**Social Listening Project (Anset)**

# **Our Team:**

The domain experts: the journalists Wissam Atwan, and Asma Hreash.

The Software engineers: Taha Nakhleh, and Amr Saleem.

The Social listening project aims to develop a digital tool based on machine learning and python codes that is able to detect certain patterns and follow up certain keywords in the info sphere to provide a visual report regarding the case

This conceptual framework will be considered the essential one to set the basis of this project:

The scope of the initial project will be focused on Facebook and Twitter as the major area of activities because these two sites hold the majority of the digital community in Palestine and the region regarding the lexical content.

Economic & Political value

**Introduction and definitions:**

**Social media listening consists in identifying and analyzing social media activity around topic(s) of interest.**

For marketers, social listening is the process of monitoring online conversations and turning them into powerful insights that drive your business strategy. The key word here is “insights”, which are attained from online conversations.

For instance, every modern business, regardless of its size, needs to know what existing and ideal customers say about:

· their brand

· their product or service

· their campaigns

· specific keywords and topics

· their competition

**Social Media Listening** focuses on those conversations outside of your control (user-generated, competitors’, industry leaders and product-related content.) On the other hand, **social media monitoring** focuses on tracking and analyzing how audiences respond to your content.

Unlike social media monitoring, social listening helps brands understand not only what your audience talks about online, but also the “why” behind all those conversations.

Here are some differences between the two:

· Social media monitoring is a daily activity that focuses on individual social media mentions, while social listening is a strategic process that focuses on identifying patterns and trends.

· Social media monitoring is the process of tracking and collecting the data – and social listening is the process of analyzing the data and extracting the meaning from it.

· Social media monitoring is reactive – it’s about monitoring conversations and reacting to comments, complaints, or praise in a timely manner. Social listening on the other hand is proactive. Its end goal is to provide businesses with audience insights that can help companies develop a more informed strategy.

In other words, social media monitoring is the first step in the social listening process. The figures and statistics gathered through monitoring of online channels help in noticing patterns and identifying gaps in your content.

However, social listening goes a step further and turns the raw data into insights that you can use to spot opportunities to develop and grow your business. It can help consolidate your position into existing markets, expand into new verticals, or improve customer care and grow your brand loyalty.

# **Problem statement:**

The lack of social listening tools that can deliver accurate insights and and sentimental analysis in Arabic language and various Palestinian dialicts present a huge challenge regarding such disciplines.

# **Solution:**

The main focus of the project is to establish an Arabic – English bilingual social listening digital platform that is able to deliver a highly accurate reports of the trends and patterns of public opinion on social media (Facebook and Twitter) in particular regarding certain issues upon the requests of customers, depending on big data analysis and visualization that primarily targets business and political sectors.

The scope of the initial project will be focused on Facebook and Twitter as the major area of activities because these two sites hold the majority of the digital community in Palestine and the region regarding the lexical content.

# **Value:**

Provide a product or utility for stakeholders interested in the Arabic audience and Palestinians in particular, to identify the trends of public opinion in social media on a particular topic at an affordable price. The application is characterized by being specific to the Arabic language and Palestinian local dialects; it will understand and analyze the social activity in Arabic language.

# **Menu of Product and services (Features):**

- Social Monitoring: Statistics about topics or keywords.

- Social Listening: Trends reports and deep insights.

- Analyze Arabic entries in affective methodology (Sentiment Analysis).

- Analyze Posts, Images, Videos Captions Comments, Reactions (Emoticons, share, report).

Patterns of public opinion towards certain issues based on ;lexical sentiment analysis.

* Numerical values of engagements and reactions towards certain issues vs. time.
* Numerical analysis of hashtags and words clouds regarding certain keywords vs. time.

# **Technologies:**

The project consists of two main parts or layers, each of them may have different technology stack:

1. Data Collector and Analyzer: we will use Python language and its related frameworks and technologies to collect data from sources, and analyze the data by Machine learning and Natural language processing (NLP) implementations.

2. Presentation layer: a web-based application (in first phase) that interact between user and first layer, we can use any technology stack to build this layer that give us a stable release in shortest time

The two parts may be hosted in Cloud solutions like Amazon AWS and Microsoft Azure clouds solutions.

# **Innovation:**

We want to suggest and test rules to analyze the data in Arabic language, these rules come from our understanding of Arab (Palestinian) community.

The suggested analysis must be built over a three-degree sentimental analysis (Negative, Neutral, Positive) according to a set of key words in every group to detect the public opinion regarding the issue of inspection. The keywords group is suggested to be around (10-20) words in every group.

# **Revenue streams:**

We can make revenue in different way for this type of product, we studied other similar products and the methods they make revenues. We can make a mix of pricing methods to make revenue:

- We can implement a participants plans approach to provide data and reports for customers based on their purchased plan.

- We can sell individual or specific reports and results based on customer requirements.

- We can sell row data for customers.

- We can create paid newsletters based on general statistics or analytics.

# **Customers / Users:**

- Business Owners: in order to get insights and statistics about their products.

- Researchers: to get reports about a specific topic after analyzing the data that related to their research topic.

- News Agencies and journalists:

- Governmental sector.

# **Needed funds for development:**

We are building a data product, so the DATA is our main fuel. We target Facebook and Twitter and in order to get the data from there we need to buy it, and it is very expensive.

In addition to that, the product needs many services and subscriptions e.g. AWS.

Targets:

Brand Names

Keywords

Hashtags

Of Stimuli::

* Posts
* Images
* Videos
* Captions
* Comments
* Reactions (Emoticons, share, report)

------

* Connotes mediated messages دلالات المحتوى المنتج من تعليقات
* Assess sentiments
* Make attributions

Data Management Plan

- Planning

- Collection

- Assuring

- Description

- Preservation

- Discovering

- Integration

- Analysis

**Data management plan basic framework :**

- **One**: Data types to be collected : Textual Lexical data in Arabic and English from public social media pages (FB) (Focused at Palestinian public pages)

- **Two**: Format and data description standards: The initial data collected will be stored in a CSV file in a defined columns and rows:

To transform Json files into structured DB (CSV or else)

The description will address:

**Page info:**

Name of the page

about

location.city

category

fan\_count

talking\_about\_count

**Facebook post**:

message

created\_time

updated\_time

**Facebook comments:**

message

created\_time

like\_count

comment\_count

**Facebook page likes :**

id

about

category

fan\_count

link

location.city

location.country

location.latitude

location.longitude

location.street

location.zip

name

start\_info.type

start\_info.date.year

start\_info.date.month

talking\_about\_count

username

website

were\_here\_count

**Facebook reactions parameters:**

comments.summary.total\_count

shares.count

reactions.summary.total\_count

like.summary.total\_count

love.summary.total\_count

wow.summary.total\_count

haha.summary.total\_count

sad.summary.total\_count

angry.summary.total\_count

- Three: Data availability: the data shall never be revealed or sold out and the product

- Four: Data reuse & attributions: We need a legal consultation to inspect the complaint of our data plan for the EU standards and regulations.

- Five: Storage, preservation, security, integration and migration. (The programmers)

**Social media listening consists in identifying and analyzing social media activity around topic(s) of interest.**

For marketers, social listening is the process of monitoring online conversations and turning them into powerful insights that drive your business strategy. The key word here is “insights”, which are attained from online conversations.

For instance, every modern business, regardless of its size, needs to know what existing and ideal customers say about:

· their brand

· their product or service

· their campaigns

· specific keywords and topics

· their competition

**Phase one: Data Collection :**

The data collection shall rely on NLP and Lexicon based sentimental analysis achieving a representative data

As for FB a list of the most engaging public pages must be provided by the domain expert Mr. Wissam Atwan a list of the most followed and interactive palestinian pages (From Socialbakers) (n=100)

The initial data collection interval is suggested to be (2018 - 2021)

Data should be stored and secured in a cloud service

Data should be structured on a timely basis attached to a timestamp for each entry.

**Phase Two: Data Analysis and processing:**

The suggested analysis must be built over a three degree sentimental analysis (Negative, Neutral, Positive) according to a set of key words in every group to detect the public opinion regarding the issue of inspection . The keywords group is suggested to be around (10-20) words in every group. With following distinctions:

* Negative
* Neutral
* Positive
* (Additional keyword groups could be added upon need).

The analysis should explore also the numerical values and their change over time regarding the main key words. In posts, comments.

The analysis should explore the numerical values of the posts that contain keywords , engagements (Include: sharing, clicks, emojis)

**Phase three: Data representation and reporting:**

The outcomes should detect the following:

* Patterns of public opinion ratings towards certain issues based on lexical sentiment analysis over a certain period of time.
* Numerical values of engagements and reactions towards certain issues vs. time.
* Numerical analysis of hashtags and words clouds regarding certain keywords vs. time.
* A combination of all previous outcomes in a highly interactive visual-numerical report that delivers very deep and specific insights regarding the issue at hand.

# 