DELIVERABLE SUBMISSION SHEET

To: Susan Fraser *(Project Officer)*

EUROPEAN COMMISSION
Directorate-General Information Society and Media
EUFO 1165A
L-2920 Luxembourg

From:

Project acronym: PHEME Project number: 611233

Project manager: Kalina Bontcheva

Project coordinator The University of Sheffield (USFD)

The following deliverable:

Deliverable title: Digital Journalism Prototype (v2)

Deliverable number: D8.3.1

Deliverable date: 28 February 2017

Partners responsible: I-Hub Limited (USH)

Status: ☒ Public ☐ Restricted ☐ Confidential

is now complete. ☒ It is available for your inspection.
☒ Relevant descriptive documents are attached.

The deliverable is:

☒ a document
☐ a Website (URL: .....)
☐ software (.....)
☐ an event
☒ other (....Prototype.....)

Sent to Project Officer:	Sent to functional mail box:	On date:
Susan.Fraser@ec.europa.eu	*CNECT-ICT*-611233*@ec.europa.eu*	02 March 2017

FP7-ICT Strategic Targeted Research Project PHEME (No. 611233)
Computing Veracity Across Media, Languages, and Social Networks

D8.3.1 Digital Journalism Prototype (v.2)

David Losada
Ivelina Nikolova
Andrey Tagarev
Geraldine Wong Sak Hoi

Abstract

FP7-ICT Strategic Targeted Research Project PHEME (No. 611233)
Deliverable D8.3.1 (WP 8)

A new version of the digital journalism dashboard prototype is now available online. This updated document contains the URL of the prototype, along with a short manual that guides users through the dashboard, plus a brief description of the technical approach to creating the prototype. It also features a description of Hercule, the web-based fact check assistant, being developed by Ontotext as a new subtask to WP8, and work achieved to date in this regard.

Keyword list: journalism dashboard, prototype, rumour verification, rumours, fact checking

Nature: Prototype and Report
Dissemination: PU
Contractual date of delivery: 28 Feb 2017
Actual date of delivery: 2 Mar 2017
Reviewed By: Kalina Bontcheva
Web links: http://pheme-jd.ushahidi.com, http://hercule.ontotext.com/

Research partially supported by the European Union under the Information and Communication Technologies (ICT) theme of the 7th Framework Programme, grant PHEME (611233)

PHEME Consortium

This document is part of the PHEME research project (No. 611233), partially funded by the FP7-ICT Programme.

University of Sheffield
Department of Computer Science
Regent Court, 211 Portobello St.
Sheffield S1 4DP, UK
Tel: +44 114 222 1930
Fax: +44 114 222 1810
Contact person: Kalina Bontcheva
E-mail: K.Bontcheva@dcs.shef.ac.uk

Universitaet des Saarlandes
Language Technology Lab
Campus
D-66041 Saarbrücken
Germany
Contact person: Thierry Declerck
E-mail: declerck@dfki.de

MODUL University Vienna GMBH
Am Kahlenberg 1
1190 Wien
Austria
Contact person: Arno Scharl
E-mail: scharl@modul.ac.at

Ontotext AD
Polygraphia Office Center fl.4,
47A Tsarigradsko Shosse,
Sofia 1504, Bulgaria
Contact person: Georgi Georgiev
E-mail: georgiev@ontotext.com

ATOS Spain SA
Calle de Albarracin 25
28037 Madrid
Spain
Contact person: Tomás Pariente Lobo
E-mail: tomas.parientelobo@atos.net

King's College London
Strand
WC2R 2LS London
United Kingdom
Contact person: Robert Stewart
E-mail: robert.stewart@kcl.ac.uk

iHub Ltd.
Ngong Road, Bishop Magua Building
4th floor
00200 Nairobi
Kenya
Contact person: Rob Baker
E-mail: robbaker@ushahidi.com

SWI swissinfo.ch
Giacomettistrasse 3
3000 Bern
Switzerland
Contact person: Peter Schibli
E-mail: Peter.Schibli@swissinfo.ch

The University of Warwick
Kirby Corner Road
University House
CV4 8UW Coventry
United Kingdom
Contact person: Rob Procter
E-mail: Rob.Procter@warwick.ac.uk

Contents

1 Introduction

The open-source digital journalism prototype aims to harness the systems being developed within PHEME and present them in a dashboard geared specifically at journalists looking to quickly locate and verify information online.

The PHEME journalism dashboard extends the open-source Ushahidi platform with PHEME technology for veracity and socio-semantic intelligence. This deliverable has been a close collaboration between technical partners (iHUB, ATOS, ONTO, USFD, UWAR and USAAR) and the use case partner (SWI).

Taking the user requirements and interface mock-ups from T8.1 as its point of departure, iHub has developed the dashboard in an iterative process involving regular discussions with project partners, adjustments based on the status of various PHEME technologies, and outcomes of user evaluations.

In addition, in Year 3 of the project, Ontotext began working on a new subtask of Work Package 8. It aims to develop a tool (dubbed Hercule) to assist professional fact checkers in locating and verifying questionable factual claims made in social media and online news.

Following this introduction and a brief look at the relevance of D8.3 to other work packages, Section 3 of this document provides a short user manual designed to guide the user through the current prototype. Section 4 describes the technical elements supporting the dashboard development. Section 5 is dedicated to a description of the Hercule dashboard in its current state of development and functionality. We conclude with directions for future work.

2 Relevance to PHEME

2.1 Relevance to project objectives

Task 8.3 is at the core of the project, as it seeks to transpose the project results into a real world setting by creating a digital journalism showcase for the modelling, identification and verification of rumourous stories, spreading on social media networks.

The latest development of the Hercule fact-checking assistant was motivated by the sleuth of false or misleading claims made during key political campaigns, and in particular, the UK EU membership referendum (ako Brexit) and the US presidential elections. As evidenced by the creation of the First Draft News initiative, fact checking and the development of new tools to assist that, have now become a core concern for both traditional news media and social media sites alike. Becoming an academic member of First Draft News and the development of Hercule is the PHEME contribution in this context.

In more detail, Hercule builds on the project technologies for semantic enrichment of social media content and provides quick access to automatically clustered stories. By interlinking the social media content to related news and Linked Open Data knowledge, it offers an extended context of the analysed content, and thus provides means for journalists to fact-check online information.

2.2 Relation to other work packages

Task 8.3 has been closely linked to, and highly dependent on, numerous other work packages that are concerned with the development of specific technologies being harnessed by the journalism dashboard and Hercule. These are in particular: WP2 (Ontological modelling, multilinguality, spatio-temporal grounding), WP3 (Contextual interpretation), WP4 (Detecting rumours and veracity), WP5 (Interactive visual analytics dashboard), and WP6 (Scalability, integration and evaluation).

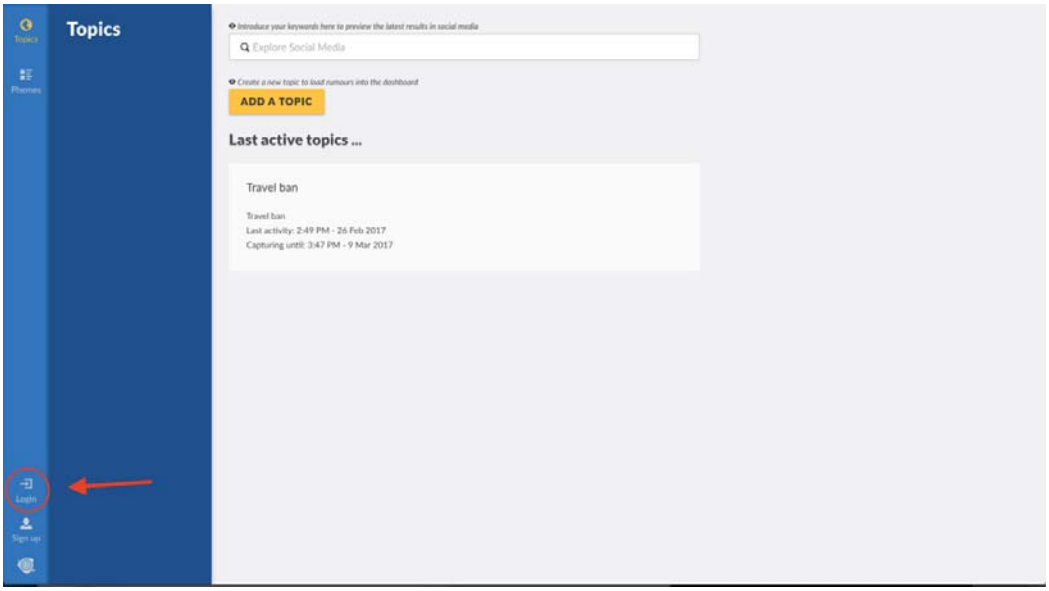
3 The digital journalism prototype: User's guide

The prototype is available at http://pheme-jd.ushahidi.com .

In order to perform any operations, the user must be logged in with the following credentials:

email address: demo@ushahidi.com
password: fp7Pheme

See the following screenshot for the location of the Login button:

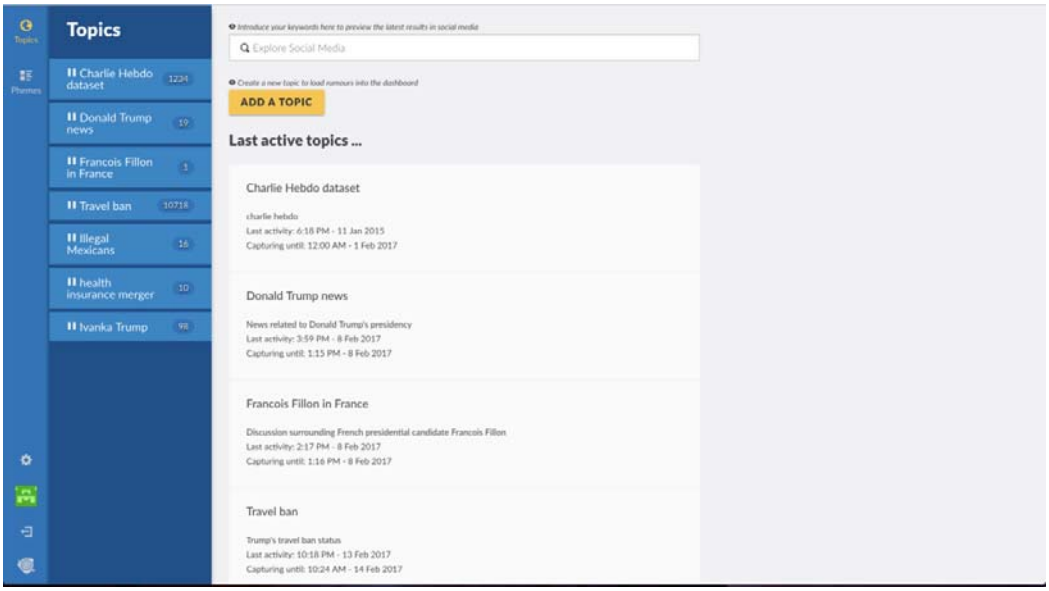


The prototype is currently fully functional for evaluation purposes.

3.1 Listing Topics

A Topic refers to a persistently running search over social media platforms. When a journalist decides to find information about a real-world event that is unfolding, he/she will create a Topic in the dashboard and provide keywords associated with the search in question.

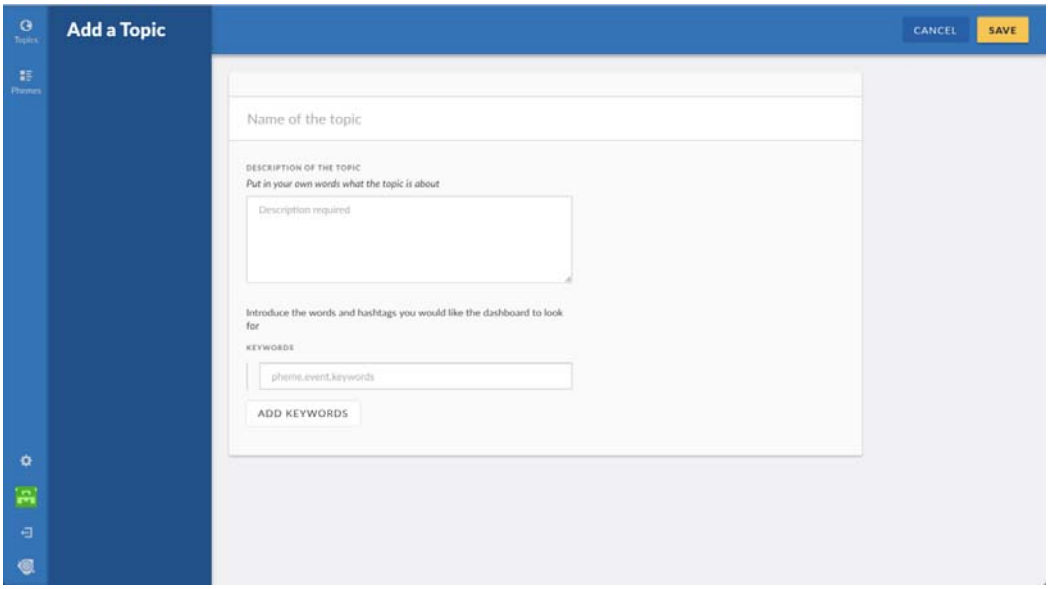
When the user first accesses the dashboard, he/she is presented with the list of Topics currently being streamed into the system.



In the above screenshot, seven Topics are registered in the system.

3.2 Creating Topics

Topics may be created from the Topic list screen by clicking on the yellow "Add a Topic" button. The following screen is presented after the click:



Please note that all fields as shown in the screenshot above must be completed.

After clicking "Save", the search parameters will be sent to the system and queries across social media will be initiated. Any messages from that moment on will be captured by the system.

3.2. Keyword groups and system resources

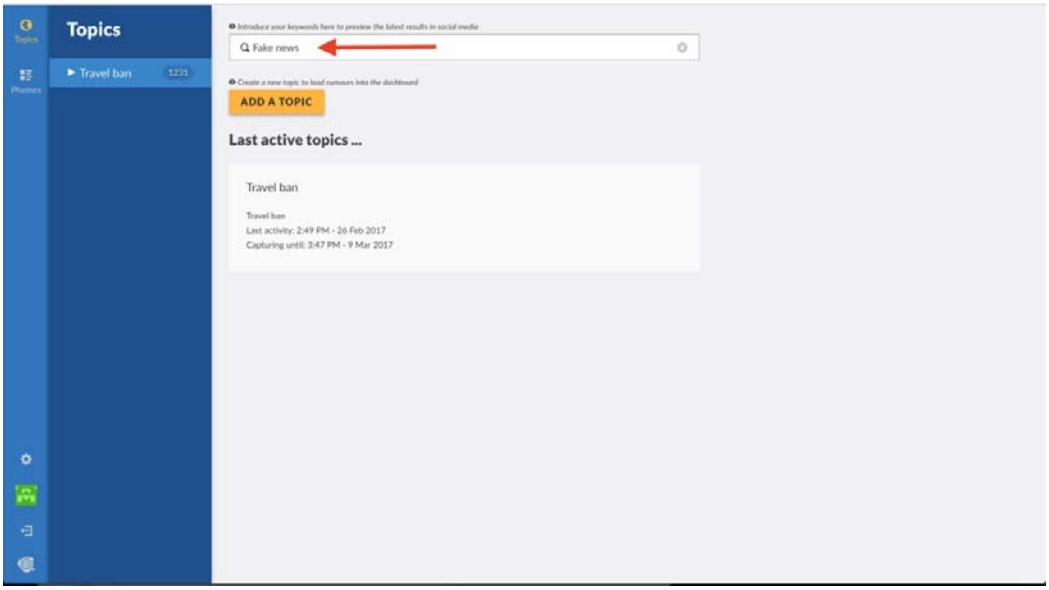
When creating a Topic, several keyword groups may be provided. It is important to understand the following:

- Writing "foo bar" in a single keyword group will create exactly one running search, matching messages containing both the words "foo" and "bar"
- Writing "foo" and "bar" in different keyword groups will create two different running searches, one matching messages with the word "foo", and another matching messages with the word "bar". Compared to the first option, the second option will result in a higher number of transactions and a higher volume of matches, consuming more resources overall.

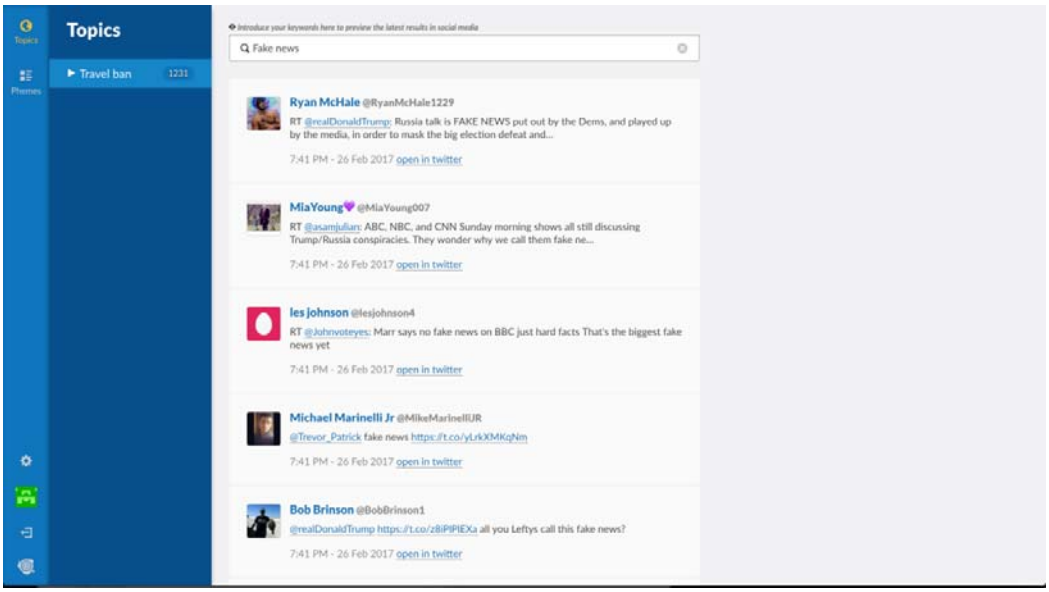
For fine-tuning the queries performed, it's useful to note that the underlying social media collection system used by the PHEME project follows Twitter's query syntax .

3.3 Previewing keyword results

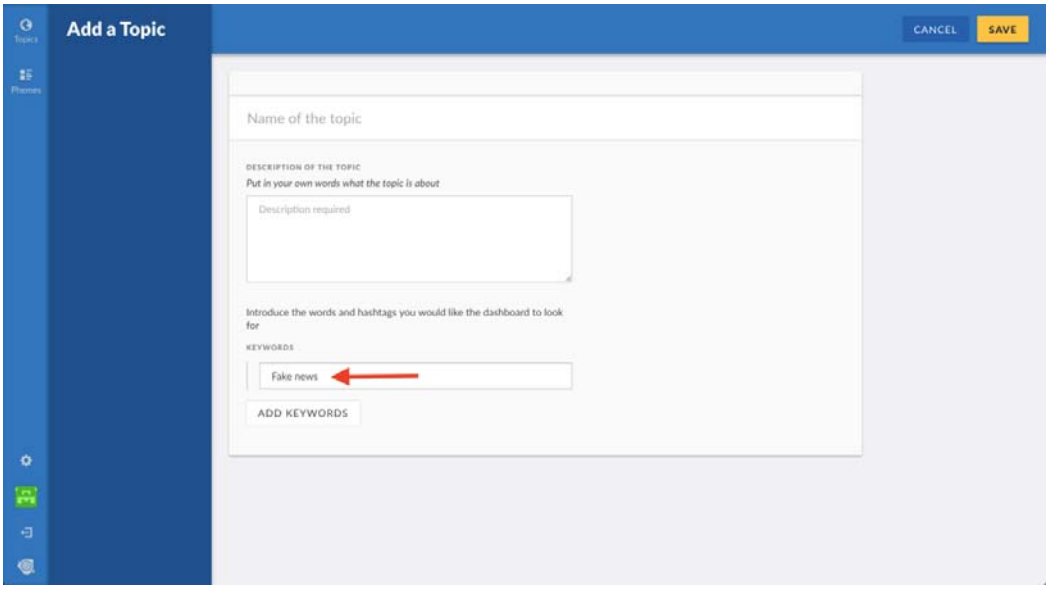
As we have seen, creating a Topic can place a significant data and processing load on the system, should the keyword(s) match many social media posts. In order for the user to evaluate keyword search results, prior to committing to creating a topic permanently, we provide the "Explore Social Media" functionality.



This immediately returns live results from Twitter:



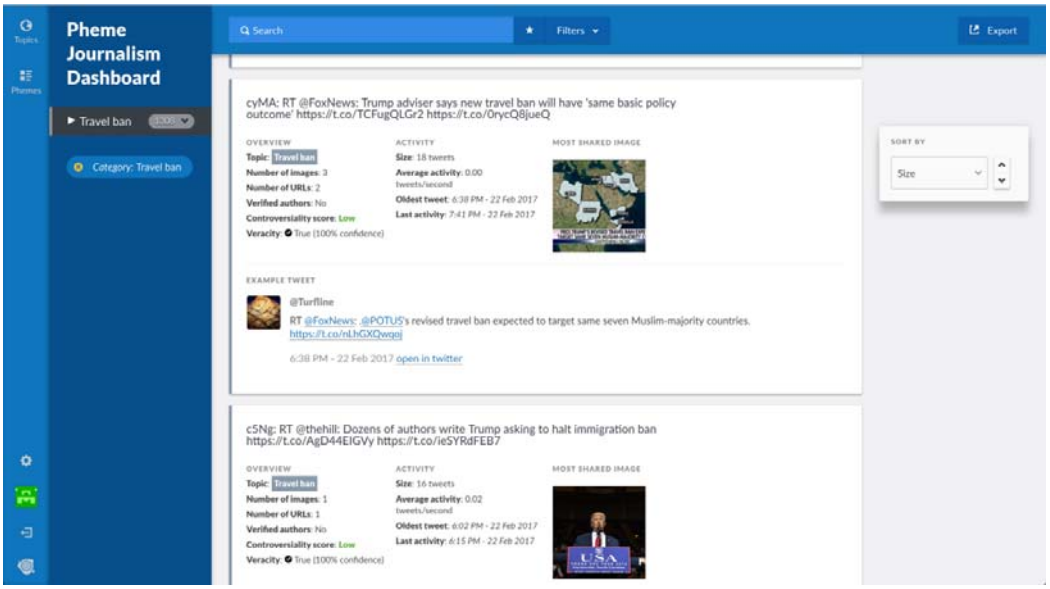
Clicking the "Add a topic" button at this point will bring over your keywords to the creation screen, so the user does not have to type them again:



3.4 Browsing Phemes detected within Topics

Phemes are groups of topically-related social media posts that have been clustered together automatically, using the tools developed in WP3. This grouping can be useful for identifying fast growing memes/trends and allowing the user to "cut through the noise" and find the more relevant and rumorous conversations faster.

To see the list of Phemes, the user clicks on the "Phemes" icon.

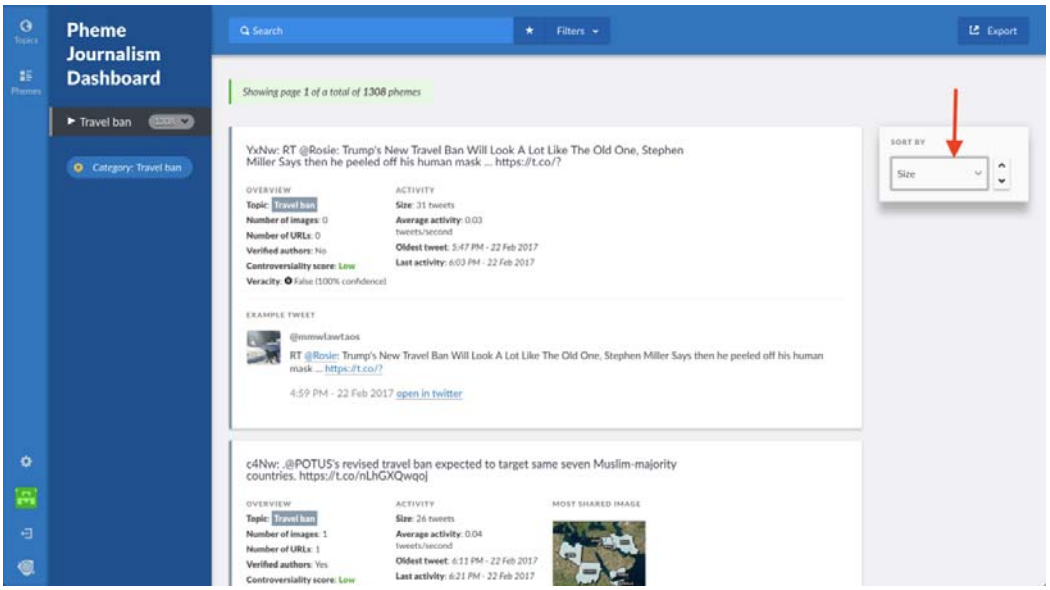


The screen displays a list of Phemes that have been detected by the system for the created Topics. For each Pheme, summary information is available. This information includes:

- *Pheme title*: a title for the Pheme, inferred from the content of all the messages contained in this set.
- *Oldest tweet*: date and time of the oldest message in the Pheme
- *Last activity*: date and time of the most recent message in the Pheme
- *Size*: number of messages in the Pheme
- *Speed*: rate at which messages have been added to the Pheme
- *Number of images*: how many links to images have been embedded in the messages
- *Number of URLs*: how many links to articles have been embedded in the messages
- *Verified authors*: presence of messages created by Twitter verified authors
- *Controversiality score (high, medium, low, uncertain)*: an indication of the controversiality of the discussions present in the Pheme.
- *Veracity*: an estimation of the veracity of the conversations included in the Pheme
- *Example twee*t: a message selected from among the group of messages as a representative account of the content of the Pheme. Message from verified, popular authors are favored here.

3.5 Sorting and filtering

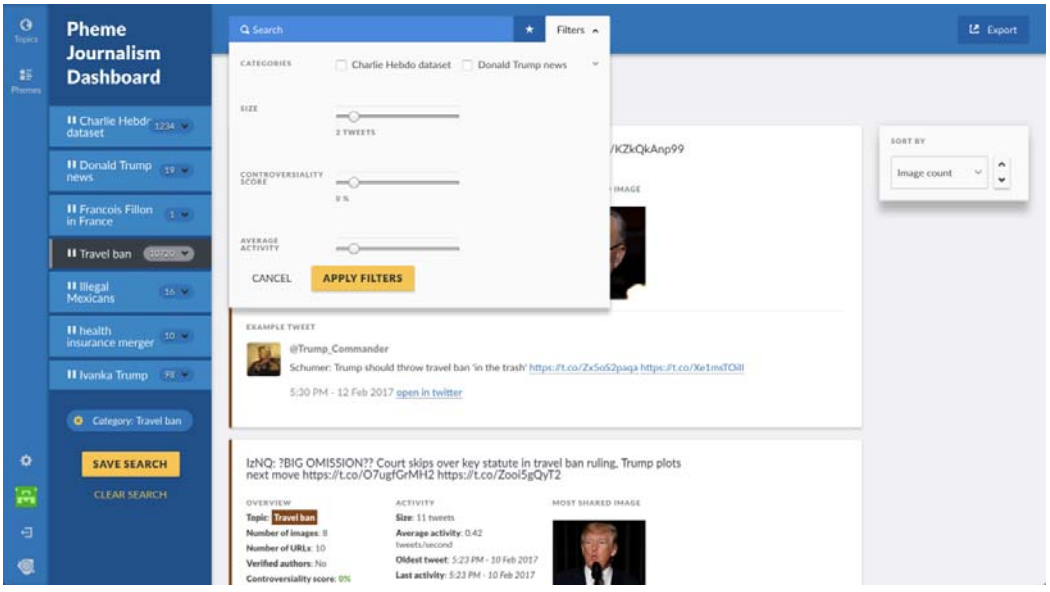
It is possible to change the order in which the Phemes appear in the list.



The current sorting options are: size of Pheme, most recently updated, controversiality, image count and URL count.

It is also possible at this point to do some filtering over the list, in order to:

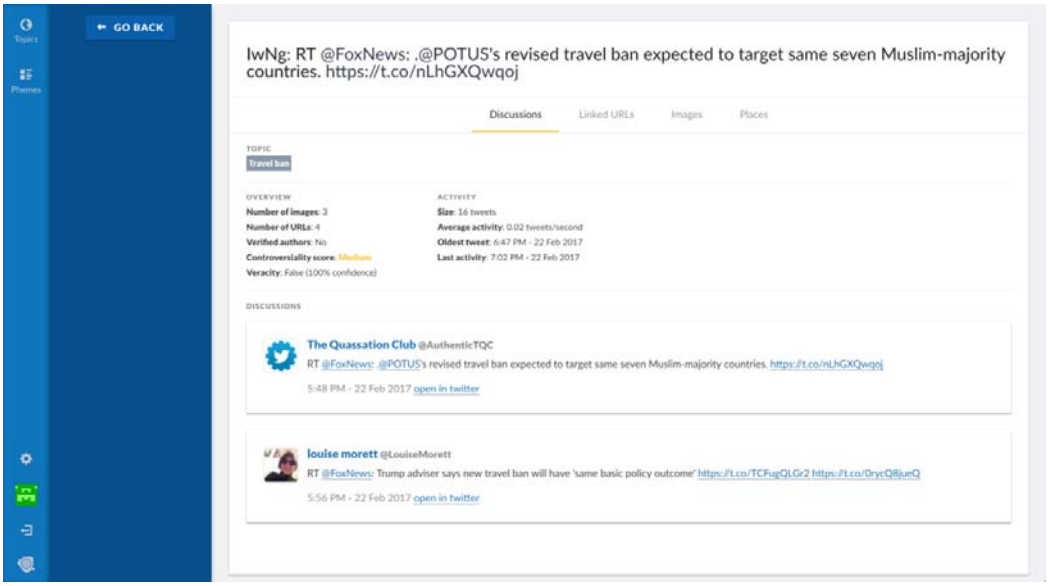
1. see the Phemes that belong to one (or a subset) of the Topics registered in the system (left hand side bar)
2. see the Phemes that comply with some criteria, like size or average activity (Filters toolbar)
3. see the Phemes that match one or several text keywords (Search bar)



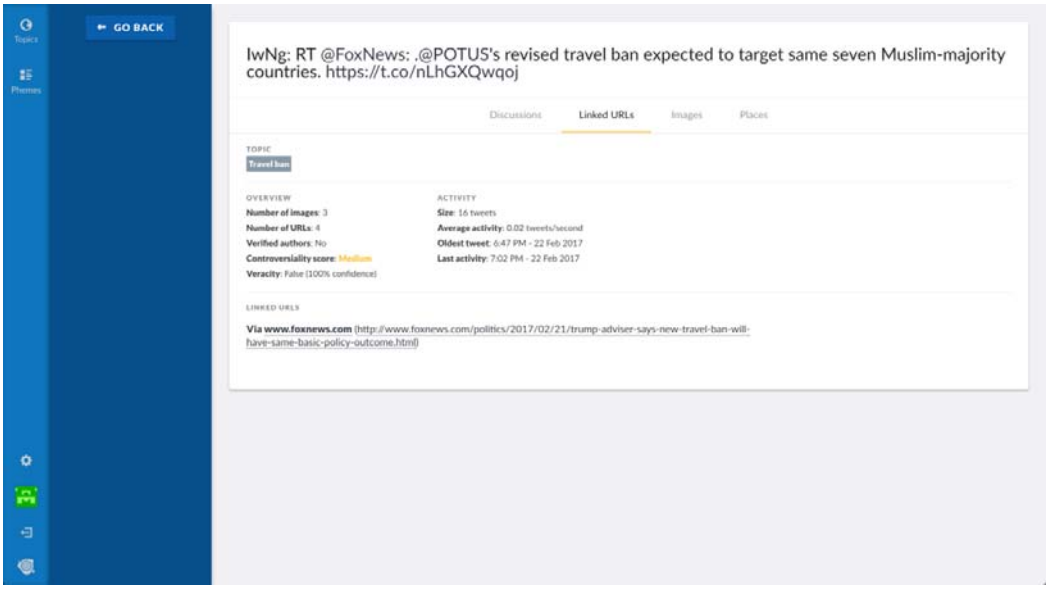
3.6 Pheme detail view

Clicking on any Pheme title in the list will open a detailed view of that Pheme. This detailed view contains several tabs.

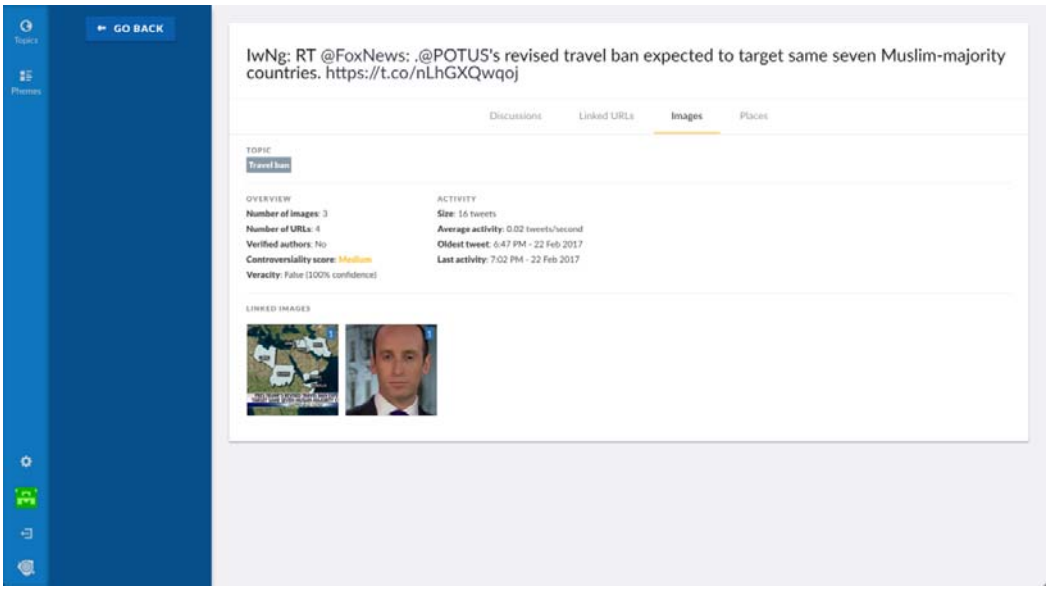
The first tab lists the discussions present in the Pheme. Each discussion is represented by its most representative tweet (favouring verified and most followed authors).



A summary of all linked URLs is available in the second tab.

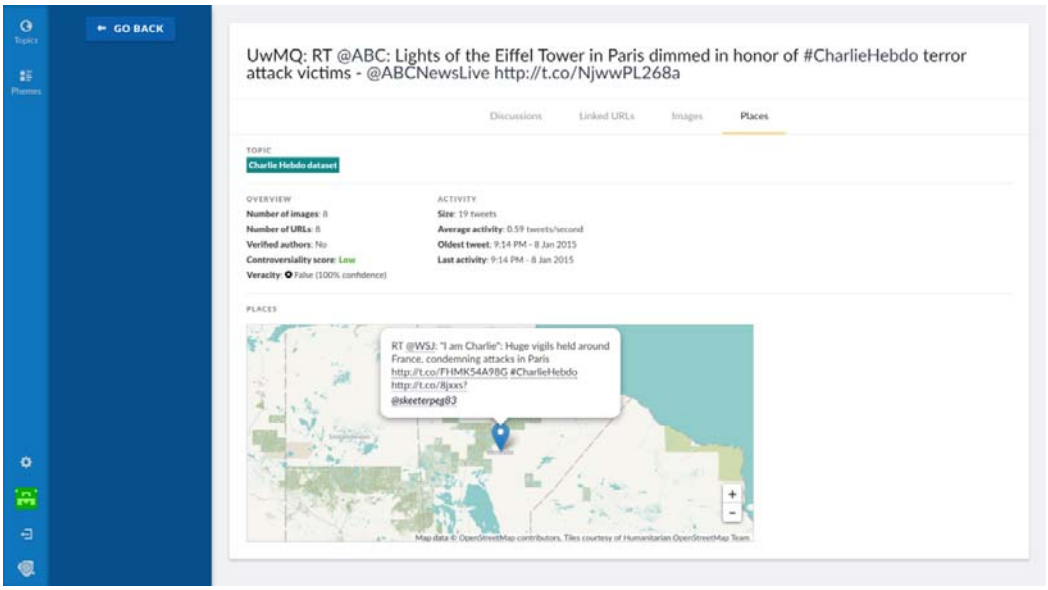


A preview of images linked in the messages is available in the third tab.



Finally, the places tab displays the locations of the authors, who posted those tweets. It should be noted that latitude/longitude information is typically available for only around 10% of all tweets.

In addition, since we try to infer locations automatically, when precise coordinates are not available, the pins may not always be placed at the correct location.



4 The digital journalism dashboard: Technical approach

4.1 Components

The components that support the operation of the digital journalism dashboard can be divided into three categories:

- **Components for conversation collection and analysis**. As delivered (WP2, WP3, WP4) and integrated (WP6) through other work packages in the project.

- **The digital journalism dashboard web application**. This is a modified version of the latest Ushahidi Platform. It consists of the following components:

    - The database server. This is a standard relational database, more specifically MySQL. It stores the data that is most frequently displayed in the dashboard as well as the profiles and information associated with users of the system.

    - The server side component (or API). This is web request processing code, coded in PHP following the MVC pattern. Modifications have been introduced to communicate with the rest of the system through the interface component detailed below.

      The source code is available at https://github.com/project-pheme/project-pheme-platform.

    - The client side application, which loads in the user browser. This is a fully dynamic web application running in the browser, coded in Javascript, HTML and CSS with AngularJS, and following the MVC pattern. It is pre-designed to work well on desktop computers with large screens, laptops and tablets.

      This component is undergoing more extensive customisation, in order to adapt the presentation of information and interaction model of the application to the necessities of the collection and analysis package, as well as those of the journalists potentially using the system.

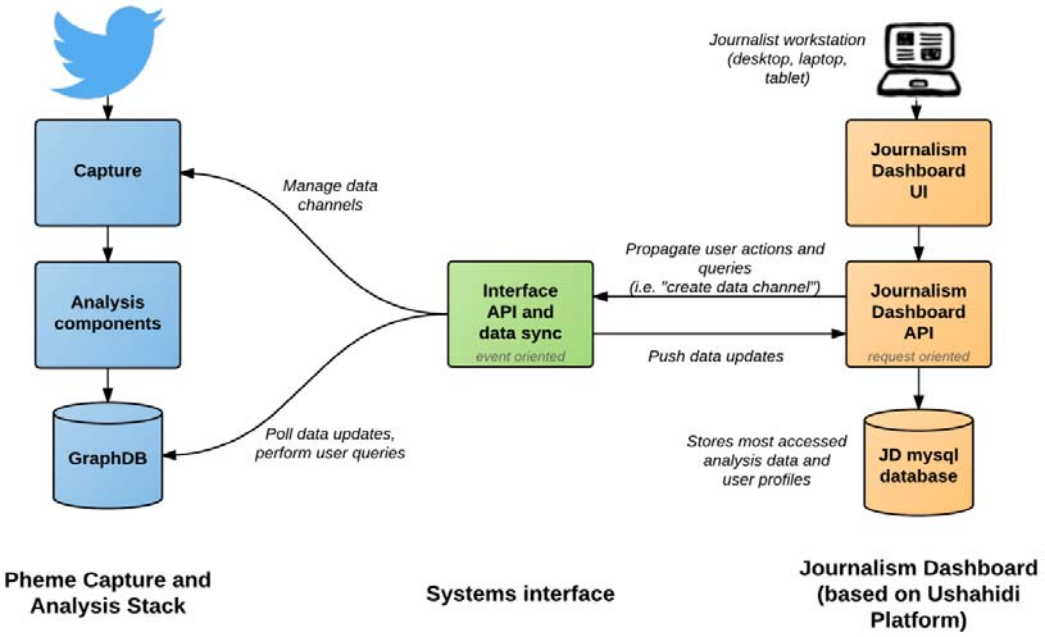
      The source code is available from https://github.com/project-pheme/project-pheme-platform-client

- **Interface components**. This component handles the communication with capture and analysis stack elements, in order to track the evolution of the events, threads and rumours. Thanks to this approach, no significant architectural changes need to be introduced into the Ushahidi platform code.

  This component is a continuously running application, built around an event loop, using Python and the Tornado framework.

  The source code is available at https://github.com/project-pheme/project-pheme-data-interface .

The following diagram illustrates the components and their relationships.



5 The Fact-Checking Assistant - Hercule

5.1. Overview

The fact-checking assistant Hercule is a web-based portal that aims to help journalists with the daily tasks of sorting and retrieving newsworthy pieces of information from social media (Twitter). Similar to the journalist dashboard, Hercule allows its users to create topics of interest by entering a set of keywords. Once created, the **topics** are automatically populated with tweets containing the same pre-defined **keywords**.

With the help of the PHEME named entity recognition and resolution tools (linking objects to the respective concept in Linked Open Data), and the application of high-confidence classifiers for rumour detection, veracity calculation and "check-worthiness" calculation, each tweet is enriched with new features. Based on all of these enrichments, the tweets within each topic are automatically clustered in the form of more focused **stories** which are then presented to the user.

The user can sort the stories in the topic according to their size, level of check-worthiness and recency. In addition, the user can choose the order in which separate tweets appear in the story based on the level of check-worthiness, rumour, veracity score or time of publication. In this way tweets which are worth checking appear at the top, so the journalist does not have to scroll through the whole story, which may sometimes comprise hundreds or thousands of tweets.

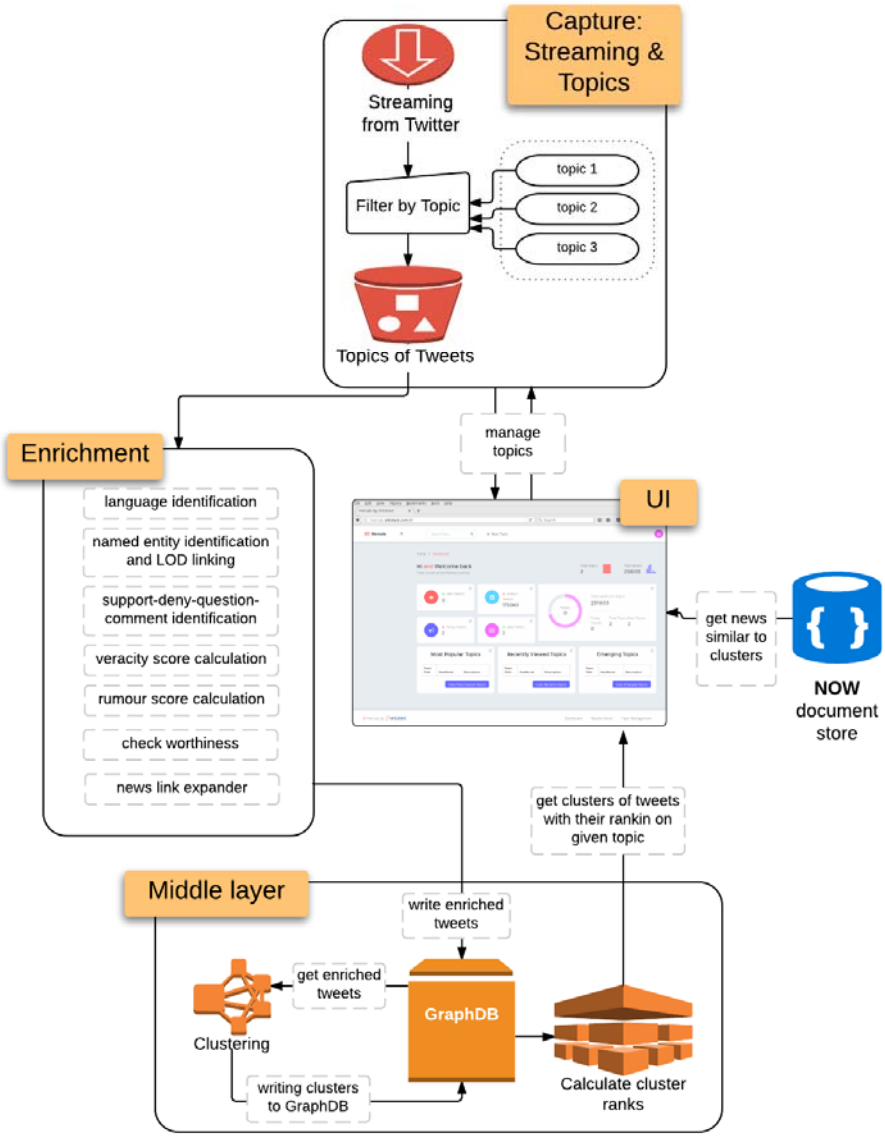
The main advantages of Hercule rest with the various features that can be used to fact-check the content presented in the dashboard. Each story can be visualised together with the **concepts** mentioned in the individual tweets (e.g., names of persons, organizations, and locations) and with **news articles** related to the story. These features provide greater context to the story, thereby facilitating the verification of claims on the social network. The user can

explore each related concept and news article with a single click and quickly obtain the information needed to help with fact-checking the tweets' content.

The demonstrator can be found at: http://hercule.ontotext.com
user: hercule
password: 0nt0text

5.2. Overall architecture

The overall architecture of Hercule is presented in the following figure:

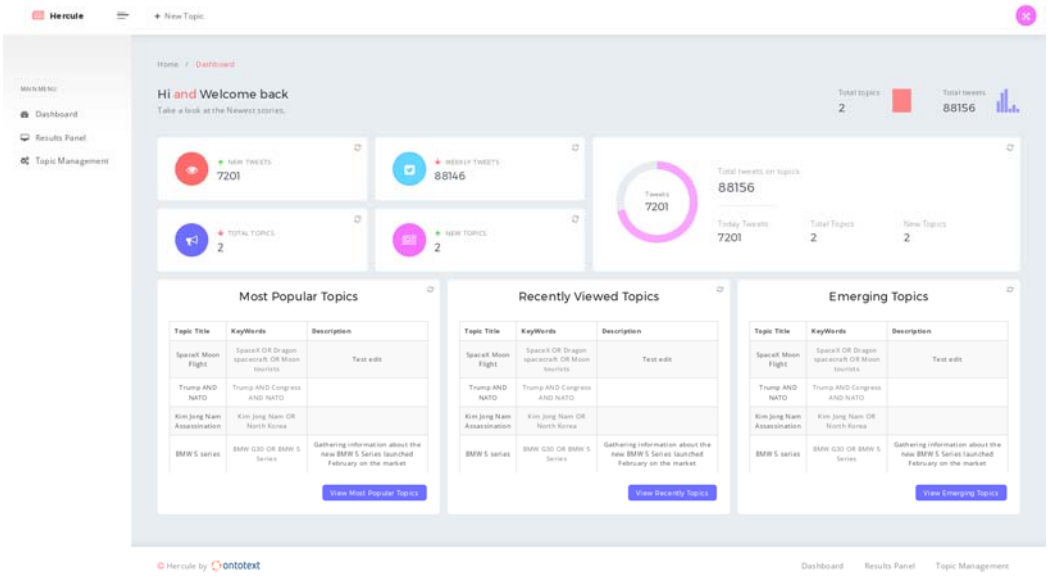


Hercule is comprised of four main components:

- *Capture component* - responsible for the collection of social media posts in topics and topic management. It monitors the web for new tweets which contain the keywords describing the topics in the system and then collects these.
- *Semantic enrichment environment* - composed of various natural language processing components for semantic enrichment of the input tweets: language detection, named entity recognition and linking, support-deny-question-comment identification, veracity score calculation, rumour score calculation, component for checking the worthiness of each tweet, and a link expander which crawls the content of articles referred to in the tweets.
- *Middle layer* - responsible for combining tweets into stories and calculating story metadata. The clustering component inserts into existing stories or creates new stories when there are no appropriate existing ones. It is based on Twitter metadata (retweet, hashtags, links) and concept enrichments.
- *UI* - the user interface component communicates with three other components upon user request:
    - when the user manages or creates new topics, the UI accesses the *Capture* component;
    - when visualising results for the user, the UI acquires tweets and stories with their rankings from the *Middle layer*, and
    - along with that it requests similar articles from the service *NOW* (http://now.ontotext.com), where news articles have already been enriched. *NOW* recommends articles based on the top concepts in a story.

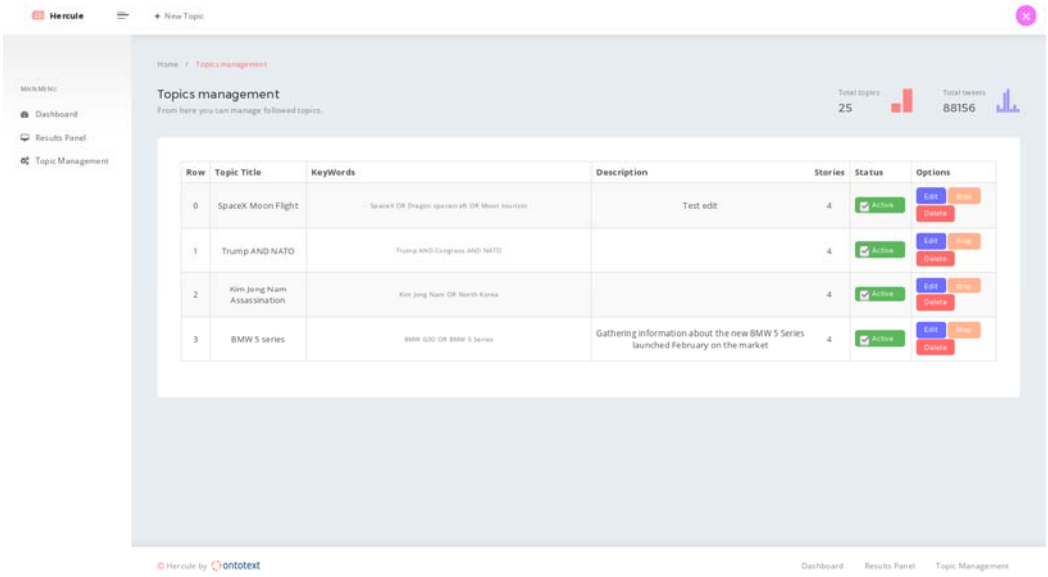
5.3 Hercule Functionality

Landing page



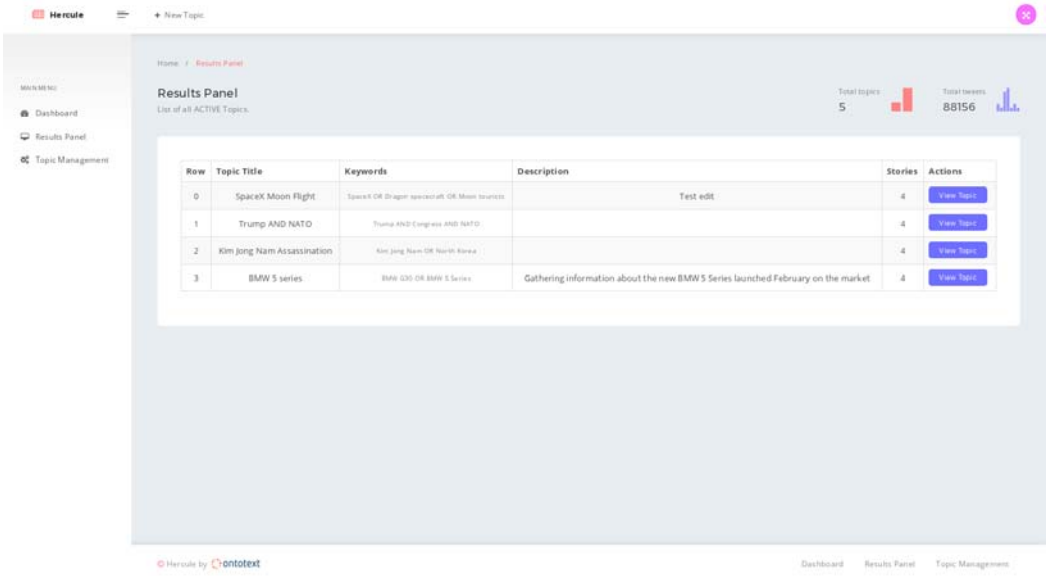
The landing page provides a quick overview of the data available in the system, including total, weekly and daily counts of ingested posts and existing topics. It also proposes a few different ways of exploring topics of interest in the system: by size or number of posts (Most Popular); by topic last viewed (Recently Viewed); and by fastest growing (Emerging Topics). The users can follow links to the Results Panel or the Topic Management page.

Topic Management



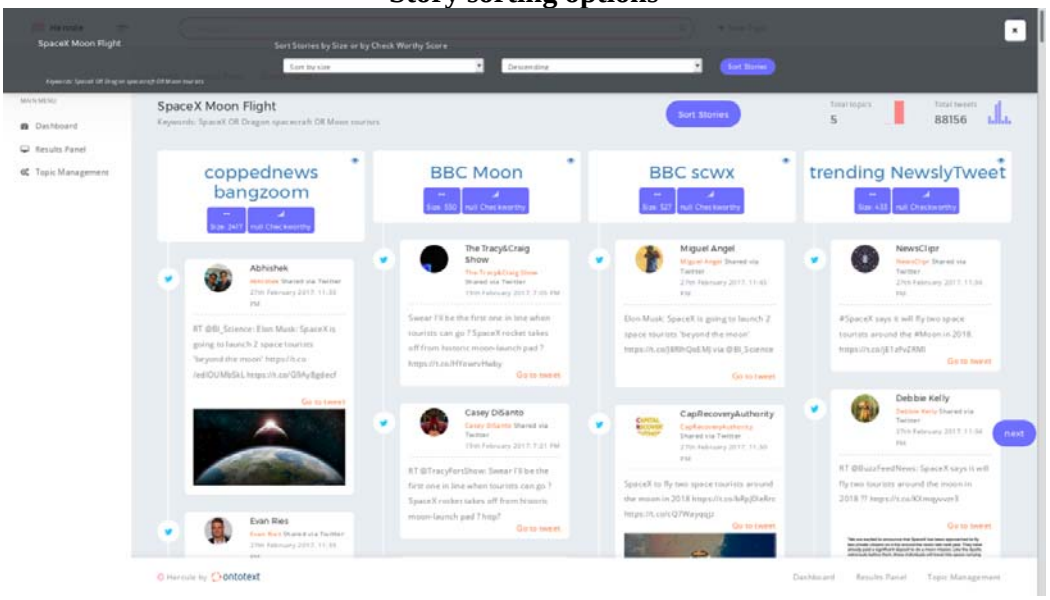
The topic management page allows users to make modifications to existing topics. This includes *editing* (change title, keywords, description), *stopping* the ingestion of new tweets, and *deleting* the topic from the system. While *delete* does not automatically remove tweets from the repository, it does make the topic permanently invisible to users.

Results Panel



This page shows a list of all topics in the system and offers users an entry point to view them in further detail. The user clicks the “View Topic” button associated with whichever topic they wish to explore.

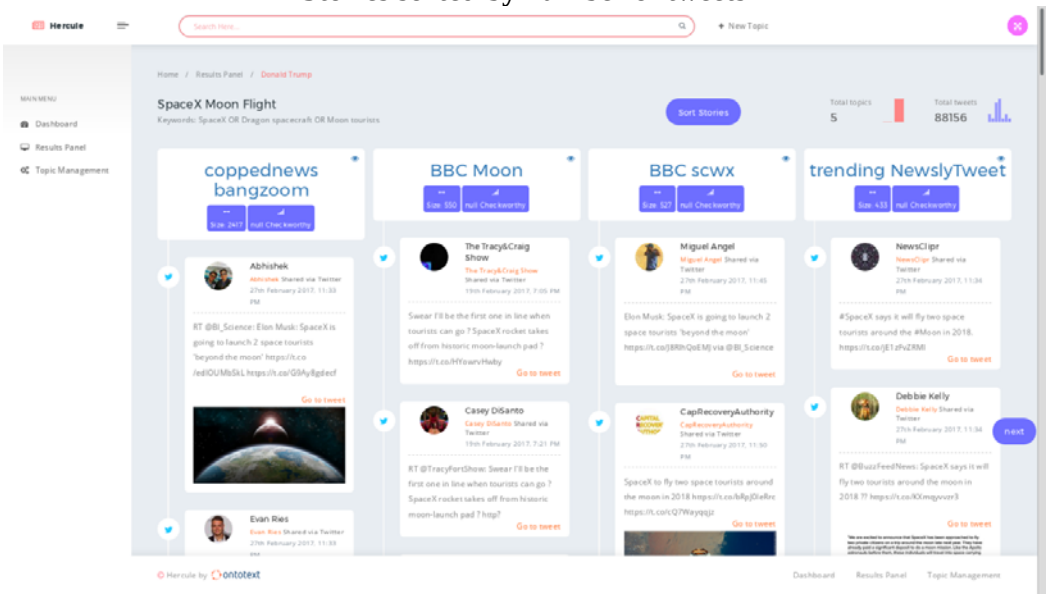
Story sorting options



The story view screen displays four stories at a time. The *next* button on the right allows users to move farther down while the *sort stories* button on top enables ordering them in different ways - by size, recency and "checkworthiness" score. For example, choosing sort by size - descending would return the largest stories first.

An important tool to highlight at this stage is the text search bar at the top of the page. This allows users to search for stories and tweets which contain certain words. These narrowed-down results are still ordered based on the sorting option specified by the user and allows the person to quickly reduce the scope of large topics.

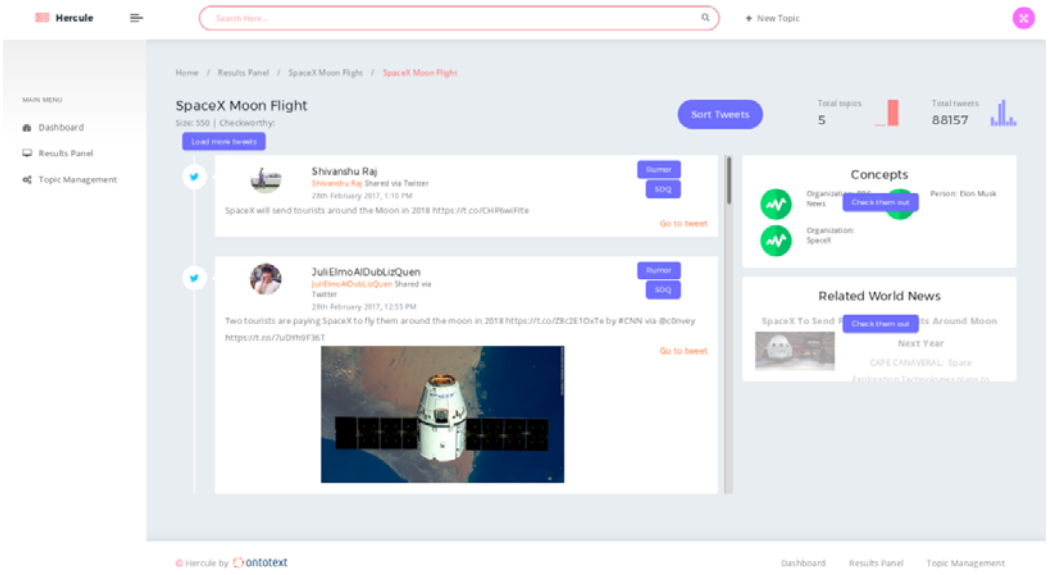
Stories sorted by number of tweets



When the user is happy with the story ordering scheme, they can look at the actual stories in the main frame. Each story on this screen has a title (big blue text at the top), size and checkworthy score and 20 representative tweets. From here a user can choose to navigate back to the topic list, open individual tweets using the *Go to tweet* button or open a specific

story by clicking on the title.

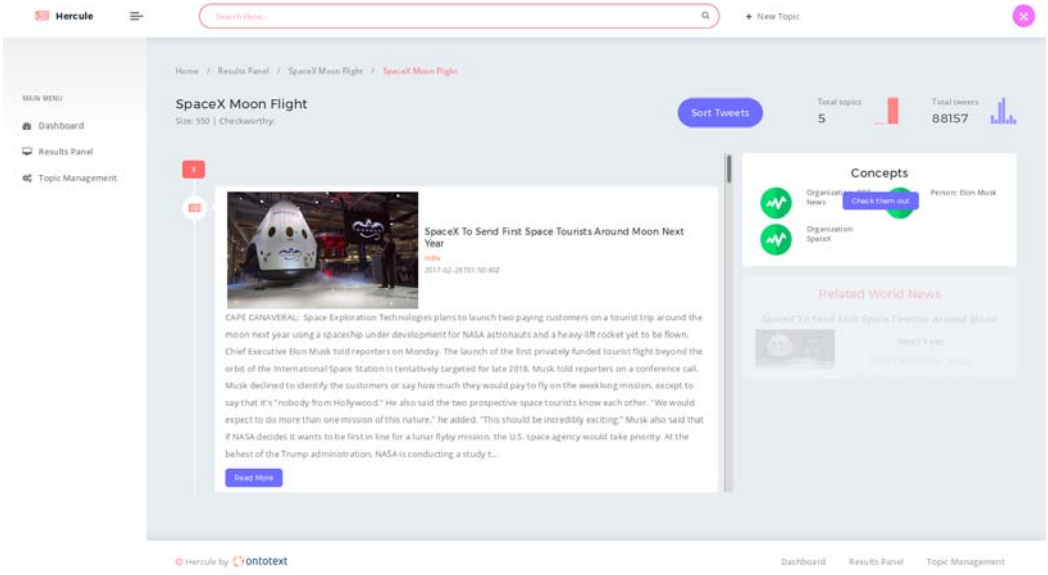
Story view



The story view has many of the same tools as the topic view: sorting of tweets (only based on recency), filtering through text search, and loading a representative sample of tweets in the story.

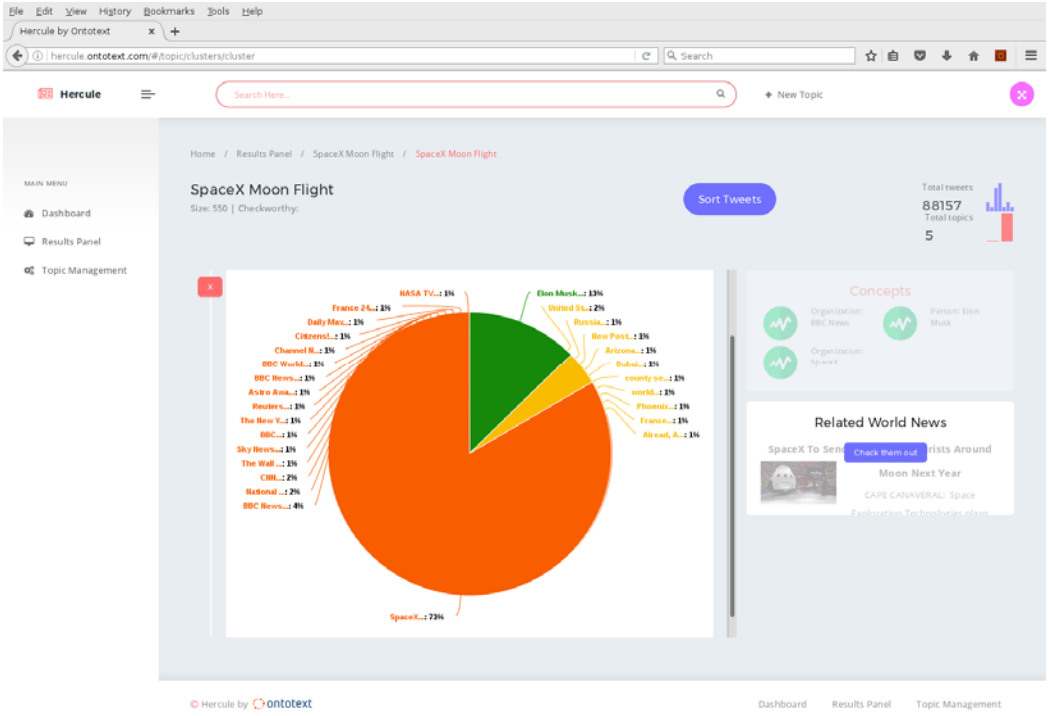
The extra space is used to display potentially valuable metadata. On a per tweet basis, this includes rumour score and SDQC (support/deny/question/comment) classification in the upper right corner of each tweet's frame. On a per story concept, this includes the collections of concepts mentioned and related news, both seen on the right-hand side of the screen.

Related news



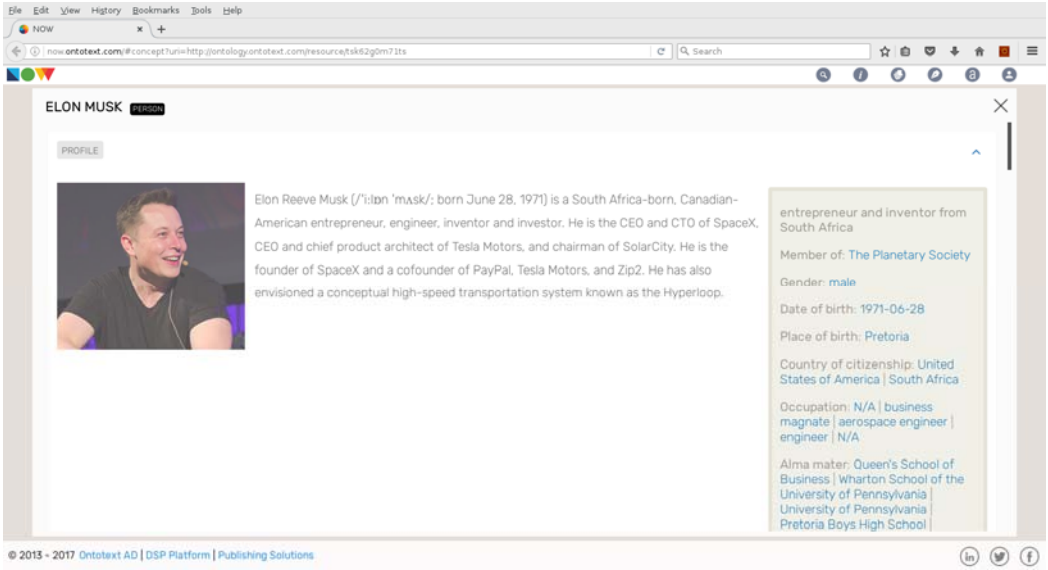
When the user clicks on a title in the Related articles panel on the right-hand side, the article content appears in the main frame.

Related concepts



When a user clicks on the Concepts panel on the right-hand side, Hercule visualises the distribution of related concepts which are mentioned in the story. In green are the concepts of type person, in yellow are the mentions of locations, and in orange are the organisations. For example, the screenshot above shows that the only person concept which was recognised in the story is Elon Musk; however, it was mentioned with relatively high frequency (about 13% of all mentions of concepts). There are occurrences of locations, such as United States and Arizona, and the highest number of mentions is of the organisation SpaceX (73% of all mentions), while there are mentions of other companies, such as BBC News .

Single Concept Exploration - Elon Musk



Through the concept visualisation panel, the user may also select a concrete concept and explore the related information available in the Hercule database (which is collected from various Linked Open Data resources, like DBpedia, Wikidata, etc.) The screenshot above shows the information available on Elon Musk, the only person mentioned in this particular story.

5.4 Hercule in relation to other fact-checking initiatives

Hercule is designed primarily as a tool that facilitates the work of professional fact checkers to find and verify questionable factual claims. The aim is to create a simple and efficient way to collect, organize, view and search large numbers of posts in fast-moving social media discussions. To this end, the system is fast (tweets are shown to users within minutes of creating a new topic) and effectively summarises large quantities of data by automatically combining tweets into news stories. The news stories allow users to read all individual tweets (good for small/new stories) or rely on data enrichment to quickly get a sense of the content (good for large stories with hundreds of tweets).

In 2016 the number of fact-checking initiatives increased by 50% worldwide[1]. Major political events like the Brexit vote and the US presidential election and the extent of misinformation during the campaigns show that the demand for fact checking journalism will only continue to grow. Large companies like Facebook[2] have also put efforts in this direction to automate the fact-checking of the their content, competitions for automating the fact-checking process have been organised[3].

The approaches to fact-checking vary between being fully automated[4] where the user input is

[1] http://www.poynter.org/2016/366-links-to-understand-fact-checking-in-2016/440618/
[2] https://www.theguardian.com/technology/2016/dec/16/facebook-fake-news-system-problems-fact-checking
[3] https://herox.com/factcheck
[4] https://www.theguardian.com/technology/2016/dec/16/facebook-fake-news-system-problems-fact-checking

not taken in account to approaches based on human judgment according to predefined rules[5]. Hercule is positioned somewhere at the mid-way point, offering automated collection of context-related information based on state-of-the art semantic technologies and allowing the human user to undertake the work of verifying the claim using that information.

6 Conclusion

The development of the digital journalism dashboard has been dependent on many factors: the primary use cases identified in the early stages of the project; the availability and status of various PHEME technologies being developed by technical partners; the outcomes of formative evaluations undertaken at SWI; and outcomes of regular discussions with partners.

A concerted effort was made in Year 3 to include an additional social media source -- Reddit -- into the journalism dashboard.

At the time of delivering this second version of the prototype, work on the German pipeline was ongoing, with the goal of having the latter up and running inside the journalism dashboard by the end of the project. Thus, only a limited set of German-language functionalities has been evaluated by journalists, with much of the evaluation focusing instead on the English pipeline. The outcomes of these user evaluations will be presented in D8.4.

The Hercule tool was designed to assist professional fact checkers in locating and verifying questionable factual claims and a live demo version has been released. It builds on the base of technologies for semantic enrichment of Twitter content developed in the project, interlinks social media content with related news and Linked Open Data thus offering an extended context of the analysed content. The tool was evaluated along the Journalism dashboard and results will be presented in D8.4.

[5] *http://www.npr.org/sections/alltechconsidered/2016/12/05/503581220/fake-or-real-how-to-self-check-the-news-and-get-the-facts*