Privacy-preserving Social Media Forensic Analysis for Preventive Policing of Online Activities

Syed Naqvi*, Sean Enderby*, Ian Williams*, Waqar Asif †, Muttukrishnan Rajarajan†, Cristi Potlog‡, and Monica Florea‡

*School of Computing and Digital Technology, Birmingham City University, United Kingdom
(Syed.Naqvi, Sean.Enderby, Ian.Williams)@bcu.ac.uk

†School of Maths, Computer Science and Engineering, City, University of London, United Kingdom
(Waqar.Asif, R.Muttukrishnan)@city.ac.uk

‡European Projects Department, SIVECO Romania SA, Romania
(Cristi.Potlog, Monica.Florea)@siveco.ro

Abstract—Social media is extensively used nowadays and is gaining popularity among the users with the increasing growth in the network capacity, connectivity, and speed. Moreover, affordable prices of data plans, especially mobile data packages, have considerably increased the use of multimedia by different users. This includes terrorists who use social media platforms to promote their ideology and intimidate their adversaries. It is therefore very important to develop automated solutions to semantically analyse given multimedia contents to assist law enforcement agencies in the preventive policing of online activities. A major challenge for the social media forensic analysis is to preserve the privacy of citizens who use online social networking platforms. This paper presents results of European H2020 project RED-Alert that aims to enable secure and privacy preserving data processing; hence the malicious content and the corresponding personality can be tracked while the privacy of innocent citizens can be preserved. We have mined seven social media channels for content and providing support for ten languages for analysis. Our proposed solution is designed to ensure security and policing of online contents by detecting terrorist material. We have used speech recognition, face and object detection besides audio event detection to extract information from multimedia files. We have applied anonymization techniques to ensure the privacy of citizens using social media. We have discussed the challenges and prospects of this work especially the need of using digital forensic techniques while respecting European and national data protection laws notably GDPR.

I. Introduction

Security and policing of social media in ethical way is very challenging due to the scale and scope of the social media [1]. Privacy and data protection requirements call for careful consideration of the techniques and algorithms for data extraction and its processing by the digital forensic analysts. Efficient solutions are needed to ensure that information is timely found with minimal chances of errors (false positives or negatives). Predicting an imminent terrorist activity without any prejudice in almost real-time is the focal point of modern policing [2] also called Policing 4.0 [3].

The European Commission has funded a project RED-Alert to deal with the challenges of preventive policing of online activities. This project is presented in the Section 2 of this paper. We have developed a Semantic Multimedia Analysis (SMA) Tool as part of the RED-Alert project. This tool is designed to extract data from multimedia contents. Its details are provided in the Section 3. The implementation challenges especially the privacy-preserving of the general public users of social media are discussed in the Section 4 together with the design details of the anonymization tool we have developed to meet data protection compliance requirements. The overall integration of these tools in the RED-Alert architecture and their testing is presented in the Section 5.

We have made a pragmatic discussion in the Section 6 about the features and issues with the multimedia forensics and how we have managed them in the design and development of SMA Tool in RED-Alert project. Finally we have drawn some conclusions together with a concise description of our future directions in the Section 7.

II. Context of this work

This work is a part of European Project RED-Alert (Realtime Early Detection and Alert System for Online Terrorist Content) [4]. This project is consist of 16 partners from European Unions Member States as well as its Associated Countries. The consortium includes academia, businesses and Law Enforcement Agencies (LEAs). The project brings data mining and predictive analytics tools to the next level, developing novel Natural Language Processing (NLP), Semantic Multimedia Analysis (SMA), Social Network Analysis (SNA), Complex Event Processing (CEP) and Artificial Intelligence (AI) technologies. These technologies are combined for the first time and validated by 6 LEAs to collect, process, visualize and store online data related to terrorist groups, allowing them to take coordinated action in real-time while preserving the privacy of citizens.

The RED-Alert project aims to mine at least seven social media channels for content, and support at least ten languages for analysis. Therefore a good quantity and variety of data is analysed by the project that requires complex and (semi-) automatic solution that can efficiently process the data within legal constraints. This implies improved accuracy and usability

of tools within the context of data privacy, as well as extended real-time and collaborative capabilities and support for further development.

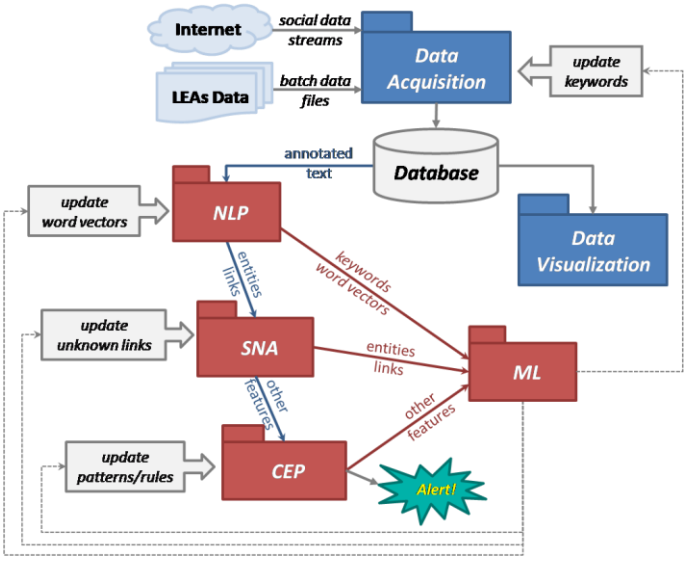

Fig. 1. High-level RED-Alert architecture

In order to meet the project objectives, these software components (CEP, NLP, SMA, SNA, etc.) are developed by project partners. The integration of these software components brings innovation and impact in the everyday investigation of modern policing. The main challenge of the integrated solution is to assess social features in communications by terrorist organizations that will imply harmonisation of theories, tools and techniques from cognitive science, communications, computational linguistics, discourse processing, language studies and social psychology. Moreover, in order that the system performance to be adapted for each component the project implements a meta-learning process to assist each software component defined processes. A high-level architecture is shown in the Figure 1.

The other major challenge for the project is to preserve the privacy of citizens who use online social networking platforms. The project aims to enable secure and privacy preserving data processing; hence the malicious content and the corresponding personality can be tracked while the privacy of innocent citizens can be preserved. RED-Alert system is required to include privacy-preserving mechanisms allowing the capture, processing and storage of social media data in accordance to European and national legislations (especially GDPR [5]).

III. Semantic Multimedia Analysis (SMA) Tool

Multimedia is extensively used in social networks nowadays and is gaining popularity among the users with the increasing growth in the network capacity, connectivity, and speed. Moreover, affordable prices of data plans, especially mobile data packages, have considerably increased the use of multimedia by different users including terrorists, who use social media platforms to promote their ideology and intimidate their adversaries. It is therefore very important to develop automated solutions to semantically analyse online multimedia contents.

The objective of the SMA Tool is to ensure security and policing of online contents by detecting terrorist material. It extracts meaningful information from multimedia contents taken from social media. The five main features of the tool are:

- Segmentation of audio streams, identifying sections of speech.
- Transcription of the segmented speech sections using an Automatic Speech Recognition (ASR) engine.
- Detection of sound events within audio streams, such as gunfire, explosions, crowd noise, etc.
- Extraction and identification of objects, such as logos, flags, weapons, faces, etc., within image and video scene elements.
- Extraction and transcription of text elements in image and video elements.

Moreover, the SMA Tool retrieves multimedia data, converts it to a uniform format and delivers the analysis results. The extraction of semantic information is the third of four stages the tool will perform. All four stages are as follows:

1) Input: Retrieval of multimedia files from disk or URL.
2) Stream Separation: Extraction of audio/video streams in multimedia files.
3) Feature Analysis: Semantic analysis of audio/image content.
4) Output: Compilation of results in a uniform JSON format.

A. Input options for the SMA tool

The SMA tool can process files which are stored locally or can optionally download media from a given URL. Currently the tool recognises the media format and type (audio/image/video) from its file extension, currently supported extensions are as follows:

- Audio File Extensions: .wav, .ogg, .mp3, .aiff, .flac
- Image File Extensions: .bmp, .jpeg, .jpg, .png, .tiff, .tif, .JPEG
- Video File Extensions: .avi, .mp4

B. Output format of the SMA tool

The SMA tool produces standard JSON formatted output for all types of multimedia input. A separate *.json* file is produced for each multimedia file processed. This file's name is simply the name of the original file with *.json* appended to it. For example, the output for the input file *Gun.jpg* would be called *Gun.jpg.json*. These output files are written into the directory from which the SMATool executable was run. All JSON output has two fields in common. These are:

- file name: The name of the file that the analysis results apply to.

- media _type: The type of media the input file represents (*audio*, *image* or *video*).

JSON output for each of the supported media type have further fields which describe the results of the analysis which apply to that type of media. The output is sent to the other key components of the project such as NLP, SNA and CEP. Component details of SMA Tool is provided hereunder.

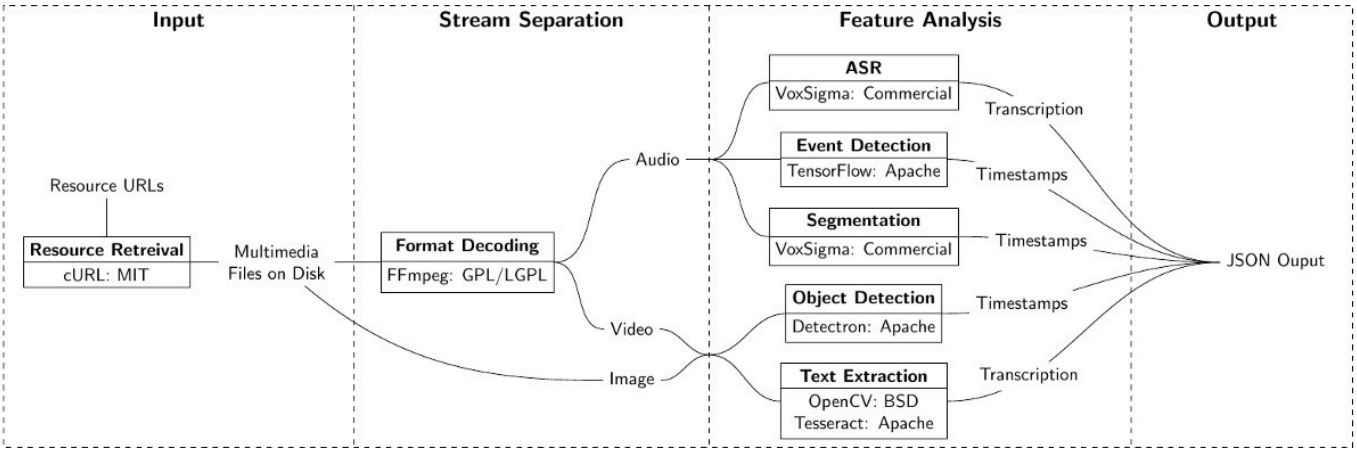

Fig. 2. SMA tool system diagram

C. Speech recognition

This component is used for audio segmentation, language detection and speech transcription. The RED-Alert project is required to support 10 languages, and be able to run offline, without having to send data to a 3rd party web API. We have consulted our LEA partners to prepare a list of 10 languages which must be supported by the speech / written text transcription elements of the SMA Tool. These languages are: Arabic, English, French, German, Hebrew, Romanian, Russian, Spanish, Turkish, and Ukrainian.

D. Face detection

The SMA tool uses a Haar-like feature based cascade classifier [6] to detect both frontal facing and profile faces in images. Haar-like features are calculated by finding the difference in average pixel intensity between two or more adjacent rectangular regions of an image. In the SMA tool, Haar cascades are used as a supplementary feature to implement simple face detection. More advanced techniques are implemented in the object detection element, which can also be used to detect people/faces.

E. Object detection

State of the art methods for detection of objects within images use large neural networks consisting of multiple subnetworks (region proposal network, classification network etc.). The SMA tools object detection utility uses the Faster R-CNN structure [7]. Faster R-CNN is constructed primarily of two separate networks: a Region Proposal Network (RPN) which produces suggestions of regions of an image which might contain objects, and a typical Convolutional Neural Network (CNN) which generates a feature map and classifies the objects in the proposed regions.

F. Audio event detection

Audio event detection is implemented in the SMA Tool by using a recurrent convolutional neural network [8]. The convolutional element classifies the short term temporal / spectral features of the audio, while the recurrent element detects longer term temporal changes in the signal. The SMA Tool apply feature extraction prior to processing by the network. This provides a more detailed representation of the audio signal to the network, meaning the first few layers can extract more meaningful information. Peak picking algorithms [9] are applied to remove any noise and only annotate the onset of any detected audio events.

IV. Implementation Challenges

This section outlines our actual experience of dealing with various challenges of implementing research results; and how we overcome them.

A. Technical challenges

First and foremost challenge was to overcome a set of technical issues to build the first prototype of the SMA Tool and to integrate it in the main RED-Alert architecture. These challenges include:

- Automatic Speech Recognition (ASR): The RED-Alert project has to provide support for minimum 10 languages. These languages are chosen by the LEAs to help them facilitate their investigations. They are: Arabic, English, French, German, Hebrew, Romanian, Russian, Spanish, Turkish, and Ukrainian. A custom-built automatic speech recognition component to cover this set of languages was challenging especially to meet the constraint of development timeline. We carried out an extensive state of the art analysis and tested a number of tools and libraries. We found that an ASR component

developed under the auspices of another European funded project DANTE [14] responds our needs including development timeline. The use of this component also provided us a real-world inter-project collaboration and better technical support.

- Reliability of the results: The critical nature of the SMA Tool and the importance of its results require extra care for the enhanced quality and reliability of the results. In other words, the keep the false positives to minimum possible level. One obvious, yet time and resource intensive, solution is to use bigger dataset to train the SMA Tool components to cover all possible attributes of multimedia content. The other possible techniques are improved object detection (GPU inference, selectable models, server like architecture, etc.) and support for more media file formats.

B. Legal considerations

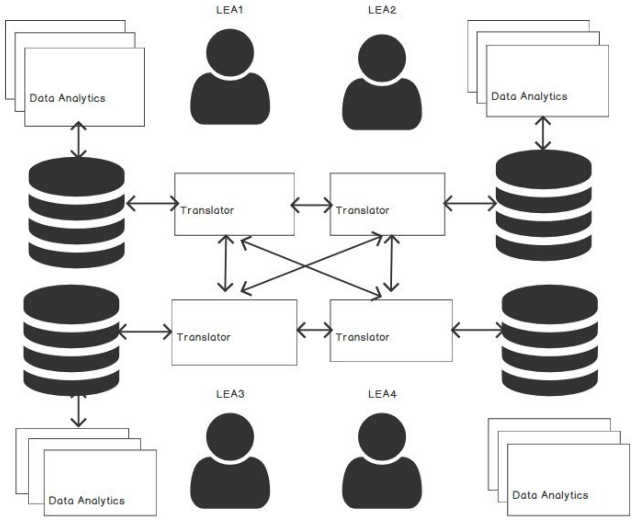

Fig. 3. Two-layer networked privacy preserving big data analytics model between coalition forces

This is imperative for us to ensure that legal requirements are met throughout the lifecycle of the project. The everincreasing data protection requirements the latest one is the GDPR requires due protection of citizens privacy. This section provides details of the privacy-preserving anonymization unit developed to ensure compliance with the data protection requirements.

The anonymization techniques applied in this project work in hiding information about all innocent individuals but it also helps terrorist organizations in hiding behind the covers. This as a result puts extra burden on the SNA, CEP, NLP modules. They adapt to working on anonymized social network data and narrow down the search of terrorist organizations. Once identified, the LEAs need to know the identity of the highlighted individuals. In order to cater for this need, a de-identification approach is also developed in this project, that takes as input the surrogate ids that are provided by the anonymization technique and provide the true identity of an individual. This de-identification algorithm only exists due to the nature of the project and where one can argue that this would make the anonymization algorithm pseudo in nature, it is key to highlight that the de-identification approach only resides with the LEAs thus limiting any adversary from actually identifying individuals and also complying with the GDPR.

Data Networked Privacy Tool

Intelligence information can be very tricky at times and the nature of this information limits LEAs located in different geographical location from sharing information. On the contrary social networks have no territorial boundaries and terrorist organization can operate from any possible location, making it harder for LEAs to track and tackle them. In order to overcome this difficulty, the Red-Alert project is equipped with a novel Inter-LEA search algorithm. It limits and controls the amount of information that LEAs located in different geographical location can share with the use of high end encryption algorithm. Under this approach, LEAs are independent in performing their own search and collecting their own intelligence information, they then are requested to populate a list of names of the individuals identified. The second LEA who is looking for a particular individual can search in the encrypted list and find out if one exists or not. The benefit of using high end encryption techniques is of limiting what else an inquiring LEA can see. The LEA inquiring only sees a response in terms of a YES or a NO, therefore hiding all other names in the database. The search query is made with taking a probability attack into consideration, thus if an LEA searches for the same name over and over again, there exists no defined pattern. This limits the first LEA (who is hosting the list) from knowing what name is being searched, thus making it a double sided blinded process.

C. Ethical issues

Besides legal issues, the scope of this project and the use of artefacts developed for this project require rigorous ethical review. Each partner of the project consortium has individually reviewed the ethical issues of their contributions and the safeguards in place to comply with the organisational policies and to ensure health and safety of the staff engaged in this research. In the context of SMA Tool, the multimedia content are provided by the LEAs. We have data sharing agreement and have developed a mechanism to ensure that neither any disturbing content is received, processed, or stored in our organisational infrastructure nor any staff member is exposed to such contents.

V. Integration and Testing

The SMA Tool is going to be fully integrated in the REDAlert architecture followed by its integration testing together with the other components of the project. This section outlines this stage of the project.

A. Integration Components

The integration component of the project RED-Alert will have some different subcomponents to ensure that internal interfaces are not exposed to the outside world and likewise new interfaces are developed to interact with the users. The main system user interface is shown in the figure 4.

1) Main System User Interface: This component provides common look-and-feel to the graphical user interface of the overall RED-Alert System, offering a portal-like user interface for the overall system with common interface placeholders, such as header and footer, main menu, and user interface components hosting through custom common APIs;

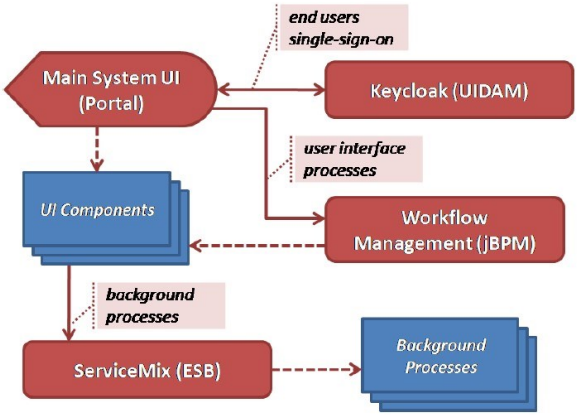

Fig. 4. Main System User Interface - Component Interactions Diagrams

2) User Identification and Access Management component: This component will be implemented based on RedHat Keycloak [10] and will provide the means for identifying users and managing their access to application components, both to front-end user interface and to back-end processes;

3) The Collaborative Workflow/Case Management component: This component is based on RedHat jBPM [11], a light-weight and extensible workflow engine, offering process management features and tools for both business users and developers. RedHat jBPM supports adaptive and dynamic processes that require flexibility to model complex, real-life situations that cannot easily be described using a rigid process;

4) Application Integration Services component: This component is built with Apache ServiceMix [12], an open-source integration container that unifies the functionalities of Apache ActiveMQ, Camel, CXF, and Karaf into a powerful runtime platform you can use to build your own integrations solutions. It provides a complete, flexible, enterprise ready ESB exclusively powered by OSGi;

5) System Interoperability Services component: This component will be built on top of the Application Integration Services, exposing selected RED-Alert systems functionalities to external system, including existing systems of LEAs;

6) Centralized Audit and Logging component: This component will be implemented using Audit4j [13], an open source auditing framework which is a full stack application auditing and logging solution for Java enterprise applications, tested on a common distributions of Linux, Windows and Mac OS, designed to run with minimum configurations, yet providing various options for customization.

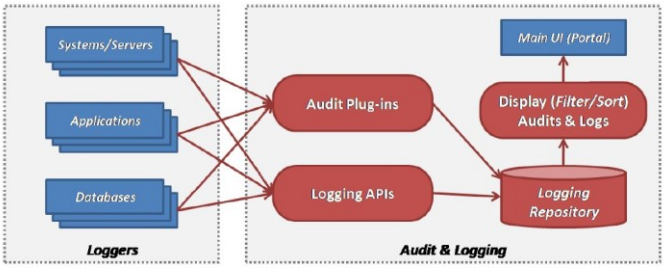

Fig. 5. Centralized Audit and Logging - Interactions Diagram

The Figure 5 shows the interactions of the Centralized Audit and Logging component with the Main System UI (Portal), by means of hosting the visual part exposed by the component, and also with the other components of the REDAlert system, by means of custom common APIs that will allow all components to log entries into a central repository.

VI. Discussions

The RED-Alert project is an ambitious project with significant impact in the modern policing of the online activities notably in predicting any imminent terrorist activity. The backend of the system requires smart and high-performance elements to ensure real-time delivery of the results. Semantic Multimedia Analysis (SMA) Tool is one of the key components of the project solution that extracts information from a sheer volume of the input data. This is like finding a needle in a haystack. The reliability of the tools results require fine-tuned object detection techniques otherwise recurrent execution of the searches will have higher computing cost and performance overheads.

Besides technical challenges and project constraints, we have the obligation to address legal and ethical issues related to data collection, processing and storage. Data anonymization and visualisation [15] components positioning and features are the key elements to ensure compliance with the data protection legislations notably GDPR. However, the overall challenge of the project remains its integrated architecture performance and overall compliance with the legal requirements. The end users LEAs have the flexibility to adapt the features to their operational requirements and tune the components to match their needs.

VII. Conclusions and Future Work

We have presented our work of developing a social media forensic tool to improve security and policing of online activities and enable the law enforcement agencies to tackle imminent terrorist activities by analysing social media files without compromising privacy of the users of social media

platforms. Although RED-Alert project is of sensitive nature, the research leading to project results could be applied in other online content analysis applications. Moreover, the project's granular design will facilitate integration of its individual components into other systems.

The multimedia and social media forensic analysis has several other commercial applications that can benefit from our work and leverage their competitiveness and product/services quality. One of the major challenges of multimedia analysis is to reduce the number of false positives and false negatives. We need to make finer grained tuning of SMA tools components by using larger dataset of a broad range of objects and audio variations. Nowadays data collection, processing and storage have become itself very challenging due to the recently enforced GDPR compliance requirements. The situation is improving with the development of new data management processes and good practices for the data protection. We aim to further improve the performance of SMA Tool and evolve it towards a comprehensive Multimedia Forensics Analysis Toolkit. Another challenge to be addressed is to develop tools for hierarchical visualization of time evolving networks, which helps the analyst in understanding the possible correlations and trends at different scales.

ACKNOWLEDGMENT

The research leading to the results presented in this paper has received funding from European Commission Horizon 2020 (H2020) Programme under Research and Innovation Action H2020-SEC-12-FCT-2016 (Technologies for prevention, investigation, and mitigation in the context of fight against crime and terrorism). Grant agreement number 740688.

REFERENCES

[1] Chris Bousquet, Mining Social Media Data for Policing, the Ethical Way, online article of Gov. Tech, 2018 http://www.govtech.com/publicsafety/Mining-Social-Media-Data-for-Policing-the-Ethical-Way.html

[2] Mohammad Tayebi, Uwe Glasser, Social Network Analysis in Predictive Policing: Concepts, Models and Methods, Springer Lecture Notes in Social Networks, 2016

[3] Policing 4.0: Deciding the Future of Policing in the UK, Deloitte report https://www2.deloitte.com/uk/en/pages/public-sector/articles/the-futureof-policing.html

[4] RED-Alert Project: http://redalertproject.eu

[5] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)

[6] P. Viola and M. Jones, Repid Object Detection using a Boosted Cascade of Simple Features, in Computer Vision and Pattern Recognition, 2001.

[7] S. Ren, K. He, R. Girshick and J. Sun, Faster R-CNN: Towards Real-Time Object Detection with Region proposal Networks, 2016.

[8] I. Goodfellow, Y. Bengio and A. Courville, Deep Learning, MIT Press, 2016.

[9] C. Southall, R. Stables and J. Hockman, Improving Peak-Picking Using Multiple Time-Step Loss Functions, in Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR), 2018.

[10] Keycloak Open Source Identity and Access Management htttp://www.keycloak.org

[11] Java Business Process Model (jBPM) Workflow Engine https://www.jbpm.org

[12] Apache ServiceMix - enterprise-class open-source distributed enterprise service bus (ESB) http://servicemix.apache.org

[13] Audit4j Auditing Framework http://audit4j.org

[14] European Commission funded Horizon 2020 Project DANTE (Detecting and analysing terrorist-related online contents and financing activities) https://www.h2020-dante.eu

[15] W. Asif, I. G. Ray, S. Tahir and R. Muttukrishnan, Privacy-preserving Anonymization with Restricted Search (PARS) on Social Network Data for Criminal Investigations, 2018.