

# Open Source Intelligence in Disaster Management

Gerhard Backfried, Christian Schmidt, Mark Pfeiffer, Gerald Quirchmayr\*, Markus Glanzer\*\*, Karin Rainer\*\*\*

SAIL LABS Technology AG, Vienna, Austria

University of Vienna, Department of Distributed and Multimedia Systems, Vienna, Austria\*

Austrian Red Cross, Vienna, Austria\*\*

INSET Research&Advisory, Vienna, Austria\*\*\*

{gerhard.backfried, christian.schmidt, mark.pfeiffer}@sail-labs.com,

[gerald.quirchmayr@univie.ac.at](mailto:gerald.quirchmayr@univie.ac.at), [markus.glanzer@roteskreuz.at](mailto:markus.glanzer@roteskreuz.at), [karin.rainer@inset-advisory.com](mailto:karin.rainer@inset-advisory.com)

*Abstract—In this paper we discuss the role of Open Source Intelligence (OSINT) in Disaster Management. In particular we present the use of the Sail Labs Media Mining System in the context of disaster relief operations and use samples to point out advantages and strengths of the MM-System. Future challenges in research and further development of this field are addressed towards the end of the paper.*

**Keywords—Information Systems, Multimedia Computing, Speech Processing, Situational Awareness, Multilinguality, Social Media, Open Source Intelligence, Disaster Management**

## I. INTRODUCTION

In situations of disaster or crisis it is essential to have access to current, up-to-date information. Irregular events require a profound and comprehensive situational understanding of the relevant circumstances and actors involved in order to form a proper policy and timely response. Particularly in today's rapidly changing information society, the access to a wide array of sources, ranging from Web-Feeds, blogs, TV- and radio-stations to social media platforms can be key in responding to critical events.

The actual information-sources vary widely in their nature, quantity, reliability and quality. The amount of information being produced increases constantly and changes almost by the second. Information is produced continuously around the globe in a multitude of languages and across a spectrum of media. New information is put on the Internet by professional news agencies or individuals, broadcast by TV- and radio stations or chatted about on social media such as Twitter and Facebook. The latter have added yet another dimension of complexity to the mix of sources. All of the above sources differ in multiple dimensions such as modality or language used and require specific kinds of processing. Language plays a crucial role in this process, as often one of the key-problems of disasters with global relevance is that information is only available from local sources and in local languages. Some of the most immediate information on an event might be produced by citizen journalists with little experience and barely any track record that could be independently vetted.

Open sources provide a cheap, fast and efficient way to assess the situation of those affected in a time of crisis or disaster. They complement, amplify and frequently even pre-date information gathered from traditional and official sources. In many cases they allow the geographic mapping of events, this being especially the case for crowd-mapped and crowd-sourced information gathering. The number and diversity of sources help organizations estimate the integrity and trust-level of the information provided, thus allowing for more targeted responses and shorter decision-cycles in the initiation of first-responder reactions. One fundamental assumption underlying the concept of Open Source Intelligence is that information is indeed available in publicly accessible sources and *just* has to be gathered and provided to the right people at the right time. But clearly, this is far from trivial and a highly complex and demanding task.

This is the background against which experiments were conducted based on the Sail Labs Media Mining System (MM-System). Features such as video analysis, speech transcription and ontology-based search improvements were combined to provide information which can be passed to analysts and fed into a situation awareness center. The system setup is targeted toward multi-modality, multi-linguality and real-time behavior and can be deployed within a matter of days.

## II. SYSTEM DESCRIPTION

The MM-System is a modular system aiming to cover the complete OSINT workflow cycle from the requirements phase to the dissemination and feedback phases [1,2]. It enables OSINT professionals to quickly extract meaningful analyses from unstructured data in a variety of formats across multiple languages and sources. Analysts are supported in their work by providing them with tools to visually explore and search large volumes of data according to their mission and fields of intelligence. Compression and condensation of information, relating sources over time and across media form the pillars on which analysts can carry out their tasks. The system is a truly complete, yet also open system, which allows for integration into existing environments. Its purpose is to support existing

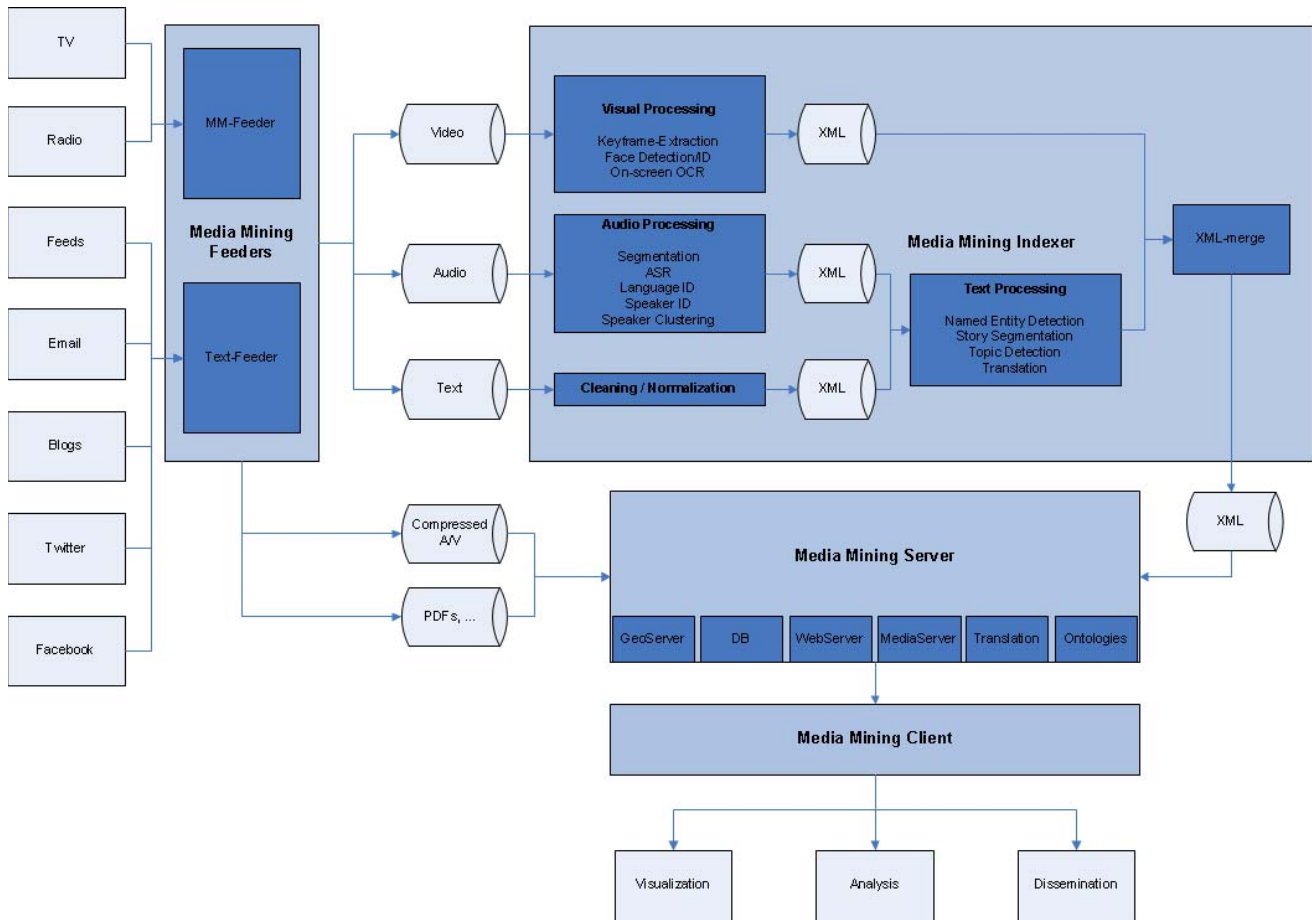


Figure 1 Overall Architecture of the MM System

processes and procedures. Existing workflows can be maintained, enhanced and simplified for users.

The MM-System consists of a set of technologies packaged into components and models and combined into a single system for end-to-end deployment. A number of toolkits allow end-users to update, extend and refine models in order to respond flexibly to a highly dynamic environment. The overall architecture of the MM-System is a server-client one and allows for deployment of the different components on multiple computers and platforms. Not all components and technologies need to be present from the start and can be added flexibly over time. Several Feeders, Indexers and Servers (also called Media Mining Feeder, -Indexer and -Server resp.) may be combined to form a complete system. Fig. 1 provides an overview of the components of the MM-System and their interaction.

#### A. Overall Data Flow

Data enter the system via so-called *Media Mining Feeders* and then pass through a series of processing steps. Multimedia data are split into audio- and video-processing tracks. Audio data are passed into the Media Mining Indexer (MMI) for

segmentation, speaker-identification and large vocabulary, automatic speech-recognition. For video-data, the processing chain comprises key-frame extraction, the detection and identification of faces and the detection and subsequent OCR of text. Textual data, including text from feeds and social media sources, as well as the output of ASR, are processed by several normalization steps before undergoing Named Entity Detection and Topic Detection processing.

The resulting documents of the individual tracks are fused at the end of processing (*late-fusion*) and uploaded, together with a compressed version of the original media files, onto the Media Mining Server (MMS), where they are made available for full-text search and retrieval.

#### B. Media Mining Feeders

Feeders represent the input interface of the MM-System to the outside world. For audio or mixed audio/video input a variety of formats can be ingested from external sources and processed by subsequent components. To handle textual input, such as data coming from Web-Pages, RSS-feeds, e-mails, blogs and social media sources, separate feeders exist which extract the relevant data from these sources and pass them on to the text processing components. Additional feeders for

further sources can be added with minimal effort. This allows responding quickly to changing environments and new data sources.

### C. Media Mining Indexer (MMI)

The MMI forms the core for the audio- and text-processing capabilities within the MM-System. It consists of a suite of technologies and associated models, which perform a variety of analyses on the audio and textual content. Processing results are combined by incrementally enriching XML-structures and can be joined across different modalities via timing information included in the annotated structures. Facilities for processing a number of natural languages exist for the components of the MMI, e.g., ASR is available for fourteen languages already. Models for languages are developed by Sail Labs in cooperation with partner organizations and customers and form an active area of development.

Audio processing includes the segmentation and classification of the incoming audio-stream, the identification of speakers (SID), and/or gender of the speaker and automatic speech recognition (ASR).

After having been converted to the appropriate format by the feeder, the audio signal is processed and segmented into homogeneous sections of speech and non-speech [3]. Speech segments are processed to determine to speaker's identity [4] and passed on to the ASR component. ASR is designed for large-vocabulary, speaker-independent, multi-lingual, real-time decoding of continuous speech and performs speech-recognition in a multi-pass, time-synchronous manner employing several sets of models [5]. Subsequently, text-normalization as well as language-dependent processing is applied to yield the final decoding result in an XML format. The ASR component per se is language independent and can be run with a variety of models created for different choices of language and bandwidth.

Text-processing includes the normalization and language-specific processing of text [6] (for text produced as output of the ASR-component or provided by text-feeders), the annotation of Named Entities, such as but not limited to persons, organizations or locations and the segmentation and classification of text-segments according to topic (Topic Detection). The NED system is based on patterns as well as statistical models defined over words and word-features and is run in multiple stages [7]. The topic-detection component (TD) first classifies sections of text according to a specific hierarchy of topics and then creates coherent stories by grouping together similar adjacent sections. The models used for TD and story segmentation are based on support vector machines (SVM) with linear kernels [8].

Visual-processing currently includes the detection and identification of faces as well as text-inserts in the video signal. These complement the information extracted from the audio stream. (All visual processing components are 3rd-party components which have been developed in co-operation with

partners and integrated into the MM-System). Regarding face detection, faces are first localized [9] and then a recognition step is performed [10]. The available temporal information is employed by running a tracker and performing the recognition on the identified image locations [11].

On-screen OCR processes individual frames, which are extracted from the live video stream at regular intervals. Stable areas which are likely to contain characters are detected, grouped together and fed to the Tesseract OCR-engine ([12]), which performs the actual text extraction. In a post-processing step, the detections are matched against a dictionary and filtered by the OCR detection confidence.

Based on time-tags, the XML-outputs of all different technologies are combined (fused) into one final result which is then uploaded to the server and made available for search and retrieval.

All processing can be configured to take place in real-time or to put emphasis on the quality of processing results at the expense of processing time. Toolkits are provided for several technologies to allow for easy extension and adaptation of models.

### D. Media Mining Server and Client (MMS and MMC)

The MMS comprises the actual server, used for storage of XML and media files, as well as a set of tools and interfaces used to update and query the contents of the database. The MMS serves as the central hub to which all data produced by MMI's are sent and from which all analysis and visualization takes place. Components integrated into the MMS include a media-server, for playback of multimedia content, a geo-server, for map image rendering and projection of underlying data-sources, a (3<sup>rd</sup>-party) translation-server, for parallel translations of transcripts and documents and Oracle 11g which provides all search and retrieval functionalities [13]. The Semantic Technologies provided by Oracle form the basis for all ontology-related operations within the MM-System.

All user-interaction takes place through the MMC. It provides a set of features to let users query, interact with, visualize and update the contents of the data stored in the MMS. Users can perform queries, download content, request translations or add annotations to the stored documents. Documents, search results and summaries can be viewed and manipulated from different angles, allowing users to focus on relevant aspects first and to iteratively drill-down on relevant issues. Combinations of visualization and queries can be used to explore data-sets and gain deeper insight.

Locations detected in documents can be plotted onto maps via geo-coordinates. Relationships between detected entities can be visualized and explored via a relationship-graph. A trend-view relates entities with their occurrences over time, allowing to observe the temporal behavior of entities. A cluster-view relates entities and sources mentioning them, allowing to contrast reports on identical topics or events. Different twists and angles of reporting may be elicited this

way, hidden agendas may be made visible by contrasted exploration of data. Information derived from ontologies is used to modify and guide searches by expanding query terms to semantically related terms and for the presentation of query results.

The results of queries are presented according to the type of document – e.g. for video/audio documents in addition to the time-aligned transcript, the names of speakers or persons detected in the image are displayed and time-synchronous playback of audio and video content is provided. For textual documents, the (extracted) text and any associated documents (PDF file for web-content) are displayed.

For monitoring purposes a further interface, the Sail Labs Crisis Room, is available which allows for the concurrent monitoring, alerting and search of information from live TV input.

In a typical setup, several (up to hundreds of) MMIs, concurrently operating on different sources and channels provide a continuous feed of input data while analysts are connected to the MMS via dozens of MMCs.

### III. SAMPLE SEARCHES

When the flood-crisis started to fully hit the capital of Thailand in the fall of 2011, the major challenge was to get an up-to-date and correct assessment of the situation. Given the confusion typically prevailing in such a crisis-situation, this information is essential for planning and launching disaster relief operations. A US Navy aircraft carrier lying idle and subsequently leaving the region without becoming active demonstrates the scale of the potential impact of not-having access to appropriate information. On the other hand, quite accurate information was available on web-pages of leading news broadcasters in the region at the same time.

*“It’s true that the US military was confused about the Thai government’s attitude about help and thus withdrew the aircraft carrier but the US military still left one ship and two helicopters for us if we need help.” [14]*

Ideally, the necessary information would have been provided to the decision makers at the right time and in the right quality. This sort of information confusion is unfortunately not limited to the above example but rather universal for crisis situations and can, in extreme cases, lead to panic. A recent example shows how essential it is to detect emerging discussion patterns.

In our example we assume that our initial information source is from the traditional media, in particular a local newspaper. The article alerts us to an earthquake close to the city of Ravenna, Italy. With this information we can now query further sources, for example Twitter. Among the tweets

found, the following one is of particular interest, as it contains a link to a Tsunami-related Web-Page:

*#terremoto a Ravenna non c’è il pericolo #tsunami http://t.co/He9yhbV Ma il Mediterraneo non è esente da questo rischio*

Following this link and carrying out further investigations quickly reveals that fear of a Tsunami in the region around Ravenna is developing.



Figure 2 Screen shot of Focus IT

In a decision situation, it might be beneficial to enforce messages that help calm down an overheated discussion such as the following laid-back tweet:

*#terremoto a largo di Ravenna ci manca pure lo #tsunami e siamo a posto Tutti sull'arca di Noè #cealpossiamofare*

Once we know that valuable information is available, a next step could be to get a better understanding of the situation by eliciting relations in a relationship-graph.

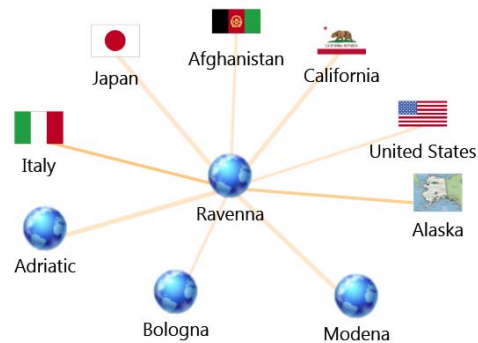


Figure 3 Relationship-Graph

As can be seen from the above graph, links can clearly be established between the regional Italian earthquake sites and

the global picture. A better understanding for whether a Twitter-based discussion like the one above is an isolated event or actually launches sustained activity can be inferred from statistics.

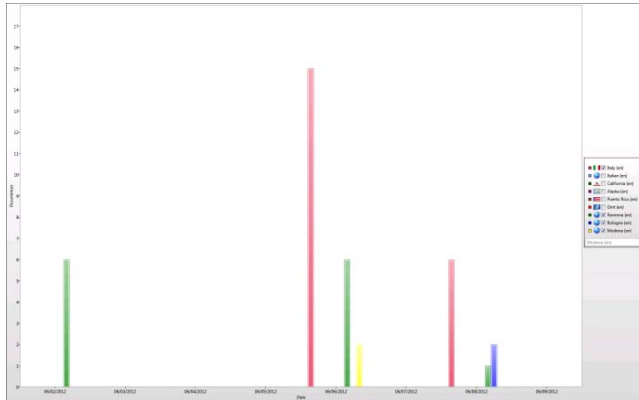


Figure 4 Trend-Graph

The above chart plotting the frequencies of entities mentioned in documents over time does confirm that there is no sustained panic initiated by the media-reports and Twitter discussions as the spikes of activity are limited to just a single day.

In case of having to deliver aid and medical supplies to a disaster-affected area, knowing that the situation is calm and that no panic is developing is essential for planning and carrying out relieve operations.

#### IV. POTENTIAL BENEFITS

One of the most important lessons learned from experiments with our system is that information is available, but the major challenge remains to prepare it for ready access by human decision makers. Reducing the flood of information and chaos which typically prevail in a crisis-situation by targeted search and the inclusion of multi-lingual and multi-media sources can lead to a significant improvement of a decision maker's basis and timely reaction. It can also help with documenting and justifying why a decision has been made.

In our example this translates to the ability to reduce the number of potential scenarios a disaster-relief operation will be confronted with. In the given case, this means that a panic can be excluded at the present state and that law-enforcement and military response units can focus on first-aid and engineering tasks. The positive psychological side-effect of disaster relief forces knowing that they (ambulances services and fire-brigades) will enter a calm and stable area is self-explanatory.

#### V. OUTLOOK AND CONCLUSIONS

With the MM –System having been deployed and tested in cooperation with several governmental organizations, we are now in a position to move on to a tighter integration with mission planning and support systems for emergency and disaster management. The most interesting research questions associated with this integration will likely be the development of efficient interfaces, data-model and semantics. From a practical perspective, technical and organizational questions related to the collaboration between different disaster relief units will be addressed in future research projects planned by the authors.

#### VI. REFERENCES

- [1] <http://www.fbi.gov/about-us/intelligence/intelligence-cycle>, 03/29/2012
- [2] Open Source Intelligence, US Army, available at <http://www.fas.org/irp/doddir/army/fmi2-22-9.pdf>, 03/29/2012
- [3] D. Liu, F. Kubala, Fast Speaker Change Detection for Broadcast News Transcription and Indexing, Eurospeech 1999
- [4] D. Liu, F. Kubala, Online Speaker Clustering, ICASSP 2003
- [5] R. Schwartz, L. Nguyen, J. Makhoul, Multiple-Pass Search Strategies, Automatic Speech and Speaker Recognition, 1996
- [6] R. Hecht, J. Riedler, G. Backfried, Fitting German into N-Gram Language Models, TSD 2002
- [7] D.Bikel, S.Miller, R.Schwartz, R.Weischedel, „Nymble: High-Performance Learning Name-Finder“, Conference on Applied Natural Language Processing, 1997
- [8] T.Joachims, “Text Categorization with Support-Vector-Machines: Learning with Many Relevant Features”, ECML, 1998
- [9] P. Viola, M.J. Jones, “Robust Real-Time face detection”, Int’l Journal of Computer Vision, 57(2):137–154, 2004
- [10] S. Kim, S. Chung, S. Jung, S. Jeon, J. Kim, S. Cho, “Robust face recognition using AAM and gabor features”, In Proc. World Academy of Science, Engineering and Technology, 2007
- [11] G. Dedeoglu, T. Kanade, and S. Baker, “The asymmetry of image registration and its application to face tracking”, IEEE Trans. on Pattern Analysis and Machine Intelligence, 29(5):807–823, 2007
- [12] <http://code.google.com/p/tesseract-ocr/>, 03/29/2012
- [13] <http://www.oracle.com/technetwork/database/options/semantic-tech/index.html>, 06/14/2012
- [14] Defence Minister Gen Yutthasak Sasiprapa) Published: 26/10/2011 at 08:16 PM, <http://www.bangkokpost.com/breakingnews/263316/yutthasak-thai-us-relation-fine>

#### ACKNOWLEDGMENTS

This work was partially supported by the FFG project MDL (818800) under the Austrian National Research Development Programme KIRAS.