



# Social Media Data Engine

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824282

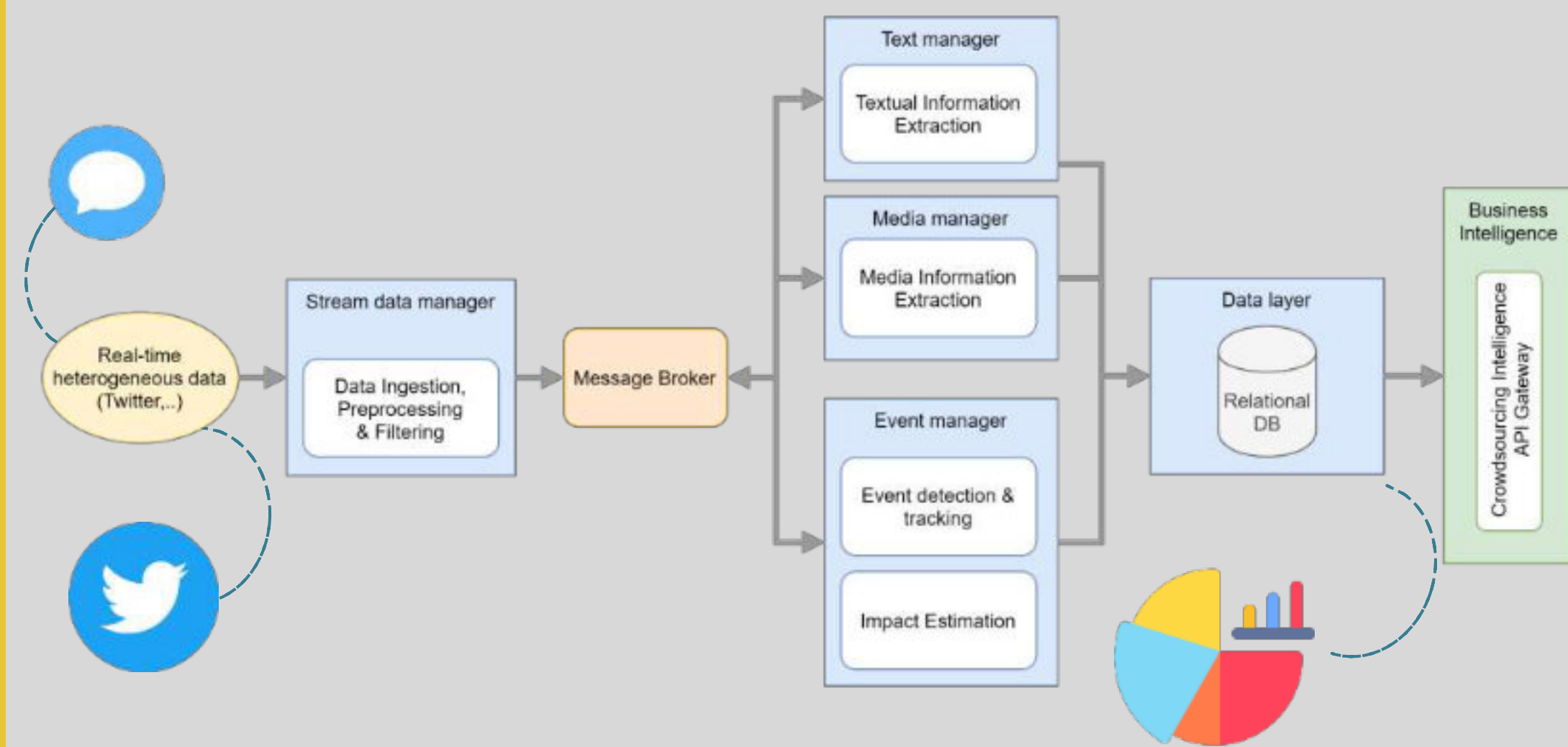
The SHELTER social media data engine was designed in order to be able to fetch in real time social media data, with a focus on Heritage Areas, and automatically classify the content using advanced text analytics, image processing and deep learning algorithms. These models extract meaningful information related to emergency situations, early warning signals, and data to complement damage assessment. The social media data engine also provides a smart dashboard for data filtering and visualization.

## How

### The real-time streaming pipeline.

The Social Media Data Engine (SMDE) automatically gathers tweets and their contents through Twitter's streaming API.

Multimedia data is processed by several ML models, to extract: **informativeness, information type, hazard type, and named entities.** This tweet-level information is lastly grouped into more meaningful clusters, which will eventually become **events.**



## Results

### Performance and numbers of the pipeline.

The platform can handle thousands of tweets in real-time, with an average F1 score of **0.88 in informativeness**, and **0.76 in information type** classification.

Since its deployment, the module detected:

- More than **104K** hazardous events around the globe
- More than **53M** individual tweets

## Functionality

### SMDE interactions with the Shelter framework.

The module mainly works through standard **REST API**, providing enriched information to the web dashboard in **detailed** (tweet-level) or **aggregated views** (events).

Additionally, the module also sends **event notifications** through the **message bus**.

