History of Information: The case of Privacy and Security in Social Media

Dimitris **Gritzalis,** Miltiadis **Kandias**, Vasilis **Stavrou**, Lilian **Mitrou**

Athens, 2014

History of Information: The case of Privacy and Security in Social Media

Dimitris **Gritzalis,** Miltiadis **Kandias**, Vasilis **Stavrou**, Lilian **Mitrou**

Information Security & Critical Infrastructure Protection Research Laboratory
Dept. of Informatics, Athens University of Economics & Business
76 Patission Ave., GR-10434, Athens, Greece
{dgrit, kandiasm, stavrouv, l.mitrou}@aueb.gr

Abstract. Web 2.0 and Social media have widened society's opportunities for communication amusement and networking, while enabling internet users to contribute online content. The information flows created, along with Open Source Intelligence techniques, can be crawled and processed for a variety of reasons from personalized advertising to behaviour prediction/profiling. In this paper, our goal is to present two horror stories from the digital world of Social Media, in order to: (a). Present an insider threat prediction method by (e)valuating a users' personality trait of narcissism, which is deemed to be closely connected to the manifestation of malevolent insiders, as it was developed via Twitter. (b). Raise user awareness over political profiling and data processing for discriminative purposes, via a methodology applied to YouTube. As both methodologies set user's privacy and dignity at stake, we provide the reader with an analysis of the legal means for each case, so as to effectively be prevented from a privacy violation threat and also present the exceptional cases, such as the selection of security officers of critical infrastructures, where methodologies to detect the insider threat could be used.

Keywords: Awareness, Information History, Narcissism, Political Profiling, Social Media, Surveillance, Twitter, Usage Deviation, User Profiling, YouTube.

1 Introduction

The advent of Web 2.0 has contributed in the transformation of the average user from a passive reader into a content contributor. Users are now able to interact with other people, create, redistribute or exchange information and opinions, and also express themselves in virtual communities. All these means of interaction, along with the exchange of user-generated content, are referred as *Social Media* [1]. Through them, users form their personal identity in the digital world, while revealing aspects of their personality.

Users tend to participate in Social Media for a variety of reasons, including professional networking, amusement and communication with other users. The observation of this constant use of the online world has led to researches which have proved that

individuals tend to transfer their offline behavior online [2]. By combining (a) the availability of such user generated information flows with (b) the ability of processing vast amounts of information, using Open Source Intelligence techniques (OSINT), one may observe that the construction of usage and users patterns is more feasible than ever. Furthermore, these patterns may be utilized for purposes ranging from profiling for targeted advertising (on the basis of analysing online features and behaviour of the users), to personality profiling and behaviour prediction.

An important feature of this valuable source of personal data, namely Web 2.0 and Social Media, is that they are available for crawling and processing, even, without the user's consent. Although, the processing of such data may be used for innocent purposes, so as to provide the user with a more personalized user experience, or a more targeted marketing profile, one may notice that user's privacy violations may occur. The ability of conducting automated psychometric evaluations and revealing personal data regarding political and religious beliefs, sets user privacy at stake and it is important to institutionalize digital rights, in order to protect the user in the online world.

Tools that use OSINT techniques and process online user data in one of the above mentioned ways can be distinguished into two perspectives, regarding the user; the privacy violation perspective and the improved and impersonalised user experience. In the first case, user's data are processed, without her consent, to expose personal data of hers for discriminative purposes, while in the second one, user is provided with a more user-friendly environment adapted to her preferences.

However, there is another perspective of the issue, which can be used to enhance security in critical infrastructures, against the insider threat [3]. This perspective includes the prediction of malevolent behaviours that may cause catastrophic results. An infrastructure is characterized as critical when a possible loss of the confidentiality, integrity, or availability of the services offered by the infrastructure may affect aspects like national security, economic prosperity or social well-being.

Insider threat is one of the most demanding problems in cyber and corporate security and solely technical countermeasures, techniques and methods proposed to tackle it down, fail to reduce the impact of this threat. In principle, the malevolent insider manifests when a trusted user of the information system behaves in a way that the security policy defines as unacceptable [4]. To this end, research suggested [5] that technical and social solutions should be implemented. Furthermore, behavior analysis leads to studying employees under the prism of predisposition towards malevolent behavior, by examining personal traits that have been proved to abut to this kind of behavior. In the past, one should utilize stiff questionnaires and psychometric evaluations, in order to examine the above mentioned personal traits. On the contrary, nowadays social media rise the opportunity to study such traits in an automated and flexible manner, making it possible to perform psychometric evaluations by utilizing the content a user has made publicly available.

In this paper we present two horror stories of how the multifaceted information shared/revealed in the (context of) social media can be utilized, in order to achieve a dual purpose:

(a). *Predict the insider threat via Narcissism*: In a previous work of ours [6], we have proposed prediction and deterrence measures against the insider threat, via Twitter. From the perspective of protecting critical infrastructures from malevolent actions, our goal was to extract conclusions over the users regarding

the personality trait of narcissism, which is a common characteristic among insiders.

(b). *Raise user awareness over profiling*: In order to raise user awareness we have developed a negative case scenario involving political affiliation profiling by processing online available date from YouTube, as presented in a previous work of ours [7]. We formed the hypothesis that this scenario is nowadays realistic, applicable to social media, and violates civil rights, privacy and freedom. Our goal was to reveal this threat and contribute in raising social awareness.

The examination of the ways that information in Social Media can be processed, may reveal the possible exploitation of available data in order to enhance the protection of a critical infrastructure or the user privacy breaches that occur. We have applied our methodologies to two popular Social Media, so as to highlight the applicability and the success rate of such techniques. We, also, exploited the available data in Social Media in order to examine the ethical and legal issues that arise from such profiling methods.

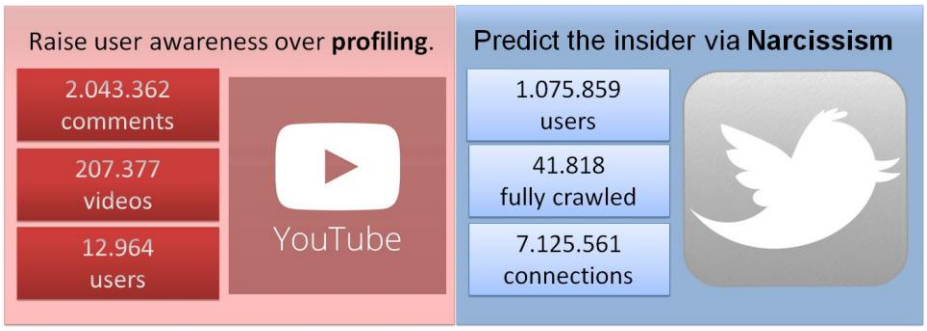

Fig. 1 Testing Environments

The paper is organized as follows: in section 2 we review the existing literature. In section 3 we describe the Twitter case study. In section 4 we describe the YouTube case study. In section 5 we discuss the ethical and legal issues that rise from the methods described in each case. In section 6 we sum up our findings.

2 Related Work

Web 2.0 and Social Media are characterized by the creation and exchange of user-generated content. As users tend to increasingly use the Social Media, one may notice that they form a valuable source of personal data, which are available for crawling and processing without the user's consent. The rise of social media usage has influenced researchers towards opinion mining and sentiment analysis [8], which constitute computational techniques in social computing [9]. As presented by King et al., social computing is a computing paradigm that involves multi-disciplinary approach in analysing and modelling social behaviour on different media and platforms to produce intelligence and interactive platform results.

User and usage profiling, along with conclusion extraction from content processing, have become more feasible and valuable than ever before. Several methods have been introduced that are able to processes online data form profile patterns. Researchers have explored ways that expressions of human moods can be measured, inferred and expressed from social media activity [10]. Additional methods that have been utilized to demonstrate the profiling threat include user behaviour characterization in online social networks [11], and analysis of the relationship between users' gratifications and offline political/civic participation [12].

Users often appear to be unaware of the fact that their data are being processed for various reasons, such as consumer behaviour analysis, personalized advertisement, opinion mining or profiling. In order to raise awareness, researchers have conducted attacks on realistic environments, consisting of Social Media communities or groups. Such realistic approaches are proposed as a reliable way to estimate the success rate of an attack in the real-world and also raise user awareness over the threat. Balduzzi et al. [13] performed automated user profiling and opinion mining to extract conclusions over a crowd of users.

For a long time academic and corporate security researchers have been trying to propose solid countermeasures to deal with the insider threat [14]. Insider threat, apart from being a demanding problem in cyber and corporate security and has also been identified as a threat in the cloud [15]. Mitigating to a cloud infrastructure may offer several benefits [16], however various threats [17] should be considered [18] and among them the potential insider threat that has to be tackled down. Regarding the area of insider threat prediction, it involves various methods and techniques that have been proposed to predict insiders [19][20]. Other approaches include scope-specific attempts in relational databases [21] or highlight the need for both technical and psychological approaches [22]. The psychosocial perspective of the insider is also referred to by Greitzer et al. [23] and Brdiczka et al. [24]. Shaw et al. [25] studied the parameter of the psychological state of an insider and presented personal and cultural vulnerabilities of a person that may indicate predisposition towards malevolent behavior.

FBI also presents the personal factors that may increase the likelihood someone to develop malevolent behavior [26]. Personal traits such as anger, revenge or greed along with certain circumstances present in an organization could lead to the manifestation of a malevolent insider. Profiling methods that utilize social media intelligence may contribute towards detection of common characteristics that insiders have been found to share as automated user profiling can lead to accurate prediction of personal information [27][28][29]. Furthermore, researchers have noted the social threat that emerges from automated user and usage profiling.

3 Twitter: Catching the insider

The success story applied to Twitter aims at enhancing the prediction front of insider threat. To this end, we have focused on a Greek community of Twitter. We use data crawled by Twitter, so as to analyze the collected users in a graph theoretic manner. In addition, we identify connections between usage patterns and define which users'

behavior could be considered as deviating from the average. The use of such method interferes with the personality and privacy rights of the affected persons. It is important to mention that data is gathered from publicly available information, which results from Twitter communication. The selected data are used for prediction and deterrence purposes, since we analyze each user under the prism of usage deviation with the tool of graph theory.

Moreover, the ability to rapidly conduct psychometrics for such a large number of people may become a social threat. Thus, we have adopted a privacy-sensitive and pro-employee attitude over the results of this work. The potentially intrusive nature of this method dictates the necessity of its confined application to certain information systems and organizations. This application may be acceptable in exceptional cases, such as for selecting security officers, as well as for personnel involved in the decision-making process within security-critical information systems and critical infrastructures.

3.1 Motivation

We have chosen Twitter to apply our methodology, as Social Media and collaborative environments have been used in the battle against the insider threat [30]. Usage deviation aids in detecting narcissistic behavior through social media popularity and intense usage of the media [31][32][33]. Interestingly, the psychosocial trait of narcissism is closely related to delinquent behavior, especially regarding the insider threat [25].

Twitter is a popular social medium. Data from Twitter are often utilized in order to find out why some users of certain communities distinct from other users of the same communities. A *distinct community* is defined as a set of users who share common characteristics. In [34], Mislove et al. present a large-scale study of multiple online social networks and refer to the small-world phenomenon [35].

An important question for the research with a social medium is *"who is influential in a social network". User influence* can be defined as the ability to convince an audience to engage in a single act. As only a minority of users has the ability to persuade others in a higher level, research shows that influential users have a certain way of expressing themselves via social media, while they decorate their tweets with appropriate sentiment [36].

3.2 Methodology and testing environment

In our work [6] we focused on a Greek community of Twitter. We provided a graph analysis of this community using specific metrics. We also defined the content and measures of user influence. Furthermore, we analyzed the different ranks of users, ordered by their influence valuation score, and presented the set of users who are outliers both in the whole graph and in smaller communities. The graph theoretic methods are utilized so as to tackle the insider threat under the prism of outlier detection, narcissism detection, and group homogeneity. Our goal is to analyze the collected

data, in order to extract useful results about the psychosocial trait of narcissism of the users, as well as the group they belong to inside an organization.

The social actions of a Twitter user can have more than one label. Thus, there are three types of user categories: (a) follower, i.e., she is followed by someone, (b) following, i.e., she follows someone, and (c) retweeter, i.e., she spreads the speech of someone else via tweets.

In order to create a community consisting only of Greeks, we crawled only those users who have published on their profiles that they are Greeks. In case someone had not published such information, we checked the language she used or her *geolocation*.

The dataset includes a social graph of: (a) 1.075.859 distinct users, (b) 41.818 fully crawled users, where we have managed to collect their full profile, and (c) 7.125.561 connections among them. The graph created on the basis that (a) each user is a node, and (b) every connection is a directed edge.

We propose three ways to measure user influence and tested these metrics with our 41.818 fully crawled users. Regarding user information, we kept each user's screen name, id, the description that she wrote about herself, her url, her language, and her *geolocation*. We also kept the state of her profile (protected or not), the number of lists she is participating, the numbers of her following and follower users, all her tweets, the number of her favorites, the number of tweets she has mentioned, and the number of retweets she has made. As a result, we ranked the users by their scores and we identified ways for distinguishing users who are outliers in a community.

3.3 The graph-theoretic approach

The main conclusions regarding the structure of the Greek Twitter community as initially described in [6] are the following:

(a). We detected that there is a *large strongly connected component* (153.121 nodes) and several others (significantly smaller). The largest strongly connected component of the graph, where everyone connects to each other, is small.

(b). We detected that 99% of the nodes are connected to someone else. We detected that 1.075.815 users are connected to someone.

(c). The *small world phenomenon* does apply to the collected Greek community, as every user of the community is 6 hops away from everyone else**.**

The term user's indegree value indicates the number of users who follow the user, while outdegree value is the number of users to whom she points. In order to identify the characteristics of the users in this Twitter community, we used the following three metrics:

(1). *Commonality valuation of the users*: We produce a statistical distribution of the *commonality* of the users, i.e., a representation of usage intensity of the specific social media. This utilizes the data collected from the public user profiles and a combinatorial function is created that aggregates the values of seven parameters that emerge from a user's profile, as follows: (a) number of followers, (b) number of followings, (c) number of tweets, (d) number of retweets, (e) number of mentions, (f) number of favorites, and (g) number of lists. We detected that there are many users

with small usage in Twitter. On the contrary, a much smaller set of users shows intense usage of Twitter.

(2). *klout score*: We reveal the *outliers* in this community by taking into account their "Klout scores" (klout.com). The *klout score* is a metric that represents someone's social media influence. We first crawled the klout scores of each user in the Twitter community, and then we ranked the users according to this score.

(3). *Influence valuation of a user*: We define someone's *influence* by seeking the set of individuals around her - considering her as a central user - rather than her usage intensity in social media. Every user has a specific subset of users which she affects. Thus, we focus on the set of users who are possible candidates to adopt her words by retweeting them. The influential set of each user consists of: (a) followers who directly learn her quotes, (b) her mentioners, (c) retweeters who mention her or repeat her word of mouth, even without following her, and (d) the followers of her last two categories, as there is a possibility to learn about her indirectly. Therefore, we identify the influential sets of each user in the graph and rank them by the size of these sets, taking into consideration only the number of distinct users in each set. We use these ranks of the Greek Twitter Community users in order to observe the minimum, the average, and the maximum values in each case. The average values can reveal the common users of social media, while the maximum values can determine the outliers.

3.4 Outlier's common characteristics and detection

The results that refer to the Greek Twitter Community indicate that this is not true, i.e., (a) the majority of the Greek users make very poor use of the medium, (b) there are a lot of normally active users, and (c) very few users are popular.

Now we propose a general taxonomy of the Twitter users, the data of whom were crawled and analyzed. According to our findings, the most influential users' influence valuation is between 942 and 3604 and usage valuation is between 21004 and 569000. Based on this, users whose sum of the previous values is higher of the threshold of 22000 are classified in a different category of the taxonomy. Furthermore, the majority of the users with usage valuation above 21000 are either real life celebrities or news media. This leads us assume that the "normal" users with high scores should belong to a different category. The proposed categories appear on Table 1.

The question that emerges is "why one should care of the usage differentiation or deviation between the users". Based on the available data, we can spot a threshold above which the users may become quite influential and perform intense medium usage. Therefore, we can define a specific point where a user turns from a normal one to a "media persona".

Table 1 The proposed Twitter user taxonomy

Category	Influence valuation	Klout score	Usage valuation
Loners	0 - 90	3.55 - 11.07	0 - 500
Individuals	90 - 283	11.07 - 26.0	500 - 4500

Known users	283 - 1011	26.0 - 50.0	4500 - 21000
News Media & Personas	1011 - 3604	50.0 - 81.99	21000 - 569000

Research has proved that individuals tend to transfer their offline behavior online. Thus, more *extravert* individuals tend to form large groups and communicate easier in the territory of social media, while *introvert* individuals tend to communicate less. Furthermore, research work has connected excessive usage of social media to the personality trait of narcissism [31][32][33]. The connection between narcissism and insider threat has been verified in a number of research works. In specific, research has demonstrated that a narcissistic personality is more vulnerable to become a malevolent insider [25].

In order to determine a user's narcissism and the level of usage differentiation, our method first utilizes a user's fully crawled Twitter profile. Then, this user's influence over a social medium, as well as his overall activity in this medium, is also evaluated. These actions may - in certain cases - be considered intrusive and violate privacy, therefore they should only be performed in the course of a legitimate action.

3.5 Group Homogeneity

Research clearly indicates that poor relations with co-workers and/or supervisors, as well as dysfunctional working groups not only facilitate the manifestation of insider threats, but also catalyzes an insider's malevolent behavior [30]. Some researchers refer to the need for thorough screening of the employees prior to employment, especially for those who are going to occupy high risk positions [25][30]. As a result, it has been suggested [30] that a background check would be essential, not only regarding a users' criminal background, but also under the prism of work group homogeneity.

Researchers have suggested that management and human resources staff should maintain awareness of users' satisfaction and well-being [37]. An important parameter of employees' satisfaction is the sense of belonging to the group, which depicts the group's homogeneity. Shaw et al. [38] explains that 90% of insider cases involve serious employment crises and personnel problems prior to the attack. Moreover, destructive group dynamics analysis has identified similar results [24].

Based on the above, a three-step process can be used for user screening and group homogeneity testing. The scope of each step is the following: (a) how well could a new user fit to an existing group, in terms of group homogeneity, (b) how homogeneous is an existing group inside the organization, and (c) how similar is the specific user's social media behavior to other users' of the same profession.

In order to evaluate how well a newcomer could fit in an existing group, one could check her Twitter account and crawl it. Then, the crawled data could be sent as input to a graph theoretic analysis that calculates her influence on the media and her overall usage valuation. Having a collection of data about the users of the group, one can define a range of acceptable media influence and usage values for the newcomers.

In case we wish to study the homogeneity of a specific group in an organization, we can utilize a fraction of the above process. In specific, we need to crawl each group member's Twitter account and feed the results to the appropriate algorithms.

Then, we should analyze the results of each user's influence and usage assessment, so as to decide over the homogeneity of the group.

4 YouTube: Exploiting user's privacy

The horror story involves political detection affiliation using the online available data from YouTube. To this end, we have experimented with an extensive Greek community of YouTube. We present a political affiliation detection method, i.e. the Panopticon, in order to raise user awareness over political profiling via social media. Furthermore, we present our findings related to political profiling as a proof-of-concept. The twofold purpose of this research is to (a) raise users' awareness over political profiling, and (b) highlight the social threat of processing users' online available data for discriminative purposes.

The dataset includes: (a) 12.964 users, (b) 207.377 videos, and (c) 2.043.362 comments. The time span of the collected data covers a period of 7 years (Nov. 2005 - Oct. 2012). In addition, data was classified into three categories: (a) user-related information, e.g., profile, uploaded videos, subscriptions, favorite videos, playlists, (b) video-related information, e.g., video's license, number of likes and dislikes received, category and tags, and (c) comment-related information, e.g., the content of the comment and the number of likes and dislikes it received.

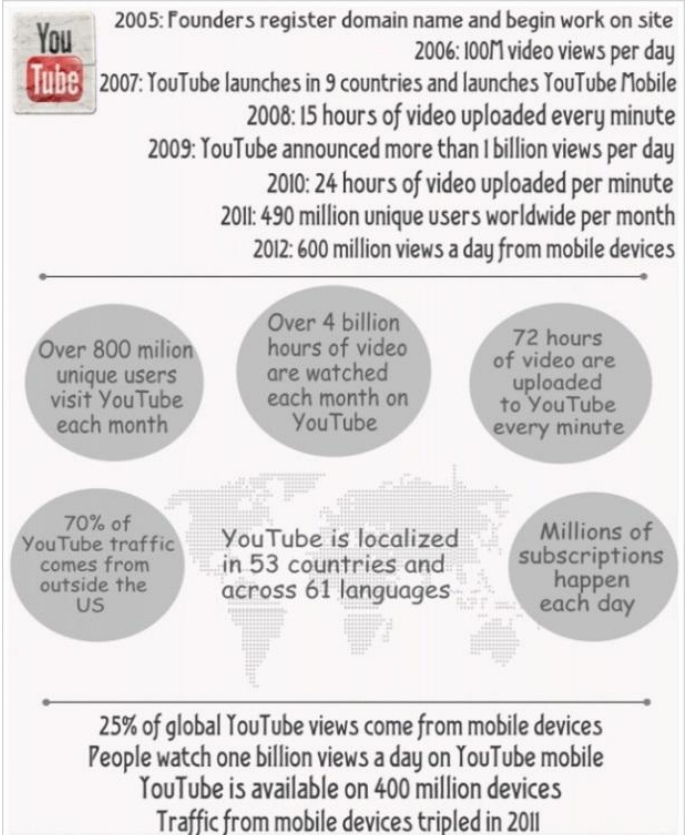

Fig. 2 YouTube Penetration

4.3 Tag Cloud Description

For better observing the axis of content of the collected data, we visualized the results in the form of a tag cloud. This is demonstrated in Fig. 2. Tags "Greece" and "greek" appear frequently in the dataset because the experimentation focuses on a Greek community of YouTube. The majority of the tag cloud tags are Greek words written in Latin (i.e. "greeklish). We have transformed the Greek tags to greeklish, in order to deal with duplicates of a word (one in Greek and one in greeklish).

The majority of videos are related to music and entertainment. The next topic that can be found on the YouTube video tags is sports. Several tags containing Greek sports teams' names are also shown in the tag cloud. One may also notice political content in the tag cloud (i.e., tags with the names of the Greek major political parties).

Fig. 3. Tag cloud of the dataset

4.4 Panopticon and YouTube

YouTube is a popular social medium. Furthermore, it grows exponentially along with user generated content that it hosts. YouTube is characterized by emotional-driven responses in its comments because of videos' emotional content. Audio-visual stimuli, combined with the anonymity offered by usernames, appear to enable users to express their feelings and opinions, regarding the content of a video. Also, YouTube users are able to interact with each other through video comments or subscriptions. Even though it is not essential for the users to have formed a real life bond, they can interact on the basis of common interests, views, hobbies, or political affiliations.

Our observations indicate that users tend to participate in the medium and generate personalized content. YouTube videos and comments contain political characteristics as presented in the tag cloud. Thus, we formed the hypothesis that political affiliation may be extracted via content analysis. Based on our observations, we consider that: (a) YouTube often contains political content, (b) users often feel free to express their opinions, especially when it comes to politics (because of the anonymity they assume and the emotional content of the medium), (c) most users join YouTube to participate, so one can reasonably expect that they will reveal, inter alia, their personal data.

4.5 Drawing Conclusions

In this section we demonstrate a method that enables the identification of a user's political affiliation via the political beliefs expressed within the comments of her videos, i.e. the Panopticon methodology. Video is YouTube basic module. As we cannot process the video itself, we draw a conclusion for each video via its comments.

We detect the political affiliation expressed in a comment by performing text classification into three main categories: (a) category R, which contains expressions related to radical affiliation, (b) category C, which contains expressions related to

conservative affiliation and (c) category N, which contains all the comments that hold a neutral political stance or have no political content.

Text classification uses machine learning techniques to classify a comment into the appropriate category. A comment is assigned to one of the three categories, so as to indicate the respective political affiliation that the specific category depicts. We further examined the efficiency of another technique. Thus, we formed a dictionary that included a series of words and phrases that indicate political affiliations to the specified categories. According to our findings, machine learning leads to a more reliable result than a simple word existence check and also, text classification performs better than scanning lists of words in a dictionary.

Comment classification enables to extract conclusions for a video's political affiliation. The conclusion drawn helps us to classify any video into one of the defined categories of political affiliations. So, by assigning a comment into a category implies that the conclusion drawn for the comment is the political affiliation expressed in the category. The same applies to a list of videos, such as favorite videos and playlists. Having the category in which a video falls into, a conclusion can be drawn for the political affiliation expressed in the list. Being able to classify user's content, we may extract conclusions for user's comments, uploaded videos, favorite videos and playlists. This way we can draw a final conclusion about user's political affiliation.

4.5.1 Comment Classification

According to the methodology we followed in [7] we formed a training set by selecting comments from the database and assigning them a proper label for each one of them, based on the category it belongs to. We consulted a domain expert (i.e., Sociologist), who could assign and justify the chosen labels on the training sets. Thus we created a reliable classification mechanism. We chose 300 comments from each category (R, C, N) of the training set for each language. The expert contributed by assigning a category label to each comment.

We performed comment classification using: (a) Naïve Bayes Multinomial (NBM), (b) Support Vector Machines (SVM), and (c) Multinomial Logistic Regression (MLR), so as to compare the results and pick the most efficient classifier. We compared each classifier's efficiency based on the metrics of precision, recall, f-measure and accuracy [40].

Table 2 presents each classifier's efficiency, based on accuracy, precision, recall, and f-score metrics. Multinomial Logistic Regression and Support Vector Machines achieve the highest accuracy. The accuracy metric is high due to the dominant number of politically neutral comments. Precision and recall are proper metrics to evaluate each classifier [40].

Table 2 Metrics comparison of classification algorithms

	Metrics								
Classifier	**NBM**			**SVM**			**MLR**		
Classes	**R**	**N**	**C**	**R**	**N**	**C**	**R**	**N**	**C**

Precision	65	93	55	75	91	74	83	91	77
Recall	83	56	85	80	89	73	77	93	78
F-Score	73	70	60	76	89	73	80	92	77
Accuracy		68			84			87	

Multinomial Logistic Regression achieves better precision value and SVM better recall value. Multinomial Logistic Regression achieves a slightly better f-score assessment. Support Vector Machines and Multinomial Logistic Regression achieve similar results regarding both recall and precision metrics. As a result, we chose Multinomial Logistic Regression because of the better f-score value achieved for each one of the categories.

4.5.2 Video Classification

Regarding the extraction of the political affiliation expressed in a video, we studied each video, based on its comments, classified to one of the three categories. Also, we know the number of likes/dislikes each comment has received. Likes and dislikes represent the acceptability a comment has from the audience, so it may be an indication of the comment's importance to the overall video's result. Thus, a comment that receives a significant number of likes should be treated differently than a comment with no likes, as the first one is acknowledged as important by more users. This assumption has been confirmed by the data mining process. Subsequently, in order to extract a conclusion for the video, we take into consideration only comments that belong either to category R or C. Neutral comments are ignored.

Each comment importance is measured via its number of likes and dislikes. In order to come to a video overall result we utilize two sums, one for category R and one for C. For every comment that belongs to categories R or C we add the following quantity to the respective aggregation:

$$1 + \{(\text{likes}/\text{total\_likes}) - (\text{dislikes}/\text{total\_dislikes})\}$$

The quantity added to each sum shows that a comment that has received more likes than dislikes should affect the overall score more than a comment with more dislikes than likes. Finally, the category with the larger sum is the category that represents video's political affiliation.

4.5.3 List Classification

The procedure followed to extract a conclusion about a list of videos is similar to the above mentioned video method. The only difference is that we utilize videos instead of comments. The two sums are also applied, one for category R and one for C. In this case, instead of likes/dislikes we used the video ones. In the end, the category with the greater sum is the result for the list's political affiliation. This procedure is applied to the "favorite videos" list, as well as to the other playlists that the user may have created.

4.5.3 User Classification

A user political affiliation can be identified based on the category she is assigned to. The procedure, as shown in Fig. 3, takes into account a user's comments, her uploaded videos, her favourite videos, and her playlists. A user is able to: (a) write a comment to express her feelings or her opinion, (b) upload a video (the content may have a distinctive meaning for her), (c) add a video to her favourites list (it may have an emotional or intellectual meaning for her), and (d) create a playlist and add videos to it.

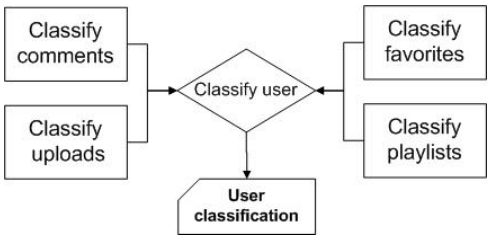

Fig. 3. User classification process

Based on these observations one may look for indications of political beliefs within the user generated content. For the needs of our experimentation, we have defined ad-hoc weights for each of the cases we examine (Table 3). Each phase of the process generates a result on the category each user belongs to. In comment classification, the result is based on the number of political comments that exhibit the highest aggregation. If the comments classified in category R are more than those classified in category C, then the result is that user tends to be Radical. The results on uploaded videos, favorite videos, and playlists are extracted as described in the list result extraction process. User comments are the most important factor to decide of a user's political affiliation.

Regarding the aggregation of the results, we utilize two sums, one for category R and the other for C. Comments, videos, and lists classified as neutral do not contribute to the aggregation. The sub-results are appropriately weighted and added to the final sums, in order to extract the final result.

5 Ethical and legal issues

The Panopticon methodology allows the automated classification of users into predefined categories, based on the content they have generated in Social Media, even without their consent. This may lead to exposure of personal data in an automated way. On the other hand, using the information available from Social Media, may contribute in tackling down the insider threat in critical infrastructures.

Thus, applying such methods in unethical manner raises several ethical and legal issues. Violating user privacy for profiling raises several emerging threats, while

using automated profiling to enhance lawful mechanisms may be a useful median in the appropriate hands.

5.1 Emerging threats

The examination of YouTube's horror story indicates that the use of methods like Panopticon may result to problems that are actually inherent in every kind of profiling. In brief, these methods may be regarded as a kind of (behavioral) profiling on the Internet, in the meaning of collecting data (recording, storing, tracking, etc.) and searching it for identifying patterns. Such profiling methods interfere with the right to informational privacy and are associated with discrimination risks.

The observation of the behavior and characteristics of individuals through mining of large quantities of data may infringe fundamental rights, let alone the determination of correlation between characteristics and patterns and the respective classification of individuals. A major threat for privacy rights derives from the fact that profiling methods can generate sensitive information "out of seemingly trivial and/or even anonymous data" [41].

By studying user's uploads it is possible to extract information related to the content, especially when it refers to areas such as political affiliation. Furthermore, a user is possible to have a private profile, however her comments could be collected from crawling random videos. Thus, a limited profile can be build based on those comments. The predominant rationales for acquiring knowledge about the political opinions and the relative sentiments seems to be either (political) research purposes or the goal of reducing risks both in the private and the public sector. However, personal data that are, by their nature, particularly sensitive and vulnerable to abuse, deserve specific protection.

Collecting and processing data about political beliefs is regarded by law as a highly exceptional situation. Many international and national laws prohibit explicitly the processing of personal data revealing political opinions (e.g. Art. 8 of the European Data Protection Directive and Art. 6 of the Convention 108 of the Council of Europe). Derogating from the prohibition on processing this "sensitive category" of data is allowed if done by a law that lays down the specific purposes and subject to suitable safeguards. Such derogations rely on a manifest public interest or the explicit, informed and written consent of the person concerned.

However, in European data protection law derogation is sometimes allowed also in the cases that "the processing relates to data which are manifestly made public by the data subject" (Art. 8, §2e of the European Data Protection Directive), which is the case if people generate content or comment on other users' content in social networks or media using their real identity and aiming at expressing their opinions publicly. According to the American theory and jurisprudence there is no "reasonable expectation of privacy if data is voluntarily revealed to others" [42]. It is "apparent", according to this theory, that one cannot retain a reasonable expectation of privacy in the case of YouTube, videos, likes, and comments left open to the public [43].

By generating content in social media users are generating information flows and aggregations. Providers and Online Social Networks encourage - also through the de-

fault settings - “producers” to publish personal information and enable anyone accessing this information thus actively contributing to shaping social media as an attractive product [44]. Does self-exposure in social media amount to freely and consciously chosen privacy abandonment?

Although, YouTube offers several privacy options to users [7], we should take into consideration individual’s general inertia toward default terms [45][46][47]. Moreover, it seems that the majority of users choose to disclose their personal data to as many users as possible, although average users do not have a clear idea about the actual reach of information they reveal or they underestimate the possible reach of their profiles visibility [44].

Users are losing control over their data and the use thereof, as they are becoming detectable and “correlatable”. The combination of all this information provides a powerful tool for the accurate profiling of users. Moreover, it is quite simple to identify a particular person, even after her key attributes (name, affiliation, address, etc.) have been removed, based on her web history.

However, even if individuals are profiled in a pseudonimised way they may be adversely influenced [48][**Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**]. Informational privacy protects individuals against practices that erode individual freedom, their capacity for self-determination, and their autonomy to engage in relationships and foster social appearance. If individuals fear that information pertaining to them might lead to false incrimination, reprisals or manipulation of their data, they would probably hesitate to engage in communication and participatory activities [48]. The autonomy fostered by informational privacy generates collective benefits because it promotes “reasoned participation in the governance of the community” [51].

Risks of misuse and errors arising out of the aggregation and data mining of a large amount of data made public for other purposes are manifest. According to the German Federal Constitutional Court the “cataloguing” of the personality through the connection of personal data for the purpose of creating profiles and patterns is not permitted (Judgment of the Bundesverfassungsgericht, 4 April 2006, 1 BvR 518/ 02, 23.05.2006). A mass profiling of persons on the base of their views expressed in social media could have intimidation effects with further impacts on their behavior, the conception of their identity and the exercise of fundamental rights and freedoms such as the freedom of speech [52]. Fear of discrimination and prejudice may result to self-censorship and self-oppression [53]. Indeed, while profiling risks are usually conceived as threats to informational privacy we should point out the - eventually more - significant and actual risk of discrimination [54]. The safeguards relating to the use of personal information aim - among others. If not principally - at preventing discrimination against persons because of their opinions, beliefs, health or social status. Studies conveyed how profiling and the widespread collection and aggregation of personal information increase social injustice and generate even further discrimination against political or ethnical minorities or traditionally disadvantaged groups [48].

Individuals may be confronted with major problems both in their workplace and in their social environment. Employers or rigid micro-societies could demonstrate marginalizing behavior against persons because of their deviating political affiliation. There are a lot of historical examples of people who have been side-lined by the

hegemonic attitude of society. One should not look for numerous examples in order to evaluate this thesis: Victor Hugo's "Les Miserables" (Section X: The Bishop in the presence of an unknown light) is the most representative evidence towards this result.

If we may generalize the above mentioned consideration to a macro environment, consequences to the deviating from the average political affiliation could lead to mass social exclusion, prejudice and discriminations. Such minorities may even be considered de facto delinquent and face a social stigma. In the context of a totalitarian/authoritarian regime, implementation of such methods could lead to massive violation of civil and human rights or even threat the life of specific individuals.

5.2 Lawful mechanisms

Online social media profiles, blogs, tweets, and online fora are increasingly monitored by employers searching for information that may provide insight on employees and prospective hires [53] [55]]. A broader and potentially less censored or more honest array of information is easily accessible on the Internet. Taking into consideration the exponentially growing participation in online social networking sites and social media, it is not surprising that employers are searching for unique information about applicants and employees not found with other selection methods. In particular in the US, recent surveys point out lifestyle concerns among the most common reasons for rejecting candidates [56]. Such findings indicate that the once clear lines between the private and the public, as well as the employee's personal and professional life, are gradually blurring as a result of : (a) the "boundary-crossing technologies" [57], (b) the transformation of workplace structure and ethos through ICT, and (c) the radical changes in mass self-communication [58]. In other words, the more private information has become easily accessible and infinitely shareable and transferable, the more monitoring may extent to private spaces, activities, and time [53][55]. Methods (such the proposed in this paper) allows employers to collect and aggregate information, which reflects behavior of the user and her interaction with other users, in order to produce relevant patterns/profiles and anticipate future behaviors and threats.

Employers are, in principle, not prohibited to consider information about a person who is documented in publicly available social media profiles, public posts, or public Twitter accounts. However, both the wide availability of private information, as well as its use beyond the initial context that this information has been produced, may have far reaching effects for the employees' rights and liberties. With regard to the findings referred to in our paper, we should take into consideration that the insider threat prediction and prevention raises ethical and legal issues concerning the protection of employees' privacy, personality, and dignity.

The openness and sharing culture that dominates the online social media reflects a population that does not construct communication on the traditional division between private and public contexts. Many argue that when one publishes something to all comers it is not reasonable to expect (current and future) employers to respect her privacy and freedom of expression and refrain from judging her based on publicly available information [57]. This is true, in particular when employers have legitimate interests to protect their secrets and reputation, and ensure a secure business environment.

On the other hand (informational), privacy responds to the requirement that everyone should be in control of the information concerning her, so as to formulate conceptions of personal identity, values, preferences, goals, and to protect her life choices from public control, social disgrace, or objectification. Individuals tailor their social identities and aim at controlling others' impressions and opinions of them through behavior and performances within particular audiences [57]. Informational privacy offers safeguards to preserve an underlying capacity for autonomous decision - and choice-making [55] and to maintain a variety of social identities and roles. Moreover, privacy is a requirement for maintaining the human condition with dignity and respect. As related to privacy, dignity summarizes, among other principles, the recognition of an individual's personality, non-interference with another's life choices, and the possibility to act and express freely in society. Employer's intrusion into an employee's personal life through "social media background checks" may lead employers to judge opinions and behaviors out of their initial context] De-contextualization is an inherent characteristic of social media that pertains to over-simplification of social relations and the wide dissemination of information [59].

However, as Nissenbaum [60] underlines, the definitive value to be protected by the right to privacy is exactly the "contextual integrity" of a given contextual-self having different behaviors and sharing different information depending on the context. Information gathered through social media analysis is normally not only unintended as application information but often job-irrelevant or, moreover, related to sensitive activities and, consequently, information of the person concerned (religion, political beliefs, etc.) [61]. We should also take into account that this information may be inaccurate or not timely, reflecting a different life-phase of the person. Due to the Internet's excessive and perpetual memory, it is becoming harder and harder for persons to escape their past.

Furthermore we should consider that the social and communication norms which dominate the social media appear not only to lead to projected identities that job applicants may not wish to be seen by potential employers [62] but also to encourage exaggeration, bravado or shameless behavior [63]. Moreover, social media screening may expose employees and candidates to discrimination [56]. Profiling with the aim to gain probabilistic knowledge from data of the past and (to) propose/ predictions and identify risks for the future may infringe privacy as a right to be a multiple personality and carries far-reaching consequences in terms of social selection and unjustified and - often - invisible discrimination. Profiling may indeed entail the risk of formatting and customization of individual behavior that affects her personal autonomy [40]. Extending monitoring to social communication relationships of employees and candidates augments the chances of employers to influence behavior and promote the "well-adjusted employee" [45]. Information gathering about employee performances outside the traditionally conceived work sphere not only increases the dependence on (future) employers but has also a chilling effect on individuals' personality and freedom of speech. This is so, as they may sacrifice "Internet participation to segregate their multiple life performance" [56] and consequently refrain from expressing themselves.

Employees are routinely asked to sacrifice privacy rights to managerial interests like productivity, prevention and detection of threats and liability risks. Given the workplace belongs to the "public sphere", scholars argue that employees, who are hir-

ed to attend company business, cannot have a (subjective) "reasonable expectation of privacy" that society (objectively) accepts and legitimizes. American Courts are reluctant to recognize a workplace privacy right: in any case reasonable expectation of privacy of employees should be judged under all the circumstances and must be reasonable both in inception and scope (Supreme Court, Case O'Connor vs. Ortega). In the employment context privacy, if any, seems to be exchanged for something of commensurate value, like taking or keeping a job [65]. Regarding privacy as a purely bargainable and alienable right ignores the dignity element, inherent in the notion of privacy. The European approach seems diametrically opposite in many respects: Privacy is not conceived as a right to seclusion and intimacy but as a phenomenon, a protectable situation that regards the relationships between a person and its environment/-other persons. The European Court of Human Rights (Niemitz v. Germany) rejected the distinction between private life and professional life. According to the Court, European employees have "a right to dignity and a private life that does not stop at the employer's doorstep".

Finally, it has been found that excessive monitoring disturbs the relationship between the employer and the employees. It has been proved that employees whose communications were monitored, suffered from higher levels of depression, anxiety and fatigue than those who were not monitored, within the same organization [53]. The panoptic effect of being constantly monitored even concerning activities that fall out of the workplace frame has negative impacts on the employer-employee relationship that should be based on mutual trust and confidence [66][67].

6 Conclusions

In this paper we dealt with a success and a horror story of how the multifaceted information shared/ revealed in the (context of) social media can be utilized, in order to achieve a dual purpose:

(a). Deal with the insider threat prediction and prevention, as malevolent insiders and predisposition towards computer crime has been closely linked to the personality trait of narcissism. We proposed a method of outlier detection in social media via influence, usage intensity, and klout score valuation, in order to detect users with narcissistic behavior. We have also proposed a method for group analysis under the prism of group homogeneity, as this homogeneity is a valuable characteristic to deter the manifestation of insider threats.

(b). Deal with the possibility of a social threat that is based on user generated content exploitation and leads to political affiliation profiling; namely a new panopticon of the digital era. Political beliefs and affiliation have been a cause for social marginalization, prejudice, and discrimination, especially in totalitarian and authoritarian regimes. Thus, we bring this issue to the fore and contribute to the debate and awareness raising. A user might want to protect personal information other than political affiliation, namely information related to sexual orientation, racial discrimination or even the health condition of the user regardless of the national scope. Such an improper information disclosure could be easily conducted via expansion

of the panopticon methodology on the condition that domain experts of each case are available to interpret the collected data and train an appropriate model.

To demonstrate the efficiency of the aforementioned horror stories, we collected a vast amount of data from Social Media. To predict the insider via narcissism we collected data from Twitter. Then, we adopted a specific graph theoretic approach to analyze the crawled data. We focused on a fraction of Twitter users, i.e., a community of 41.818 Greek users.

To raise user awareness via YouTube, we collected a number of 12.964 users, 207.377 videos and 2.043.362 comments from YouTube. Afterwards, we conducted content and graph theoretic analysis of the dataset in order to verify that it is possible to extract conclusions over users' political affiliation, using the panopticon methodology. Our results confirmed the initial hypothesis that YouTube is a social medium that can support the study of users' political affiliation, namely audio-visual stimuli along with the feeling of anonymity enables users to express their political beliefs, even the most extreme ones.

Regarding the ethical and legal issues that rise from both horror stories, in the case of insider threat prediction, privacy violations may occur, in case someone chooses to apply the proposed method in an illegal or unethical manner. Users' privacy and dignity may be at stake if someone uses the method to promote employee/user discrimination and careless punishment. Therefore, the method should be utilized in the course of a legitimate action. Due to the nature of the employment relationship, in which there is an inherent asymmetry of power, reliance on consent for monitoring and screening is highly questionable. Consent should be confined only to, the very few, cases where the employee has a genuine free choice and is subsequently able to withdraw the consent without detriment [68][69]. The employer's monitoring policy should be tailored to the type and degree of risk the employer faces and the level of tolerated privacy intrusion depends on the nature of the employment as well as on the specific circumstances surrounding and interacting with the employment relationship [67]. Thus, the dialectic over the issue should follow the road of selecting the appropriate field of application for these profiling techniques, similarly to other proactive detection techniques, even of differentiated scope [70][71][**Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**][72]. To further elaborate on this issue, we propose that the method is applied for security officers and for personnel involved in the decision making process within security critical information systems and critical infrastructures.

Following, we highlighted possible consequences of an alleged implementation of the described panopticon method. Regardless of the scope of the implementation, the resulting threats include working place discriminations, social prejudice or even stigma and marginalization of the victims. These phenomena could be identified even in a democratic and stable society, not to mention the threats one could face in a military or totalitarian regime. Thus, we adopted a pro-privacy attitude and included a legal point of view in our analysis, along with the emergence of the demand for raising social awareness over this threat and the necessity for institutionalization of digital rights.

References

1. Kaplan, A., Haenlein, M.: Users of the world, unite! The challenges and opportunities of Social Media. In: Business horizons, vol. 53, no. 1, pp. 59--68 (2010).
2. Amichai-Hamburger, Y., Vinitzky, G.: Social network use and personality. In: Computers in Human Behavior, vol. 26, pp. 1289--1295. (2010).
3. Randazzo, M., Keeney, M., Kowalski, E., Cappelli, D., Moore, A.: Insider threat study: Illicit cyber activity in the banking and finance sector, CMU/SEI-2004-TR-021. (2005).
4. Theoharidou, M., Kokolakis, S., Karyda, M., Kiountouzis, E.: The insider threat to information systems and the effectiveness of ISO17799. In: Computers & Security, vol. 24, no. 6, pp. 472--484. (2005).
5. Lee, J., Lee, Y.: A holistic model of computer abuse within organizations. In: Information Management & Computer Security, vol. 10, no. 2, pp. 57--63. (2002).
6. Kandias, M., Galbogini, K., Mitrou, L., Gritzalis, D.: Insiders trapped in the mirror reveal themselves in social media. In: 7th International Conference on Network and System Security, pp. 220--235. Springer. (2013).
7. Kandias, M., Mitrou, L., Stavrou, V., Gritzalis, D.: Which side are you on? A new Panopticon vs. privacy. In: 10th International Conference on Security and Cryptography (SECRYPT-2013), Iceland, pp. 98--110. (2013).
8. Pang, B., Lee, L.: Opinion mining and sentiment analysis. In: Foundations and Trends in Information Retrieval, vol. 2, no. 1-2, pp. 1--135. (2008).
9. King, I., Li, J., Chan, K.: A brief survey of computational approaches in social computing. In: International Joint Conference on Neural Networks, pp. 1625--1632. (2009).
10. De Choudhury, M., Counts, S.: The nature of emotional expression in social media: measurement, inference and utility. In: 2012 Human Computer Interaction Consortium (HCIC) Workshop. (2012).
11. Benevenuto, F., Rodrigues, T., Cha, M., Almeida, V.: Characterizing user behaviour in online social networks. In: 9th ACM SIGCOMM Conference on Internet Measurement, pp. 49--62. ACM. (2009).
12. Park, N., Kee, K., Valenzuela, S.: Being immersed in social networking environment: Facebook groups, uses and gratifications, and social outcomes. In: CyberPsychology & Behavior, vol. 12, no. 6, pp. 729--733. (2009).
13. Balduzzi, M., Platzer, C., Holz, T., Kirda, E., Balzarotti, D., Kruegel, C.: Abusing social networks for automated user profiling. In: Recent Advances in Intrusion Detection, pp. 422--441. Springer. (2010).
14. Theoharidou, M., Gritzalis, D.: A Common Body of Knowledge for Information Security. In: IEEE Security & Privacy, vol. 4, no. 2, pp. 64--67. (2007).
15. Kandias, M., Virvilis, N., Gritzalis, D.: The Insider Threat in Cloud Computing. In: 6th International Conference on Critical Infrastructure Security, pp. 93--103. Springer (LNCS 6983). (2013).
16. Tsalis, N., Theoharidou, M., Gritzalis, D.: Return on security investment for Cloud platforms. In: Economics of Security in the Cloud Workshop, United Kingdom. IEEE Press. (2013).
17. Theoharidou, M., Papanikolaou, N., Pearson, S., Gritzalis, D.: Privacy risks, security and accountability in the Cloud. In: 5th IEEE Conference on Cloud Computing Technology and Science, pp.177--184, United Kingdom. IEEE Press. (2013).
18. Theoharidou, M., Tsalis, N., Gritzalis, D.: In Cloud we Trust: Risk-Assessment-as-a-Service. In: 7th IFIP International Conference on Trust Management, pp. 100--110, Spain. Springer (AICT 401). (2013).

19. Magklaras, G., Furnell, S.: Insider threat prediction tool: Evaluating the probability of IT misuse. In: Computers & Security, vol. 21, no. 1, pp. 62--73. (2001).
20. Magklaras, G., Furnell, S., Brooke, P.: Towards an insider threat prediction specification language. In: Information Management & Computer Security, vol. 14, no. 4, pp. 361--381. (2006).
21. Yaseen, Q., Panda, B.: Knowledge acquisition and insider threat prediction in relational database systems. In: International Conference on Computational Science and Engineering, pp. 450--455. IEEE. (2009).
22. Kandias, M., Mylonas, A., Virvilis, N., Theoharidou, M., Gritzalis, D.: An insider threat prediction model. In: 3rd Trust, Privacy and Security in Digital Business Conference, pp. 26--37. (2010).
23. Greitzer, F., Kangas, L., Noonan, C., Dalton, A., Hohimer, R.: Identifying At-Risk Employees: Modeling Psychosocial Precursors of Potential Insider Threats. In: 45th Hawaii International Conference on System Science, pp. 2392--2401. IEEE. (2012).
24. Brdiczka, O., Liu, J., Price, B., Shen, J., Patil, A., Chow, R., Ducheneaut, N.: Proactive Insider Threat Detection through Graph Learning and Psychological Context. In: IEEE Symposium on Security and Privacy Workshops, pp. 142--149. IEEE. (2012).
25. Shaw, E., Ruby, K., Post, J.: The insider threat to information systems: The psychology of the dangerous insider. In: Security Awareness Bulletin, vol. 2, no. 98, pp. 1--10. (1998).
26. Federal Bureau of Investigation: The Insider Threat: An introduction to detecting and deterring an insider spy. (2012). Available at http://www.fbi.gov/about-us/investigate/ counterintelligence/ the-insider-threat.
27. Kosinski, M., Stillwell, D., Graepel, T.: Private traits and attributes are predictable from digital records of human behavior. In: National Academy of Sciences, vol. 110, no. 15, pp. 5802--5805. (2013).
28. Kandias, M., Stavrou, V., Bozovic, N., Mitrou, L., Gritzalis, D.: Can we trust this user? Predicting insider's attitude via YouTube usage profiling. In: 10th IEEE International Conference on Autonomic and Trusted Computing, pp. 347--354, Italy. IEEE Press. (2013).
29. Kandias, M., Stavrou, V., Bosovic, N., Gritzalis, D.: Predicting the insider threat via social media: The YouTube case. In: 12th Workshop on Privacy in the Electronic Society, pp. 261--266, Germany. ACM Press. (2013).
30. Chen, Y., Nyemba, S., Zhang, W., Malin, B.: Leveraging social networks to detect anomalous insider actions in collaborative environments. In: IEEE International Conference on Intelligence and Security Informatics, pp. 119--124. IEEE. (2011).
31. Skues, J., Williams, B., Wise, L.: The effects of personality traits, self-esteem, loneliness, and narcissism on Facebook use among university students. In: Computers in Human Behavior. (2012).
32. Buffardi, L., Campbell, W.: Narcissism and social networking web sites. In: Personality and Social Psychology Bulletin, vol. 34, no. 10, pp. 1303--1314. (2008).
33. Mehdizadeh, S.: Self-presentation 2.0: Narcissism and self-esteem on Facebook. In: Cyberpsychology Behavior, and Social Networking, vol. 13, no. 4, pp. 357--364. (2010).
34. Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P., Bhattacharjee, B.: Measurement and analysis of online social networks. In: 7th ACM SIGCOMM Conference on Internet Measurement, pp. 29--42. ACM. (2007).
35. Travers, J., Milgram, S.: An experimental study of the small world problem. In: Sociometry, pp. 425--443. (1969).
36. Quercia, D., Ellis, J., Capra, L., Crowcroft, J.: In the mood for being influential on twitter. In: Privacy, Security, Risk and Trust. In IEEE 3rd International Conference on Social Computing, pp. 307--314. IEEE. (2011).

37. Frank, L., Hohimer, R.: Modeling human behavior to anticipate insider attacks. In: Journal of Strategic Security, vol. 4, no. 2, pp. 3. (2011).
38. Shaw, E., Fischer, L.: Ten tales of betrayal: The threat to corporate infrastructure by information technology insiders analysis and observations. Defense Personnel Security Research Center, USA. (2005).
39. Sebastiani, F.: Machine learning in automated text categorization. In: ACM Computing Surveys, vol. 34, no. 1, pp. 1--47. (2002).
40. Manning, C., Raghavan, P., Schütze, H.: Introduction to Information Retrieval. Cambridge University Press. (2008).
41. Hildebrandt, M.: Who is profiling who? Invisible visibility. In: Reinventing Data Protection, pp. 239--252. (2009).
42. Solove, D.: A taxonomy of privacy. In: University of Pennsylvania Law Review, vol. 154, no. 3, pp. 477. (2006).
43. Henderson, S.: Expectations of Privacy in Social Media. In: Mississippi College L. Rev., pp. 31 (Symposium Edition)). (2012). Available at: http://works.bepress.com/stephen_hender son/10
44. Ziegele, M., Quiring, O.: Privacy in social network sites. In: Privacy Online: Perspectives on privacy and self-disclosure in the Social Web, pp. 175--189. Springer. (2011).
45. Mitrou, L., Gritzalis, D., Katsikas, S., Quirchmayr, G.: Electronic voting: Constitutional and legal requirements, and their technical implications. In: Secure Electronic Voting, pp. 43--60. Springer. (2003).
46. Mitrou, L.: The Commodification of the Individual in the Internet Era: Informational self-determination or "self-alienation". In: 8th International Conference of Computer Ethics Philosophical Enquiry, pp. 466--485. (2009).
47. Lambrinoudakis, C., Gritzalis, D., Tsoumas V., Karyda, M., Ikonomopoulos, S.: Secure Electronic Voting: The current landscape. In: Secure Electronic Voting, pp. 101--122. Springer. (2003).
48. Schermer, B.: The limits of privacy in automated profiling and data mining. In: Computer Law and Security Review, vol. 27, pp. 45--52. (2011).
49. Mylonas, A., Kastania, A., Gritzalis, D.: Delegate the smartphone user? Security awareness in smartphone platforms. In: Computers & Security, vol. 34, pp. 47-66. (2013).
50. Mitrou, L.: The impact of communications data retention on fundamental rights and democracy: The case of the EU Data Retention Directive. In: Haggerty/Samatas, pp. 127--147. (2010).
51. Cohen, J.: Examined Lives: Informational Privacy and the subject as object. In: Stanford Law Review, vol. 52, pp. 1373--1438. (2000).
52. Cas, I.: Ubiquitous Computing, Privacy and Data Protection: Options and limitations to reconcile the unprecedented contradictions. In: Computers, Privacy and Data Protection: An Element of Choice, pp. 139--170. Springer. (2011).
53. Fazekas, C.: 1984 is Still Fiction: Electronic Monitoring in the Workplace and US Privacy Law. In: Duke Law & Technology Review, pp. 15--15. (2004).
54. Gutwirth, S., De Hert, P.: Regulating profiling in a democratic constitutional State. In: Profiling the European citizen cross-disciplinary perspectives, pp. 271--302. Springer. (2008).
55. Mitrou, L., Karyda, M.: Employees' privacy vs. employers' security: Can they be balanced? In: Telematics and Informatics, vol. 23, no. 3, pp. 164--178. (2006).
56. Broughton, A., Higgins, T., Hicks, B., Cox, A.: Workplaces and Social Networking - The Implications for Employment Relations. Institute for Employment Studies, Brighton. (2009).

57. Abril-Sánchez, P., Levin, A., Del Riego, A.: Blurred Boundaries: Social Media Privacy and the 21st Century Employee. In: American Business Law Journal, vol. 49, no. 1, pp. 63--124. (2012).
58. Castells, M.: Communication Power. Oxford University Press. (2009).
59. Dumortier, F.: Facebook and Risks of "De-contextualization" of Information. In: Data Protection in a Profiled World, pp. 119--137. (2010).
60. Nissenbaum, H.: Privacy as Contextual Integrity. In: Washington Law Review, vol. 79, pp. 119--157. (2004).
61. Davison, K., Maraist, C., Hamilton, R., Bing, M.: To Screen or Not to Screen? Using the Internet for Selection Decisions. In: Employ Response Rights, vol. 24, pp. 1--21. (2012).
62. Smith, W., Kidder, D.: You've been tagged! (Then again, maybe not): Employers and Facebook. In: Business Horizons, vol. 53, pp. 491--499. (2010).
63. Slovensky, R., Ross, W.: Should human resource managers use social media to screen job applicants? Managerial and legal issues in the USA, vol. 14, Is. 1, pp. 55--69. (2012).
64. Simitis, S.: Reconsidering the premises of labor law: Prolegomena to an EU regulation on the protection of employees' personal data. In: European Law Journal, vol. 5, pp. 45--62. (1999).
65. Lasprogata, G., King, N., Pillay, S.: Regulation of electronic employee monitoring: Identifying fundamental principles of employee privacy through a comparative study of data privacy legislation in the European Union, US and Canada. In: Stanford Technology Law Review 4 (http://stlr.stanford.edu/STLR/Article?04_STLR_4). (2004).
66. UK Information Commissioner: The Employment Practices Data Protection Code. (2003).
67. Data Protection Working Party. Opinion 8/2001 on the processing of personal data in the employment context (5062/01/Final). (2001).
68. Mitrou, L., Karyda, M.: Bridging the gap between employee's surveillance and privacy protection. In: Social and Human Elements of Information Security: Emerging Trends and Countermeasures, IGI Global, New York, pp. 283--300. (2009).
69. Mitrou, L.: The Commodification of the Individual in the Internet Era: Informational Self-determination or "Self-alienation"? In: 8th International Conference of Computer Ethics Philosophical Enquiry, Greece, pp. 466--485. (2009).
70. Mylonas, A., Meletiadis, V., Tsoumas, B., Mitrou, L., Gritzalis, D.: Smartphone Forensics: A Proactive Investigation Scheme for Evidence Acquisition. In: 27th International Information Security and Privacy Conference, pp. 249--260. Springer. (2012).
71. Mylonas, A., Meletiadis, V., Mitrou, L., Gritzalis, D.: Smartphone sensor data as digital evidence. In: Computers & Security, vol. 38, pp. 51--75. (2013).
72. Mylonas, A., Tsoumas, B., Dritsas, S., Gritzalis, D.: A secure smartphone applications roll-out scheme. In: 8th International Conference on Trust, Privacy & Security in Digital Business, pp. 49-61. Springer. (2011).
73. Virvilis, N., Dritsas, S., Gritzalis, D.: A cloud provider-agnostic secure storage protocol. In: 5th International Conference on Critical Information Infrastructure Security, pp. 104-115. Springer. (2010).