A REVIEW OF BEHAVIOURAL RESEARCH ON DATA SECURITY

Martin Maguire, Nathan Stuttard, Andrew Morris, Eleanor Harvey

Abstract

Protection of confidential information or data from being leaked to the public is a growing concern among organisations and individuals. This paper presents the results of the search for literature on behavioural and security aspects of data protection. The topics covered by this review include a summary of the changes brought about by the EU GDPR (General Data Protection Regulation). It covers human and behavioural aspects of data protection, security and data breach or loss (threats), IT architectures to protect data (prevention), managing data breaches (mitigation), risk assessment and data protection audits. A distinction is made between threats and prevention from within an organisation and from the outside.

1. Introduction

The General Data Protection Regulation (GDPR) harmonizes data protection laws in the EU that are fit for purpose in the digital age. By introducing a single law, the EU believes that it will bring better transparency to help support the rights of individuals and grow the digital economy.

The GDPR imposes new rules on companies, government agencies, non-profits, and other organizations that offer goods and services to people in the

EU, or that collect and analyse data tied to EU residents. Even organizations outside Europe need to be compliant, or otherwise face significant penalties.

The primary objective of the GDPR is to give citizens back control of their personal data. From an economic standpoint, the GDPR aims to simplify the regulatory environment for international business by unifying the regulation within the EU.

The GDPR reinforces the need for organisations to collect, manage and protect their data in responsible and secure manner. There are many aspects of this activity that directly relates to human behaviour, whether it is employees being aware of data security policy and putting it into practice and adopting procedures that protect the organisation from data loss through carelessness or from data breaches due to planned and targeted behaviour.

This paper reviews literature on the behavioural aspects of data security within an organisation and by its employees. It can form a basis for thinking about data security and developing procedures for keeping data secure.

2. Human Aspects of Data Protection

It has been shown that very often, people are the weak link in a data protection system. Not only are they prone to error and poor decisions, but their behavioural patterns can be tracked making them targets for attacks. They are unpredictable and unique which makes design solutions challenging. They vary widely in their knowledge and perceptions of security. Their behaviours contradict their attitudes. They are ill-informed of the dangers. They are apathetic because of the pace of change. They are deceived by services they trust. They are easily persuaded to share data for small gains like popularity. The privacy paradox, behavioural targeting, and motivations for sharing information are key topics.

2.1. Awareness and understanding of potential data loss

Information security awareness

In general, awareness of information security policies within organisations is questionable. For example, one study found that only 15% of their sample of university employees were aware of the contents of their information security policy, while only 25% of employees practiced appropriate information security behaviours during work (Moquin and Wakefield, 2016). Indeed, more recent evidence revealed that 40% of employees, when questioned about their company's security policy, stated that they knew nothing about the policy and recommended that information security awareness training be implemented.

Information security awareness training has been used to boost employee's engagement with information security policies. For example, Alkalbani et al (2015) found that awareness training significantly influenced compliance towards adopting information security compliance in organisations. It has also been demonstrated that when employees are made aware of a company's information security policy, they tend to be more competent with regards to carrying out cybersecurity behaviours, compared to employees who have not been made aware.

Aspects of previously described theory are also relevant with regards to raising awareness of information security policies. For example, as this lecture has already explored, threat appraisal is important with regards to compliance with information security practices. However, in the same vein, awareness of security threats drives this formation of beliefs. Therefore, both protection motivation and awareness are important for compliance with information security policies.

Information security awareness training can also alter specific computer related behaviours. Creating longer and stronger passwords is one outcome of information security awareness training. Other approaches include using phishing education programs. Phishing refers to an attempt to gain electronic access to sensitive information such as passwords by masquerading as a seemingly trustworthy entity, and usually occurs in the form of a malicious link within an email. If such a link is clicked on, it may allow logon details/ passwords to be extracted. Phishing education programs aim to spread awareness of this threat and provide subsequent training in how to deal with such problems. In one study, the number of times a phishing link was clicked on decreased as the study went on, suggesting that users' were learning about the potential threat (Jansson and Solms, 2013).

Social networks

The privacy risks of social networks have grown exponentially since inception. Even an account which does not list any information still reveals a social graph which can be analysed to infer personal information (Akcora *et al.*, 2012).

Risks ina number of areas of social media have been researched including: bullying (Akcora *et al.*, 2012), crime (fraud and burglary), stalking, regret (Ghiglieri *et al.*, 2014) helicopter parenting, micromanagement of staff (Cheung, 2014), vulnerable to malware and even cyber-warfare (Crossler and Bélanger, 2014), threats to organisations if employees post work-related information of social media (Molok *et al.*, 2010)

Health terms (i.e., Pregnancy, Depression, Breast Cancer), Job terms (i.e., Analyst, Senior Analyst in New York), Travel terms (i.e., traveling from Napoli Capodichino to New York (JFK) and travel dates) have also been considered as sensitive information (Malandrino *et al.*, 2013)

In the default settings, friending on Facebook affords that 'friend' full access to the user's profile, and 'friends of friends' almost as much – on average this amounts to 16,900 strangers. Would users comfortably share this much personal information with this many friends and strangers offline? Furthermore, the risk increases when considering the substantial number of connections through social media who have only met virtually, not physically. (Akcora *et al.*, 2012). Facebook is good at mitigating threats from outsiders but poor within a user's existing friend network (Spiliotopoulos and Oakley, 2013)

Unfortunately, studies have shown that users are not used to specifying privacy settings and very rarely change the permissive default settings (Akcora *et al.*, 2012) – they do not know who can access their content or how to change this (Bergström, 2015; Ghiglieri *et al.*, 2014). Matters are made worse by providers who constantly change privacy settings to make personal data more visible (Ghiglieri *et al.*, 2014) like Facebook Places which arguably never obtained valid consent to share these data and does not provide a way to disable it (Cheung, 2014).

'Even after numerous press reports and widespread disclosure of leakages on the Web and on popular Online Social Networks, many users appear not be fully aware of the fact that their information may be collected, aggregated and linked with ambient information for a variety of purposes'. Free OSN services are paid for by advertising, which in turn is paid for by users with increasing amounts of their personal data, allowing firms to target their campaign (Malandrino *et al.*, 2013). Most of the economic value in data is in purposes other than that which it was collected for (Bartolomeo *et al.*, 2013).

People may be aware of the privacy of content of messages or posts they send, but usually forget the metadata attached to these which could allow a third-party to track their location and social connections (Bartolomeo *et al.*, 2013)

Many users are unaware that photos taken on their smart phones are geotagged with a location which is then made available to third-parties when they upload these images to social media, making them vulnerable to attacks (Cheung, 2014)

"According to a 2012 nationwide survey of 2254 adults by the Pew Research Centre in the US, 54% of app users decided not to install cell phone app after they discovered how much personal data they would need to share to use it" (Boyles *et al.*, 2012).

Advertising companies claim to aggregate data, so individuals are not identifiable, but this was not found to be the case (Cheung, 2014). The GDPR recognises the need for anonymisation and also provides guidance on pseudonymisation where personally identifiable information fields within a data record are replaced by one or more artificial identifiers, or pseudonyms.

Prevention

Tools on social networking sites are emerging which support users' more privacy-awareness (such as controlling access based on relationships or setting privacy preferences). However, (Akcora *et al.*, 2012) argue these tools do not do enough to raise awareness of the potential risks that should form the basis of user decisions about disclosure.

Data Loss Prevention (DLP) is well-established in enterprise environments to prevent leaks due to human error, but in a personal context DLP is an emerging field. While enterprise DLP scans content, warning of and blocking unauthorised communication, in personal DLP the final choice lies with the user making it more important to make them aware of the risks (Ghiglieri *et al.*, 2014).

Hull (2015) argues putting self-management in such a prominent position in privacy law allows privacy as a commodity that can be bought and sold and actually facilitates its abuse and erosion.

Attitude to risk varies between individuals so tools have been developed based on users behaviours. (Akcora *et al.*, 2012) developed an algorithm which learns users' attitude to risk based on their responses to questions about sample profiles and sets their privacy preferences accordingly. (Ghiglieri *et al.*, 2014) developed a tool which supports self-assessment of Facebook posts based on learning from users' previous activity.

Bartolomeo *et al.* (2013) calls for flexible, timely and efficient ways for users to change their privacy preferences and stresses the link between social and technical aspects of data protection - the emphasis should be on how the data is used, not the data itself. In line with GDPR, he proposes a system of user-friendly 'licences' which allow users to subscribe to a 'privacy package' and re-evaluate it when it expires – simplifying the processes and giving users more options than a binary tick box.

(Terada *et al.*, 2016) analysed PC logs to identify behaviours and psychological characteristics of people most likely to experience a cyber-attack: For instance people who spent the shortest amount of time reading the terms and conditions have a strong benefit perception (bias) whereas those who avoid key presses when the PC freezes are risk averse. These data could be used to predict risk, target interventions, and select employees.

Challenges

Protection Motivation Theory is a model describing the psychology that motivates people to protect their data (Floyd, *et al.* 2000 cited in (Crossler and Bélanger, 2014). Similar to risk homeostasis (Wilde, 1982) when people feel vulnerable they take actions to reduce the threat.

One of the problems of alerting people about behaviours which could compromise security is how this is presented to the user without disruption or annoyance (Maurer *et al.*, 2011). (Stoll *et al.*, 2008) found many existing security alerts are designed for experts not users: They are text-based and threat-specific, and often mean the alerts are ignored, whereas a well-designed GUI would support non-experts in making better informed security decisions.

(Leon *et al.*, 2012) found serious usability flaws which prevent users opting out of behavioural advertising: Confusion between multiple tracking services, inappropriate default settings, communication problems, lack of feedback, reduced functionality, and confusing interfaces.

Cloud computing is quickly becoming the 5th utility service – after water, electricity, gas, and telephony – but 'the current state of the art is not specifically tuned to understand the behaviour of internal (employees) or external users (customers) of IaaS' (Infrastructure as a Service) (Khan *et al.*, 2012).

Many infrastructure providers have designed fully automated cloud systems, with no allowances for human behaviour, meaning once they are provisioned human errors cannot be controlled. This is a significant risk and one many people are unaware of: Google promises 0.01% for data outages but this does not account for data loss due to human error (Khan *et al.*, 2012). Human behaviour is unpredictable which poses a problem when creating scenarios to test user behaviours with traditional host-based intrusion detection systems (Khan *et al.*, 2012).

2.2. Human error

Human error can roughly be thought of as a failure to carry out a specific job or task, which results in the disruption of scheduled operations.

There is little data on human error of end-users, most statistics are about losses from organisations. The IT policy compliance group suggests that 75% of all data loss is due to user error (cited in Khan *et al.*, 2012). For incidents involving economic loss, 65% are caused by human error compared to 3% external threats. Human error cannot be controlled by technical solutions alone so individuals need to made sure their data is secure (Crossler and Bélanger, 2014)

The majority of data loss is due to human error. According to a survey conducted by IBM, 47% of all information leakages are due to internal fraud or malicious attacks [this includes being fooled by phishing as a type of human error], 25% are due to human error, and 29% are due to system vulnerabilities (Terada *et al.*, 2016).

Human error has different definitions. However, one way in which error has been defined was by James Reason (1990). Reason provides different definitions of error.

Reason defines error in relation to intentional actions.

– A slip occurs when an everyday, familiar task fails to go as the person intended. An example of this could be operating the wrong switch while operating a crane.

– Lapses cause a user to forget to carry out an action/lose track of a task, due to distraction or interruptions.

– A mistake occurs due to an error from intended actions not achieving their desired outcome, in other words, the goal or plan was wrong. Thus, a mistake occurs when a user does the wrong thing, believing it to be correct. A mistake involves a mismatch with regards to mental processing. An example of a mistake could be a pilot switching off the wrong engine in a plane, believing it to be the correct engine to switch off. In this situation, the plan is wrong.

Reason also describes two types of mistakes – rule-based and knowledge based.

A rule-based mistake occurs when a behaviour or action occurs, based on previously remembered rules or procedures. As an example of a rule-based mistake, a vessel could be constructed, with sails designed for rules and regulations concerning the wind speeds of the North Sea. However, that same vessel could then later be deployed in the arctic sea, where the wind speeds are significantly different. The subsequent capsizing of this vessel could be viewed as a rule-based mistake.

A knowledge-based mistake is when insufficient knowledge concerning a task or procedure results in a solution which is inappropriate for the problem. In the case of the vessel described previously, the captain may not have known how to properly evacuate the ship, resulting in the drowning of his crew.

Human error is an extremely important issue with regards to data and GDPR. At some estimates, 75% of all data loss is caused by human error. IBM estimates that around 46% of all information leakages are caused by internal fraud or malicious attack.

Social engineering (where someone is tricked into complying with malicious instructions) account for 97% of malware attacks. It is important to address the problem of human error, in order to mitigate the potential losses incurred by a data breach. This is especially important, given the increased powers for fines from the GDPR regulations.

Human error can occur in a variety of different ways. For instance, Evans et al (2019) analysed a number of information security incidents, using a newly developed tool, in order to explore the causes and circumstances of information security. The study found that 90% of the human errors identified were due to commission, or in other words, the person completed the required task, but did not complete it correctly.

One of the most broadly cited reasons for human error concerns the organisation itself. For instance, Linginlal et al (2009) analysed a number of interviews with IT privacy officers in healthcare organisations. The authors identify that the top three causes of data breaches relate to organisational factors – organisational limitations (workload, turnover), an inefficient business process (management) and poor monitoring (little or no enforcement or penalties for not following information security guidelines).

Research has also shown that there may be a technological slant to data related errors. For instance, security related technostress (which refers to stress caused by working with computer technology on a daily basis) has been shown to negatively affect compliance with data security policies, as well as impacting the errors made by information security staff. This has implications, certainly for the persons approach to human error.

Equally, aspects such as security culture and policy implementation, and communication have been shown to be important in explaining the cause of information security violations.

Workload is also relevant with regards to human error in information security. For instance, Albrechtsen et al (2007) explored the reasons for information security related errors, with a sample of information security personnel. This study found that workers who had a high information security workload experienced conflict with regards to conforming to information and data security guidelines. These findings have been replicated across other information security related studies (Kraemer *et al.*, 2006; Kraemer and Carayon, 2007).

In terms of privacy related data breaches, research has shown that a number of different factors are relevant with regards to causation. For instance, in cases where computer equipment and sensitive data has been stolen, the cause attributed to this is a lack of under emphasis/ understanding of data protection policies (Linginal *et al.,* 2009). Other causes identified are inappropriate skill with regards to using computer related equipment, as well as insufficient monitoring by IT management. Social engineering and employee manipulation are other cited causes of privacy breach incidents.

A prominent theory concerning human error was proposed by James Reason, in 1990. This model breaks down the problem of human error into two main approaches – the Person Approach and the system approach. The Persons Approach focuses on the individual person within the scenario – to what extent did their actions, e.g. forgetfulness, inattention or moral weakness, contribute to the incident in question.

In contrast, the Systems approach concentrates on the conditions under which individuals work and attempts to create defences to avert errors or reduce their effects. The systems approach places less emphasis on the individual in question; instead the institution or organisation as a whole are highlighted as being a cause of the error.

Another prominent model within the study of human error is the Swiss Cheese Model, also developed by Reason. This model posits that in any system, there are defences, barriers and safeguards put in place. In the model, each slice represents such a defence – for instance using physical defences such as locked doors or alarms, as well as human factors such as password usage and backups. The model describes how an incident occurs when all the holes simultaneously 'line up', meaning that at each individual level of the system, there has been a lapse or error. The lapses at each stage of the model can be attributed to one or both of the following factors – active failures and latent conditions.

Active Failures: These are unsafe acts, which are generally committed by people who are in direct contact with the system in question

Latent conditions: This describes conditions which may remain inactive until the actual event occurs. Latent conditions may translate into workplace conditions which invite error and may leave weakness within the system.

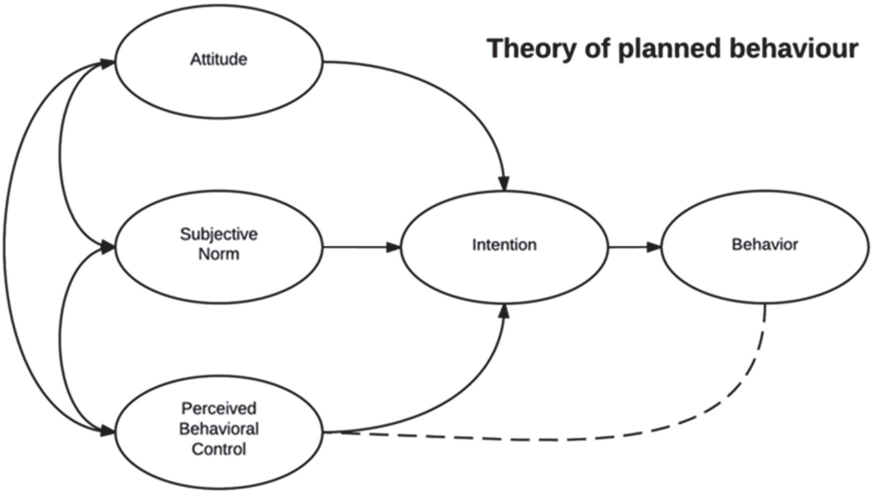
As an example, in data management, a person may not use a password correctly (an active failure), which may be brought about by the conditions of their workplace, i.e. the information security policy may be poorly enforced.

In terms of reducing human error in data and information handling, there is no clear cut solution, or solutions. There are a number of ways in which human error might be prevented or reduced. One way is through promoting the information security policy of the company. Research has shown that employees who read and understand the information security policy have an enhanced and better understanding of it. This has implications for management, as human related errors and risk may be mitigated, if such education and awareness is promoted. Another way to reduce the errors in information security is through organization-focused strategies include implementing stringent administrative measures, reengineering workflows, and creating better work environments aimed at facilitating both preventive and corrective cognitive interventions. A paper by Sommestad et al (2014) concisely summarises different organisational interventions and measures which may be useful for ensuring compliance with information security policies. For instance, the types of training are important, as are perceived benefits and involvement. The remaining focus of this lecture will be on compliance and motivation, in order to explore how these Human Factors relate to data and information security.

2.3. Theory of Planned Behaviour

The theory of planned behaviour was originally developed through the work of Fishbein and Ajzen (1975), where it was originally referred to as the Theory of Reasoned action. The main assumption of this theory is that a person's reasoned decision to engage in a particular behaviour is predicted by their attitude, subjective norms and perceived behavioural controls.

Figure 1. – Representation of the Theory of planned behaviour



Within the model, intention represent a person's motivation, that is, their plan to exert effort to exercise a decision. 'Subjective norms' refer to a person's belief about how others would view their behaviour, should they engage in it.

'Perceived behavioural control' refers to an individual's perception of the ease or difficulty of performing a specific behaviour.

'Attitude' refers to how the person views the behaviour, or in other words, the more favourable an attitude towards a behaviour is, the more likely they are to engage in that specific behaviour.

The theory of planned behaviour also makes reference to social norms. Social norms are essentially the unwritten laws which dictate how a person should behave.

The theory of Planned Behaviour can explain compliance with information security policies in the following way: If an employee perceives that they possess sufficient capacity to carry out a security related task (perceived behavioural control), has a positive attitude towards complying with the policy (attitude) and also observes other people within the organisation behaving in the same way (perceived subjective norms), they are likely to comply with these information security policies.

Research has demonstrated that several of the tenets of the Theory of Planned behaviour are relevant for improving employee compliance towards the information security policy of an organisation. For example:

The theory of planned behaviour has indicated that a number of specific information security behaviours can be enhanced. These include making stronger passwords and the intention to comply with information security regulations. Additionally, other information security specific behaviours have been shown to be linked to the theory of planned behaviour. For example, Glaspie and Karwowski (2017) found that users who perceived others to be compliant with information security policies, as well as to be control themselves, were more likely to comply with certain behaviours. For example, users did not click on links sent to them by strangers, updating their malware settings more often and were more likely to pay attention to security training.

Furthermore, Bauer and Bernroider (2017) found that improvements in attitude, personal and social norms are related to an increase in intention to comply with information security behaviour. Similarly, Belanger et al (2017) found that attitude plays a critical role in affecting the intention and eventually the enactment of information security conformance behaviour, specifically the usage of stronger passwords.

Foth (2016) examined the role of attitude, subjective norms and the perceived behavioural controls on employee’s intention to conform with data protection regulations. The results showed that these variables are important for ensuring compliance with data protection regulations, with subjective norms being especially relevant and important.

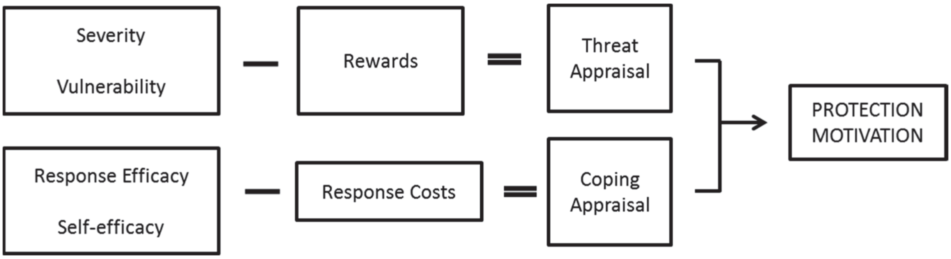
Similarly, workgroup norms have also been shown to be important with regards to users engaging in non-malicious security violations (Guo *et al.,* 2011). This study found that if users perceive that they are following the similar behaviours to their peers, they are more likely to engage in that behaviour, whether it is correct or incorrect.

The findings surrounding social norms and subjective norms are particularly important, as they suggest that it is the internal culture of the organisation which is important with regards to compliance with information security regulations. Thus, in any kind of intervention, the people themselves, the ‘human factor’ should be of paramount importance. An intervention should also target the group, not simply individuals, in order to ensure that all the group are familiar with the regulations. As GDPR brings new updates to policy, this is particularly important.

2.4. Protection motivation theory

Protection Motivation Theory was originally created by R.W Rogers in 1975 and was further developed in 1983. The basic idea of protection motivation theory is that an individual's motivation to protect themselves against threats is based on two components, with two parts in each component. These two components are Threat appraisal and coping appraisal.

Figure 2. – Representation of the Protection Motivation Theory



Threat Appraisal consists of:

[1] The perceived severity of an event. The judgement an individual will have about how severe a potential event will be for them.

[2] The perceived probability of an event occurring (this is sometimes referred to as vulnerability).

Coping Appraisal consists of:

[1] Response Efficacy. Response efficacy refers to the ability to produce a desired result or the belief about the perceived benefits of an action.

[2] *Perceived Self Efficacy*. The level of confidence an individual has in their ability to undertake a specific behaviour to prevent an event occurring.

[3] *Response Costs.* The response costs are the costs associated with the carrying out the recommended behavior (time/monetary).

Protection motivation refers to the result of the threat appraisal and coping appraisal. In other words, perceptions of threat and coping appraisals explain engagement in specific behaviour. In the context of information security compliance, research has shown that facets of protection motivation theory are relevant for explaining compliance.

The research findings surrounding Protection Motivation Theory and information security compliance are mixed. For example, a recent meta-analysis of information security behaviour with respect to protection motivation theory suggests that the coping appraisal variables of response efficacy and self-efficacy were the most important with regards to explaining information security behaviour (Mou et a, 2017). Therefore, from a Human Factors perspective, an individual may be more likely to comply with information security guidelines based on their own volition.

The findings of this meta-analysis are supported by other studies which have investigated Protection motivation theory and compliance behaviour. For instance, Burns et al (2017) found that coping appraisal was much more influential in terms of the development of protection motivation, compared to threat appraisal. Research has also highlighted that response efficacy is particularly relevant for explaining compliance to information security guidelines. This suggests that if an employee believes that their input to a particular system is beneficial for the system as a whole, they are more likely to conform to that specific behaviour type.

However, threat appraisal has also been shown to influence information security compliance. Research has demonstrated that threat appraisal is important with regards to password strength; this suggests that making users aware of the potential dangers of misusing data and information may be beneficial with regards to enhancing security related behaviours.

In another study, threat appraisal was investigated within a cloud-based backup scenario. The purpose of this study was to examine whether the perceived threat of data loss affected the intention to regularly back up data to the cloud services. It was found that this perceived threat affected the user's intention to store data securely – those who felt that they were more susceptible to a threat were more likely to back up their files regularly, compared to those who did not feel the same way. These findings been supported by other research studies (Sommestad *et al.,* 2015).

GDPR brings with it a new system of fines and powers of regulation. Thus, making users and employees aware of this and the potential threat which it poses should be of importance to employers.

2.5. Self-efficacy

Self-efficacy refers to an individual's own belief in their capabilities to successfully engage in a specific task. In terms of information security, if a person has a high level of self-efficacy, they will be more confident in their own ability to carry out a specific information security related task, for instance, using stronger passwords. In the information security literature, self-efficacy is sometimes referred to as Self-Efficacy in Information Security (SEIS).

Research has shown that employees who identify as having high levels of Self-efficacy also score higher on measures of information security awareness.

For instance, a study by Safa et al (2016) found that information security awareness was strongly predicted by a number of variables, including self-efficacy, perceived behavioural control and subjective norms. Specifically, the results suggested that individuals with high levels of Self-efficacy believed that they possessed the expertise and skills to protect private data, as well as self-belief that they could prevent information security violations. These research findings have been replicated elsewhere (Son, 2011; Vance, 2012). Users who possess high levels of Self-efficacy have also been shown to be more compliant with information security policies.

Self-efficacy is also related to the adoption of specific information security behaviours. For example, users who possess high Self-efficacy use more security software and features designed to protect information and data from cyber-attacks. Users with high self-efficacy in Information security also backed up their data more, use stronger passwords and checked the website encryption before entering it. Other information security behaviours have also been shown to be positively impacted by possessing self-efficacy. For example, adherence to smartphone security policy in the workplace has been shown to be predicted by self-efficacy, as well as compliance with bring your own device policies.

2.6. Deterrence theory and Sanctions

General deterrence theory proposes that an individual will be disinclined to behave in a certain way, if they are faced with a potential punishment for their behaviour. General Deterrence Theory has three main principles:

[1] Certainty. The punishment must be certain.

[2] Celerity. The punishment given must be given swiftly.

[3] Severity. The punishment must be severe enough to ensure that the person judges that the costs of punishment outweigh the gains.

Within the context of information security research, general deterrence theory is often referred to alongside perceived sanctions, which refers to a perceived punishment for a behaviour.

General deterrence theory has been shown to be influential with regards to changing individual’s information security behaviour. For example, research has shown that when punishment for a specific information security deviant behaviour is both certain and severe, employees are less likely to engage in that specific behaviour. Thus, this suggests that if employees feel like they are more likely to be caught and punished for a specific behaviour, they may be less inclined to commit such an act. These findings are supported by research which has found that perceived sanctions can also influence an employees’ intention to follow information security policies concerning USB usage. In addition, linked to the previous research findings concerning protection motivation theory, threat appraisal stemming from perceived legal (fines or imprisonment) and organisational (demotion, penalties) sanctions can impact on compliance attitude. In particular, this research showed that organisational sanctions are particularly relevant for compliance. This suggests that the notion that a person may be punished for an action is important for shaping future compliance attitudes.

Overall, research has generally shown that sanction effects impact the cognitive attitude of users to develop either a positive or negative viewpoint with regards to information security.

Awareness of information security policies is crucial towards influencing compliance with such policies. This is important as GDPR provides a number of new stipulations with regards to how data is handled. For instance, GDPR now defines personal data differently – it now includes online identifiers such as IP addresses and mobile ID devices. GDPR also defines genetic and biometric data under the term "sensitive data".

Information security policies should aim to raise awareness of these changes, to ensure that compliance occurs.

These policies should raise general and threat awareness:

[1] In order to fulfil the requirements of 'privacy by design', awareness training should aim to incorporate the relevant aspects of the new GDPR regulations. For instance, it should make it clear that personal data now included IP addresses, or that biometric data is now sensitive data. Raising the awareness of this should, according to the literature findings, help boost compliance with such policies.

[2] Organisations should also make it clear to employees and users that there is a real threat to their data. Where this threat originates from should be illustrated using two perspectives. First, the potential threat from cyber-attacks such as phishing and social engineering should be made clear – this should be backed up by the appropriate statistics. Second, the stipulations of GDPR should also be included here. The sanctions brought in by the GDPR are much higher than the original Data Protection Act (1998) – the GDPR states that fines of up to 4% of annual turnover may be administered, or up to 20 million euros.

In this section, it should be emphasised that fines have drastically increased. For instance, TalkTalks fine of £400,000 in 2016 would have cost them £59 million under the new GDPR regulations.

Aside from the potential threat stemming from monetary fines, the Information Commissioner's office has a range of powers, including the ability to issue warnings and reprimands, imposing temporary bans on data processing or suspending data transfers. Again, the literature supports the notion that sanction based awareness is a key component of compliance and so this should be a key approach with regards to implementing the data protection reform act.

2.7. Trust and Acceptance: How people feel about their personal data and its uses

Technology, psychology, economics, management, and law all play a role in our view of privacy (Pelteret and Ophoff, 2016)

Attitudes and behaviours are incongruent and sometimes contradictory (Bergström, 2015). Very high levels of mistrust can damage customer relationships as consumers feel pressured into sharing data reluctantly and they may disengage from e-commerce.

The Privacy Paradox – people care about their privacy but in practice are willing to give away personal information for small rewards (e.g. likes and followers on social media). In a key paper, behavioural economics have unveiled the role of cognitive biases and heuristics in decision making. The type of data and the type of perceived threat also affect people's privacy behaviour, but more research is needed to fully understand it. Kokolakis (2017) cites the problem of illiteracy and lack of awareness of the risks. The relationship is not simple – people differentiate between the type of data and the type of risk. (Taddicken, 2014)

Hull, (2015) argues people do care about their privacy but services have vested interests in making it difficult for them to protect themselves from invasion.

Bergström, (2015) includes data from a survey looking at age, level of education, internet usage, trust in others and some political questions, and how concerned they were about using different types of internet usage (social media, credit card, searching). How trusting was the only significant factor in people's privacy preferences, although slight increases in concern were seen with age and leftward political views. Users are more bothered by the privacy of services that handle personal data (social media) than generic data (search results).

Similarly (Taddicken, 2014) found "willingness to disclose information" [which could be synonym for trust] as the personality trait that most influenced people's willingness to self-disclose. Introverts share less than extraverts.

People are more likely to adopt online banking services if they have a good relationship with the offline/physical bank (Chiou and Shen, 2012). Tweens (young people between 10 and 12) both care about and take measures to protect their privacy online. Yet, these measures are sometimes haphazardly employed and some youth display uncertainty or ignorance about how to protect their privacy. This is a trend also seen in teens and adults. The focus of education

("Stranger danger") is too narrow, it doesn't address their fears or equip them with the knowledge they need today (Davis and James, 2013). Preibusch (2013) provides a review of methods used to measure privacy concern.

Among over 65s "Higher ICT use was associated with self-perceived socio-personal characteristics such as being ‘‘satisfied with activities'', ‘‘persevering'', ‘‘physically and emotionally independent'' and having a ‘‘positive outlook''. Whereas, the majority of non-users reported that their activities did not change across time and that they felt ‘‘intimidated'' and ‘‘anxious'' with technology. The performance of ICT-based activities and/or the desire to perform them were significantly associated with the perceived importance of the activities. The older population's age, education, attitudes, and personalities influence how they approach ICT" (Vroman *et al.*, 2015).

2.8. Security Behaviours: How people treat their personal data and the problems that could arise

‘50% of mobile users regularly provide consent without effectively reading agreements, because of lack of time and excessive language complexity. And even if modern operating systems notify users about the permissions each new installed application needs, users are only provided with very limited options – either give consent or not use the service at all' (Bartolomeo *et al.*, 2013).

"In our online experiment, we found that of our 995 participants, many reported finding their social security numbers (20%), credit/debit card numbers (16 and 17%, respectively), bank account numbers (26%), birth dates (46%), email passwords (30%), and/or home addresses (76%) stored in their email accounts" (Egelman *et al.*, 2014)

A survey found password issues are the second most likely human error to impact an information system. Requirements are put in place to increase strength but then overload human memory capabilities: Thus human factors need to be considered. Participants with more than 8 passwords to remember forgot their passwords every 2 weeks. However, these trends are complicated by other factors – unemployed people, for example, were found to have more passwords to remember than those in work were are less likely to forget them. Women are twice as likely to write down their passwords as men (Carstens *et al.*, 2004).

A qualitative study which interviewed people about their motivations to lock or not lock their phones found around a third of users do not lock their phones. Many believe there is no sensitive data stored on them. Only 61% of these were concerned about identity theft online, but less than 25% connected this with offline identity theft, financial loss, or loss of sensitive information even though all participants accessed email on their phone. Those who did not lock their phones said they did so for convenience, so that they could be reunited with a lost phone, or someone could make a call from their phone in an emergency (Egelman *et al.*, 2014) – also offers design suggestions.

Mobile devices also allow people to access sensitive data (their own or work-related) in any physical location where they could be overlooked; (Tarasewich *et al.*, 2006) proposes designing 'blinders' to obscure sensitive parts of the screen but other solutions could be considered.

Using social media for sharing work-related information (Molok *et al.*, 2010) and using personal phones for business use (Goode, 2010) create new expose points were organisations are vulnerable to external threats. These need to be addressed through company policy and training.

Location tracking embedded in mobile devises has created a state of constant 'uberveillance'. Location-based services (LBS) in wireless mobile devices allow real-time tracking which is widely used by consumers and the data now reveals a person's direction of travel, trajectory, or even his or her predicted movements. Lawsuits have been filed against Apple and Microsoft for tracking users despite their request not to be and even when location software was turned off. Geo-location services are considered sensitive data under GDPR (Cheung, 2014).

2.9. Personality theory

Personality is loosely defined as a set of characteristics which determine how an individual thinks, feels and behaves.

The big five factor model was originally developed by Tupes and Christal in 1961, though the theory built heavily on work carried out by Cattell (1943). The model was then validated by Costa and McCrae in 1986, who provided evidence as to the validity of the measures. The findings of Costa and McCrae have been validated across a number of different societies and cultures, suggesting that it is an accepted and valid measure of personality.

The big five model of personality consists of five main traits. These are:

[1] Extraversion. This refers to how outgoing and social a person is. The polar opposite of extraversion is introversion, which refers to a shy and reserve individual.

[2] Emotional Stability/ Neuroticism. This refers to a person's ability to remain stable and balanced. Neuroticism refers to insecurity, anxiousness and hostility. The terms 'emotional stability' and 'neuroticism' are often used inter-

changeably, as emotional stability is the inverse of neuroticism.

[3] Agreeableness. This is a person's tendency to be compassionate and co-operative toward others and obedient to regulations.

[4] Conscientiousness. This is a person's tendency to act in an organized or thoughtful way.

[5] Openness to experience. This is the extent to which a person is open to experiencing a variety of activities.

The five-factor model has been studied extensively with regards to information security compliance. Generally, this is because it is the most widely accepted measure of personality. In the next section of this lecture, each personality dimension of the big five will be discussed in detail with regards to what research has shown in terms of data and information security.

Conscientiousness

This refers to a person's tendency to act in an organized or thoughtful way. A person who is conscientious will feel that they have a sense of duty towards others and this will be reflected in their actions.

Conscientiousness impacts on how a person handles data and information within an organisation. One way which this has been demonstrated is through examining the typical characteristics of top performing security personnel compared to non-security personnel. In doing so, light has been shed on the key differences in terms of personality between these two groups. Top performing security personnel not only score higher on the workplace values of theoretical (discovering the truth through systematic thinking and reasoning) and economic (an interest in usefulness and practicality), but also score higher on the personality trait of openness. These findings suggest that IT personnel who can perform a job well also tend to behave in an organised and focused way. Other research into personality and information security awareness has found that conscientious and agreeable individuals tend to have higher information security awareness, compared to other personnel.

Not only are workplace values correlated with conscientiousness, but it has been shown that individuals with higher levels of conscientiousness generate stronger passwords and also do not reuse/recycle passwords between accounts. Furthermore, the relationship between intent and actual use of security programs for data and information is moderated by conscientiousness and agreeableness (Shropshire, Warkentin and Sharma, 2015).

Research has also shown that the security knowledge of management within information security can be influenced by personality. Specifically, conscientiousness is significantly related to information security executive's attitude towards the management of information security. 2.7 Implications for GDPR

Extraversion

Extraversion generally refers to how outgoing and social a person is. To be clear, the opposite of extraversion is introversion, which refers to being reserved and shy.

Extraversion has been linked to a number of information security behaviours. For instance, extraversion has been correlated with the likelihood of securing a device which contains sensitive data (Gratien et al 2018). This suggests that those with outgoing personalities are more likely to secure their device. Similarly, employees who create stronger passwords also tend to score higher on extraversion measures. The same behaviour also correlated with higher scores on neuroticism, which is the opposite trait of emotional stability.

As GDPR introduces legislative changes, this will require the information security policy of the company to be updated. Part of this process may involve providing feedback to employees. How well this feedback is received may depend upon the personality of the individual. Research has shown that IT personnel who score highly on measures of extraversion are more responsive to feedback concerning their behaviour, compared to individuals who score highly on personality traits such as neuroticism. Similarly, individuals who are classed as extraverted have also been shown to be better at detecting phishing emails, which is surprising given that extraverted users may be more likely to be trusting, rather than less so.

Agreeableness and Emotional Stability

Agreeableness refers to being compassionate and cooperative toward others and obedient to regulations. In the context of information security, research has shown that agreeableness has a positive influence on information personnel’s concerns for information security and that agreeable individuals have stronger information sensitivity. Furthermore, for these individuals, agreeableness alongside extraversion was influential in forming data privacy concerns.

A further personality trait is Emotional Stability. Users who are emotionally stable will be generally be able to remain stable and balanced, especially under pressure. Halevi et al (2016) examined the cyber security habits of IT professionals, in relation to cultural, personality and demographic variables. The study found that neuroticism (the opposite of emotional stability) was inversely correlated towards self-efficacy beliefs in ability to use security software correctly. McCormac et al (2017) found that emotional stability significantly explained the variance in information security awareness.

Psychological capital, locus of control and impulsivity

Psychological Capital refers to a construct which is made up of work-related tenets of positive psychology: Hope, optimism, resilience and Self efficacy. Research has shown that psycap relates positively with the positive coping mechanism of security response efficacy

Another personality dimension to be discussed is Locus of control. This is defined as an individual's belief system regarding the causes of his or her experiences and the factors to which that person attributes success or failure. Research has shown that locus of control is related to information security. For instance, Hadlington et al (2018) found that individuals who scored higher on the work locus of control scale, and thus exhibited a greater degree of externality. Workman et al (2008) found that coping with information security threats was dependent on whether people perceived that the threat was preventable in the first place (LOC).

A further personality dimension is impulsivity. Impulsivity refers to a tendency to act spontaneously without reflecting on an action and its consequences. Research has examined awareness and engagement in good cybersecurity behaviours, including using different passwords for device securement and found that impulsivity was negatively correlated with good information security behaviours. Therefore, this trait may help to predict if a person will engage in risky security behaviour. Similar findings have also been found with regards to detecting phishing emails.

Clearly, personality affects how an individual behaves around data. As the new GDPR regulations have brought in new rules concerning both how data must be stored as well as new definitions of data, one consideration could be the screening of job candidates. The previous slides have illustrated that certain personality characteristics are associated with positive information security behaviours, i.e. password usage/backups. Therefore, in light of the new regulations, it makes sense that potential hires should be compliant with information security policies – and this could easily be gauged through using personality type testing. This could come in form of identifying potential members of staff which might be more at risk of dangerous data behaviours. This could be done by subtle personality screening, perhaps using a condensed big five scale.

3. Security and data loss (threats)

3.1. Within an organisation

Colwill (2009) examines some of the key issues relating to insider threats to information security and the nature of loyalty and betrayal in the context of organisational, cultural factors and changing economic and social factors. It is recognised that insiders pose security risks due to their legitimate access to facilities and information, knowledge of the organisation and the location of valuable assets.

Insiders will know how to achieve the greatest impact whilst leaving little evidence. The paper describes a practitioner's view of the issue and the approaches used by BT to assess and address insider threats and risks. Proactive measures need to be taken to mitigate against insider attacks rather than reactive measures after the event. A key priority is to include a focus on insiders within security risk assessments and compliance regimes. This requires a focus on human factors, education and awareness and greater attention on the security 'aftercare' of employees and third parties.

Khan and Oriol (2012) state that as cloud computing is transparent to both the programmers and the users, it induces challenges that were not present in previous forms of distributed computing. Furthermore, cloud computing enables its users to abstract away from low-level configuration such as configuring IP addresses and routers. This is also true for security services, for instance automating security policies and access control in cloud, so that individuals or end-users using the cloud only perform very high-level (business oriented) configuration.

Koops and Leenes (2014) also promote the privacy by design, employing the principle that principle that privacy should be promoted as a default setting of every new ICT system and should be built into systems from the design stage. The paper discusses what the proposed legal obligation for 'privacy by design' implies in practice for online businesses. In terms of the regulatory tool-box, privacy by design should be approached less from a 'code' perspective, but rather from the perspective of 'communication' strategies.

Nurse et al (2014) emphasise that threat that insiders pose to businesses, institutions and governmental organisations continues to be of serious concern. They propose a novel conceptualisation grounded in insider threat case studies, existing literature and relevant psychological theory. It acts as a platform for general understanding of the threat, and also for reflection, modelling past attacks and looking for useful patterns that can lead to addressing the threats.

Stamati-Koromina et al (2012) investigate the different types of insider threats and their implications to the corporate environment, with specific emphasis to the special case of data leakage. They propose a design for a forensic readiness model, which is able to identify, prevent and log email messages, which attempt to leak information from an organisation with the aid of steganography.

Urquhart and McAuley (2018) state that security incidents such as targeted distributed denial of service (DDoS) attacks on power grids and hacking of factory industrial control systems (ICS) are on the increase. They argue the industrial IoT brings four security challenges, namely: appreciating the shift from offline to online infrastructure; managing temporal dimensions of security; addressing the implementation gap for best practice; and engaging with infrastructural complexity. Their goal is to surface risks and foster dialogue to avoid the emergence of an Internet of Insecure Industrial Things.

Borgesius (2016) states that many behavioural targeting companies say that, as long as they do not tie names to data they hold about individuals, they do not process any personal data, and that, therefore, data protection law does not apply to them. However European Data Protection Authorities take the view that a company processes personal data if it uses data to single out a person, even if it cannot tie a name to these data. This paper argues that data protection law should indeed apply to behavioural targeting and so appropriate measures need to be taken to address this issue.

3.2. Outside an organisation

Alneyadi et al (2014) have carried out a survey on data leakage prevention systems. Traditionally, confidentiality of data has been preserved using security procedures such as information security policies along with conventional security mechanisms such as firewalls, virtual private networks and intrusion detection systems. Unfortunately, these mechanisms lack pro-activeness and dedication towards protecting confidential data, and in most cases, they require predefined rules by which protection actions are taken. This can result in serious consequences, as confidential data can appear indifferent forms indifferent leaking channels. Therefore, there has been an urge to mitigate these drawbacks using more efficient mechanisms. Recently, data leakage preventions systems (DLPSs) have been introduced as dedicated mechanisms to detect and prevent the leak age of confidential data in use, in transit and at rest. DLPs use different techniques to analyse the content and the context of confidential data to detect or prevent the leakage. Although DLPs are increasingly being designed and developed as standalone products by IT security vendors and researchers, the term is still ambiguous. In their paper, the authors have carried out a comprehensive survey on the current DLPs mechanisms. They explicitly define DLPs and categorise active research directions in this field. In addition, they suggest future directions towards developing more consistent DLPs that can overcome some of the weaknesses of the current ones.

Al-Sayid and Aldlaeen (2013) conduct a survey of database security threats. They highlight threat types and their impacts on sensitive data, and present different security models. The assumption underlying this study is that by understanding the weaknesses and the threats facing databases, database administrators can then begin to create a security plan to better protect their databases.

Gupta and Sharman (2012) provide a categorization of data breaches using empirical investigation. The article categorizes the incident and loss trends into different dimensions including the industry, victim type, data type, and threats such as stolen computer, hacking and unauthorized access. The research can aid individuals and organizations understand the data breach, trends and evaluate their own risks in handling personal information, which will help them to make better and informed decisions to protect against data breaches.

Lechler et al (2011) identify and evaluate the threat of transitive information leakage in healthcare systems. Using a case study of a major hospital in the NY/NJ metropolitan area they demonstrate the complexity of the healthcare system and its inherent information security risks. The study shows that only a systemic perspective allows identifying all potential risks and providing solutions for improved information security. The main conclusion of this study is that transitive information risks have major implications for healthcare organizations and regulators. Identifying these risks will significantly improve information security in the healthcare environment.

Sabillon et al (2016) acknowledge that cybercrime is growing rapidly around the world, as new technologies, applications and networks emerge. In addition, the Deep Web has contributed to the growth of illegal activities in cyberspace. As a result, cybercriminals are taking advantage of system vulnerabilities for their own benefit. Their article presents the history and conceptualization of cybercrime, explores different categorizations of cybercriminals and cyberattacks, and sets forth a cyberattack typology, or taxonomy. Common categories include where the computer is the target to commit the crime, where the computer is used as a tool to perpetrate the felony, or where a digital device is an incidental condition to the execution of a crime. They conclude their study by analysing lessons learned and future actions that can be undertaken to tackle cybercrime and harden cybersecurity at all levels.

4. IT architectures to protect data (prevention)

4.1. Within an organisation - Prevention

Software assurance methodologies aim to assist developers in improving the security of software production. The Security Development Lifecycle (SDL) is one such example developed by Al-Fedaghi and Alkandari (2011). Their method uses a flow-based methodology and focuses on the system requirements relating to threats, risks

Work by Bolognini and Bistolfi (2017) discusses the move towards the collection of personal data as part of "Big Data analysis". They state that the opportunities arising from the analysis of such information need to be balanced with the risks for the data protection of individuals. While anonymization is a traditional technique for protecting personal data, they argue that pseudonymisation can be used both to reduce the risks of re-identification and help data controllers and processors to maintain their personal data protection obligations by keeping control over their activities.

Cavoukian's (2011) 7 principles for 'Privacy by Design' developed in the 1990's advances the view that the future of privacy cannot be assured solely by compliance with regulatory frameworks; rather, privacy assurance must ideally become an organization's default mode of operation. Ensuring privacy and gaining personal control over one's information and, for organisations, gaining a sustainable competitive advantage can be achieved by following the 7 principles that she describes.

Chow *et al.* (2009) considers the problem of how sensitive content is revised in order for it to be distributed. In order to do this redaction, is often used, the "blacking out" of sensitive words and phrases. Redaction has the side effect of reducing the utility of the content, often so much that the content is no longer useful. Consequently, government agencies and others are increasingly exploring the revision of sensitive content as an alternative to redaction that preserves more content utility. This approach is called sanitisation or pseudonymisation. Chow et al's have developed a Web tool to automatically identify sensitive words and phrases to help quickly evaluate revision for sensitivity. The authors warn against "slippery slope" where the usefulness and apparent authoritativeness of the tool may lead users to neglect their own judgment in favour of the tool's.

Colesky et al's (2016) paper looks at design strategies to achieve 'Privacy by Design' to meet the requirements of the GDPR. Their paper improves the strategy definitions and suggests an additional level of abstraction between strategies and privacy patterns called 'tactics'. They identify a collection of such tactics based on an extensive literature review, in particular a catalogue of surveyed privacy patterns. They explore the relationships between the concepts that they introduce, and similar concepts used in software engineering. This paper aims to helps bridge the gap between data protection requirements set out in law and help system development practice.

Heurix et al (2015) consider privacy-enhancing technologies (PETs) belong to a class of technical measures which aim at preserving the privacy of individuals or groups of individuals. They describe a universal taxonomy of PETs and a tool for the systematic comparison of them. To demonstrate its applicability, the author's proposed taxonomy is applied to a set of key technologies covering different disciplines such as data anonymization, privacy-preserving data querying, communication protection, and identity hiding.

Hildebrand (2014) discusses the new smart environments that anticipate our future behaviours and adapt their own behaviours to accommodate our inferred preferences. They provide us with a ubiquitous artificial intelligence that uproots the common sense of our Enlightenment heritage that matter is passive and mind active. She argues the new regulation could be a game changer. It establishes a new incentive structure and is based on a salient understanding of law's need for effective [not theoretical] technology neutrality.

Distributed usage control is concerned with how data may or may not be used in distributed system environments after initial access has been granted. If data flows through a distributed system, there exist multiple copies of the data on different client machines. Usage constraints then have to be enforced for all these clients. Kelbert and Pretschner (2013) develop a distributed usage control enforcement infrastructure that generically and application-independently extends the scope of usage control enforcement to any system receiving usage-controlled data.

Khan *et al.* (2012) discuss the security risks and their management in cloud computing. Cloud computing provides outsourcing of resources bringing economic benefits. The outsourcing however does not allow data owners to outsource the responsibility of confidentiality, integrity and access control, as it still is the responsibility of the data owner. The paper presents a risk inventory which documents the security threats identified in terms of availability, integrity and confidentiality for cloud infrastructures in detail for future security risks. They also propose a methodology for performing security risk assessment for cloud computing architectures presenting some of the initial results.

From an information privacy perspective, King and Raja (2012) analyse how well the regulatory frameworks in place in Europe and the United States help protect the privacy and security of sensitive consumer data in the cloud. It makes suggestions for regulatory reform to protect sensitive information in cloud computing environments and to remove regulatory constraints that limit the growth of this vibrant new industry.

Two exploratory studies were conducted by Kraemer and Carayon (2005) to identify the various dimensions of CIS (Computer and Information Security) culture. One study included an industry workgroup consisting of six CIS managers and specialists. The second study consisted of individual interviews with eight CIS managers and managers and eight network administrators. The workgroup and CIS managers and network administrators provided a preliminary list of elements in CIS culture dimensions, including: employee participation, training, hiring practices, reward system, management commitment, and communication and feedback.

Personas represent a powerful tool for designing systems to meet user needs. In a paper by Lewis and Coles Kemp (2014) comic strips are used as an approach to align personas and narrative scenarios; the resulting visual artefact was tested with information security practitioners, who often struggle with wider engagement. It offers ways in which different professional roles can work together to share understanding of complex topics such as information security. It also offers user-centred design practitioners a way to reflect on, and participate with, user research data.

Luger et al (2015) recognises the heightened role of designers in the regulation of ambient interactive technologies. They developed and tested a series of data protection ideation cards with teams of designers to help make emerging European data protection regulations more accessible to the design community.

Lukosch et al (2015) report on different scenarios from the security domain in which augmented reality (AR) techniques are used to support information exchange. A combination of quantitative and qualitative evaluation showed that AR can improve the distributed situational awareness of a team.

McCauley-Bell and Crumpton (1998) consider the human factors issues in information security. They provide an overview of possible risks that users can pose to information systems and of the information security issues impacted by human interaction that may or may not play a role in promoting system security.

Othmane and Leszek (2009) offer a solution to protecting data privacy called Active Bundles. The technique protects sensitive data from their disclosure to unauthorised parties and from unauthorised dissemination (even if started by an authorized party). The protocol uses buddies to provide anonymity to senders and receivers. Evaluation of the solution indicates that: (i) the percentage of sensitive data that reaches unauthorized hosts during dissemination can be high, (ii) the apoptosis mechanism protects sensitive data from dissemination to unauthorized hosts and (iii) the Active Bundles solution provides a level of anonymity to hosts while it does not decrease significantly the throughput of buddies.

Romanou (2018) examines the extent to which Privacy by Design can safeguard privacy and personal data within a rapidly evolving society. This paper explains the theoretical concept and the general principles of Privacy by Design, as laid down in the General Data Protection Regulation. Then, by indicating specific examples of the implementation of the Privacy by Design approach, it demonstrates why the implementation of Privacy by Design is a necessity in a number of sectors where specific data protection concerns arise (biometrics, e-health and video-surveillance) and how it can be implemented.

Tahboub and Saleh (2014) state that traditional security approaches such as firewalls can’t protect data from leakage. Data leakage/loss prevention ( ) systems are solutions that protect sensitive data from being in non-trusted hands. Their paper surveys DLP systems and compares them with other security and data protection approaches.

Inadvertent data disclosure by insiders is considered as one of the biggest threats for corporate information security Wűchner and Pretschner (2012). Data loss prevention systems typically try to cope with that problem by monitoring access to confidential data and preventing their leakage or improper handling. However, they are limited when enforcing more complex security policies that for instance specify temporal or cardinal constraints on the execution of events. Their paper presents UC4Win, a data loss prevention solution based on the concept of data-driven usage control to allow such a fine-grained policy-based protection. UC4Win is capable of detecting and controlling data-loss related events at the level of individual function calls and can track the flows of confidential data through the system.

Halboob et al (2015) examined computer forensics and privacy protection fields and observed that there are two conflicting directions in computer data security. In the other words, computer forensics tools try to discover and extract digital evidences related to a specific crime, while privacy protection techniques aim at protecting the data owner's privacy. As a result, finding a balance between these two fields is a serious challenge. Existing privacy-preserving computer forensics solutions consider all data owner's data as private and, as a result, they collect and encrypt the entire data. However, this increases the investigation cost in terms of time and resources. Therefore, there is a need for having privacy levels for computer forensics so that only relevant data are collected data and then only private relevant data are encrypted. Halboob et al propose privacy levels for computer forensics starting with classifying forensic data, and then analysing all data access possibilities within computer forensics. They also define several privacy levels based on access possibilities which lead to more efficient privacy-preserving computer forensics solutions. They conclude that development of a privacy-preserving computer forensics framework based on these privacy levels is required and that further work is required for defining levels and policies for network forensics and supporting fully automatic data selection for selecting the private and/or relevant data.

Version control is also a technique that can be used to avert risks. Version control is the process by which different drafts and versions of a document or record are managed. It is a "tool" which tracks a series of draft documents, culminating in a final version and it provides an audit trail for the revision and update of these finalised versions. Version control is important for data documents that undergo a lot of revision and redrafting and is particularly important for electronic documents because they can easily be changed by several different users. These changes may not be immediately apparent. Knowing which version of a document is being examined is important, for example, in order to find out which version of a policy is currently in force, or which version of a policy was in use at a time. There are several software tools that can be used to manage version control safely and effectively. One of the main issues regarding version control is the requirements to destroy previous versions of documents once the final version has been approved (although retaining previous versions is sometimes useful in documenting the thinking that led to the final version.) In some cases, Data Protection and Freedom of Information Acts determine that the information held, including drafts of previous versions may be subject to disclosure.

Li et al (2017) argue that having multiple copies of data has high reliability but also has the disadvantage of high redundancy storage and low space utilization. They demonstrate a data backup method based on XOR checksum being suitable for storing hot temporary data, which first splits the data into two parts and then performs the XOR operation of the two parts to generate another part of the data. Finally, the XOR checksum stores the three data parts into different nodes. The checksum not only ensures the security of data but also saves the storage space, thus improving the performance of reading and writing. This strategy achieves a mutual backup between the three nodes in order to ensure data security. Because there is only one copy of original data in the system, this model resolves the data inconsistency problem reasonably and simplifies the data version control existing in the redundancy backup model.

Mao et al (2104) state that cloud storage is an ideal way of data storage and has been widely used. However, it must be ensured that data is not lost or damaged. Most cloud storage systems use replica redundancy technique to solve the problem of the loss of data. Commonly used strategies include: consistency strategy based on replica chain, consistency strategy based on multi-version control, and consistency strategy based on timestamp. They discuss their research into consistency strategies as a focus for cloud storage data security.

Conway et al (2107) also consider the data security aspects of cloud computing. They state that there is still a lot to learn about the adoption and use of cloud, including issues such as security, data protection, interoperability, service maturity, and return on investment. They describe an assessment model developed by the Innovation Value Institute (IVI) using a multi method, two-phased approach. The first phase involved a review of the current academic and practitioner literature in the area of cloud. The second phase employed the principles of design science and open innovation to pilot, test, validate, and refine the cloud adoption assessment in collaboration with industry-based practitioners. By using the assessment model, the level of maturity will identify areas of strength and weakness within the organization and serve as the basis for an improvement roadmap, to ensure the successful adoption and on-going management of cloud.

Scharnick et al (2016) discuss that investing in data security that is intangible is often not seen as priority expenditure as it brings no Return on Investment nor contributes to expanding the business. However, the newly enacted Protection of Personal Information (POPI) Act in South Africa, requires businesses to re-evaluate their stance on information security and data storage protection as POPI requires that ‘appropriate and reasonable security measures’ be put in place to effectively protect all personal information that large organisations as well as smaller businesses process and more importantly store. However, the lack of comprehensive controls found within any one information security approach (information security standard, best practice or framework) to fully address the requirements of the POPI act, leaves businesses exposed to legislative action under POPI. Their paper analyses widely implemented information security approaches in the context of POPI compliance. An evaluation of the comprehensiveness of these approaches and their proposed mechanisms for protecting data within businesses is conducted.

4.2. Outsiders – prevention

Al-Fedaghi (2011) presents a theoretical framework for DLP in terms of a flow-based conceptual model. It proposes that DLP can be oriented toward the entire information lifecycle and incorporate diverse processes in the course of business by using this framework as a foundation for specification and design of DLP. To demonstrate feasibility of the approach, the model is applied to the Records and Information Management system of the U.S. Internal Revenue Service.

According to Blasco et al (2012) insider threats are an increasing concern for most modern organizations. Data Leakage Protection (DLP) systems have been developed to tackle this issue and work by tracking sensitive information flows and monitoring executed applications to ensure that sensitive information is not leaving the organisation. However, current DLP systems do not fully consider that trusted applications represent a threat to sensitive information confidentiality. In their paper, they demonstrate how to use common trusted applications to evade current DLP systems. They analyse the proposed evasion technique from the malicious insider point of view and discuss some possible countermeasures to mitigate its use to steal information.

Borders and Prakash (2009) present an approach for quantifying information leak capacity in network traffic. Instead of trying to detect the presence of sensitive data—an impossible task in the general case—their goal is to measure and constrain its maximum volume. They take advantage of the insight that most network traffic is repeated or determined by external information, such as protocol specifications or messages sent by a server. By filtering this data, we can isolate and quantify true information flowing from a computer. When applied to real web browsing traffic, the algorithms were able to discount 98.5% of measured bytes and effectively isolate information leaks.

Costante et al (2016) state that to confront the problem of data loss (i.e. the unauthorized/unwanted disclosure of data, Data Loss Protection (DLP) solutions either employ patterns of known attacks (signature-based) or try to find deviations from normal behaviour (anomaly-based). While signature-based solutions provide accurate identification of known attacks and can prevention them, they cannot cope with unknown attacks, nor with attackers who follow unusual paths (like those known only to insiders) to carry out their attack. They offer a DLP protection framework uses an anomaly-based engine that automatically learns a model of normal user behaviour, allowing it to flag when insiders carry out anomalous transactions.

Gessiou et al (2011) state that the traditional approach for detecting information leaks is to generate fingerprints of sensitive data, by partitioning and hashing it, and then comparing these fingerprints against outgoing documents. Unfortunately, this approach incurs a high computation cost as every part of document needs to be checked. As a result, it is not applicable to systems with a large number of documents that need to be protected. Additionally, the approach is prone to false positives if the fingerprints are common phrases. They propose an improvement for this approach to offer a much faster processing time with less false positives. The core idea of their solution is to eliminate common phrases and non-sensitive phrases from the fingerprinting process.

Jung et al (2008) describe the design and implementation of Privacy Oracle, a system that reports on application leaks of user information via the network traffic that they send. Privacy Oracle treats each application as a black box, without access to either its internal structure or communication protocols. This means that it can be used over a broad range of applications and information leaks (i.e., not only Web traffic content or credit card numbers). To accomplish this, they develop a differential testing technique in which perturbations in the application inputs are mapped to perturbations in the application outputs to discover likely leaks; they leverage alignment algorithms from computational biology to find high quality mappings between different byte-sequences efficiently.

Kim and Kim (2010) address company monitoring of their employee’s behaviour using a DLP (Data Loss Prevention) solution to protect their information assets from internal attackers. During the monitoring process, it is inevitable that private information is disclosed to recognize the violation of internal regulations for handling the company’s critical information. In their paper, the authors suggest a data loss prevention method considering the privacy violation level. They consider a method of quantifying the degree of privacy violation based on the data units which are exposed when the employee’s data handling is monitored. At the same time, they suggest a method of quantifying the degree of importance of data units which are monitored.

Lui and Khun (2010) provide an introductory paper on data loss prevention. It covers the data loss problem, the need to address the problem (driven by government and industry requirements), the data loss prevention approach covering loss modes, solution capabilities and best practices.

Oetzel and Spiekermann (2014) propose a methodology to systematically consider privacy issues by using a step-by-step privacy impact assessment (PIA). Existing PIA approaches cannot be applied easily because they are improperly structured or imprecise and lengthy. They argue that companies that employ their PIA method can achieve ‘privacy-by-design’, which is widely heralded by data protection authorities. The contribution of the artefacts they created is twofold: First, they provide a formal problem representation structure for the analysis of privacy requirements. Second, we reduce the complexity of the privacy regulation landscape for practitioners who need to make privacy management decisions for their IT applications.

5. Managing data breaches (mitigation)

Detection of data breaches is a critical area. (Shu *et al.,* 2016) utilise sequence alignment techniques for detecting complex data-leak patterns. Their algorithm is designed for detecting long and inexact sensitive data patterns. This detection is paired with a comparable sampling algorithm, which allows one to compare the similarity of two separately sampled sequences. The system achieves good detection accuracy in recognizing transformed leaks. In implementing the system, the authors demonstrate the high multithreading scalability of their data leak detection method required by a large organization.

Solutions exist for detecting inadvertent sensitive data leaks caused by human errors and to provide alerts for organizations. A common approach is to screen content in storage and transmission for exposed sensitive information. Such an approach usually requires the detection operation to be conducted in secrecy. However, this requirement is challenging to satisfy in practice, as detection servers may be compromised or outsourced. Shu et al (2015) have devised a privacy-preserving data-leak detection (DLD) solution to solve the issue where a special set of sensitive data digests is used in detection. The advantage of their method is that it enables the data owner to safely delegate the detection operation to a semi-honest provider without revealing the sensitive data to the provider. They describe how Internet service providers can offer their customers DLD as an add-on service with strong privacy guarantees. Evaluation has shown that the method can support accurate detection with very small number of false alarms under various data-leak scenarios.

Data backup operation is an essential part of common IT system administration. Since the backup and restore operations accrue downtime overhead or performance degradation, they have to be designed to ensure the data reliability while minimising the performance and availability overhead. This is especially important if data needs to be recovered if it is compromised by a virus, malware or hardware failure. Yin et al (2012) study the impacts of different backup policies on availability measures such as storage availability, system availability, and user-perceived availability. The author's studies show the effectiveness of the combination of full back-up and partial back-up in terms of user-perceived data availability and data loss rate. Sensitivity ranking can also help improve the availability measures.

"The three categories that are used to classify in-formation security risks are confidentiality, integrity, and accessibility or availability of information (U.S. Department of Homeland Security, 2002).

- A security breach in confidentiality is defined as sources not intended to have knowledge of the information have been provided with this knowledge. An example of this category would be sending sensitive data to the wrong person.
- A security breach in integrity is an incident where there is an unauthorized or incorrect change made to an information source. An example of this category is a financial accounting error causing the information in the database to be inaccurate.
- A security breach in accessibility occurs when either access for those entitled to a system is denied or access is given to those who are not authorized to access the system. An example of this category would be an authorized user of a system who is unable to access a system due to forgetting their password." Cited in (Carstens *et al.*, 2004)

6. Risk assessment and data protection audit

An important internal audit function is evaluating the effectiveness and efficiency of an organization's control processes. These control processes include the policies, procedures and activities in place within an organization for managing risk and achieving organizational objectives.

In the standard, developed by the International Organization for Standardization (ISO), "control processes" are defined as an organization's management system. To conform to the ISO standard, organizations are required to establish and maintain management system processes. Organizations are also required to establish internal audit programs. The guidelines for understanding these internal audit requirements are set out in ISO 19011, Guidelines for Auditing Management Systems (ISO 19011, 2011).

ISO 19011:2011 provides guidance on auditing management systems, including the principles of auditing, managing an audit programme and conducting management system audits, as well as guidance on the evaluation of competence of individuals involved in the audit process, including the person managing the audit programme, auditors and audit teams. The standard is applicable to all organizations that need to conduct internal or external audits of management systems or manage an audit programme. The application of the standard to other types of audits is possible, provided that special consideration is given to the specific competence needed.

Acknowledgement

The work undertaken for this paper was conducted as part of the EU funded TAtoDPR project - Training activities for the implementation of the Data Protection Reform.

List of acronyms

DLP	Data leakage prevention/Data loss prevention
DPO	Data Protection Officer
EDPA	European Data Protection Act
EDPS	European Data Protection Supervisor
GDPR	EU General Data Protection Regulation
PIA	Privacy Impact Assessment
POPI	Protection of Personal Information Act 4, 2013
SA	Supervisory Authority

References

Akcora, C., Carminati, B., Ferrari, E., 2012. Privacy in Social Networks: How Risky is Your Social Graph?, in: 2012 IEEE 28th International Conference on Data Engineering. IEEE, 9-19. doi: 10.1109/ICDE.2012.99.

Albrechtsen, E. 2007. A qualitative study of users ‘ view on information security, 26(167), 276-289. https://doi.org/10.1016/j.cose.2006.11.004.

Al-Fedaghi, S., 2011. A Conceptual Foundation for Data Loss Prevention. Int. J.

Digit. Content Technol. its Appl. 5, 293-303. doi: 10.4156/jdcta.vol5.issue3.29.

Al-Fedaghi, S., Alkandari, A., 2011. On Security Development Lifecycle: Conceptual Description of Vulnerabilities, Risks, and Threats. Int. J. Digit. Content Technol. its Appl. 5, 296-306. doi: 10.4156/jdcta.vol5.issue5.32.

Alkalbani, A., Deng, H., & Kam, B. 2015. Investigating the Role of Socio-organizational Factors in the Information Security Compliance in Organizations. Australasian Conference on Information Systems, (2010).

Alneyadi, S., Sithirasenan, E., Muthukkumarasamy, V., 2016. A survey on data leakage prevention systems. J. Netw. Comput. Appl. 62, 137-152.

Al-Sayid, N.A., Aldlaeen, D., 2013. Database security threats: A survey study, in: 2013 5th International Conference on Computer Science and Information Technology. IEEE, 60-64. doi: 10.1109/CSIT.2013.6588759.

Barnard-Wills, D., Pauner Chulvi, C., De Hert, P., 2016. Data protection authority perspectives on the impact of data protection reform on cooperation in the EU. Computer Law and Security Review, 32, 587-598.

Bartolomeo, G., Frisiello, A., Petersen, F., 2013. Redesigning personal data protection: User requirements and guidelines for an interoperable solution. 65-73.

Bauer, S., & Bernroider, E. W. N. 2017. From Information Security Awareness to Reasoned Compliant Action. ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 48(3), 44-68. https://doi.org/10.1145/3130515.3130519.

Beckett, P., 2017. GDPR compliance: your tech department’s next big opportunity. Comput. Fraud Secur. 2017, 9-13. doi: 10.1016/S1361-3723(17)30041-6.

Bélanger, F., Collignon, S., Enget, K., & Negangard, E., 2017. Determinants of early conformance with information security policies. Information and Management, 54(7), 887-901. https://doi.org/10.1016/j.im.2017.01.003.

Berendt, B., Engel, T., Ikonomou, D., Le Métayer, D., Schiffner, S. (Eds.), (2014) Privacy Technology and Policy. Springer Verlag, Luxembourg.

Bergström, A., 2015. Online privacy concerns: A broad approach to understanding the concerns of different groups for different uses. Comput. Human Behav. 53, 419-426. doi: 10.1016/j.chb.2015.07.025.

Blasco, J., Hernandez-Castro, J.C., Tapiador, J.E., Ribagorda, A., 2012. Bypassing information leakage protection with trusted applications. Comput. Secur. 31, 557-568.

Bolognini, L., Bistolfi, C., 2017. Pseudonymization and impacts of Big (personal/anonymous) Data processing in the transition from the Directive 95/46/EC to the new EU General Data Protection Regulation. Comput. Law Secur. Rev. 33, 171-181.

Borders, K., Prakash, A., 2009. Quantifying information leaks in outbound web traffic, in: 2009 30th IEEE Symposium on Security and Privacy. IEEE, 129-140. doi: 10.1109/SP.2009.9.

Boyles, J.L., Smith, A. and Madden, M., 2012. Privacy and Data Management on Mobile Devices Pew Internet Reports, 5 September 2012, http://pewinternet.org/

Reports/2012/Mobile-Privacy.aspx accessed 27January 2013. Computer Law and Security Review, 30 (2014) 41e5446).

Burns, A.J., Posey, C., Roberts, T.L., Benjamin Lowry, P., 2017. Examining the relationship of organizational insiders' psychological capital with information security threat and coping appraisals. Computers in Human Behavior, 68, 190-209. https://doi.org/10.1016/j.chb.2016.11.018.

Carolan, E., 2016. The continuing problems with online consent under the EU's emerging data protection principles. Computer Law and Security Review, 32, 462-473. doi: 10.1016/j.clsr.2016.02.004.

Carstens, D.S., McCauley-bell, P.R., Malone, L.C., Demara, R.F., 2004. Evaluation of the Human Impact of Password Authentication Practices on Information Security. Informing Sci. Int. J. an Emerg. Transdiscipl. 7, 67-85.

Cattell, R.B., 1943. The description of personality: Basic traits resolved into clusters. The journal of abnormal and social psychology, 38(4), 476.

Cavoukian, A., 2011. The 7 Foundational Principles. Ontario.

Champion, M., Jariwala, S., Ward, P., Cooke, N.J., 2014. Using cognitive task analysis to investigate the contribution of informal education to developing cyber security expertise, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Human Factors an Ergonomics Society Inc., 310-314. doi: 10.1177/1541931214581064.

Cheung, A.S.Y., 2014. Location privacy: The challenges of mobile service devices. Comput. Law Secur. Rev. 30, 41-54.

Chiou, J.-S.S., Shen, C.-C.C., 2012. The antecedents of online financial service adoption: the impact of physical banking services on Internet banking acceptance. Behav. Inf. Technol. 31, 859-871. doi: 10.1080/0144929X.2010.549509.

Chow Parc, R., Oberst, I., Staddon Parc, J., Chow, R., Oberst, I., Staddon, J., 2009. Sanitization's Slippery Slope: The Design and Study of a Text Revision Assistant, in: Proceedings of the 5th Symposium on Usable Privacy and Security - SOUPS '09. ACM Press, New York, New York, USA, p. 1. doi: 10.1145/1572532.1572550.

Colesky, M., Hoepman, J.H., Hillen, C., 2016. A Critical Analysis of Privacy Design Strategies, in: 2016 IEEE Security and Privacy Workshops (SPW). Institute of Electrical and Electronics Engineers Inc., 33-40. doi: 10.1109/SPW.2016.23.

Colwill, C., 2009. Human factors in information security: The insider threat – Who can you trust these days? Inf. Secur. Tech. Rep. 14, 186-196. doi: 10.1016/J.ISTR.2010.04.004.

Conway, G., Doherty, E., Carcary, M., 2017. Enterprise cloud adoption-cloud maturity assessment model. Proceedings of the 11th European Conference on Information Systems Management, ECISM 2017, 56-63.

Costante, E., Fauri, D., Etalle, S., Hartog, J. Den, Zannone, N., 2016. A Hybrid Framework for Data Loss Prevention and Detection, in: 2016 IEEE Security and

Privacy Workshops (SPW). Institute of Electrical and Electronics Engineers Inc., 324-333. doi: 10.1109/SPW.2016.24.

Crossler, R., Bélanger, F., 2014. An extended perspective on individual security behaviors: Protection motivation theory and a unified security practices (USP) instrument. ACM SIGMIS Database 45, 51-71. doi: 10.1145/2691517. 2691521.

Cumbley, R., Church, P., 2013. Is "Big Data" creepy?, Computer Law and Security Review. Elsevier Advanced Technology.

Davis, K., James, C., 2013. Tweens' conceptions of privacy online: implications for educators. Learn. Media Technol. 38, 4-25. doi: 10.1080/17439884. 2012.658404

de Hert, P., Papakonstantinou, V., 2016. The new General Data Protection Regulation: Still a sound system for the protection of individuals? Computer Law and Security Review, 32, 179-194.

De Hert, P., Papakonstantinou, V., Malgieri, G., Beslay, L., Sanchez, I., 2017. The right to data portability in the GDPR: Towards user-centric interoperability of digital services. Computer Law and Security Review, doi: 10.1016/j.clsr. 2017.10.003.

Egelman, S., Jain, S., Portnoff, R.S., Liao, K., Consolvo, S., Wagner, D., 2014. Are you ready to lock? Understanding user motivations for smartphone locking behaviors, in: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security - CCS '14. Association for Computing Machinery, New York, New York, USA, 750-761. doi: 10.1145/2660267.2660273.

Fishbein M., Ajzen I., 1975. Belief, attitude, intention, and behavior: an introduction to theory and research. Reading, MA: Addison- Wesley.

Foth, M., 2016. Factors influencing the intention to comply with data protection regulations in hospitals: Based on gender differences in behaviour and deterrence. European Journal of Information Systems, 25(2), 91-109. https://doi.org/10.1057/ejis.2015.9.

Gellert, R., 2018. Understanding the notion of risk in the General Data Protection Regulation, Computer Law and Security Review.

Gessiou, E., Vu, Q.H., Ioannidis, S., 2011. IRILD: An Information Retrieval Based Method for Information Leak Detection. 2011 Seventh Eur. Conf. Comput. Netw. Def. 33-40. doi: 10.1109/EC2ND.2011.21.

Ghiglieri, M., Stopczynski, M., Waidner, M., 2014. Personal DLP for facebook, in: 2014 IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS). IEEE Computer Society, 629-634. doi: 10.1109/PerComW.2014.6815279.

Glaspie, H.W., & Karwowski, W. 2017, July. Human factors in information security culture: A literature review. In International Conference on Applied Human Factors and Ergonomics, 269-280. Springer, Cham.

Gonçalves, M.E., 2017. The EU data protection reform and the challenges of big data: remaining uncertainties and ways forward. Inf. Commun. Technol. Law 26, 90-115. doi: 10.1080/13600834.2017.1295838.

Goode, A., 2010. Managing mobile security: How are we doing? Network Security 2010, 12-15. doi: 10.1016/S1353-4858(10)70025-8.

Gratian, M., Bandi, S., Cukier, M., Dykstra, J., Ginther, A., 2018. Correlating human traits and cyber security behavior intentions. computers & security, 73, 345-358.

Gray, A., 2013. Conflict of laws and the cloud. Computer Law Security Rev. 29, 58-65.

Guo, K.H., Yuan, Y., Archer, N.P., & Connelly, C.E., 2011. Understanding nonmalicious security violations in the workplace: A composite behavior model. Journal of management information systems, 28(2), 203-236.

Gupta, M., Sharman, R., 2012. Determinants of Data Breaches: A Categorization-Based Empirical Investigation. J. Appl. Secur. Res. 7, 375-395. doi: 10.1080/19361610.2012.686098.

Hadlington, L., Chivers, S., 2018. Segmentation Analysis of Susceptibility to Cybercrime: Exploring Individual Differences in Information Security Awareness and Personality Factors. Policing: A Journal of Policy and Practice, 1-14. https://doi.org/10.1093/police/pay027.

Halboob, W., Mahmod, R., Abulaish, M., Abbas, H., Saleem. K., 2015. Data warehousing based computer forensics investigation framework. 12th International Conference on Information Technology - New Generations, IEEE Computer Society, 163-168. doi 10.1109/ITNG.2015.31.

Heurix, J., Zimmermann, P., Neubauer, T., Fenz, S., 2015. A taxonomy for privacy enhancing technologies. Comput. Secur. 53, 1-17.

Hildebrandt, M., 2014. The Public(s) Onlife A Call For Legal Protection by Design, in: Floridi, L. (Ed.), The OnLife Manifesto: Being Human in a Hyperconnected Era. SpringerOpen, Oxford, 181-193. doi: 10.1007/978-3-319-04093-6_19.

Huang, D.L., Rau, P.L.P., Salvendy, G., Shang, X.L., Liu, Y., Wang, X., 2008. Perception of information security and its implications for mobile phone. 1650-1654.

Hudic, A., Islam, S., Kieseberg, P., Rennert, S., Weippl, E.R., 2013. Data confidentiality using fragmentation in cloud computing. Int. J. Pervasive Comput. Commun. 9, 37-51. doi: 10.1108/17427371311315743.

Hull, G., 2015. Successful failure: what Foucault can teach us about privacy self-management in a world of Facebook and big data. Ethics Inf. Technol. 17, 89-101. doi: 10.1007/s10676-015-9363-z.

ISO 19011:2011. Guidance on auditing management systems, International Organisation for Standards, Geneva, Switzerland : ISO.

Jansson, K., & Von Solms, R. 2013. Phishing for phishing awareness. Behaviour and Information Technology, 32(6), 584-593. https://doi.org/10.1080/0144929X.2011.632650.

Jianjiang Li, J., Zhang, P., Li, Y., Chena, W., Liua, Y.,Wang, L., A data-check based distributed storage model for storing hot temporary data, Future Generation Computer Systems 73 (2017) 13-21, doi.org/10.1016/j.future.2017.03.019.

Jørgensen, R.F., Desai, T., 2017. Right to Privacy Meets Online Platforms: Exploring Privacy Complaints against Facebook and Google. Nord. J. Hum. Rights 35, 106-126. doi: 10.1080/18918131.2017.1314110.

Jung, J., Sheth, A., Greenstein, B., Wetherall, D., Maganis, G., Kohno, T., 2008. Privacy Oracle: a System for Finding Application Leaks with Black Box Differential Testing Jaeyeon, Proceedings of the 15th ACM conference on Computer and communications security - CCS ‘08. ACM Press, New York, New York, USA. doi: 10.1145/1455770.1455806.

Kelbert, F., Pretschner, A., 2013. Data usage control enforcement in distributed systems, in: Proceedings of the Third ACM Conference on Data and Application Security and Privacy - CODASPY ‘13. ACM Press, New York, New York, USA, 71-82. doi: 10.1145/2435349.2435358.

Khan, A.U., Oriol, M., Kiran, M., Jiang, M., Djemame, K., 2012. Security risks and their management in cloud computing, in: 4th IEEE International Conference on Cloud Computing Technology and Science Proceedings. IEEE, 121-128. doi: 10.1109/CloudCom.2012.6427574.

Kim, J., Kim, H.J., 2010. Design of internal information leakage detection system considering the privacy violation, in: 2010 International Conference on Information and Communication Technology Convergence (ICTC). IEEE, 480-481. doi: 10.1109/ICTC.2010.5674800.

King, N.J., Raja, V.T.T., 2012. Protecting the privacy and security of sensitive customer data in the cloud. Comput. Law Secur. Rev. 28, 308-319.

Kokolakis, S., 2017. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. Comput. Secur. 64, 122-134. doi: 10.1016/j.cose.2015.07.002.

Koops, B.J., Leenes, R., 2014. Privacy regulation cannot be hardcoded. A critical comment on the “privacy by design” provision in data-protection law. Int. Rev. Law, Comput. Technol. 28, 159-171. doi: 10.1080/13600869.2013.801589.

Kraemer, S., & Carayon, P. 2007. Human errors and violations in computer and information security: The viewpoint of network administrators and security specialists. Applied Ergonomics, 38(2), 143-154. https://doi.org/10.1016/j.apergo.2006.03.010.

Kraemer, S., Carayon, P., 2005. Computer and information security culture: Findings from two studies. 1483-1487.

Kraemer, S., Carayon, P., Clem, J.F., 2006. Characterizing violations in computer and information security systems. Proceedings of the 16th …. Retrieved from http://cqpi.engr.wisc.edu/system/files/IEA.pdf.

Lachaud, E., 2016. Why the certification process defined in the General Data Protection Regulation cannot be successful. Computer Law and Security Review, 32, 814-826. doi: 10.1016/j.clsr.2016.07.001.

Larson, R.G., 2013. Forgetting the first amendment: How obscurity-based privacy and a right to be forgotten are incompatible with free speech. Commun. Law Policy, 18, 91-120. doi: 10.1080/10811680.2013.746140.

Lechler, T., Wetzel, S., Jankowski, R., 2011. Identifying and Evaluating the Threat of Transitive Information Leakage in Healthcare Systems, in: 2011 44th Hawaii International Conference on System Sciences. IEEE, 1-10. doi: 10.1109/HICSS. 2011.230.

Leon, P.G., Ur, B., Balebako, R., Cranor, L.F., Shay, R., Wang, Y., 2012. Why Johnny can't opt out: A usability evaluation of tools to limit online behavioral advertising, in: Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems - CHI '12. ACM Press, New York, New York, USA, 589-598. doi: 10.1145/2207676.2207759.

Lewis, M., Coles-Kemp, L., 2014. Who says personas can't dance? The use of comic strips to design information security personas, in: Proceedings of the Extended Abstracts of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI EA '14. Association for Computing Machinery, New York, New York, USA, 2485-2490. doi: 10.1145/2559206.2581323.

Liginlal D., Sim I., Khansa L., 2009. How significant is human error as a cause of privacy breaches? An empirical study and a framework for error management. Computer Security 2009;28:215e28.

Liu, S., Kuhn, R., 2010. Data Loss Prevention. IT Prof. 12, 2008-2011. doi: 10.1109/MITP.2010.52.

Luger, E., Urquhart, L., Rodden, T., Golembewski, M., 2015. Playing the legal card: Using ideation cards to raise data protection issues within the design process, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15. Association for Computing Machinery, New York, New York, USA, 457-466. doi: 10.1145/2702123.2702142.

Lukosch, S., Lukosch, H., Datcu, D., Cidota, M., 2015. On the spot information in augmented reality for teams in the security domain, in: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15. Association for Computing Machinery, New York, New York, USA, 983-988. doi: 10.1145/2702613.2732879.

Malandrino, D., Petta, A., Scarano, V., Serra, L., Spinelli, R., Krishnamurthy, B., 2013. Privacy Awareness about Information Leakage: Who knows what about me?, in: Proceedings of the 12th ACM Workshop on Workshop on Privacy in the Electronic Society - WPES '13. ACM Press, New York, New York, USA, 279-284. doi: 10.1145/2517840.2517868.

Malatras, A., Sanchez, I., Beslay, L., Coisel, I., Vakalis, I., D'Acquisto, G., Sanchez, M.G., Grall, M., Hansen, M., Zorkadis, V., 2017. Pan-European personal data breaches: Mapping of current practices and recommendations to facilitate cooperation among Data Protection Authorities. Comput. Law Secur. Rev. 33, 458-469. doi: 10.1016/j.clsr.2017.03.013.

Mantelero, A., 2013. The EU Proposal for a General Data Protection Regulation and the roots of the "right to be forgotten." Comput. Law Secur. Rev. 29, 229-235. doi: 10.1016/J.CLSR.2013.03.010.

Mantelero, A., 2014. The future of consumer data protection in the E.U. Rethinking the "notice and consent" paradigm in the new era of predictive analytics. Comput. Law Secur. Rev. 30, 643-660. doi: 10.1016/J.CLSR.2014.09.004.

Mao, H.-X., Huang, K., Shu, X.-L., 2014. Research of cloud storage and data consistency strategies based on replica redundant technology, 2014 International Conference on Computer, Intelligent Computing and Education Technology, CICET (Computer, Intelligent Computing and Education Technology) 2014, 2, 1053-1056.

Maurer, M.E., De Luca, A., Hussmann, H., 2011. Data type based security alert dialogs, in: Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '11. ACM Press, New York, New York, USA, 2359-2364. doi: 10.1145/1979742.1979903.

McCauley-Bell, P.R., Crumpton, L.L., 1998. Human factors issues in information security: What are they and do they matter? 1, 439-443.

McCormac, A., Zwaans, T., Parsons, K., Calic, D., Butavicius, M., & Pattinson, M., 2017. Individual differences and Information Security Awareness. Computers in Human Behavior, 69, 151-156. https://doi.org/10.1016/j.chb.2016.11.065.

McCrae, R.R., & Costa Jr, P.T., 1986. Personality, coping, and coping effectiveness in an adult sample. Journal of personality, 54(2), 385-404.

Molok, N.N.A., Chang, S., Ahmad, A., 2010. Information leakage through online social networking: Opening the doorway for advanced persistence threats. Proc. 8th Aust. Inf. Secur. Manag. Conf. 70-80. doi: 10.4225/75/57b673cf34781.

Moquin, R., & Wakefield, R. L. 2016. The roles of awareness, sanctions, and ethics in software compliance. Journal of Computer Information Systems, 56(3), 261-270. https://doi.org/10.1080/08874417.2016.1153922.

Mou, J., Cohen, J., & Kim, J., 2017. A Meta-Analytic Structural Equation Modeling Test of Protection Motivation Theory in Information Security Literature. Thirty Eighth International Conference on Information Systems, 1-20. https://doi.org/10.1007/s10800-015-0795-2.

Mowbray, M., Pearson, S., 2009. A client-based privacy manager for cloud computing, in: Proceedings of the Fourth International ICST Conference on Communication System Software and middlewaRE – COMSWARE '09. ACM Press, New York, New York, USA, p. 1. doi: 10.1145/1621890.1621897.

Nurse, J.R.C., Buckley, O., Legg, P.A., Goldsmith, M., Creese, S., Wright, G.R.T., Whitty, M., 2014. Understanding insider threat: A framework for characterising attacks, in: 2014 IEEE Security and Privacy Workshops. Institute of Electrical and Electronics Engineers Inc., 214-228. doi: 10.1109/SPW.2014.38.

Oetzel, M.C., Spiekermann, S., 2014. A systematic methodology for privacy impact assessments: A design science approach. Eur. J. Inf. Syst. 23, 126-150. doi: 10.1057/ejis.2013.18.

Othmane, L. Ben, Lilien, L., 2009. Protecting privacy in sensitive data dissemination with active bundles, in: 2009 World Congress on Privacy, Security, Trust and the Management of E-Business. IEEE, 202-213. doi: 10.1109/CON GRESS.2009.30.

Pelteret, M., Ophoff, J., 2016. A review of information privacy and its importance to consumers and organizations 19, 277-301.

Preibusch, S., 2013. Guide to measuring privacy concern: Review of survey and observational instruments. Int. J. Hum. Comput. Stud. 71, 1133-1143. doi: 10.1016/j.ijhcs.2013.09.002.

Preuveneers, D., Joosen, W., Ilie-Zudor, E., 2016. Data protection compliance regulations and implications for smart factories of the future, in: 2016 12th International Conference on Intelligent Environments (IE). Institute of Electrical and Electronics Engineers Inc., 40-47. doi: 10.1109/IE.2016.15.

PWC, 2015. 2015 Information Security Breached Survey. London, UK.

Reason, J., 1990. Human Error. Cambridge University Press, New York.

Rogers, R.W., 1975 and 1983. A protection motivation theory of fear appeals and attitude change1. The journal of psychology, 91(1), 93-114.

Romanou, A., 2017. The necessity of the implementation of Privacy by Design in sectors where data protection concerns arise. Comput. Law Secur. Rev. 34, 99-110. doi: 10.1016/j.clsr.2017.05.021.

Sabillon, R., Cavaller, V., Cano, J., Serra-Ruiz, J., 2016. Cybercriminals, cyberattacks and cybercrime. 2016 IEEE International Conference on Cybercrime and Computer Forensic, ICCCF 2016, Simon Fraser University's Harbour Centre and Work Centre for Dialogue, Vancouver; Canada; 12-14 June 2016.

Safa N. S., Von Solms R., Furnell S., 2016. Information security policy compliance model in organizations. Comput Secur 2016; 56: 70-82.

Scharnick, N., Gerber, M., Futcher, 2016. Review of data storage protection approaches for POPI compliance. 2016 Information Security for South Africa - Proceedings of the 2016 ISSA Conference, 30 December 2016, Article number 7802928, 48-55.

Schultz, E.E., 2012. Human Factors and Information Security, in: Salvendy, G. (Ed.), doi: 10.1002/9781118131350.ch45.

Shropshire, J., Warkentin, M., Sharma, S., 2015. Personality, attitudes, and intentions: Predicting initial adoption of information security behavior. Computers and Security, 49, 177-191. https://doi.org/10.1016/j.cose.2015.01.002.

Shu, X., Yao, D., Bertino, E., 2015. Privacy-preserving detection of sensitive data exposure. IEEE Trans. Inf. Forensics Secur. 10, 1092-1103. doi: 10.1109/TIFS.2015.2398363.

Shu, X., Zhang, J., Yao, D., Feng, W.C., 2016. Fast detection of transformed data leaks. IEEE Trans. Inf. Forensics Secur. 11, 528-542. doi: 10.1109/TIFS.2015.2503271.

Snedaker, S., Rima, C., 2013. Business Continuity and Disaster Recovery Planning for IT Professionals: Second Edition, 2nd ed. Elsevier Inc. doi: 10.1016/C2012-0-06206-0.

Sommestad, T., Hallberg, J., Lundholm, K., & Bengtsson, J. 2014. Variables influencing information security policy compliance: a systematic review of quantitative studies. Information Management & Computer Security, 22(1), 42-75.

Sommestad, T., Karlzén, H., Hallberg, J., 2015. A Meta-Analysis of Studies on Protection Motivation Theory and Information Security Behaviour. International Journal of Information Security and Privacy, 9(1), 26-46. https://doi.org/10.4018/IJISP.2015010102.

Son, J.Y., 2011. Out of fear or desire? Toward a better understanding of employees' motivation to follow IS security policies. Information and Management, 48(7), 296-302. https://doi.org/10.1016/j.im.2011.07.002.

Spiliotopoulos, T., Oakley, I., 2013. Understanding Motivations for Facebook Use: Usage Metrics, Network Structure, and Privacy, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13. ACM Press, New York, New York, USA, 3287-3296. doi: 10.1145/2470654.2466449.

Stamati-Koromina, V., Ilioudis, C., Overill, R., Georgiadis, C.K., Stamatis, D., 2012. Insider threats in corporate environments: A case study for data leakage prevention, in: Proceedings of the Fifth Balkan Conference in Informatics on - BCI '12. ACM Press, New York, New York, USA, 271-274. doi: 10.1145/2371316.2371374.

Stolfo, S.J., Bellovin, S.M., Keromytis, A.D., Hershkop, S., Smith, S.W., Sinclair, S. (Eds.), (2007) The Insider Attack and Cyber Security: Beyond the Hacker.

Stoll, J., Tashman, C.S., Edwards, W.K., Spafford, K., 2008. Sesame: Informing user security decisions with system visualization, in: Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems - CHI '08. ACM Press, New York, New York, USA, 1045-1054. doi: 10.1145/1357054.1357217.

Sullivan, C., Burger, E., 2017. "In the public interest": The privacy implications of international business-to-business sharing of cyber-threat intelligence. Comput. Law Secur. Rev. 33, 14-29.

Taddicken, M., 2014. The "Privacy Paradox" in the Social Web: The Impact of Privacy Concerns, Individual Characteristics, and the Perceived Social Relevance on Different Forms of Self-Disclosure. J. Comput. Commun. 19, 248-273. doi: 10.1111/jcc4.12052.

Tahboub, R., Saleh, Y., 2014. Data leakage/loss prevention systems (DLP), in: 2014 World Congress on Computer Applications and Information Systems (WCCAIS). Institute of Electrical and Electronics Engineers Inc., 1-6. doi: 10.1109/WCCAIS.2014.6916624.

Tankard, C., 2016. What the GDPR means for businesses. Netw. Secur. 2016, 5-8. doi: 10.1016/S1353-4858(16)30056-3.

Tarasewich, P., Gong, J., Conlan, R., 2006. Protecting private data in public, in: CHI '06 Extended Abstracts on Human Factors in Computing Systems - CHI EA '06. ACM Press, New York, New York, USA, 1409-1414. doi: 10.1145/1125451.1125711.

Taylor, M., Haggerty, J., Gresty, D., Almond, P., Berry, T., 2014. Forensic investigation of social networking applications. Netw. Secur. 2014, 9-16. doi: 10.1016/S1353-4858(14)70112-6.

Terada, T., Katayama, Y., Torii, S., Tsuda, H., 2016. Security measures based on human behavior characteristics 52, 78-84.

Tikkinen-Piri, C., Rohunen, A., Markkula, J., 2017. EU General Data Protection Regulation: Changes and implications for personal data collecting companies. Comput. Law Secur. Rev. 34, 134-153. doi: 10.1016/j.clsr.2017.05.015.

Tupes, E.C., Christal, R.E., 1961. Recurrent personality factors based on trait ratings. USAF ASD Tech. Rep. 61-97.

Urquhart, L., McAuley, D., 2018. Avoiding the internet of insecure industrial things, Computer Law & Security Review.

Van der Auwermeulen, B., 2017. How to attribute the right to data portability in Europe: A comparative analysis of legislations. Comput. Law Secur. Rev. 33, 57-72.

van Lieshout, M., Friedewald, M., Wright, D., Gutwirth, S., 2013. Reconciling privacy and security. Innovation 26, 119-132. doi: 10.1080/13511610.2013.723378.

Vance A, Siponen M, Pahnila S., 2012. Motivating IS security compliance: insights from habit and protection motivation theory. Inform Manage 2012; 49(3): 190-198.

Vroman, K.G., Arthanat, S., Lysack, C., 2015. "Who over 65 is online?" Older Adults' dispositions towards information communication technology. Comput. Human Behav. 43, 156-166. doi: 10.1016/j.chb.2014.10.018.

Weber, R.H., 2015. The digital future - A challenge for privacy? Comput. Law Secur. Rev. 31, 234-242. doi: 10.1016/J.CLSR.2015.01.003.

Witzleb, N., Lindsay, D., Paterson, M., Rodrick, S. (Eds.), 2015. Emerging Challenges in Privacy Law. Cambridge University Press, Cambridge, UK.

Workman M., Bommer W. H., Straub D. W., 2008. Security lapses and the omission of information security measures: a threat control model and empirical test. Computers in Human Behavior 2008;24(6):2799e816.

Yin, X., Alonso, J., Machida, F., Andrade, E.C., Trivedi, K.S., 2012. Availability modeling and analysis for data backup and restore operations, in: 2012 IEEE 31st Symposium on Reliable Distributed Systems. IEEE, 141-150. doi: 10.1109/SRDS.2012.9.

Zuiderveen Borgesius, F.J., 2016. Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation. Comput. Law Secur. Rev. 32, 256-271.