by a substantial growth in references to state initiatives like ‘Vision 2030.’ Within this specific dataset, the ecosystem exhibits ‘fragile positivity’ (an overwhelmingly high positive-to-negative descriptive ratio), indicating a ‘spiral of silence’ regarding critical engagement. Furthermore, user profiling uncovers an ‘elite oligarchy’ dominated by middle-aged technocrats (35–49) and highly educated individuals (e.g., PhD holders), while Generation Z remains largely marginalized. The findings suggest the Arabic Twitter sphere has evolved into a platform for ‘digital statism.’ Here, entrepreneurship functions less as a market disruptor and more as an instrument to consolidate the legitimacy of state-led modernization projects. ©**authors**.

►**Citation:** Vahabi Abyaneh, M., Adib, A., & Mobini Dehkordi, A. (2026). Entrepreneurship Discourse in the Arabic Twitter Sphere: A Sentiment and Content Analysis. *International Journal of Knowledge Processing Studies (KPS)*, 6(1): 12-30. Doi: 10.22034/kps.2026.580547.1275

1. Introduction

At the turn of the third millennium, the Middle East and North Africa (MENA) region has confronted a wave of fundamental structural, demographic, and technological transformations that have challenged the traditional socio-economic order. Having relied for decades on rentier economies and oil revenues, the region is now undergoing a paradigmatic transition, striving to move beyond traditional models toward a knowledge-based economy and the diversification of revenue sources (Nurunnabi, 2017). At the epicenter of these shifts, the concept of entrepreneurship has evolved from a merely commercial activity into a *central signifier* and a strategic imperative for addressing chronic regional challenges, such as high youth unemployment, structural inflation, and gender inequalities. This strategic pivot is explicitly reflected in the region's overarching national development frameworks; notably, Saudi Arabia's *Vision 2030* emphasizes the pivotal role of the digital economy and the development of Small and Medium Enterprises (SMEs) as the primary engines of the post-oil era, reframing entrepreneurship not merely as an economic necessity, but as a national value and a component of identity (Kingdom of Saudi Arabia, 2016; Alsaaidi, 2020).

Concurrently with this economic transition, the Arab world has witnessed a digital revolution and the pervasive penetration of social platforms. In this landscape, Twitter (X) has occupied a distinct and strategic vantage point. Unlike platforms predicated on friendship networks, Twitter in the Arab world , particularly in the Gulf Cooperation Council (GCC) states , functions as a *digital public sphere* where elites, policymakers, entrepreneurs, and citizens interact and debate in a relatively open environment (Al-Jenaibi, 2016). The confluence of state pressure for economic modernization and the expansion of cyberspace has given rise to a complex phenomenon described by scholars as the digital double bind (Zayani & Khalil, 2024). On one hand, governments invest heavily in Information and Communication Technology (ICT) infrastructure to cultivate an entrepreneurial culture; on the other, these very tools have transformed into spaces for reflecting the authentic voice of society, expressing skepticism, and even challenging the efficacy of official policies.

Despite the critical socio-economic importance of this subject, empirical research detailing how state-led economic visions translate into public digital discourse remains underexplored. While existing literature extensively covers the macroeconomic impacts and policy challenges of entrepreneurship in developing startup ecosystems (Mirahmadi et al., 2025), and separate computational studies focus purely on algorithmic optimizations for Arabic sentiment analysis without deep sociological interpretation (Musleh et al., 2023), there is a distinct interdisciplinary gap. Specifically, the precise mechanisms through which top-down state narratives are adopted, negotiated, or contested by organic users on platforms like Twitter represent a critical blind spot.

To address this gap, this study systematically deconstructs the mechanisms of constitution, dissemination, and perception of what we operationally define as Entrepreneurial Discourse—the socially constructed narratives, vocabularies, and sentiments surrounding the act of entrepreneurship and startup culture among Arabic-speaking users. Adopting the tweet and its generating user as the primary units of analysis, this research seeks to answer two fundamental questions: First, how is the emotional structure governing entrepreneurial discourse distributed? Does this distribution reflect a genuine, realistic acceptance, or does it signal a form of performative 'fragile positivity'? Second, what are the core thematic clusters shaping the identity of this discourse, and how do they correlate with official state narratives such as Vision 2030?

The significance of this study lies in its policy implications; for policymakers, a precise understanding of public sentiment regarding entrepreneurial initiatives is vital for the success of economic reform programs. If analyses reveal a neutral or negative attitude toward emerging sectors, or indicate that the ecosystem has devolved into an elitist echo chamber, a revision of communication and support strategies will be inevitable. Theoretically, by employing textual content analysis and advanced deep learning models, this study contributes to the literature on digital entrepreneurship and economic nationalism, demonstrating how entrepreneurship is reproduced in the digital sphere not merely as an economic activity, but as an identity and an ideology.

2. Literature Review

This section provides a critical and structured review of the existing literature to contextualize our primary unit of analysis: *Entrepreneurial Discourse*. To systematically unpack this phenomenon, the review is organized across three fundamental domains: 1) the theoretical foundations of entrepreneurial discourse and its intersection with neoliberalism in the digital age; 2) the specific coordinates of this discourse in the Arab world, particularly how it encompasses dimensions such as state narratives and nation branding; and 3) the technical and methodological evolution in Arabic sentiment analysis, which provides the computational tools to measure this discourse.

2.1. The Conceptual Construction of Entrepreneurial Discourse: Beyond the Economic Paradigm

In contemporary studies, the concept of entrepreneurship has transcended a purely economic process focused on firm formation or profit generation. Scholars in the social sciences and management analyze entrepreneurship as a **narrative** and **discourse** that possesses ideological and identity-forming functions. Gartner (1988), in his seminal article, argued that the locus of inquiry must shift from personality traits to behaviors and words. Extending this approach, Steyaert (2007) developed the concept of entrepreneurship as discourse, demonstrating how language creates a space where certain activities are legitimized while others are marginalized. More recent research, such as Welter (2011), emphasizes the importance of context, positing that entrepreneurial discourse does not form in a vacuum but is the product of a dialectical interaction between individual narratives and macro-institutional structures. In the digital sphere, this discourse is reproduced through linguistic mechanisms and metaphors , such as the metaphor of war or a journey for business (Clarke & Holt, 2017).

2.1.1. Neoliberalism and the Formation of the 'Entrepreneurial Subject'

One of the most dominant theoretical frameworks in the critical analysis of entrepreneurial discourse is the neoliberal paradigm. According to this view, the promotion of entrepreneurship is part of the neoliberal project to transfer social and economic responsibilities from the state to the individual (Scharff, 2016). In this paradigm, the neoliberal subject is an individual who manages their existence akin to a business enterprise, interpreting failures not as the result of unjust structures, but as a consequence of a lack of personal effort or skill (Bröckling, 2016).

In the Arab world, this discourse has been aggressively promoted by governments and international institutions to combat the Youth Bulge crisis and the public sector's inability to absorb the labor force (Sukarieh & Tannock, 2015). Case studies in countries such as Jordan and Egypt indicate that youth empowerment programs, utilizing concepts like active citizenship and resilience, seek to channel the political energy of the youth into apolitical economic activities (Murphy, 2018).

2.1.2. Hustle Culture and Theological Dimensions

Extending the neoliberal discourse, the global phenomenon of Hustle Culture has found a reflection in the Arab digital sphere. This culture, which emphasizes relentless work and the sacrifice of personal life for professional success, is propagated on social media through influencers (Nurazizah et al., 2024). A distinctive feature within the Arab context is the amalgamation of this culture with religious concepts; specific content creators selectively utilize Quranic verses and Hadiths to present an interpretation of Islam that aligns with capitalist values (Adly, 2020). This phenomenon, which can be termed epistemological distortion, seeks to provide divine legitimation for material success (Nurazizah et al., 2024).

2.1.3. Narrative Entrepreneurship and Legitimation

The theory of Narrative Entrepreneurship is predicated on the principle that entrepreneurs need to create compelling stories about their businesses to attract capital (Garud et al., 2014; Martens et al., 2007). In high-risk environments, it is these narratives that confer legitimacy upon entrepreneurial activities (Lounsbury & Glynn, 2001). Twitter, with its interactive nature, provides an ideal platform for this storytelling, and studies have shown that the expression of positive emotions and passion in these narratives directly correlates with success in resource acquisition (Cardon et al., 2009).

2.2. Regional Context: Entrepreneurship in the Arab Public Sphere

2.2.1. Twitter as a Digital Public Sphere

Following the events of 2011, Twitter evolved into a vital pillar of debate and dialogue in the Arab world. Although initial optimism regarding democratization has faded (Lynch, 2015), the platform continues to function as a barometer of public opinion (Al-Jenaibi, 2016).

However, the phenomenon of Astroturfing, the manufacture of artificial grassroots support, poses a serious challenge to the analysis of regional discourses (Jones, 2019).

2.2.2. Nation Branding and Vision 2030

A central axis of entrepreneurial discourse in the region is its intimate linkage to macro-projects of Nation Branding. Within this framework, entrepreneurship is no longer merely a micro-economic activity but part of the state's soft power to reinvent its international image. The *Vision 2030* document, published by the Kingdom of Saudi Arabia (2016), explicitly designates entrepreneurship as the primary driver of the transition from an oil economy. The document sets specific quantitative targets, including increasing the contribution of Small and Medium Enterprises (SMEs) to GDP from 20% to 35% and reducing the unemployment rate to 7% (Kingdom of Saudi Arabia, 2016). Research indicates that the state uses these quantitative goals to construct a redemptive narrative in which entrepreneurs are portrayed as new national heroes guiding the country toward modernity (Alsaaidi, 2020). Discourse analysis suggests that the ruling establishment utilizes specific linguistic strategies to forge national solidarity around this vision, effectively framing criticism of economic policy as opposition to national interests (Al-Khatib, 2001; Almaghlouth, 2022).

2.2.3. Women and Digital Agency

Social media has enabled Arab women to circumvent spatial restrictions and manage their businesses (Al-Dajani & Marlow, 2010). Research indicates that female entrepreneurs on platforms like Twitter strategically construct their identities to simultaneously align with traditional norms and modern expectations (Tlaiss & McAdam, 2021; Kemp et al., 2021).

2.3. Challenges in Arabic Natural Language Processing (NLP) and Methodological Evolution

Sentiment analysis of Arabic texts faces unique structural challenges. The phenomenon of Diglossia, the gap between Modern Standard Arabic (MSA) and local dialects (Ferguson, 1959; Shendy, 2022), diminishes the efficacy of traditional tools. Furthermore, the prevalence of Arabizi (Romanized Arabic) among the youth constitutes a significant obstacle for standard algorithms (Aboelezz, 2009; Duwairi et al., 2016).

2.4. Transition from Lexicon-Based Approaches to Large Language Models

Lexicon-based tools such as VADER face severe limitations in the Arabic context, often erroneously classifying tweets as neutral due to out-of-vocabulary terms (Hutto & Gilbert, 2014; Rusydiana & As-Salafiyah, 2022). A significant methodological shift occurred with the emergence of Transformer-based architectures, such as AraBERT and MARBERT. By leveraging massive pre-training on diverse Arabic corpora, these models capture deep contextual text representations and manage dialectal variations far more effectively than traditional methods (Antoun et al., 2020). Recent comparative literature indicates that Transformer models demonstrate state-of-the-art performance and significant empirical advantages in Arabic sentiment analysis benchmark datasets compared to rule-based approaches (Alosaimi et al., 2024; Alsugair & Alghamdi, 2024).

Table 1 provides a comparative analysis of these two approaches.

Table 1. *Comparative Analysis of Sentiment Analysis Methods in Arabic Texts*

Comparison Criteria	Lexicon-Based Approach (e.g., VADER)	Transformer Approach (e.g., AraBERT, MARBERT)
Mechanism	Matching words with predefined lists (Hutto & Gilbert, 2014)	Learning deep contextual text representations (Devlin et al., 2019)
Context Understanding	Limited (Analyzes words independently without contextual weighting)	High Capability (Captures bidirectional context and semantic nuances)
Dialect Handling	Low Coverage (Struggles with out-of-vocabulary dialectal terms)	Robust (Specifically optimized via pre-training on dialectal tweets)
Accuracy	Baseline Performance (Typically 60–70% in complex texts)	State-of-the-Art (Often exceeding 90% in benchmark tests)

2.5. Critical Synthesis and Identification of Research Gaps

A critical synthesis of the reviewed literature reveals a profound epistemological fragmentation across the three established domains. While critical management scholars have effectively deconstructed the ideological underpinnings of the 'neoliberal entrepreneurial subject' (Scharff, 2016; Bröckling, 2016), their analyses remain predominantly localized and reliant on traditional, small-scale qualitative methods that fail to capture the macro-dynamics of digital public spheres. Conversely, computational and Arabic NLP literatures demonstrate high algorithmic sophistication in sentiment classification and dialect identification (Alosaimi et al., 2024; Alsugair & Alghamdi, 2024; Matrane et al., 2023), yet they operate within a sociological vacuum, treating textual data as mere linguistic tokens rather than artifacts of political identity and nation-branding strategies.

This literature gap creates a significant double bind in understanding the Arab digital ecosystem. Existing frameworks implicitly assume a Westernized, bottom-up model of digital expression, thereby ignoring how state-led modernization agendas—such as Vision 2030—actively shape, formalize, and direct public sentiment through digital statism (Zayani & Khalil, 2024). Consequently, prior research has failed to establish a coherent scientific argument regarding how abstract political ideologies are organically internalized and amplified by real digital actors. By transitioning from a mechanical accumulation of citations to an integrated Computational Social Science framework, this study bridges this epistemic divide. Rather than an isolated technical exercise, the calibrated ensemble learning framework was designed a priori, directly guided by these critical concepts to structurally map how the interplay between institutional power and digital agency constructs the contemporary Arabic entrepreneurial discourse.

3. Method

3.1. Research Design and Approach

This study adopts an exploratory data-driven design situated within the framework of Computational Social Science (CSS). Rather than traditional qualitative mixed-methods—which rely on manual thematic coding—this research employs an advanced computational pipeline to extract macro-sociological patterns from massive unstructured textual data. On the quantitative dimension, the research utilizes Big Data Analytics and descriptive statistics to extract frequency distributions. On the semantic dimension, the study employs Deep Learning architectures to perform scalable sentiment and pragmatic analysis, identifying hidden semantic layers without the subjectivity of manual coding.

3.2. Data Source and Collection Strategy

The microblogging platform Twitter (X) was selected as the data source due to its pivotal role in shaping the digital public sphere in Arab countries. Data collection was executed using the Twitter API v2 (Academic/Enterprise tier). To avoid the narrow constraints of single-keyword sampling, a comprehensive boolean query was constructed. The query anchored on the core term "ريادة الأعمال" (Entrepreneurship), but was expanded to include morphological variations, relevant hashtags (e.g., #رواد_الأعمال, #مشاريع_ناشئة), and dialectal synonyms. The API parameters were strictly configured to filter out non-Arabic text (lang:ar) and handle rate limits through pagination. The data collection covered a six-month period (June 30 to December 27, 2025), yielding a final corpus of 48,841 content units across Original Tweets, Retweets, Replies, and Quotes.

3.3. Data Preprocessing

Given the noisy nature of social media text and the morphological challenges of the Arabic language, a multi-stage pipeline was implemented for data refinement (Rusydiana & As-Salafiyah, 2022; Farghaly & Shaalan, 2009):

- **Structural Cleaning**

Removal of URLs, usernames (@mentions), and non-textual characters.

- **Normalization**

Unification of Arabic orthography (e.g., converting أ/إ to ا and ة to ه) to enhance lexical matching accuracy.

- **Bot and Spam Detection**

To ensure data validity and eliminate algorithmic noise, automated accounts and promotional spam were filtered out. Given the specific syntactic challenges of Arabic spam tweets, we employed an established machine learning classification approach utilizing a Random Forest (RF) ensemble model, which has demonstrated superior accuracy in identifying Arabic spam behaviors based on metadata features (e.g., high URL-to-text ratio, hyper-frequency posting, and duplicate promotional phrasing) (Hantom & Rahman, 2024). Consequently, the downstream analysis focused strictly on Organic Users.

- **Linguistic Filtering**

Isolation of Arabic tweets and **Dialect Identification** to distinguish Modern Standard Arabic (MSA) from local dialects (Gulf, Levantine, Egyptian).

3.4. Research Variables and Analytical Dimensions

To achieve a multi-dimensional analysis, extracted data were operationalized into 21 performance indicators across four main layers:

A) Content and Semantic Layer

- **Sentiment Analysis**

Classification of tweet polarity into positive, negative, and neutral categories using a *Hybrid approach* combining lexicon-based methods and machine learning (Al-Ayyoub et al., 2019).

- **Emotional Model and Speech Acts**

Analysis of deeper emotional layers (e.g., joy, admiration, caution) and pragmatic analysis of Speech Acts, including declarative, directive, expressive, and assertive propositions, to understand user intent (Austin, 1962).

- **Thematic Classification and Entities**

Utilization of *Topic Modeling* to categorize content (Economic, Cultural, Social, Technological) and *NER* techniques to identify and categorize *Mega-entities* (Organization, Location, Person).

B) User Profiling Layer

- **Demographics**

Attributes such as age, gender, and education level were computationally inferred. Since Twitter does not mandate explicit demographic disclosures, these variables were estimated using a combination of self-declared profile metadata (e.g., bios stating "PhD candidate" or graduation years) and algorithmic linguistic inference. We acknowledge these are computational estimations rather than absolute census data.

- **Psychographics**

Psychological profiling was conducted using the Big Five Personality Traits model (OCEAN).

Given the semantic limitations of traditional lexicon-based methods (e.g., LIWC) in capturing complex Arabic nuances, this study utilized GPT-4 as a zero-shot personality inference engine. Recent empirical validations have demonstrated that state-of-the-art LLMs like GPT-4 can reliably infer Big Five traits from social media texts, achieving significant alignment with self-reported ground-truth scores and outperforming older algorithms (Peters & Matz, 2024). To execute this, a 'Super-document' comprising the last 100 tweets of each user (averaging 2,500 words) was constructed to provide sufficient context. The prompts instructed GPT-4 to computationally estimate the user's inclination across the five traits, acting as an exploratory psychometric tool.

C) Technical and Behavioral Layer

- **Platform and Media:** Analysis of operating systems (Android, iOS, Web) and media types (Text, Image, Video) to understand user digital behavior.
- **Activity Patterns**

Examination of user temporal activity graphs and calculation of interaction metrics, including post volume, **Influence Rate**, and **Engagement Rate**.

3.5. Fine-tuning Pre-Trained Language Models (PLMs)

Given the context-dependent nature of social media text, the use of context-aware language models is essential. Transformer architectures such as BERT (Devlin et al., 2019) have demonstrated significant success. In this study, three advanced Arabic language models were **fine-tuned** for the classification task:

- **AraBERT**

Antoun et al. (2020) trained this model on 24GB of text from Wikipedia and Arabic news, achieving competitive results in most NLP tasks.

- **MARBERT**

Abdul-Mageed et al. (2021) pre-trained this model with a focus on transfer learning in Arabic Dialects, utilizing 6 billion tweets, making it highly effective for Twitter text analysis.

- **QARiB**

Developed by Abdelali et al. (2021), this model is trained on a combination of MSA and dialectal resources (including 420 million tweets and 180 million news sentences).

3.6. Large Language Models (LLMs) and Prompt Strategy

Large Language Models (LLMs) have recently become vital tools in NLP (Alyafeai et al., 2023). This study utilized GPT-4 (OpenAI, 2023) with 1.76 trillion parameters, possessing high capabilities in sentiment and emotion analysis (Wang et al., 2023). Two approaches were employed:

- **Zero-shot Approach**

Utilizing prompt engineering to guide the model as a **Labeling Expert** without prior training. Instructions included precise category definitions and output formatting.

- **Few-shot Approach**

Based on findings by Brown et al. (2020) regarding the superiority of few-shot learning, we utilized GPT-4 by providing 9 training examples. To ensure intelligent example selection (rather than random selection), we used the sentence-transformers library and the distiluse-base-multilingual-cased-v2 model to extract tweet embeddings. The selection strategy was based on **Maximum Dissimilarity** using cosine distance to guarantee sample diversity. Additionally, given data imbalance, the number of minority class examples was increased in the prompt.

3.7. Confidence-Calibrated Ensemble Approach

In the final step, we integrated three distinct classifiers: the fine-tuned MARBERT, the GPT-4 Zero-shot, and the GPT-4 Few-shot models. A recognized challenge in such architectures is the potential violation of statistical independence; utilizing GPT-4 twice could theoretically introduce a 'double-voting' bias. To mitigate this, we relied on the concept of *Prompt-induced Diversity*. Recent advancements in LLM ensembles demonstrate that distinct prompting paradigms (zero-shot vs. in-context few-shot learning) activate different parametric reasoning pathways, yielding pseudo-independent outputs (Roumeliotis & Margaris, 2026). To ensure this empirically, we measured the inter-model agreement using Cohen’s Kappa (k= 0.68), which confirmed sufficient statistical variance and independence between the two GPT-4 configurations.

Furthermore, to address calibration and prevent unweighted dominance, we transitioned from a naive majority voting system to a *Confidence-Weighted Ensemble*. The prediction probabilities of the MARBERT model were calibrated using Platt Scaling, while GPT-4's confidence was mathematically extracted via token *logprobs* (Wang et al., 2024). In the standard voting process, if two out of three models agree on a label, that consensus is selected. However, in rare cases of Complete Disagreement (a 1-1-1 split), the system no longer arbitrarily defaults to a specific model. Instead, the conflict is algorithmically resolved by selecting the prediction with the highest *calibrated confidence score*. Preliminary evaluations demonstrated that the Few-shot model frequently produced the highest confidence scores in complex linguistic scenarios (e.g., sarcasm), which scientifically justifies its higher success rate in tie-breaking instances.

3.8. Final Analytical Procedure

Following data labeling by the ensemble model, quantitative data were analyzed using descriptive and inferential statistics. For supplementary textual data analysis, NLP techniques were used to extract Mega-hashtags and perform semantic clustering. The final output includes a set of visual charts and data-mined tables explaining the relationships between independent variables (user characteristics) and dependent variables (sentiment and discourse).

3.9. Human Validation

To ensure the validity of the automated sentiment classification and to reject the hypothesis of algorithmic bias (e.g., misclassifying sarcasm as positive), a rigorous Manual Spot-check was conducted. Rather than analyzing a single class, a stratified random sample of 500 tweets (proportionally representing Positive, Negative, and Neutral outputs) was extracted. These tweets were independently reviewed by two human annotators fluent in Arabic linguistics and regional dialects. The Inter-Annotator Agreement (IAA) was measured using Cohen’s Kappa, yielding a substantial agreement score (k= 0.82). The human validation confirmed the model's overall Precision at 95.8%. Furthermore, the manual review verified that the overwhelmingly high ratio of positive sentiment was genuine, with only a negligible fraction (less than 3%) attributed to undetected complex sarcasm or promotional spam, confirming that the statistical distribution is highly reliable.

3.10. Ethical Considerations

This research adhered to the ethical principles of Big Data research. All data were extracted from publicly available information, and user identifiers were anonymized during processing and reporting to protect privacy (Townsend & Wallace, 2016).

Here is the translation of the Results and Discussion section. I have ensured the terminology aligns with Computational Social Science (CSS) and Discourse Analysis standards, maintaining the sophisticated academic tone required for high-impact journals.

4. Findings

This section presents the empirical findings derived from the computational analysis of 48,841 content units and the profiling of 1,888 active users. To address methodological rigor and explicitly separate objective data from sociological inference, each subsection is structurally divided. We first report the empirical quantitative distributions (Descriptive Findings) and subsequently provide their contextualization (Discursive Interpretation). This structured approach aims to map the multi-layered dynamics governing the discourse of Entrepreneurship.

4.1. Structural Dynamics and Traffic Behavior

In this section, we examine the **skeletal framework** of the discourse: data volume, user interaction types, and temporal activity patterns.

4.1.1. Quantitative Findings and Data Distribution

Based on the analysis of the crawled data (a total of 48,841 content units over a six-month period), the content distribution, temporal trends, and user devices are illustrated in the figures below:

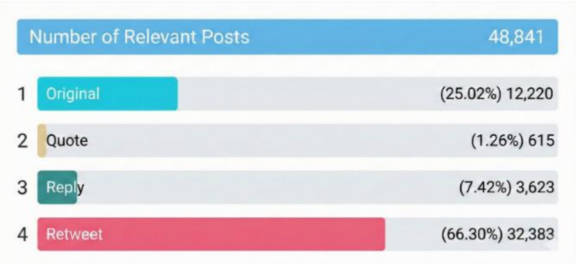

Figure 1. *Post Statistics*

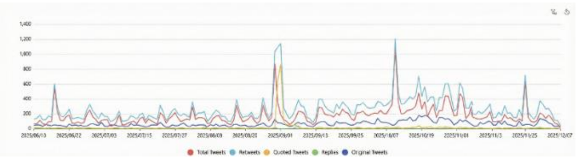

Figure 2. *User Temporal Patterns*

Note: This graph illustrates sinusoidal fluctuations in activity and traffic peaks on specific days, indicating a temporal alignment with external events rather than continuous organic engagement.

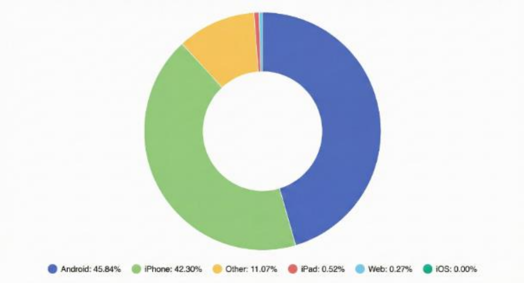

Figure 3. *Platform and Device Distribution*

4.1.2. Discursive Interpretation and Analysis

A) The Phenomenon of Amplifier Culture

In the Arabic Twitter entrepreneurship ecosystem, the general body of users primarily functions as amplifiers of content originating from official institutions, influencers, and state accounts. The negligible share of Quotes—typically utilized for adding critical or personal commentary—suggests that this discursive space is predominantly oriented toward affirmation and dissemination rather than critical dialectical exchange.

B) The Event-Driven Attention Economy

The sinusoidal and oscillating pattern observed in the temporal graph proves that entrepreneurial discourse in this context lacks an organic existence. This implies that users do not engage in discussions about entrepreneurship on ordinary days unless prompted by an external stimulus, such as the *Biban* conference, Global Entrepreneurship Week, or government announcements. This suggests the ecosystem has not yet reached a stage of maturity where innovation is integrated into the daily vernacular of citizens.

C) iPhone Hegemony: Entrepreneurship as a Luxury Lifestyle

The prevalence of iPhone users *(42%)*, visible in the platform chart, transcends a simple technological preference and carries deep sociological implications:

- **High Purchasing Power**

  In the MENA region, the iPhone is a symbol of financial affluence and the upper-middle class. This data confirms that the audience for entrepreneurial discourse is not the general mass, but a demographic enjoying relative prosperity, unburdened by basic subsistence anxieties.

- **Elitist Ecosystem**

The usage of iOS overlaps with findings regarding education levels (high density of PhD and Bachelor's degrees). This indicates we are facing a Digital Club whose members possess high levels of both cultural capital (degrees) and symbolic capital (iPhones).

4.2. Content and Semantic Ecology

In this section, the analytical focus shifts from the container (traffic and tools) to the content (meaning, language, and theme) to identify underlying semantic and discursive layers.

4.2.1. Quantitative Findings and Thematic Distribution

Based on text clustering and Natural Language Processing (NLP) analysis, three key indicators, Thematic Distribution, Word Clouds, and Linguistic Landscape, were extracted.

The distribution of thematic clusters is shown in Figure 4. As observed, contrary to traditional expectations in business studies, Cultural themes top the list of user concerns, accounting for nearly a quarter of all content, followed by economic and social themes.

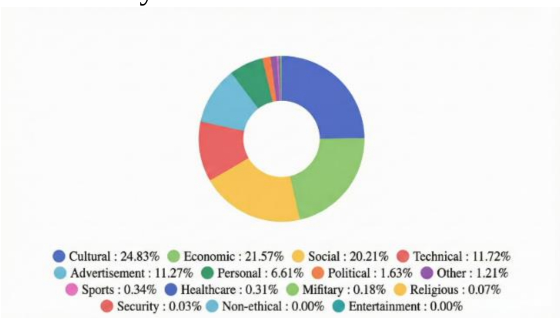

Figure 4. *Thematic Share of Tweets*

Next, the frequency analysis of vocabulary and Named Entity Recognition (NER) is plotted in Figure 5. This chart indicates a high frequency of national and state-centric keywords (such as Saudi Arabia and 2030).

Within the temporal boundaries of the collected dataset, descriptive tracking reveals a 532% relative increase in the monthly frequency of 'Vision 2030' references, calculated by comparing the baseline volume of the initial observation month (June 2025) to the peak activity month (December 2025).

Furthermore, the entity analysis highlights a clear predominance of Organizations and Locations over Persons.

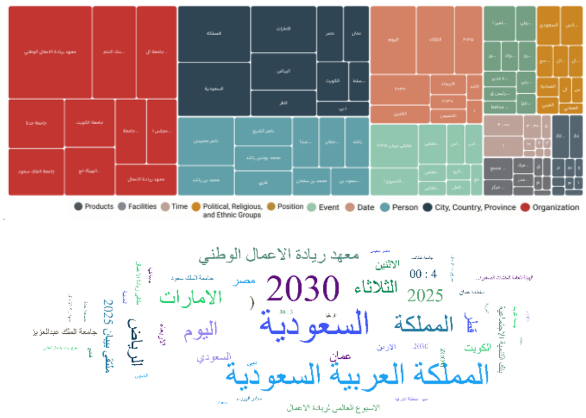

Figure 5. *Word Cloud and High-Frequency Entities*
Note: Focus on national-state keywords and high growth rate of concepts related to Vision 2030.

Finally, the linguistic and dialectal status of the tweets is presented in Figure 6. The data indicate the absolute dominance of *Modern Standard Arabic (MSA)* and a negligible share of local dialects (less than 1.5%) in content production within this domain.

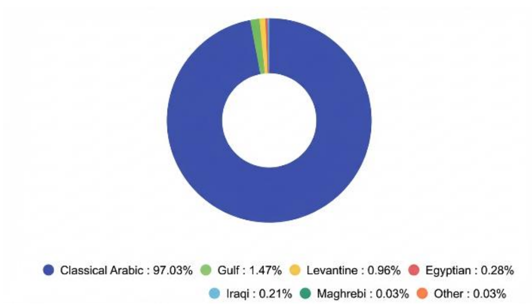

Figure 6. *Linguistic Landscape*

4.2.2. Discursive Interpretation and Analysis

A) Paradigmatic Transition: Entrepreneurship as a Cultural Project

Data from Figure 4 reveal an unexpected finding: the prioritization of cultural content over economic content. While classical theories view entrepreneurship as a profit-and-market-based phenomenon, in the Arabic Twitter context, it has transmuted into a Cultural Movement. This reflects a transition in regional societies (particularly the GCC) from a Rentier Culture (reliance on government employment) to a Productive Culture. The priority of the current discourse is not teaching business models, but rather effecting a Mindset Shift, promoting risktaking, and legitimizing private sector activity.

B) Digital Nationalism and Statism

The examination of high-frequency vocabulary in Figure 5 (heavy presence of national and temporal symbols) suggests that entrepreneurship is not perceived as a liberal private-sector activity, but as a National Duty. The descriptive data, notably the 532% growth rate of the #2030 hashtag, strongly suggests that the discourse is closely aligned with state-delineated developmental goals.

C) Hegemony of the Standard Language: Exclusion of the Everyday

The 97% dominance of formal Arabic reflected in Figure 6 indicates an unconscious effort towards the Sanctification and Formalization of the entrepreneurial field. Unlike entertainment sectors managed through intimate (vernacular) dialects, here users don a linguistic suit and tie. While this approach elevates the discourse's prestige, it creates a gap with the general public, increasing the risk of the ecosystem becoming an Elite Island detached from the realities of the street.

4.3. Emotional and Psychological Atmosphere

In this section, the hidden layers of Collective Sentiment are dissected to explicate users' emotional reactions to the concept of entrepreneurship.

4.3.1. Quantitative Findings and Sentiment Distribution

Based on the analysis of 48,841 tweets, the prevailing emotional state of the discourse was extracted. The distribution of sentiment polarity and the emotional gap is displayed in Figure 7.

As evident in the chart, the ecosystem faces a stark imbalance; positive sentiments occupy over 70% of the space, while negative sentiments do not reach even 1%. Based on the available data, the Positive-to-Negative Ratio is a staggering 94:1. This implies that for every single critical or negative tweet, there are 94 supportive tweets.

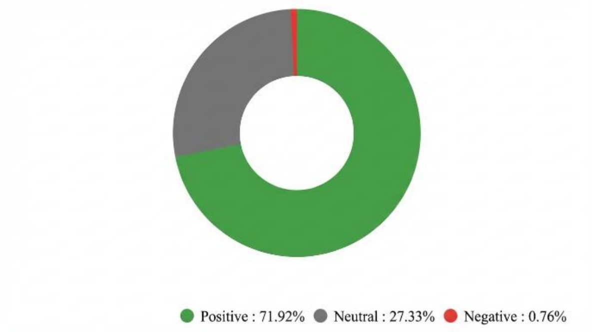

Figure 7. *Sentiment Polarity Distribution*

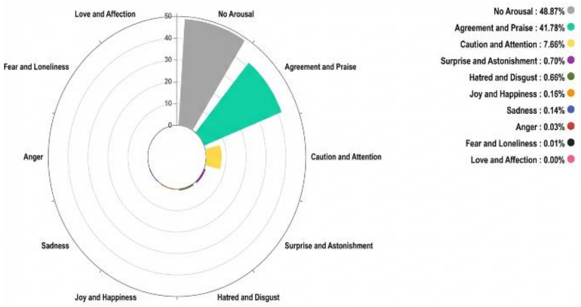

Figure 8. *Emotional Orientation and Affective Model*

4.3.2. Discursive Interpretation and Analysis

A) The Dictatorship of Positivity and the Spiral of Silence

A 94:1 ratio in a market inherently associated with risk and bankruptcy is a statistical anomaly and a sign of Toxic Positivity. This suspicious homogeneity is the product of political alignment and a culture of face-saving. According to the Spiral of Silence theory, when critical voices are in the absolute minority (0.7%), users prefer to remain silent. The result is a retouched image of the ecosystem that reflects only achievements.

B) Paradigmatic Transition:

From Festival of Joy to Allegiance of Trust

The data in Figure 8 represent the study's most critical psychological finding. The negligible share of Joy (0.16%) versus the massive share of Trust (41.78%) proves that we are not dealing with a space of entertainment or excitement, but with an Economic Allegiance. Users do not experience false euphoria; rather, they feel Institutional Security and reliance on governance structures (such as Vision 2030). This Capital of Trust is the primary fuel for the engine of state entrepreneurship.

C) Strategy of Hope and Waiting

The significant presence of the emotion of Anticipation (7.66%) indicates that the society's temporal orientation is Future-facing. Users are in a state of Hopeful Readiness for the realization of economic promises and Gigaprojects.

D) Digital Repression of Fear

The complete elimination of the emotion of Fear (0.01%) in Figure 8 indicates the Censorship of Anxiety. While fear of failure is a constant companion of real entrepreneurs, there is no space for its expression on Arabic Twitter. This Sanitization of the space, while attractive for foreign investors, deprives domestic policymakers of understanding the hidden pains and genuine anxieties of entrepreneurs.

4.4. Demographic and Social Profile

In this section, the analytical focus shifts from the Message to the Sender. To map the precise visage of ecosystem actors, a sample cluster of *1,888 Unique Users* was selected and analyzed using advanced profiling techniques.

4.4.1. Quantitative Findings and User Persona

Based on the analysis of demographic (age, education, location) and psychographic data, the characteristics of this statistical population are illustrated in the following charts.

The age distribution of users is visible in Figure 9. Data indicate the absolute dominance of the *Middle-aged* group (35–49 years) over the ecosystem and the marginal presence of the younger generation (Generation Z, under 24).

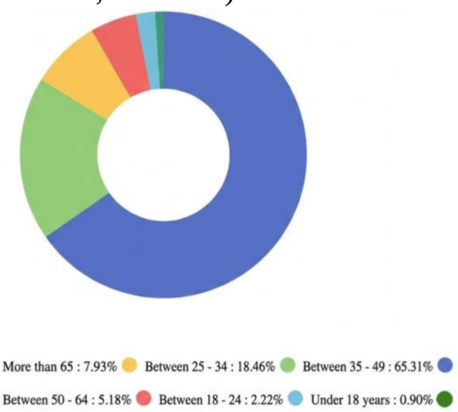

Figure 9. *User Age Pyramid*

Note: 65.31% concentration in the 35–49 age range and a negligible 3% share for Gen Z.

Furthermore, user education levels are presented in Figure 10. Statistics show a high level of academic literacy in this community, with over 50% holding a Bachelor's degree and over *17% holding a PhD.*

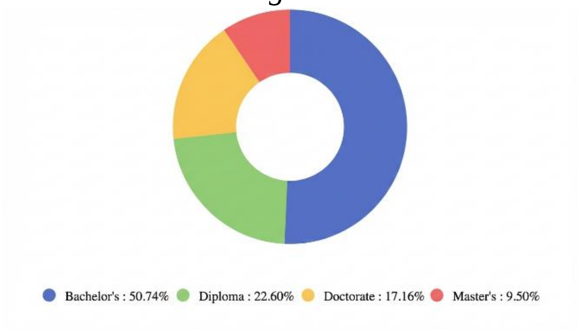

Figure 10. *Education Level Distribution*

Network health and user authenticity (Bot Detection) are displayed in Figure 11. Results show that the vast majority of the network (over 95%) consists of human, legitimate users, with very limited bot penetration.

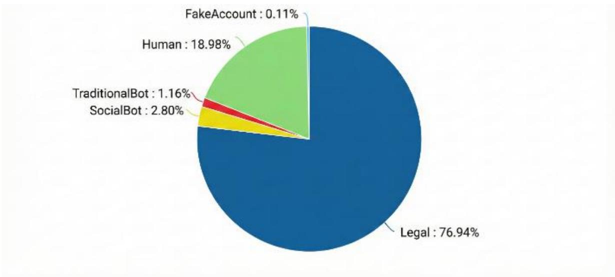

Figure 11. *Network Authenticity Assessment - Human Behavior Dominance*

Note: Minimal share of social and traditional bots (totaling less than 4%).

Finally, the psychological profile of users based on the Big Five personality model is plotted in Figure 12, indicating high scores in Conscientiousness and Agreeableness.

Figure 12. *User Personality Map*

4.4.2. Deep Interpretation and Sociological Analysis

A) Oligarchy of Middle-Aged Elites

Data from Figure 9 (65% dominance of the 35–49 group) reveal a crucial reality: The Arabic Twitter entrepreneurship ecosystem is not young. Contrary to the prevalent Western stereotype associating startups with 20-year-olds, here the primary field players are Middle-aged Technocrats. These individuals enter this space after accumulating capital and experience in state systems. The absence of Gen Z suggests that this discourse has not yet generated sufficient appeal for the new generation.

B) Educational Disconnect

The presence of *17% PhD holders* in Figure 10 is an astonishing figure (given that the PhD average in the general population is less than 1%). This Elite Density suggests Twitter has become a platform for the dialogue of Intellectuals and Academics. While this high level of literacy elevates the quality of debate, it carries the risk of creating an Elite Bubble whose language and concerns are alienated from the realities of the actual market and traditional merchants.

C) Network Health and the Agency Paradox

According to findings in Figure 11, the negligible share of bots (less than 4%) indicates the Technical Authenticity of user accounts; meaning the actors in this space are real humans, not automated algorithms. However, a precise distinction must be drawn between Account Authenticity and Organic Discourse. Although users are not bots, their convergent behavior and maximal retweeting of official content (lacking creative critique) indicate a form of Organized Behavior. in other words, we are dealing with real citizens who, either voluntarily or influenced by organizational norms, act as echo chambers for the state narrative. Thus, the ecosystem is healthy in terms of User Identity but convergent and managed in terms of Discursive Diversity.

D) Personality Profiling: The Conscientious Conservatives

The dominance of Conscientiousness and Agreeableness in Figure 12 aligns with previous findings (positivity and formal language). The typical user of this space is not a Disruptor seeking to upend structures, but a Model Citizen seeking to achieve success within existing structures (*Vision 2030*) through order and discipline.

5. Discussion

The synthesis of twenty-one variables extracted from the analysis of 48,841 content units, combined with the precise profiling of 1,888 active actors, allows us to move beyond descriptive statistics to map the inner logic and distinct identity of the entrepreneurial discourse within the Arabic context. The trajectory of this multi-layered analysis reveals six macro-conceptual dimensions that fundamentally distinguish this ecosystem from prevalent global models.

5.1. Digital Statism and the Paradigm of 'Entrepreneurial Nationalism'

The most prominent facet of this ecosystem is its profoundly State-centric nature. While classical literature, often rooting from Silicon Valley, defines entrepreneurship as a disruptive, bottom-up force often at odds with the established order, our data indicate the emergence of a phenomenon best described as Digital Statism.

The strong, significant correlation between the concept of entrepreneurship and state-centric keywords (specifically *Vision 2030*), coupled with the dominance of retweets (66%), signals a transition from the classic concept of a startup ecosystem to a Dirigiste Ecosystem (State-directed ecosystem). To rigorously operationalize these constructs, "Digital Statism" is empirically indexed in our corpus by the overwhelming dominance of programmatic amplification (the 66.30% Retweet volume) combined with the extreme high-frequency replication of institutional terminology. Concurrently, "Entrepreneurial Nationalism" is quantified via the 532% relative growth rate of the official hashtag #Vision2030 alongside the absolute hegemony of Modern Standard Arabic (97.03%), which systematically shifts the discourse from a localized market discussion to a formalized civic obligation.

In this paradigm, entrepreneurship has transmuted from a purely market-based economic activity into a Political-Identity Project. We are witnessing the rise of Entrepreneurial Nationalism, where establishing a startup is reproduced not merely for personal profit, but as a patriotic act and a symbol of loyalty to the state's modernization agenda. Although the network body consists of real humans, the Discursive Agenda is determined centrally and vertically, with users acting as normative stabilizers rather than creators of new intellectual currents.

5.2. The Discursive Turn: From Economic Pragmatism to Cultural Engineering

The second distinctive feature is the precedence of Culture over Economy. Research findings show that culture-oriented content (24.8%) significantly outweighs technical and economic content (21.5%). This signifies a Paradigmatic Transition suggesting that the digital ecosystem is currently in a phase of Normative Stabilization.

The platform functions less as a venue for commercial knowledge exchange and more as a public classroom for Social Engineering. The priority is to effect a Mindset Shift, facilitating the societal transition from the identity of *Homo Petroleus* (the oil-dependent human) to *Homo Entrepreneurialis* (the entrepreneurial human). Consequently, technical business solutions have been relegated to secondary importance in favor of narrative construction.

5.3. Fragile Positivity and the Spiral of Silence

The emotional structure governing the discourse is characterized by a *Dictatorship of Positivity*. The overwhelming dominance of positive sentiments (99% combined positive and neutral) and the reduction of dissenting voices to less than one percent (0.7%) has created an atmosphere of Performative Positivity.

While this appears promising on the surface, it reveals a lack of deep Critical Thinking. This monochromatic atmosphere reinforces the Spiral of Silence, where critiquing the ecosystem or admitting failure is stigmatized as a taboo. Consequently, users are driven toward self-censorship and the reproduction of success clichés, creating a Retouched Reality that masks the genuine risks of the market and potentially leads policymakers into cognitive errors regarding ecosystem health.

5.4. Elite Oligarchy and the New Digital Divide

Contrary to global narratives that associate digital entrepreneurship with the younger generation, our demographic analysis reveals an Elite Oligarchy dominated by middle-aged technocrats (35–49 years old) and the highly educated (with a 17% density of PhD holders). This demographic, armed with high-end tools (iPhones) and formal language, indicates that the discourse is Non-inclusive.

The significant absence of Generation Z (under 24) and the exclusion of local dialects from the mainstream conversation unveil a new Digital Divide. This gap suggests that entrepreneurship has not yet evolved into a mass, organic movement but remains enclosed within formal, elitist circles, creating invisible Barriers to Entry for marginalized groups.

5.5. Amplifier Culture and Passive Consumption

The interaction patterns, relied upon maximal retweeting and minimal original content creation, depict an ecosystem defined by an *Amplifier Culture*. In this space, the flow of knowledge moves strictly Top-down. Users function primarily as distribution nodes amplifying official narratives rather than as active agents engaged in critical dialectic or original knowledge production. This reflects a mode of passive consumption where the audience validates the message without necessarily engaging with its substance.

5.6. The Paradox of Organic Authenticity

Finally, technical evidence confirms network health and a lack of widespread bot infiltration (Technical Authenticity); however, this does not equate to divergence of opinion. Real users, driven by social pressures and organizational imperatives, exhibit highly convergent and conservative behavior. This paradox has resulted in a Corroborating Echo Chamber, where the authenticity of user identity coexists with a homogenized, managed discourse. The organic users, in effect, voluntarily replicate the function of state media apparatuses.

6. Conclusion

The exploration of the semantic structure and emotional atmosphere governing the

Arabic Twitter sphere in this study offers a novel vantage point for understanding the phenomenon of the *social construction of entrepreneurship* in the emerging markets of the Middle East. The integration of quantitative data mining with deep discourse analysis and user profiling has yielded insights that transcend mere description, exposing the inner logic of this ecosystem.

Theoretically, the findings challenge the universal applicability of prevalent Western models that conceptualize entrepreneurship as a purely individualistic, market-driven phenomenon independent of the state. The evidence points to the emergence and consolidation of a State Entrepreneurship paradigm within the digital context. Unlike its classical Schumpeterian function of creative destruction in the West, social media here has mutated into an instrument for bolstering national cohesion and legitimizing economic reform agendas. Furthermore, by injecting components of affect and culture into the digital entrepreneurship literature, this study demonstrates how abstract concepts like hope and positivity are produced, distributed, and consumed as strategic commodities within the digital economy.

In terms of practical implications, the analysis indicates that while the phase of public awareness has been successfully navigated, the dearth of constructive critique and the significant absence of the younger generation constitute a serious warning for the ecosystem's sustainability. To prevent this ecosystem from ossifying into a closed elite club, policymakers must transcend the model of official monologues and facilitate an environment for critical dialogue and the inclusion of diverse voices. Simultaneously, practitioners and entrepreneurs must remain cognizant of the sharp distinction between the harsh reality of the market and the airbrushed Twitter performance; reliance on this hyper-positive space for strategic decision-making can be misleading. Moreover, the instability resulting from an excessive focus on ephemeral events highlights the necessity of revising ecosystem-building strategies towards the production of deep educational content.

Methodologically, while the synthesis of big data analytics and qualitative interpretation enabled a transition from surface to depth, it must be acknowledged that the exclusive focus on Twitter , a platform with a predominantly elitist user base , may have obscured lived realities on more mass-market platforms. The inherent limitations of sentiment analysis tools in detecting linguistic nuances, such as sarcasm, remain a challenge that must be factored into the interpretation of results.

To expand the frontiers of knowledge in this domain, future research could adopt a comparative approach to investigate intergenerational gaps by contrasting Twitter with new-generation platforms. Furthermore, longitudinal examinations of discursive dynamics during economic downturns versus booms, or an inquiry into the phenomenon of organizational silence among failed startup founders, could provide a more holistic view of this biome.

Ultimately, the Arabic Twitter sphere functions as a mirror reflecting a utopian and sanitized image of entrepreneurship, an image that, while essential for motivation and the mobilization of national resources, remains insufficient for achieving genuine economic maturity and resolving tangible challenges.

Acknowledgement

Scientific research is an interminable voyage toward the discovery of truth; a journey that, while often beginning with the tentative steps of inquiry, ultimately reaches its destination through the symphony of collective wisdom and the unwavering support of kindred spirits. The completion of this study, which represents an interdisciplinary endeavor to bridge the nuances of social sciences with the complexities of data science, would not have been possible without the grace of the Almighty and the invaluable assistance of distinguished institutions and individuals. We consider it our duty to extend our deepest gratitude and appreciation to those who served as guiding lights along this arduous path.

First and foremost, we extend our sincere appreciation to the *Faculty of Entrepreneurship at the University of Tehran*. The dynamic scientific atmosphere, the research facilities, and the elite intellectual environment governed by this faculty provided a secure and inspiring foundation, allowing us to boldly explore novel subjects at the frontiers of knowledge. The spiritual and structural support provided by this academic institution has been the cornerstone of this research's inception and development.

We are also deeply indebted to our esteemed professors and research colleagues who, through their compassionate critiques and profound insights, assisted us in a more precise explication of the Entrepreneurial Discourse and in grasping the subtleties of the Arab Digital Ecosystem. Their wisdom in interpreting complex findings and their insistence on avoiding superficiality guided us in transcending raw data to reach meaningful sociological inferences.

Furthermore, we wish to express our special gratitude to the *Anonymous Reviewers* of this manuscript. Their time, meticulous scrutiny, and exemplary attention to detail provided valuable points that significantly enhanced the quality of this work. In particular, their constructive critiques regarding the clarification of the personality assessment methodology and the necessity for more rigorous validation of the language models led to fundamental revisions that strengthened the validity of the research. Undoubtedly, the final quality of this work owes much to the sharp vision and rigorous standards of these scholars.

Given that this research is founded on big data analytics, it is necessary to acknowledge the *Open Source Community* and the providers of advanced language models, specifically the developers of *MARBERT* and *AraBERT*. Open access to these powerful tools has made sentiment analysis and natural language processing accessible to independent researchers with unparalleled precision, democratizing the tools of modern science.

In the era of digital transformation, the boundaries between human and artificial intelligence are becoming increasingly blurred, and this research has benefited from this synergy. In this regard, we explicitly and specially acknowledge the intelligent assistant, *Gemini AI*, developed by Google. Far beyond a simple translation machine, this advanced tool served as an intelligent linguistic assistant throughout the process of translating the manuscript, selecting specialized terminology, and structurally editing English sentences. The high precision of this AI in understanding the context of scientific texts and suggesting vocabulary consistent with academic literature assisted us in presenting the research findings with a fluency and eloquence worthy of international standards. This experience stands as a testament to the fact that constructive interaction between human and machine can open new horizons in scientific production.

Finally, we dedicate our heartfelt thanks to our patient families and empathetic friends; those who, during the long nights of data mining and the high-pressure days of writing, provided the emotional support and understanding necessary to maintain the focus required for this scientific endeavor. It is our hope that the results of this research serve as a step, however small, toward a more precise understanding of regional developments and the advancement of entrepreneurship knowledge.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abdelali, A., Hassan, S., Mubarak, H., Darwish, K., & Samih, Y. (2021). Pre-training BERT on Arabic tweets: Practical considerations. *arXiv*. https://arxiv.org/abs/2102.10684

Abdul-Mageed, M., Elmadany, A., & Nagoudi, E. M. B. (2021). ARBERT and MARBERT: Deep bidirectional transformers for Arabic. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics* (pp. 7088–7105). Association for Computational

Linguistics. https://doi.org/10.18653/v1/2021.acl-long.551

Aboelezz, M. (2009). Latinized Arabic and connections to bilingual ability. In *Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching* (Vol. 3, pp. 1–23). Lancaster University.

Abu Farha, I., & Magdy, W. (2019). Mazajak: An online Arabic sentiment analyser. In *Proceedings of the Fourth Arabic Natural Language Processing Workshop* (pp. 192–198). Association for Computational Linguistics. https://doi.org/10.18653/v1/W19-4620

Adly, A. (2020). *Cleft capitalism: The social origins of failed market making in Egypt*. Stanford University Press. https://doi.org/10.1515/9781503612785

Al-Ayyoub, M., Nuseir, A., Alsmadi, K., Jararweh, Y., & Gupta, B. (2019). A comprehensive survey on Arabic sentiment analysis. *Information Processing & Management*, *56*(2), 320–342. https://doi.org/10.1016/j.ipm.2018.07.006

Al-Badrashiny, M., Eskander, R., Habash, N., & Rambow, O. (2016). Automatic transliteration of Romanized dialectal Arabic. In *Proceedings of the Conference on Computational Natural Language Learning* (pp. 30–38). https://doi.org/10.18653/v1/W14-1604

Al-Dajani, H., & Marlow, S. (2010). Impact of women's home-based enterprise on family dynamics: Evidence from Jordan. *International Small Business Journal*, *28*(5), 470–486. https://doi.org/10.1177/0266242610369745

Al-Jenaibi, B. (2016a). The Twitter revolution in the Gulf countries. *Journal of Creative Communications*, *11*(1), 61–83. https://doi.org/10.1177/0973258616630217

Al-Jenaibi, B. (2016b). The use of social media in the Arab world: The case of Twitter. *Journal of Arab & Muslim Media Research*, *9*(1), 33–54. https://doi.org/10.1386/jammr.9.1.33_1

Al-Khatib, M. A. (2001). The pragmatics of letter writing in Arabic. *Journal of Pragmatics*, *33*, 1789–1810. https://doi.org/10.1016/S0378-2166(01)00037-3

Almaghlouth, S. (2022). Environmental sustainability in the online media discourses of Saudi Arabia: A corpus-based study. *PLOS ONE*, *17*(11), e0277253. https://doi.org/10.1371/journal.pone.0277253

Alosaimi, S., Saleh, H., Hamzah, A., El-Rashidy, M., Alharb, A., Elaraby, N., & Mostafa, A. (2024). Arabic sentiment analysis using transformer models. *Frontiers in Artificial Intelligence*, *7*, 1408845. https://doi.org/10.3389/frai.2024.1408845

Alsaaidi, H. Z. (2020). *Nation branding and the case of Saudi Vision 2030 and the use of Twitter* (Master's thesis). Rochester Institute of Technology.

Alshahrani, M. A., Yaqub, M. Z., & Alsabban, A. (2025). Navigating the entrepreneurial landscape in Saudi Arabia: A systematic literature review. *Asian Journal of Interdisciplinary Research*, *8*(4), 146–172. https://doi.org/10.54392/ajir2549

Alsugair, A. M., & Alghamdi, N. S. (2024). Sentiment analysis of Arabic tweets using AraBERT as a fine-tuner and feature extractor. In *Proceedings of the International Conference on Computing* (pp. 1–6).

Alwakid, G., Osman, T., & Hussein, T. (2017). Challenges in sentiment analysis for Arabic social networks. *Procedia Computer Science*, *117*, 89–100. https://doi.org/10.1016/j.procs.2017.10.097

Alyafeai, Z., Alshaibani, M. S., AlKhamissi, B., Luqman, H., Alareqi, E., & Fadel, A. (2023). Taqyim: Evaluating Arabic NLP tasks using ChatGPT models. *arXiv*. https://arxiv.org/abs/2306.16322

Antoun, W., Baly, F., & Hajj, H. (2020). AraBERT: Transformer-based model for Arabic language understanding. In *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools* (pp. 9–15). https://doi.org/10.18653/v1/2020.osact-1.2

Austin, J. L. (1962). *How to do things with words*. Oxford University Press.

Bröckling, U. (2016). *The entrepreneurial self: Fabricating a new type of subject*. Sage. https://doi.org/10.4135/9781473921283

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901.

Cardon, M. S., Wincent, J., Singh, J., & Drnovsek, M. (2009). The nature and experience of entrepreneurial passion. *Academy of Management Review*, *34*(3), 511–532. https://doi.org/10.5465/amr.2009.40633190

Darwish, K., Abdelali, A., & Mubarak, H. (2021). A survey on Arabic dialect processing. *ACM Transactions on Asian and*

Low-Resource Language Information Processing. https://doi.org/10.1145/3476290

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT* (pp. 4171–4186). https://doi.org/10.18653/v1/N19-1423

Downing, S. (2005). The social construction of entrepreneurship: Narrative and dramatic processes in the coproduction of organizations and identities. *Entrepreneurship Theory and Practice*, *29*(2), 185–204. https://doi.org/10.1111/j.1540-6520.2005.00076.x

Duwairi, R. M., & El-Orfali, M. (2014). A study of the effects of preprocessing strategies on sentiment analysis for Arabic text. *Journal of Information Science*, *40*(4), 501–513. https://doi.org/10.1177/0165551514534143

Duwairi, R. M., Alfaqeh, M., Wardat, M., & Alrabadi, A. (2016). Sentiment analysis for Arabizi text. In *2016 International Conference on Information and Communication Systems (ICICS)* (pp. 127–132). https://doi.org/10.1109/ICICS.2016.7476098

Farghaly, A., & Shaalan, K. (2009). Arabic natural language processing: Challenges and solutions. *ACM Transactions on Asian Language Information Processing*, *8*(4), 1–22. https://doi.org/10.1145/1644879.1644881

Ferguson, C. A. (1959). Diglossia. *Word*, *15*(2), 325–340. https://doi.org/10.1080/00437956.1959.11659702

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, *59*(7), 96–104. https://doi.org/10.1145/2818717

Gartner, W. B. (1988). Who is an entrepreneur? Is the wrong question. *American Journal of Small Business*, *12*(4), 11–32.

Garud, R., Schildt, H. A., & Lant, T. K. (2014). Entrepreneurial storytelling, future expectations, and the paradox of legitimacy. *Organization Science*, *25*(5), 1479–1492. https://doi.org/10.1287/orsc.2014.0915

Guellil, I., Azouaou, F., & Valitutti, A. (2020). Sentiment analysis of Arabizi on social media: A comparative study. *Social Network Analysis and Mining*, *10*, 1–17. https://doi.org/10.1007/s13278-020-00667-w

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14596

Jones, M. O. (2019). Propaganda, fake news, and fake trends: The weaponization of Twitter bots in the Gulf crisis. *International Journal of Communication*, *13*, 1389–1415.

Kemp, L. J., Beesley, S., & Gallant, M. (2021). Women entrepreneurs in the United Arab Emirates: Their agency and values. In B. D. Metcalfe et al. (Eds.), *Women entrepreneurship development in the Middle East* (pp. 51–68). Routledge. https://doi.org/10.4324/9780429057403-4

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). Computational social science. *Science*, *323*(5915), 721–723. https://doi.org/10.1126/science.1167742

Lounsbury, M., & Glynn, M. A. (2001). Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. *Strategic Management Journal*, *22*(6–7), 545–564. https://doi.org/10.1002/smj.188

Lynch, M. (2015). After the Arab Spring: How Islamists hijacked the Middle East revolts. *Perspectives on Politics*, *13*(1), 55–73. https://doi.org/10.1017/S153759271400313X

Martens, M. L., Jennings, J. E., & Jennings, P. D. (2007). Do the stories they tell get them the money they need? The role of entrepreneurial narratives in resource acquisition. *Academy of Management Journal*, *50*(5), 1107–1132. https://doi.org/10.5465/amj.2007.27169488

Mubarak, H., Hassan, S., & Chowdhury, S. A. (2023). Emojis as anchors to detect Arabic offensive language and hate speech. *Natural Language Engineering*, 1–22.

Murphy, E. C. (2018). The in-betweeners: The youth bulge and the Arab Spring. *Middle East Critique*, *27*(1), 7–25. https://doi.org/10.1080/19436149.2017.1417030

Nielsen, S. L., & Lassen, A. H. (2012). Identity in entrepreneurship effectuation theory: A supplementary framework. *International Entrepreneurship and Management Journal*, *8*(3), 373–389. https://doi.org/10.1007/s11365-011-0180-5

Nurunnabi, M. (2017). Transformation from an oil-based economy to a knowledge-based economy in Saudi Arabia: The direction of Saudi Vision 2030. *Journal of the Knowledge Economy*, *8*, 536–564. https://doi.org/10.1007/s13132-017-0479-8

Ogbor, J. O. (2000). Mythicizing and reification in entrepreneurial discourse: Ideology-critique of entrepreneurial studies. *Journal of Management Studies*, *37*(5), 605–635. https://doi.org/10.1111/1467-6486.00206

OpenAI. (2023). *GPT-4 technical report*. arXiv. https://doi.org/10.48550/arXiv.2303.08774

Radcliffe, D., & Lam, A. (2018). *Social media in the Middle East: The story of 2017*. University of Oregon.

Sarkhoh, N., & KhosraviNik, M. (2020). Social media discourses of Arabism and the negotiation of self in the Middle East. *World Englishes*, *39*(4), 1–16. https://doi.org/10.1111/weng.12502

Scharff, C. (2016). The psychic life of neoliberalism: Mapping the contours of entrepreneurial subjectivity. *Theory, Culture & Society*, *33*(6), 107–122. https://doi.org/10.1177/0263276415590164

Schwartz, H. A., Eichstaedt, J., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M., & Ungar, L. (2013). Personality, gender, and age in the language of social media. *PLOS ONE*, *8*(9). https://doi.org/10.1371/journal.pone.0073791

Steyaert, C. (2007). 'Entrepreneuring' as a conceptual attractor? *Entrepreneurship & Regional Development*, *19*(6), 453–477. https://doi.org/10.1080/08985620701671759

Sukarieh, M., & Tannock, S. (2015). *Youth rising? The politics of youth in the global economy*. Routledge. https://doi.org/10.4324/9781315884660

Tlaiss, H. A., & McAdam, M. (2021). Unexpected lives: The intersection of Islam and Arab women's entrepreneurship. *Journal of Business Ethics*, *171*, 253–272. https://doi.org/10.1007/s10551-020-04443-y

Townsend, L., & Wallace, C. (2016). *Social media research: A guide to ethics*. University of Aberdeen.

Wang, Z., Xie, Q., Ding, Z., Feng, Y., & Xia, R. (2023). Is ChatGPT a good sentiment analyzer? A preliminary study. *arXiv*. https://doi.org/10.48550/arXiv.2304.04339

Welter, F. (2011). Contextualizing entrepreneurship: Conceptual challenges and ways forward. *Entrepreneurship Theory and Practice*, *35*(1), 165–184. https://doi.org/10.1111/j.1540-6520.2011.00427.x

World Economic Forum. (2017). *The future of jobs and skills in the Middle East and North Africa*. World Economic Forum.

Zayani, M., & Khalil, J. F. (2024). *The digital double bind: Change and stasis in the Middle East*. Oxford University Press. https://doi.org/10.1093/oso/9780197508626.001.0001

Hantom, W. H., & Rahman, A. (2024). Arabic spam tweets classification: A comprehensive machine learning approach. *AI*, *5*(3), 1049–1065. https://doi.org/10.3390/ai5030052

Mansour, O., Aboelela, E., Talaat, R., & Bustami, M. (2025). Transformer-based ensemble model for dialectal Arabic sentiment classification. *PeerJ Computer Science, 11*, e2644. https://doi.org/10.7717/peerj-cs.2644

Mirahmadi, S. M. M., Jahanbakht, M., & Rohban, M. H. (2025). Mitigating entrepreneurship policy challenges in developing countries' startup ecosystems through machine learning analysis. *Economies*, *13*(10), 295. https://doi.org/10.3390/economies13100295

Musleh, D., Alkhwaja, I., Alkhwaja, A., & Abdulqader, M. M. (2023). Arabic sentiment analysis of YouTube comments using NLP-based machine learning approaches. *Big Data and Cognitive Computing*, *7*(3), 127. https://doi.org/10.3390/bdcc7030127

Peters, H., & Matz, S. C. (2024). Large language models can infer psychological dispositions of social media users. *PNAS Nexus*, *3*(6), pgae231. https://doi.org/10.1093/pnasnexus/pgae231

Roumeliotis, K. I., & Margaris, D. (2026). A large-scale empirical study of LLM orchestration and ensemble strategies for sentiment analysis in recommender systems. *Future Internet*, *18*(2), 112. https://doi.org/10.3390/fi18020112

Mettam, G. R., & Adams, L. B. (1994). How to prepare an electronic version of your article. In B. S. Jones & R. Z. Smith (Eds.), *Introduction to the electronic age* (pp. 281–304). E-Publishing Inc.