



# Social Media, AI and Data in Action: Case Studies

Practical Cases of Applied AI in our Social  
Media Analytics Solutions

*BREAKING NEWS ...*

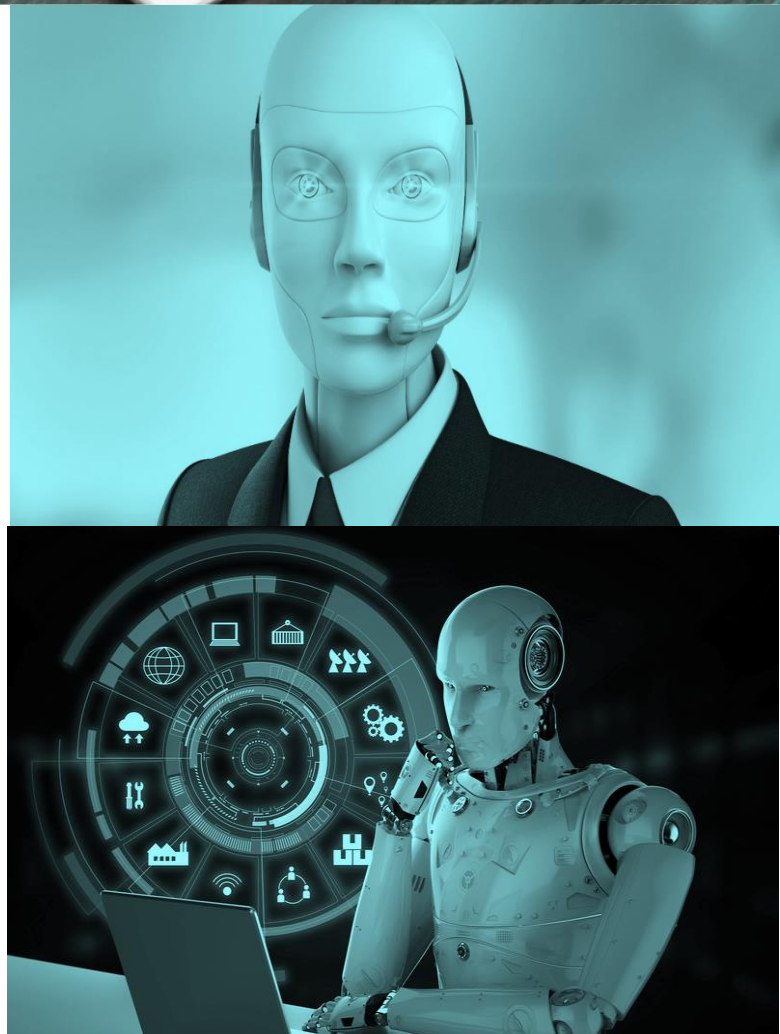
## THE AI IN AL JAZEERA NEWSROOM!

BICC  
1st floor, Buzwair Holding,  
C Ring, Doha  
Qatar



**“A computer would deserve to be called intelligent if it could deceive a human into believing that it was human”.**

**Alan Turing**



# Executive Summary of Somax Social Media Analytics

Social media has become one of the mainstream channels for broadcasting information. The user base of most of the social media channels have already transcended billion milestone. This gives business or service organizations an entry to massive number of their current and future customers.

Social media routes information with lightning speed within user communities. News, advertisement, political campaign, etc. spread within massive user graphs much faster than any existing channels. The friend of friend network helps business and organizations reach at both known and unknown communities.

Massive user base and fast information broadcasting have made social media channels critically significant to business and public service organizations.

Exploiting the full potential of social media data is not trivial. It needs unique skillset and cutting-edge AI technologies to develop custom analytics which extract meaningful intelligence from data that are enormously crowded, inconsistent, and contain irrelevant and falsified information (which are sometime spread as deliberately)

*Somax is multi-lingual social media analytics powered by Microsoft Azure with cognitive capabilities for extracting business specific insights.*

Somax is an industry neutral, customizable, extensible solution developed through extensive technical and market research to serve businesses.

# Solution Highlights

**Multi-modal Data Connection:** Somax has Multimodal Data connectors for collecting data from a multitude of sources including Twitter, Facebook, Instagram, YouTube, and Web Portal.

**Multi-layered Data Filtering:** Somax has the most advanced filtering mechanism that produce the most representative tweet set through rigorous data engineering methods.

**Know Your Customer:** With Somax business service organizations can listen to the voice of their customers.

**Know Your People:** With Somax government can listen to the voice of their customers. **Know your competitors:** Somax empowers business organizations to know what people are talking about their competitors.

**Sentiment and Opinion Mining:** Somax extracts sentiments (positive, negative, and neutral) about persons, products, services, competitor organizations.

**Trend Analysis:** Somax discovers trendy hashtags and topics.

**User Engagement:** Somax carries out statistical analysis and produce reports on user engagement.

**Entity Analysis:** Somax extracts and analyzes entities mentioned in data.

**Key Phrases Analysis:** Somax extracts and analyzes key phrases mentioned in data.

**Smart Device Integration:** Users can connect with Somax using smart phone and smart device.

**Dashboard:** It offers comprehensive dashboards for business users making critical decision.

**Scalability:** With Somax, the users gain data scalability and operational scalability.

**Customizability:** Somax can be customized to fit for business-specific needs of organization.



## Case Study

# Automated Journalism - Intelligent Media Platform

### Case Synopsis

Every minute, hundreds of thousands of information items are shared over news platforms, especially social media. In today's society, more people are getting their news instantly from online news websites, social media platforms, blogs, among others. These tools, especially social media – have become a *hotbed* of fake news, with fake news being distributed in large volumes and flooding the cyberspace. This has affected the global media industry.

News and desk editors receiving information over social media have limited opportunity to test its authenticity as they normally do with other sources. The task of filtering news is complex and burdensome. Assessing the credibility of information has become a very challenging task due to the proliferation of news, the defragmentation of sources, and biased social cues. It is a citizen's right to receive high-quality information, and the medias' role to ensure that this information is factual.

To align with this role, Al Jazeera has initiated a project to transform the media data management platform into an AI-powered platform. This project called **"Intelligent Media Platform"** is part of that initiative.

**Using AI in identifying fake news is looked upon as the cornerstone to separate the "truth" from "fake" in the news field.**

Al Jazeera is one of the largest media networks with 80 stations around the world, headquartered in Doha, Qatar. The company gathers news from many sources including social media. The journalists compose news using these sources and verify news to detect fake ones. Al Jazeera broadcasts and publishes breaking and regular news on social media, especially on Twitter.

# AL Jazeera Newsroom and the Challenge of Misinformation

## Existing Challenges and Pain Points

Detect Breaking News



Gather News



Verify News



“I want to know what’s Happening in the world, gather as much relevant data on it, and Verify it for our journalists as fast as possible!”

### What Are Missing in Running Solution of Al Jazeera?

The existing solution of Al Jazeera did not keep up with its above vision, and ended up highlighting 4 pitfalls of it:

- Data was not filtered efficiently
- On demand translation not facilitated
- It is not adequately efficient in terms of time-performance
- News verification is not facilitated

*Al Jazeera needed a solution aligned with its expectations of a competitive performance.*

### Somax for Al Jazeera

Here comes in our ultimate proposal which we thought is ideal to Al Jazeera’s business needs and can transform their pain-points to strengths:

#### Al Jazeera Challenges

Data not filtered efficiently  
 On demand translation not facilitated  
 Not adequately efficient in terms of  
 time performance  
 News verification is not facilitated

#### Somax Solutions

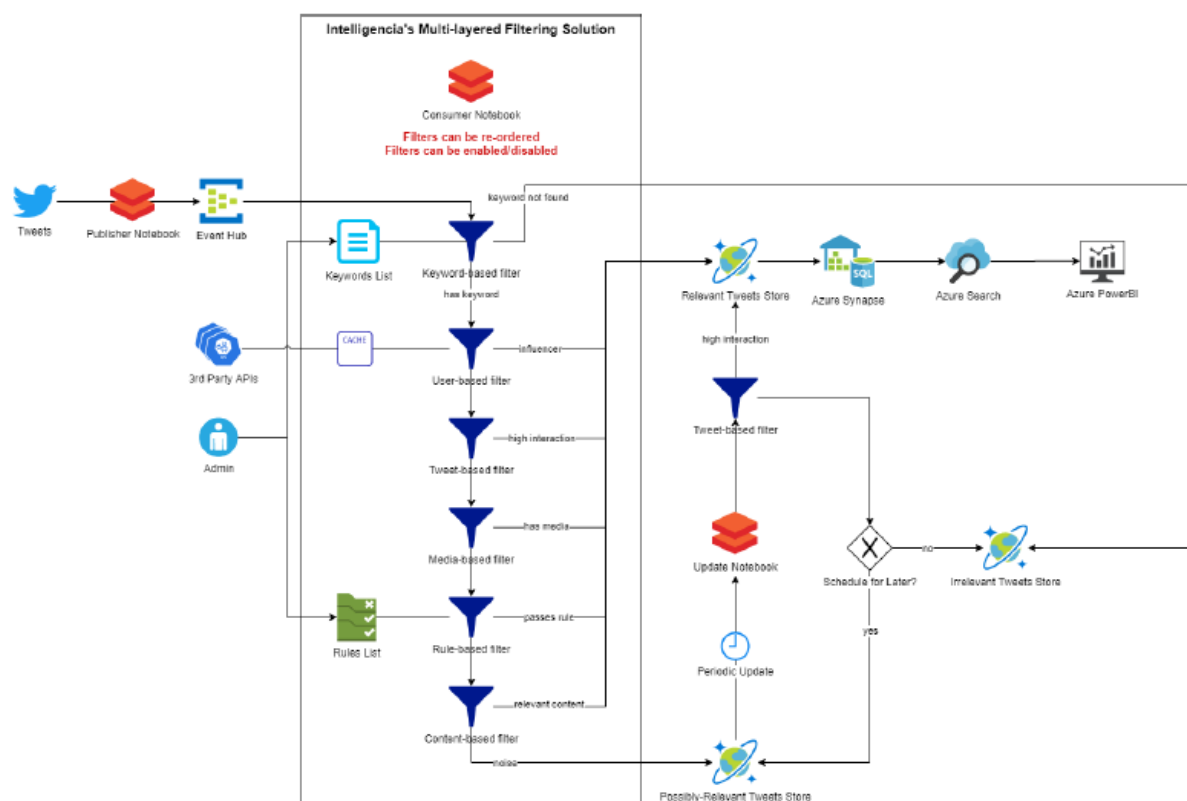
*Efficient methods for filtering news*  
*On-demand translator using MS Azure cognitive service*  
*Indexing mechanisms that help in a faster search of tweets relevant to a piece of news*  
*News verification solution using the APIs developed by QCRI*

### *Brings Data from Twitter*

Several robust and sustainable connectors have been integrated to bring data from all leading social media channels including Facebook, Twitter, and Web Portals.

### *News Content Extraction Using Multi-Layered Filtering Mechanism*

A wide variety of built-in components to prepare data efficiently and produce high quality data that are critical for decision intelligence. More importantly, for Al Jazeera offered Somax’s unique feature that is a multi-layered filtering framework (See the Figure below) for selecting the tweets that are the most representative ones for considering as news.



This feature has led a huge transition for Al Jazeera journalists who today can use tweets to prepare news contents.

### *News Content Analytics*

Al Jazeera has a solution called “Tanbih” developed in collaboration with QCRI. Tanbih was developed specifically for verifying tweets. The BICC team implemented a backend system for integrating Tanbih within the running newsroom BICC-intelligent platform ecosystem.

# Somax for Al Jazeera Media Network Features

In addition, the backend system enables the Al Jazeera team to integrate reverse image search API or any other API Al Jazeera. The solution is currently being tested. The solution is yet to be productionized

## Comprehensive Executive Dashboards

Somax offers a list of dashboards including (the Figure below shows hashtag analysis dashboard) The dashboard provides a comprehensive visual report about tweets, users, and hashtags related to Al Jazeera.



Also, it shows the location of the hashtags, trendy hashtags, hashtag timeline, hashtag statistics, tweets count location, and original tweets.



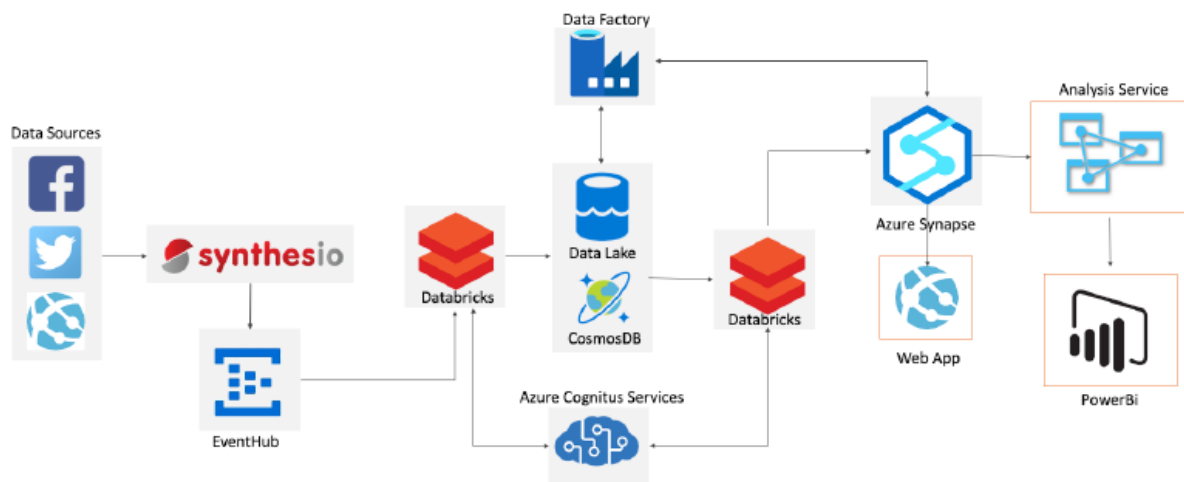
# Somax for Al Jazeera Media Network Architecture

## *Eventhub Grabs & Ingests Tweets in Seconds*

Eventhub, the Microsoft's real-time streamer collect tweets and ingest them in data engineering and analytics platform in seconds.

## *Spark of Azure Data Bricks is the Key to Develop Realtime Analytics*

Somax is powered by state-of-the-art Microsoft Azure technologies. It is grounded Microsoft Azure Data Bricks Spark Framework which is an extremely parallel computation technology and a *de facto* choice for real-time industrial data engineering and analytic. Al Jazeera required a real-time data Analytics engine.



**Figure 1:** The High-Level Architecture Somax for Al Jazeera

## *Azure Cognitive Service for Extracting Intelligence*

Cognitive service is used for Al Jazeera use case to extract entities and key-phrases mentioned in tweets. Cognitive services are an umbrella of various APIs include Computer Vision, Text Analytics, Language Translators etc. We used text analytics and translator for Al Jazeera use case. The translator would translate English to Arabic and Vice Versa

# Somax for Al Jazeera Media Network Features

*PowerBI Gives the Best Experience of Comprehensive Dashboard and Visual Analysis*

PowerBI is the ultimate choice visual data analysts and high-end users such as executives to view analysis results and make critical decision. Somax offers AL Jazeera executives, journalists, and data analytics highly rich and easy to use dashboards developed using PowerBI.