... ZAOULI N°10, Vol. 3, Juillet 2025, pp. 624-656 ISSN : 2788-9343

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

Natural Language Processing on Social Media Platforms: A Socio-Literary Perspective

DJAGHBALLOU Hamza
Mohamed Bachir AlIbrahimi
University BordjBouArreridj
hamza.djghballou@univ-bba.dz
Orcid : https://orcid.org/0009-0004-3441-5068

NAÏM Ben Ahmed
Mohamed Bachir AlIbrahimi
University BordjBouArreridj
benahmed.galbois@gmail.com
Orcid : https://orcid.org/0009-0000-7533-5700

HOURIA Nasri
Mohamed Bachir AlIbrahimi
University BordjBouArreridj
houria.nasri@univ-bba.dz
Orcid : https://orcid.org/0009-0007-9132-9547

Abstract:
This field represents a new area of research, as it combines methods used across various domains, including natural language and social processing, machine learning, diagnostics, and information retrieval. It also incorporates semantic technologies. Leveraging the social nature of interactions that occur among individuals on social media platforms offers fertile ground for exploration, enabling participants in this space to express their opinions and play central roles in virtual communities and worlds.

This virtual interaction provides a nurturing environment for numerous and complex social processes. It somewhat reflects elements of classical social dynamics such as cooperation and conflict – but at a distance, using a specific type of language. The vocabulary and topics discussed continue to expand, especially when considering the time factor. This leads users to adopt abbreviations, signs, and symbols as intelligent communicative strategies based on semantics derived from experience and a shared semantic field. These semantic codes quickly spread among users, forming a kind of social language born in the virtual environment.

Keywords: *Linguistic communication, Semantic Web, Social media platforms.*

Traitement du langage naturel sur les plateformes de réseaux sociaux : une perspective socio-littéraire

Résumé :
Ce domaine représente un nouveau champ de recherche, car il combine des méthodes utilisées dans divers domaines, notamment le traitement du langage naturel et social, l'apprentissage automatique, le diagnostic et la recherche d'informations. Il intègre également les technologies sémantiques. L'exploitation de la nature sociale des interactions qui se produisent entre les individus sur les plateformes de réseaux sociaux offre un terrain fertile pour l'exploration, permettant aux participants de cet espace d'exprimer leurs opinions et de jouer un rôle central dans les communautés et les mondes virtuels.
Cette interaction virtuelle offre un environnement propice à de nombreux processus sociaux complexes. Elle reflète en quelque sorte des éléments de la dynamique sociale classique, tels que la coopération et le conflit, mais à distance, en utilisant un type de langage spécifique. Le vocabulaire et les sujets abordés ne cessent de s'étendre, en particulier si l'on tient compte du facteur temps. Cela conduit les utilisateurs à adopter des abréviations, des signes et des symboles comme stratégies de communication intelligentes basées sur la sémantique dérivée de l'expérience et d'un champ sémantique partagé. Ces codes sémantiques se répandent rapidement parmi les utilisateurs, formant une sorte de langage social né dans l'environnement virtuel.

Mots-clés *: Communication linguistique, Web sémantique, Plateformes de réseaux sociaux.*

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

Introduction

Language emerges as a fundamental tool of social communication—a crucible in which cultural and civilizational elements are fused. It serves as a medium of interaction among individuals, groups, and societies. Indeed, we cannot envision civilizational progress without language. Linguistic communication is the means through which individuals exchange experiences and knowledge, and through which cultural and civilizational elements are transmitted from one society to another. It is also a measure of development and progress; the more a society is capable of communicating with others, the more elevated its cultural status becomes.

This is why linguistic communication attracts considerable interest from specialists. The linguistic interaction between social actors enables individuals to influence and be influenced in the pursuit of meeting their needs and achieving their objectives. Communication, in all its cultural and civilizational dimensions—what we refer to as *linguistic communication*—represents a vital element in the exercise of dominance. As Ibn Khaldun put it, "the vanquished is always enamored with imitating the victor."

Sociolinguistics has been defined as "*the study of language in its relation to society*" (1). Therefore, the value of sociolinguistics lies in its ability to clarify the nature of language in general and to highlight specific characteristics of a particular language. Accordingly, scholars of society recognize that linguistic facts can enhance their understanding of the social fabric. In this context, the sociology of language can be defined as "*the study of society*

in relation to language" (2), which is essentially the reverse of the definition of sociolinguistics.

The difference between sociolinguistics and the sociology of language does not lie in the elements themselves but rather in the focus of interest. This distinction depends on whether the researcher prioritizes language or society, as well as their proficiency in analyzing both linguistic structures and social elements. In practice, there is a considerable overlap between the two fields.

The social nature of language has become clearly evident. Antoine Meillet highlighted this aspect in his famous article "*How Words Change Their Meanings*", where he defined language as a social phenomenon, relying on the ideas of the renowned sociologist Émile Durkheim (3). Social phenomena— which constitute the subject matter of sociology—are characterized by several key features, the most important of which are:

- They are manifested in general systems that individuals of a given society collectively follow, organizing their communal life and coordinating their relationships with one another and with others.
- They are not created by individuals, but rather emerge from the nature of social life; they are "*products of the collective consciousness*" (4).
- When an individual deviates from any such system (given their external and coercive nature), society reacts with resistance, either through material or moral punishment.

As for the functions of social language, it serves both the individual and the community in many ways, including:

a. Facilitating communication among actors and the exchange of knowledge and experiences.

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

b. Expressing the individual's various needs (5).
c. Intellectual development, which is linked to linguistic growth—learning spoken or sign language gives rise to mental images and conceptual thinking.
d. Associating language with cultural and civilizational frameworks and reference systems.
e. Psychological function: language allows the individual to vent emotions and alleviate internal pressures, especially in moments of emotional intensity.

In light of the evolution of social media tools and platforms, an important question arises:
Can languages be processed on social media networks? How is this done, and to what extent does it affect the effectiveness of communication in the future?

We will attempt to answer this question by defining the concept of communication, identifying the main forms of social media, discussing the relationship between language and artificial intelligence, and exploring the concept of natural language processing (NLP) and the social semantics of the web.

1. Defining Concepts

1.1. The Concept of Communication

The concept of *communication* originates from the Latin word *communis*, which means "common" or "shared." Accordingly, communication as a process involves participation or mutual understanding regarding an idea, feeling, attitude, behavior, or action (6). It is one of the oldest aspects of human activity and plays a significant role in every individual's life. Communication influences people in various ways and is considered one of the fundamental

traits of humanity—whether it takes the form of images, music, direct or indirect interaction, informative or persuasive content, frightening or entertaining messages, clear or ambiguous, intentional or random. Communication is the channel that connects us to humanity.

There are many definitions of communication. For example, Carl Hovland defines it as: "*Communication is the process by which the communicator presents stimuli (usually linguistic symbols) in order to modify the behavior of other individuals (the message receivers)*." (7)

What is notable across the many definitions (many of which we have omitted here) is that the term *communication* is used in various contexts and carries multiple connotations. In its singular form (*communication*), it refers to the transmission of ideas, information, and attitudes between individuals or groups. In its plural form (*communications*), it refers to the means or channels that carry the content of communication.

Specialized studies and research in communication theory began in the United States during the 1940s. Various studies contributed to shaping a theoretical framework around communication systems. After multiple attempts, communication theory managed to define its subject and establish new perspectives (8). Linguistic communication became one of the specific branches studied within this theory, and it involved precise definitions of various concepts. With the involvement of mathematicians and communication engineers, the field defined its focus as a reflective inquiry into the distinctive features of each sign system used between two entities (either living beings or technological systems) aiming at communicative goals.

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

> Based on this, Charles Cooley describes communication as: **"**That mechanism through which human relations exist and develop. It involves the growth and sharing of symbolic meanings through the dissemination of these symbols across space and their persistence through time. It includes facial expressions, gestures, signs, tone of voice, words, writing… everything that serves to transmit a message**."** (9)

1.2. Digital Communication / The Concept of New Media

There is no consensus among theorists regarding a unified definition of *new media*. Initially, the term referred to the technical evolution that brought about synchronization between text, image, and sound. Over time, it evolved to encompass all tools used by individuals within society to communicate. The *High-Tech Dictionary* defines new media as: **"***The convergence of computers, computer networks, and multimedia***."** (10)

Another definition describes it as: **"***A set of media practices that emerged from media platforms operating within a dynamic communication environment, shaped by information and communication technologies***."**

Additionally, *PC Magazine* defines new media as: **"***Forms of communication in the digital world, including publication on CDs, DVDs, and more importantly, on the Internet***."** (11)

1.3. The Concept of Social Media Networks

Social media networks refer to those websites that emerged with the so-called second generation of the web (*Web 2.0*), enabling users to interact within virtual communities formed based on shared interests or affiliations (such as

university, country, journalism, company, etc.). These platforms allow for direct communication through services such as messaging, participating in others' profiles, and accessing shared updates and personal information. (12)

Many technology experts confirm that **social media platforms currently dominate 71% of the global media and communication market**. Each platform has its own distinct characteristics and strengths in delivering content. However, they all share a common feature: **the ability to facilitate human communication without spatial or temporal constraints**. Based on this, some researchers categorize *social media* as one of six types of *social communication media*. (13)

English Translation:

Sociolinguistics has been defined as "*the study of language in its relation to society*" (1). Therefore, the value of sociolinguistics lies in its ability to clarify the nature of language in general and to highlight specific characteristics of a particular language. Accordingly, scholars of society recognize that linguistic facts can enhance their understanding of the social fabric. In this context, the sociology of language can be defined as "*the study of society in relation to language*" (2), which is essentially the reverse of the definition of sociolinguistics.

The difference between sociolinguistics and the sociology of language does not lie in the elements themselves but rather in the focus of interest. This distinction depends on whether the researcher prioritizes language or society, as well as their proficiency in analyzing both linguistic structures and social elements. In practice, there is a considerable overlap between the two fields.

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

Let me know if you'd like to include the references (1) and (2) in a specific style (APA, Chicago, etc.) or translated into English.

English Translation:

The social nature of language has become clearly evident. Antoine Meillet highlighted this aspect in his famous article "*How Words Change Their Meanings*", where he defined language as a social phenomenon, relying on the ideas of the renowned sociologist Émile Durkheim (3). Social phenomena— which constitute the subject matter of sociology—are characterized by several key features, the most important of which are:

- They are manifested in general systems that individuals of a given society collectively follow, organizing their communal life and coordinating their relationships with one another and with others.
- They are not created by individuals, but rather emerge from the nature of social life; they are "*products of the collective consciousness*" (4).
- When an individual deviates from any such system (given their external and coercive nature), society reacts with resistance, either through material or moral punishment.

As for the functions of social language, it serves both the individual and the community in many ways, including:

a. Facilitating communication among actors and the exchange of knowledge and experiences.
b. Expressing the individual's various needs (5).
c. Intellectual development, which is linked to linguistic growth—learning spoken or sign language gives rise to

mental images and conceptual thinking.
d. Associating language with cultural and civilizational frameworks and reference systems.

e. Psychological function: language allows the individual to vent emotions and alleviate internal pressures, especially in moments of emotional intensity.

In light of the evolution of social media tools and platforms, an important question arises: Can languages be processed on social media networks? How is this done, and to what extent does it affect the effectiveness of communication in the future?

We will attempt to answer this question by defining the concept of communication, identifying the main forms of social media, discussing the relationship between language and artificial intelligence, and exploring the concept of natural language processing (NLP) and the social semantics of the web.

2. Defining Concepts

2.1. The Concept of Communication

The concept of *communication* originates from the Latin word *communis*, which means "common" or "shared." Accordingly, communication as a process involves participation or mutual understanding regarding an idea, feeling, attitude, behavior, or action (6). It is one of the oldest aspects of human activity and plays a significant role in every individual's life. Communication influences people in various ways and is considered one of the fundamental traits of humanity—whether it takes the form of images, music, direct or indirect interaction, informative or persuasive content, frightening or entertaining messages,

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

clear or ambiguous, intentional or random. Communication is the channel that connects us to humanity.

There are many definitions of communication. For example, Carl Hovland defines it as: "*Communication is the process by which the communicator presents stimuli (usually linguistic symbols) in order to modify the behavior of other individuals (the message receivers)*." (7)

What is notable across the many definitions (many of which we have omitted here) is that the term *communication* is used in various contexts and carries multiple connotations. In its singular form (*communication*), it refers to the transmission of ideas, information, and attitudes between individuals or groups. In its plural form (*communications*), it refers to the means or channels that carry the content of communication.

Specialized studies and research in communication theory began in the United States during the 1940s. Various studies contributed to shaping a theoretical framework around communication systems. After multiple attempts, communication theory managed to define its subject and establish new perspectives (8). Linguistic communication became one of the specific branches studied within this theory, and it involved precise definitions of various concepts. With the involvement of mathematicians and communication engineers, the field defined its focus as a reflective inquiry into the distinctive features of each sign system used between two entities (either living beings or technological systems) aiming at communicative goals.

Based on this, Charles Cooley describes communication as:

> "That mechanism through which human relations exist and develop. It involves the growth and sharing of symbolic meanings through the dissemination of

these symbols across space and their persistence through time. It includes facial expressions, gestures, signs, tone of voice, words, writing… everything that serves to transmit a message." (9)

2.2. Digital Communication / The Concept of New Media

There is no consensus among theorists regarding a unified definition of *new media*. Initially, the term referred to the technical evolution that brought about synchronization between text, image, and sound. Over time, it evolved to encompass all tools used by individuals within society to communicate. The *High-Tech Dictionary* defines new media as: **"***The convergence of computers, computer networks, and multimedia***."** (10)

Another definition describes it as: **"***A set of media practices that emerged from media platforms operating within a dynamic communication environment, shaped by information and communication technologies***."**

Additionally, *PC Magazine* defines new media as: **"***Forms of communication in the digital world, including publication on CDs, DVDs, and more importantly, on the Internet.***"** (11)

2.3. The Concept of Social Media Networks

Social media networks refer to those websites that emerged with the so-called second generation of the web (*Web 2.0*), enabling users to interact within virtual communities formed based on shared interests or affiliations (such as university, country, journalism, company, etc.). These platforms allow for direct communication through services such as messaging, participating in others' profiles, and accessing shared updates and personal information. (12)

Many technology experts confirm that **social media platforms currently dominate 71% of the global media and**

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

communication market. Each platform has its own distinct characteristics and strengths in delivering content. However, they all share a common feature: **the ability to facilitate human communication without spatial or temporal constraints**. Based on this, some researchers categorize *social media* as one of six types of *social communication media*. (13)

Based on the analysis, **social media** can be defined as digital platforms (websites or applications) where a user begins by creating a personal account on one of the available social networking platforms such as Facebook, Twitter, etc., within the global Internet network (Web). This allows the user to build a **database, a digital presence, and a virtual identity (Profile)** through which they can publish data, comments, documents, photos, and videos. (14)

This process progresses into **networking and sharing**, where users add friends, form groups, or join pre-existing networks of subscribers and users. These may start with circles such as friends, classmates, neighbors, colleagues, or family members. Users then engage in real-time communication by exchanging comments, opinions, and media content. These interactions occur **instantaneously, at any moment the user chooses**, across the network. (15)

This invention has **profoundly reshaped the principles of freedom of expression and publication**, reinforcing democratic thinking, human rights, and other political, social, and commercial values. Communities have formed around these platforms due to their **ease of use**, **low cost**, and **lack of need for technical expertise**, leading some to believe that such tools could spur **a global shift in thought and behavior**. Studies indicate that modern individuals, particularly in the West, **rely daily on Facebook for social**

interaction and Google for knowledge, commerce, marketing, and advertising.

3. Most Prominent Social Media Platforms

1. **Facebook** (16)
2. **Twitter**
3. **Google+**
4. **LinkedIn**
5. **Pinterest**

Other widely used platforms include:

- **YouTube**
- **WhatsApp**
- **Telegram**
- **Instagram**

Classification of Social Media Tools

1. **Web-Based Platforms (Online Media)**
    - Examples: Facebook, Twitter, blogs, chat sites, email.
    - In the context of media, they are seen as **the fourth system**, complementing the traditional three (print, radio, television).
2. **Mobile-Based Applications**
    - These include smartphone apps and personal digital assistants.
    - Mobile devices are considered **the fifth media system**. (17)

4. Other Types Based on Traditional Media Platforms:

There are also types of social media based on **traditional media platforms**, such as **radio and television**. Many channels, radio stations, and programs have launched their own social media platforms, to which features like **interactivity, digital integration, and on-demand access**

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

have been added. These developments have brought such platforms closer to the broader concept of **new media**.

After having explored the concept of social media, we now turn our attention to another key concept—**Artificial Intelligence (AI)** and its relation to **Natural Language Processing (NLP)**.

5. Artificial Intelligence and Natural Language Modeling:

Artificial Intelligence (AI) represents a vast interdisciplinary research field that spans various domains such as scientific inquiry, industrial application, education, and more. The scope of AI includes:

- **Industrial aspects**: The development of intelligent tools and devices.
- **Software-based aspects**: Programming systems that respond to individual, collective, and societal needs.
- **Philosophical and social aspects**: The increasingly critical question of the **reciprocal control between human and machine**.

The philosopher **Herbert Marcuse** addressed this issue in his renowned book *One-Dimensional Man* (1964), where he warned of the potential for machines to surpass human control. He envisioned a scenario where humans, who are naturally both **agents and recipients** of influence, might be reduced to mere recipients—**"one-dimensional beings"**.

Despite the relatively primitive state of AI in his time, Marcuse foresaw the emergence of technologies—such as AI and social engineering—that could reshape social relations. In this context, **language** plays a pivotal role in constructing **virtual social networks**, which in turn shape **individual behaviors**, and eventually influence **social, political,**

economic, and cultural dynamics. This process becomes a strategic entry point for forming, guiding, or even **generating public opinion**.

Marcuse cautioned against humans continuing to interact with these technologies passively, without asserting their role as active agents (from the network analysis concept of **IGO – "I Go" or "I Act"**).

Natural Language Processing (NLP) within AI:

One of the most important branches of AI is **Natural Language Processing (NLP)**. This domain aims to enable machines to **understand natural human language**—both written and spoken—through comprehensive multi-level analysis and generation processes. NLP facilitates the development of **interactive environments** that enable seamless **communication between humans and machines**. (19)

This capability paves the way for analyzing **public opinion** at local, regional, and even continental scales.

The Arabic Language and NLP:

Among the world's languages, **Arabic** has gained increasing attention in the field of AI, as it is:

- One of the **most widely spoken languages** globally.
- Ranked **fourth** in terms of number of speakers—after Chinese, Hindi-Urdu, and English.

Arabic is highly **processable by computational systems** and is capable of achieving advanced levels of integration—**if sufficient resources, funding, and institutional support** are made available. This positions the Arabic language as a **prime candidate** for application in modern AI technologies,

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

particularly in **social media, digital communication, and public opinion modeling**.

6. Methodologies Used in Natural Language Processing (NLP):

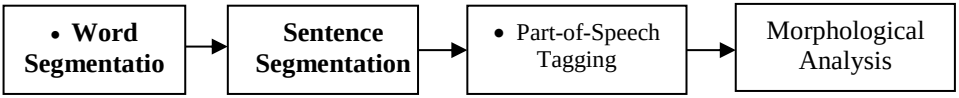
There are **two main methodologies** commonly employed in natural language processing tasks: one is **knowledge-based**, and the other is **learning-based**—with the possibility of combining both approaches. (20)

The **knowledge-based methodology** relies on **manually written rules**, crafted by experts in the field of NLP. This approach requires a deep understanding of **linguistic rules**, **language structures**, and **intuition** in addition to technical expertise.

The **preprocessing stages** of natural language tasks typically include the following core components, as illustrated in the figure:

- **Word segmentation**
- **Sentence splitting**
- **Part-of-speech tagging**
- **Morphological analysis**
- **Syntactic parsing and text segmentation**

Figure 01: Represents the components involved in the **preprocessing tasks of Natural Language Processing (NLP)**



Source (21): Diana Maynard et al., Natural Language Processing for the Semantic Web, trans. Khalid bin Abdulrahman Al-Miman, Wujuh Publishing and Distribution, Kingdom of Saudi Arabia, 1st ed., 2019, p. 41.

It is important to point out here that the process requires familiarity with terms and concepts related to computers, operating systems, and data analysis (in English), which may pose difficulties for those who are not proficient. This includes understanding the concept of the NLP pipeline and its core components, as well as some widely used open-source tools. It should also be noted that, while the performance of low-level NLP tasks is generally high, tools vary in their efficiency—not only in terms of accuracy but also in the way tasks are performed and in the nature of the outputs. These differences are largely due to the distinct linguistic theories that each tool is based on (22).

We can thus conclude that it is impossible for a mathematical model to fully replicate a naturally modeled problem. The best mathematical model that can be developed is merely a simplified and approximate representation of it.

The main challenges faced in computational modeling of natural languages can be summarized as follows:

a. The challenge of linking language as a system of symbols (sounds, words, structures...) to its real-world referents and meanings (23).

b. Imagination is an essential component of human thought. Therefore, the insurmountable challenge lies not

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

only in constructing an ontology of the real world, but also in building an ontology of the imaginary.

7. Sentiment Analysis / Opinion Mining

One of the most essential aspects of text comprehension is the identification and classification of opinions, sentiments, and emotions. This task may range from categorizing user reviews on specific products to understanding the sentiments conveyed in tweets, tracking opinions over time, identifying the views of influencers and leaders, and generating summaries based on collective opinions (24).

Sentiment analysis within text is the process of analyzing text to understand people's opinions. This excludes the analysis of emotions expressed through other media such as images or videos, as those do not fall under the scope of natural language processing (NLP) techniques (25). In its simplest form, this process involves identifying whether a person is speaking positively or negatively about something. However, opinions can take on more ambiguous forms and express a wide range of emotions that vary in intensity (e.g., liking something, fear, shock, anger, comfort, pleasant surprise...). Emotions may also relate to specific aspects of a product or event, leading to degrees of contradiction (e.g., liking some elements while disliking others).

Sentiment analysis tools can be extremely useful across almost all industrial sectors. While customer reviews and opinions may be the most obvious targets for opinion mining tools, there are numerous other applications, such as analyzing political and social sentiments toward governments, events, elections, and so forth. Opinion mining tools take a piece of text as input and produce output

in the form of information that determines whether the text contains opinions, the nature of those opinions (positive, negative, etc.), the intensity of the opinion, and other relevant features (26).

At first glance, opinion mining may seem like a simple task—where a basic system looks for the presence of positive or negative words (such as "hate," "good," "bad," etc.) and then generates an opinion accordingly. However, in practice, sentiment analysis is much more complex due to the inherently ambiguous and nuanced nature of natural language. This complexity is especially pronounced on social media platforms, where people often use unconventional language to express their emotions. They may include negative expressions in their posts, ignore grammar and spelling rules, employ sarcasm or irony, and assume that the reader has prior knowledge that helps decode the meaning (27).

One major challenge in opinion mining is distinguishing between an opinion about a person or thing and an event related to that person or thing. For example, expressions of sadness or shock may not necessarily indicate a negative sentiment, even if the content of the message is negative. Many sentiment analysis tools fail to make this distinction. Another significant difficulty lies in dealing with sarcasm, which is widespread in social media content. First, the system must detect the presence of sarcasm—an inherently challenging task. Then, it must understand how sarcasm or irony affects the polarity of the opinion. This is why the U.S. intelligence community has reportedly announced plans to purchase real-time surveillance software for social media users, with specific capabilities to detect sarcasm (28).

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

7.1. Subtasks of Opinion Mining:

From the preceding discussion, it is clear that there are several issues that need to be addressed, including:

Polarity Detection: This task involves determining whether a given statement is positive, negative, or neutral.

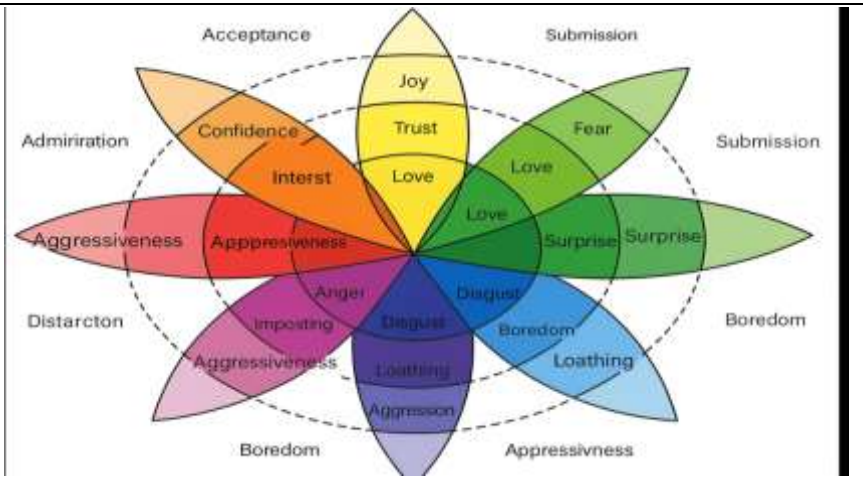
Opinion Target Identification: Often, it is not enough to know whether an opinion is positive or negative unless we also know specifically what the opinion is about.

Opinion Holder Identification: Similar to identifying the opinion target, this task involves determining who holds the opinion in question (29).

Sentiment Aggregation: Sentiments can be identified at various levels, typically at the sentence/phrase level or the document/post level. Tweets, for instance, are usually composed of a single sentence, but sometimes consist of multiple sentences. Generally, opinion detection occurs at the tweet level, but is implemented using sentence-level methodologies (30).

Emotion Detection: Practical opinion mining tools are increasingly moving away from traditional tools that merely detect positive/negative sentiments and are instead adopting emotion-based approaches. These approaches classify opinion-bearing texts based on the emotions they express. However, the task of defining a clear and comprehensive set of emotions is a difficult one. Several attempts have been made to establish standards, but there is no current consensus on a core set of emotions. Most efforts derive their frameworks from Plutchik's Wheel of Emotions, illustrated in the following figure:

Figure 02: Plutchik's Wheel of Emotions



Source (31): Diana Maynard et al., Natural Language Processing for the Semantic Web, trans. Khalid bin Abdulrahman Al-Maiman, Wujuh Publishing and Distribution, Kingdom of Saudi Arabia, 1st ed., 2019, p. 171.

Recently, opinion mining methods have increasingly focused on social media platforms, with a new trend toward applying these techniques proactively rather than reactively. Understanding public opinion in this way could have implications for forecasting future events for governments and media outlets seeking to predict responses to events and policies, as well as for individuals looking to anticipate stock performance and many other areas (32). However, adapting these tools to handle social media networks is by no means an easy task. As a result, there remain many challenges facing the development of opinion mining tools, and

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

performance levels remain low compared to many other natural language processing tasks.

8. Youth Language on Social Media

Social networking platforms have introduced a new form of writing, where the Arabic language is written using Latin letters—either in a French or English accent—commonly referred to as "Arabtini" or "Arabizi." This refers to the mixing of Arabic with English or French while speaking or writing, as in saying: bonne nuit or ok (33).

The use of Latin letters instead of Arabic has emerged in digital messaging and online chatting, especially with the rapid spread of computers and mobile phones among youth and younger generations. "Although it is currently difficult to measure the full extent of this phenomenon, observers cannot ignore its widespread presence among the new generation throughout the Arab world. There is, however, an impression that a return to the Arabic script is beginning to take shape in social media platforms" (34).

Writing in Arabic Script:

The Arabic script ranks second after the Latin alphabet in terms of the number of languages that use it. Many languages adopt the Arabic script writing system. However, the Arabtini phenomenon generally remains confined to text messaging, computer chatting, and certain social media platforms such as WhatsApp, Facebook, Twitter, Google, and the like. The primary reason appears to be that, in the early stages, these devices did not support Arabic keyboards—though this has since changed.

8. Natural Language Processing in Social Media Networks

Leveraging the social nature of interactions between individuals and groups forms the fundamental basis for the widespread dissemination of social media platforms. These platforms enable people to express their opinions, play roles in a virtual society, and collaborate remotely. For instance, in the realm of microblogging, Twitter alone has more than 300 million active users who post millions of tweets daily (35).

Currently, the dynamic engagement with these high-value, large-scale, short-lifecycle media streams has become a daily challenge for both institutions and individuals. Consequently, the need for intelligent, semantics-based methods to process and extract useful information from this data is steadily increasing. This domain represents a novel research field that intersects with several disciplines, including natural language processing (NLP), machine learning, diagnostics, information retrieval, and semantic technologies.

Traditional search methods are no longer sufficient to address the information-seeking behaviors within social media networks. These behaviors have evolved toward meaning-making, learning, investigation, and social inquiry. Semantic technologies possess the potential to assist humans in better adapting to the overwhelming information generated by social media content, thereby supporting the interpretation of information and informed decision-making in the context of vast and constantly changing media resources (36).

Unlike news articles and other online texts that are carefully crafted, social media content poses unique challenges to semantic technologies due to its massive volume, noisy and unstructured nature, and inherent social characteristics.

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

8.1 Extracting Public Opinion Trends (Tracking Opinion Evolution on Social Media)

Understanding natural language is considered one of the greatest challenges of artificial intelligence—or even the complete problem of AI (Schank & Abelson, 1977). Extracting opinions from text and identifying which parts of the text contain subjective content is a key problem in natural language processing. Opinion mining, sentiment extraction, and viewpoint analysis are often used interchangeably to refer to this area.

Opinion mining is a field of research that aims to enable automated systems to detect human opinions expressed in written (or eventually spoken) natural language. It involves tracking and identifying perspectives embedded in text (37). The domain also overlaps with computational linguistics, information retrieval (IR), text mining, natural language processing, machine learning, statistics, and predictive analytics, among others.

A wide range of applications has been developed to enable opinion mining, which has been the focus of substantial research in recent years. These applications have achieved high classification accuracy using various techniques that heavily rely on statistical sciences, artificial intelligence, machine learning, and natural language processing.

8.2 Social Media Streams: Characteristics, Challenges, and Opportunities

Social media networks allow users to connect with one another to share content (such as website links, photos, and videos), experiences, and professional information, in

addition to social interaction with friends. In essence, this technology enables communication in various forms—spontaneous and informal, semi-formal, or even formal. Users create posts or updates, and social networks disseminate them within the user's social circle. The key distinction between social networks and traditional web pages lies in the fact that users of social media are not merely passive consumers of information; rather, many of them are active content producers.

Social media platforms can be categorized according to different spectrums—based on the type of communication between users, the mode of information exchange, or how users interact with media streams:

Interest-based platforms such as Twitter encourage users to form connections with others who share common interests, regardless of whether they know each other in real life. These connections are often one-way and do not require mutual consent. The shared content typically appears in reverse chronological order as a stream of messages (38).

Social networking sites like Facebook encourage users to connect with people they have real-life relationships with. These platforms allow for the exchange of information and interaction through comments and responses on shared content.

Professional networking platforms such as LinkedIn aim to facilitate professional acquaintance in a career context, where having a connection with a specific individual often serves as a form of endorsement or recommendation (39).

Content-sharing and discussion services include blogs, video-sharing platforms (YouTube, Vimeo), presentation-sharing platforms (SlideShare), and discussion or review forums (e.g., CENT). Some blogs are even configured to

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

automatically broadcast updates to users' Facebook and Twitter accounts. While these characteristics pose significant challenges, they also offer opportunities to develop new methodologies that are better tailored to the social media environment:

Short messages (microtexts): Tweets and Facebook messages are very brief, which makes semantic-based approaches more effective when enriched with additional contextual information—such as embedded links and hashtags.

Noisy content: Social media content often includes unconventional spellings (e.g., 2moro instead of tomorrow), irregular use of uppercase letters, emoticons, and distinctive abbreviations. Therefore, specific techniques have been developed to normalize such text to its standard form.

Temporal dimension: In addition to linguistic analysis, social media content can be analyzed based on temporal sequences—a challenge that has not yet received sufficient scholarly attention.

Social Context:

The social context is essential for accurately interpreting social media content. Semantic-based methods should exploit social media data to automatically derive semantic models of networks, assess user authority, and group users with similar characteristics into clusters (40).

Multilingualism:

Social media content is characterized by a high degree of multilingualism. Therefore, automatic language identification is a crucial step for distinguishing between

different types of social media content, allowing them to be processed using appropriate algorithms tailored to specific language groups.

Conclusion

The field of technological development and its impact on language and society constitutes a new and fertile area of research that intersects linguistic techniques, symbolic repertoires, and technological advancement, including automated diagnostics and information retrieval. These technologies increasingly rely on semantic elements, thereby shifting the role of the social actor from the notion of instrumental reason to that of communicative reason—as conceptualized by Jürgen Habermas. In this context, the social dimension of interactions among virtual actors becomes a foundational and rich domain for exploration.

Social media networks serve as bridges between real social formations and the technologies that transform them into virtual structures whose behaviors can be predicted and orientations controlled. The purpose behind such control often depends on the dominant actors within these technologies—sometimes driven by commercial and marketing objectives, and at other times by narrow personal interests. More profoundly, these networks can help define virtual social relationship systems that mirror and influence the fabric of real-world society. They enable the construction of specific mechanisms for shaping and directing public opinion, which, in itself, has become a new form of the public sphere, as discussed by Habermas in his seminal

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

work The Structural Transformation of the Public Sphere (1989).

References

1) Juliette Garmaud. Sociolinguistics. Translated by Khalil Ahmad Khalil. Beirut: Dar Al-Tali'a for Printing and Publishing, 1st ed., 1990.
2) Hudson. Sociolinguistics. Translated by Mahmoud Ayad. Cairo: Alam Al-Kutub Publications, 2nd ed., 1990.
3) Antoine Meillet. Comment les mots changent de sens. In: Linguistique historique et linguistique générale. Paris: Champion, 1982.
4) Louis-Jean Calvet. Sociolinguistics. Translated by Mohamed Yahyaten. Algiers: Dar Al-Qasba Publishing, 2006.
5) Mohamed Al-Bakkaa. Media and Language (Levels and Applications). Damascus: Ninawa Center for Studies and Publishing, 1st ed., 2009.
6) Hassan Emad Mekawy and Laila Hussein El-Sayed. Communication and Its Contemporary Theories. Cairo: Egyptian-Lebanese House, 1st ed., 1998.
7) Abdelkader Al-Ghazali. Linguistics and Communication Theory. Lattakia: Dar Al-Hiwar for Publishing and Distribution, 1st ed., 2003.
8) Atef Adly Al-Abd. Theories of Media and Public Opinion, Vol. 1. Cairo: Dar Al-Fikr Al-Arabi, 2002.
9) Hussein Mahmoud Hatimi. Public Relations and Social Media Networks. Amman: Osama Publishing House, 1st ed., 2015.

10) Ali Abdel-Fattah Kana'an. Media and Society. Amman: Al-Yazouri Publishing and Distribution, 1st ed., 2014.
11) Jamal Sanad Al-Suwaidi. Social Media and Its Role in Future Transformations (From the Tribe to Facebook). UAE: Jamal Al-Suwaidi Publishing, 4th ed., 2014.
12) Sultan Khalaf Al-Mutairi. Social Media Networks and Their Relationship to Achieving Social Security. Master's thesis, College of Strategic Sciences, Naif Arab University for Security Sciences, Riyadh, 2015.
13) Khaled Ghassan Yousef Al-Muqaddadi. The Social Networks Revolution. Amman: Dar Al-Nafa'is for Publishing and Distribution, 1st ed., 2013.
14) Al-Mu'taz Billah Al-Saeed et al. Arabic and Artificial Intelligence. KSA: Wujuh Publishing and Distribution, 1st ed., 2019.
15) Herbert Marcuse. One-Dimensional Man. Translated by George Tarabishi. Beirut: Dar Al-Adab, 3rd ed., 1988.
16) Diana Maynard et al. Natural Language Processing for the Semantic Web. Translated by Khalid bin Abdulrahman Al-Maiman. KSA: Wujuh Publishing and Distribution, 1st ed., 2019.
17) Mohsen Rashwan, Al-Mu'taz Billah Al-Saeed et al. Automated Processing of Arabic Texts. KSA: Wujuh Publishing and Distribution, 1st ed., 2019.
18) Research Group. Youth Language in Modern Communication Media. Saudi Arabia: King Abdullah Bin Abdulaziz Center for Arabic Language Services, 2014.

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

Appendix

Footnotes:

1. Juliette Garmaud, Sociolinguistics, trans. Khalil Ahmad Khalil, Dar Al-Tali'a for Printing and Publishing, Beirut, Lebanon, 1st ed., 1990, p. 23.
2. Hudson, Sociolinguistics, trans. Mahmoud Ayad, Alam Al-Kutub Publications, Cairo, 2nd ed., 1990, p. 17.
3. Antoine Meillet, Comment les mots changent de sens, in: Linguistique historique et linguistique générale, Paris: Champion, 1982, p. 230.
4. Louis-Jean Calvet, Sociolinguistique, trans. Mohamed Yahyaten, Dar Al-Qasba Publishing, Algeria, 2006, p. 11.
5. Mohamed Al-Bakkaa, Media and Language (Levels and Applications), Ninawa Center for Studies and Publishing, Damascus, Syria, 1st ed., 2009, pp. 12–13.
6. Hassan Emad Mekawy, Laila Hussein El-Sayed, Communication and Its Contemporary Theories, Egyptian-Lebanese House, Cairo, 1st ed., 1998, p. 23.
7. Ibid., p. 24.
8. Abdelkader Al-Ghazali, Linguistics and Communication Theory, Dar Al-Hiwar for Publishing and Distribution, Lattakia, Syria, 1st ed., 2003, p. 23.
9. Atef Adly Al-Abd, Media and Public Opinion Theories, Part I, Dar Al-Fikr Al-Arabi, Egypt, 2002, p. 12.
10. Hussein Mahmoud Hatimi, Public Relations and Social Media Networks, Osama Publishing House, Amman, Jordan, 1st ed., 2015, p. 64.

11. Ali Abdel-Fattah Kana'an, Media and Society, Al-Yazouri Publishing and Distribution, Amman, Jordan, 1st ed., 2014, p. 162.
12. Ibid., pp. 169–170.
13. Jamal Sanad Al-Suwaidi, Social Media and Its Role in Future Transformations (From the Tribe to Facebook), Jamal Al-Suwaidi Publishing, United Arab Emirates, 4th ed., 2014, p. 21.
14. Sultan Khalaf Al-Mutairi, Social Media Networks and Their Relationship to Achieving Social Security, (Master's Thesis), College of Strategic Sciences, Naif Arab University for Security Sciences, Riyadh, 2015, p. 48.
15. Ali Abdel-Fattah Kana'an, Media and Society, pp. 171–179.
16. Khaled Ghassan Yousef Al-Muqaddadi, The Social Networks Revolution, Dar Al-Nafa'is for Publishing and Distribution, Amman, Jordan, 1st ed., 2013, p. 56.
17. Al-Mu'taz Billah Al-Saeed et al., Arabic and Artificial Intelligence, Wujuh Publishing and Distribution, Kingdom of Saudi Arabia, 1st ed., 2019, p. 13.
18. Herbert Marcuse, One-Dimensional Man, trans. George Tarabishi, Dar Al-Adab, Beirut, 3rd ed., 1988.
19. Diana Maynard et al., Natural Language Processing for the Semantic Web, trans. Khalid bin Abdulrahman Al-Maiman, Wujuh Publishing and Distribution, KSA, 1st ed., 2019, p. 39.
20. Ibid., p. 41.
21. Ibid., pp. 45–65.
22. Ibid., p. 73.
23. Ibid., p. 159.
24. Ibid., p. 160.

Soumission : **02/01/2025** **Acceptation : 03/03/2025** **Publication** : **15/08/2025**

25. Ibid., p. 161.
26. Ibid., p. 163.
27. Mohsen Rashwan, Al-Mu'taz Billah Al-Saeed et al., Automated Processing of Arabic Texts, Wujuh Publishing and Distribution, KSA, 1st ed., 2019, p. 159.
28. Ibid., p. 169.
29. Ibid., p. 171.
30. Ibid., p. 174.
31. Research Group, Youth Language in Modern Communication Media, King Abdullah bin Abdulaziz Center for Arabic Language Services, Saudi Arabia, 2014, p. 381.
32. Diana Maynard et al., op. cit., p. 171.
33. Ibid., p. 183.
34. Ibid., p. 184.
35. Saudi Arabia, 2019, p. 159.
36. Diana Maynard et al., op. cit., p. 186.
37. Ibid., pp. 184–185.
38. Mohsen Rashwan, Al-Mu'taz Billah Al-Saeed et al., Automated Processing of Arabic Texts, 1st ed., Wujuh Publishing and Distribution, KSA.
39. Ibid., p. 187.
40. Diana Maynard et al., op. cit., p. 186.
41. Ibid., p. 187.