



D4.6 FINAL VERSION OF THE CONTENT EXTRACTION AND ANALYSIS MODULE FOR ONLINE SOCIAL PLATFORMS



Grant Agreement nr	611057
Project acronym	EUMSSI
Start date of project (dur.)	December 1st 2013 (36 months)
Document due Date :	July 31st 2016 (+60 days) (32 months)
Actual date of delivery	September 8 th
Leader	UPF
Reply to	Jens.grivolla@upf.edu
Document status	Submitted

Project co-funded by ICT-7th Framework Programme from the European Commission



Project ref. no.	611057
Project acronym	EUMSSI
Project full title	Event Understanding through Multimodal Social Stream Interpretation
Document name	EUMSSI D4.6 Final version of the content extraction and analysis module for online social platforms_20160908
Security (distribution level)	PU – Public
Contractual date of delivery	July 31st 2016 (+60 days)
Actual date of delivery	September 8th
Deliverable name	D4.6 Final version of the content extraction and analysis module for online social platforms
Type	P – Prototype
Status	Submitted
Version number	1
Number of pages	5
WP /Task responsible	UPF
Author(s)	Jens Grivolla (UPF)
Other contributors	
EC Project Officer	Mrs. Alina Lupu Alina.LUPU@ec.europa.eu
Abstract	The deliverable should report on the final version of the content extraction and analysis module for online social platforms, but its content has been merged with D4.7.
Keywords	Social media, social network analysis, Twitter, visualization
Circulated to partners	N/A
Peer review completed	N/A
Peer-reviewed by	N/A
Coordinator approval	Yes

EUMSSI_D4.6_Final version of the content extraction and analysis module for online social platforms



Table of Contents

[BACKGROUND](#)



1. BACKGROUND

Deliverable D4.6 is described as follows, in the DoW:

Final version of the module that extracts publications from online social networks, blog/news providers and social news websites, together with a module that analyses the influence and impact of the extracted contents.

D4.6 is the third and last deliverable corresponding to Tasks 4.4. and 4.5, following up on D4.2 delivered in Month 12, and D4.4 delivered in Month 24.

D4.6 is a prototype deliverable consisting of the final version of the module that extracts publications from online social networks, blog/news providers and social news websites, together with a module that analyses the influence and impact of the extracted contents.

During the third and last year of the project, more content from Twitter and the Wikipedia Events portal has been collected and integrated into the platform. All this content, which amounts to over 10 million documents already, is conveniently analysed and indexed in the EUMSSI platform so that it can be efficiently queried in real time, allowing for interactive analysis of the data, even in direct response to user live interactions.

Moreover, additional sources have been added to the collection, such as comments from the YouTube site, and new visualization devices have been designed for the demonstrators, such as the tweets timeline.

D4.6 is intended to document the implementation of this module and is followed by D4.7 in month 34, which is a deliverable of type Report that contains a description and evaluation report of the social and text analysis components. D4.6 being of type Prototype, and being followed just two months later by D4.7 of type Report, we have decided to include all the descriptive part in the latter for clarity purposes and to avoid unnecessary repetitions.

The reader is referred to D4.7 for the final documentation of the social and text analysis components of the EUMSSI platform.

EUMSSI_D4.6_Final version of the content extraction and analysis module for online social platforms