Deliverable D2.1: Social Media Filtering and Extraction (confidential document - only a summary is publicly available)

Lyndon Nixon, Svitlana Vakulenko / MODUL Technology

30/06/2016

Work Package 2:   Media Selection and Analysis

InVID - In Video Veritas: Verification of Social Media Video Content for the News Industry

Dissemination level	CO
Contractual date of delivery	30/06/2016
Actual date of delivery	30/06/2016
Deliverable number	D2.1
Deliverable name	Social Media Filtering and Extraction (confidential document - only a summary is publicly available)
File	InVID_D2.1_v1.0.tex
Nature	Report
Status & version	Final & V1.0
Number of pages	26
WP contributing to the deliverable	2
Task responsible	MODUL
Other contributors	
Author(s)	Lyndon Nixon, Svitlana Vakulenko / MODUL Technology
Quality Assessors	Gerhard Rudinger / APA
EC Project Officer	Miguel Montarelo Navajo
Keywords	Topic Detection, Event Detection, First Story Detection, Breaking News Detection, Social Media, Social Networks, Social Web, Social Media Collection, Social Media Extraction, Social Media Filtering, Social Media Query, Social Network APIs, Media Retrieval

Abstract

This deliverable presents the approach taken for the first version of the InVID project's techniques for social media extraction and annotation as well as topic detection over social media streams.

We outline the functionality to be provided in the InVID project for the purposes of selection of social media out of the billions of postings daily and its annotation to support the subsequent phases of verification and rights management. The functionality also includes topic detection, so that activity on social media can be used to identify new news events occurring in the world.

We consider the current state of the art and available approaches together with the requirements of the InVID project and thus make an initial proposal for social media extraction and topic detection. This will be implemented and integrated into the first version of the InVID platform and it will be evaluated and refined iteratively in the next phases of the project.