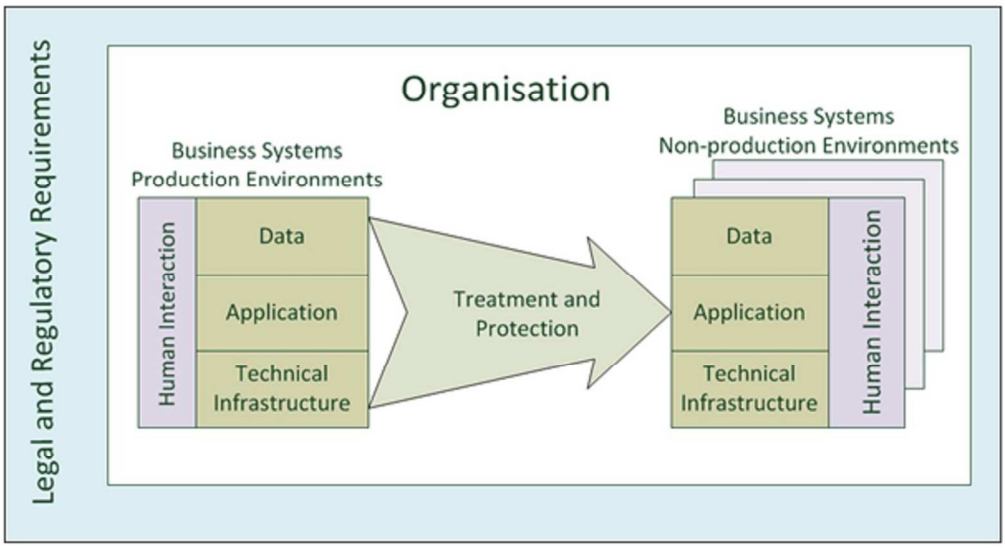
Emerald Information Management and Computer Security

On Data Leakage from Non-production Systems

Journal:	*Information and Computer Security*
Manuscript ID	ICS-02-2017-0004.R1
Manuscript Type:	Original Article
Keywords:	Data protection, framework, data leakage, de-identification, data masking

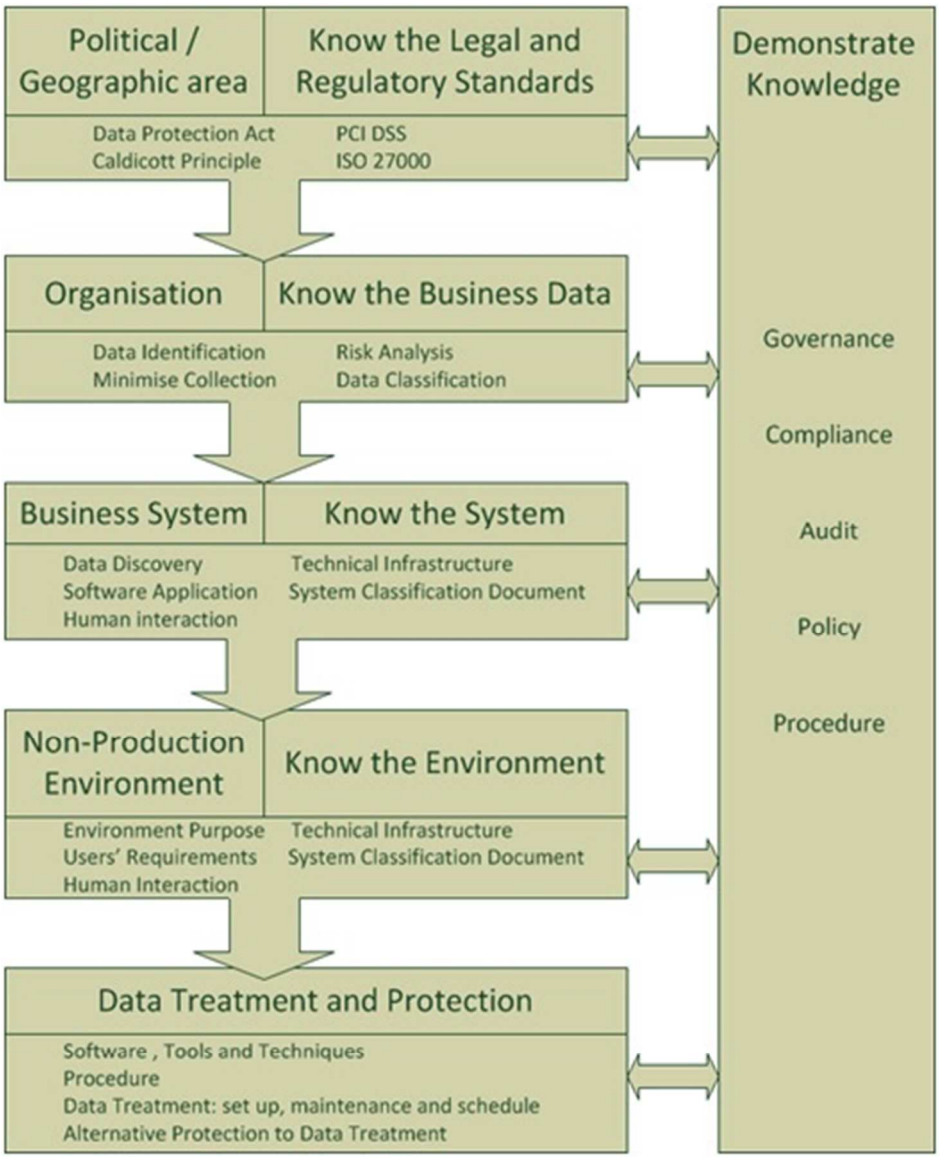
SCHOLARONE™ Manuscripts

Information and Computer Security

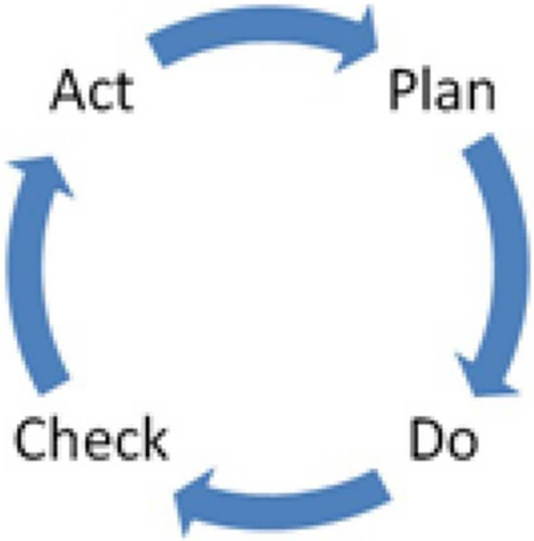


Conceptual Business Model showing the Interaction of Elements Affecting the Organisation
Conceptual Business Model show
319x173mm (96 x 96 DPI)

Information and Computer Security



Framework for Minimising Data Leakage from Non-Production Systems
Framework for Minimising Data
124x152mm (96 x 96 DPI)



Plan Do Check Act
Plan Do Check Act
78x76mm (96 x 96 DPI)

Computer Security

	Acceptable Practice	Best Practice
Know the Legal and Regulatory Standards.	Copies of relevant legislation, regulation and guidance are available.	Copies of relevant legislation, regulation and guidance are up to date and available electronically for reference.
	Employees are aware of how to access the documents above.	Employees receive guidance of how to make appropriate use of the documents.
Know the Business Data.	Business data meets legal and regulatory standards.	Evidence of business data meeting legal and regulatory standards reviewed and audited regularly.
	Data retention policy in place.	Data retention policy includes regular scheduled reviews and audits.
Know the System.	System data retention policy conforms to business requirements.	System data retention policy supported by archiving or data destruction procedures.
	System technical and control documentation is available.	Documentation is regularly maintained. A recognised change process is in place.
Know the Environment.	Environment Technical and control documentation is available.	Employees know the use and purpose of each environment. Controls are in place and actively monitored.
Data Treatment and Protection.	Documented data treatment and protection plan is in place.	Data treatment and protection plan is reviewed at regular scheduled intervals, and re-evaluated during system and environment changes.
Demonstrate Knowledge	Confirmation and evidence that acceptable or best practice achieved at each stage	

Information and Computer Security

On Data Leakage from Non-production Systems.

Abstract

This study is an exploration of areas pertaining to the use of production data in non-production environments. During the software development lifecycle, non-production environments are used to serve various purposes to include unit, component, integration, system, user acceptance, performance and configuration testing. Organisations and third parties have been and are continuing to use copies of production data in non-production environments. This can lead to personal and sensitive data being accidentally leaked if appropriate and rigorous security guidelines are not implemented. This paper proposes a comprehensive framework for minimising data leakage from non-production environments. The framework was evaluated using guided interviews and was proven effective in helping organisation managing sensitive data in non-production environments.

Keywords

Framework, simplified business model, data leakage, de-identification, data masking

1. Introduction

Data leakage is a term describing the unauthorised transfer of personal and / or sensitive data from a computer, system or data centre to the outside world[1].Organisations are expected to protect their data. A data security policy does not necessarily guarantee compliance.Access to sensitive or personal data in production environments may be controlled by user role or based on business function. Policies for production dataare not necessarily applied to non-production environments containing copies of live data, increasing the risk of data loss. The practice of cloning or refreshing non-production (support, test or development) systems from a production data set may give a further route for data leakage. It may be perceived by some as an efficient method of creating a comprehensive data set for use in development and testing[2, 3], but does not address the risk to the data.A strict prohibition policy, for example a refusal to allow any refresh from production systems, may be more secure than allowing production data to be copied. It may not, however, satisfy the requirements of analysts, developers or application support personnelwho create, deploy and test patches and configuration; or perform acceptance testing. The quantity and variety of sensitive data and the needs of the non-production system users should be considered in selecting the most appropriate method of data leakage prevention. Applying additional access controls and monitoring non-production data sets may be an appropriate alternative.Existing work in this area tends to deal with industry or government standards, and specific areas such as aspects of de-identification.

This paper proposes a comprehensive framework for minimising data leakage from non-production environments. Advice, guidelines and frameworks from vendor specific and vendor independent sources are used to identify good practice, forming the basis of a novel framework to provide practical assistance in populating and managing a non-production environment. This work is important because it brings together disparate sources to form a practical solution to the decision making process and providing documentary evidence for assurance and audit purposes. The contributions of this paper include: (i) A critical review of advice, guidelines and frameworks from vendor specific and vendor independent sources, including legal and regulatory standards for handling personal and sensitive data (Sect. 2); (ii) A comprehensive framework for minimising data leakage in non-production environments (Sect. 3); (iii) The evaluation of the proposed framework using guided interviews highlights its usefulness in real-world organisation settings.

Information and Computer Security

The paper is organised as follows: Section 2 provides critical analysis from different viewpoints, and identifies points required for a practical solution considering guidelines and advice, using vendor specific and vendor independent sources. In section 3 the framework for minimising data leakage from non-production systems is presented and evaluated in Section 4. Section 5 concludes paper.

1.1. Motivation

To reduce data leakage from non-production environments, this article proposes asimplified business model and framework to consider the use and protection of such data from a number of viewpoints. There is a range of both vendor specific [4, 5, 6,7, 8] and vendor independent [9, 10] advice relating to the treatment of data in non-production environments. These bring forward particular aspects, such as data identification [6, 10] and classification [11] which are important to the de-identification process. The vendor specific sources are by their nature constrained to proprietary tools, whereas the vendor independent sources focus on specific aspects. Numerous sources discuss the application, benefits and drawbacks of specific de-identification techniques [12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22], the understanding of techniques is useful, but the selection of one or more techniques cannot be considered adequately without an appreciation of the anticipated use of the data. Legal requirements [2, 3 23, 24] and regulatory standards [25] also impact the use of data in both production and non-production systems.

1.2. Contribution

This article brings together information from the disparate sources above to propose a simplified business model and framework. The proposed model diagrammatically describes the interactions of elements affecting the organisation. It highlights how non-production environments may be perceived as separate from the business systems, but despite the perceptions is still subject to the same legal requirements and constraints. It shows the interdependency of data, software, technical infrastructure and human interaction and how the change of one element may affect the others. The proposed framework describes the process flow and forms a practical solution in assisting the decision making process and providing documentary evidence for assurance and audit purposes. It looks at the requirements of the non-production system in relation to the legal and regulatory constraints as well as the organisational requirements, and business systems. The impact of human factors on the data is also considered to bring a holistic approach to the protection of non-production environments.

2. Related Work

This section identifies basic terminology, definitions and examples used in the classification of data within the UK. Terms such as personal, personally identifiable, confidential and sensitive take on special meaning when used in different contexts. Some differences in application and conflict within areas such as health and finance are highlighted below.The Data Protection Act 1998 (DPA), Caldicott principles and Payment Card Industry (PCI) standards are used as a basis.

The DPA[23] has eight principles of good practice for handing personal information[26] and gives detailed definitions. The Information Commissioner's Office (ICO)is responsible for providing advice relating to the enforcement of the DPA[27].Principle 7 of the DPA includes the requirement to protect "against unauthorised or unlawful processing of personal data and against accidental loss of . . . personal data" [28].DPA defines personal data as relating to or identifying a living individual. Where an individual may be identified partly by the data held and partly by other information, then the data

Information and Computer Security

held is still classed as personal data. The ICO provides further guidance[29, 30].Caldicott principles for the governance of Health and Social Care data use DPA as a basis and extend the definition of personal data to include data relating deceased people [31].Payment Card Industry Data Security Standards(PCI DSS) uses the term 'Personally Identifiable Information' (PII) [32] as the equivalent of DPA's personal data. DPA further defines sensitive personal data[23] as relating ethnic origin, political opinions, religious belief, trade union membership, physical or mental health, sexual life and criminal or alleged behaviour. The information could be used in a discriminatory way and so should be treated with greater care[27]. Caldicott principles use the term 'personal confidential data' to include the DPA definition of 'sensitive', and also information given in confidence [31]. PCI DSS cardholder details of debit or credit account number (primary account number or PAN) with the addition of cardholder name, expiration date and/or service code are defined as sensitive data. A further level of 'sensitive authentication data' includes personal identification number (PIN) which is used to authenticate the cardholder and as authorisation for payments [32].The classification of data depends, not only on the data, but the reason for the data being held with complexity increasing with combinations of data elements covered bydifferent governance.

2.1. Reasons for Protecting Data

This section is an overview of the legal and business reasons for protecting data, introducing the complexity of legal and industry standards, and a brief analysis of the impact of human factors. Data is a business asset. There is currently no global legal standard for protecting data. Legal requirements differ both between EU countries and countries outside Europe. The current UK legislation is the DPA[23]which enacts EU Data Protection Directive 95/46/EC [24],and may change to incorporate the General Data Protection Regulation (GDPR)[33].The DPA regulates how the personal data of living individuals is processed and stored by government, organisations and businesses, in line with eight data protection principles [34].The use of data[30], and its transfer outside of the European Economic area (EEA) is covered by principle eight[35] which requires assurances of an adequate level of protection for data transferred outside of the EEA.Exceptions are made for countries outside the EEA which provide sufficient protection [36], or havean information sharing agreement such as Privacy Shield the EU – US framework for data transfer[37].

There are further requirements for the protection of personal data. Organisations with access to patient data require further protection. The Caldicott Report [38]defined principles for the governance of Health and Social Care data and the appointment of Caldicott Guardians to take responsibility for protecting the confidentiality of patient and service user information. It has been reviewed [31] and includesfurther recommendations to improve data quality, confidentiality and governance whilst supporting data sharing that is necessary for patient care. Organisations dealing with financial transactions by credit or debit card are recommended to comply with the global industry standard guidance to improve financial transaction security, prevent the theft of payment card data and avoid security breaches [32].PCI DSS v3.1 is a framework for data security, risk reduction and incident response [25]. It includes standards for software, hardware vendors and resellers [39].

GDPR[40] is expected to be transposed into the law of EU member states by May 2018[41, 42].Itsintention isto improve individuals' rights over their data, allow appropriate data sharing and adapt to newer technologies and also includes controls on consent, data processing practice, 'privacy by design', and management of data breaches. In addition to legal, regulatory and business reasons for protecting data, there is also the effect of human factors on the security of data. Pricewaterhouse Coopers (PwC) 2015 data breaches report [43] identifies a significant rise in staff related data breaches, with half of the worst breaches being caused by inadvertent human error. Deliberate misuse

Information and Computer Security

of systems, whilst lower, still contributes to the number and size of data breaches. Hurran [44] categorises the actions of insiders as "non-malicious and unintentional"; "non-malicious and intentional"; and "malicious and intentional".He describes the threat from the non-malicious unintentional group as a 'substantial' part of the problem. These employees have no intention of doing wrong, but through a failure in the recruitment or training process are making mistakes.The threat from the non-malicious intentional group is seen as a bigger problem, because it is often caused by employees circumventing existing processes, such as copying data or sharing passwords, in an attempt to work more efficiently. Hurran attributes the fault to the organisation as a business process failure.The risk of malicious intentional insiders can be reduced by removing factors within the organisation which enable the insider's action to take place. Silowash[45] focusses on groups within an organizationfor insider threat mitigation, and outlines best practices including knowing the organisations assets and considering both insider and external threats.Evans [46] investigates human behavioural practices and their effects on cyber security assurance. He questions the effectiveness and suitability of existing assurance methods, as compliance does not seem to be preventing numerous high profile breaches. Evans highlights a possible connection between the increase in data breaches and the treatment of sensitive data being a compliance exercise rather than an ethical consideration.

The protection and security of data in systems is a part of the requirement. This should be balanced with the use of data within non-production environments. The next section looks briefly at the uses and often contradictory advice for using production data in non-production environments.

2.2. Uses of Production Data in a Non-Production Environment

A production environment may also be known as a live or operational environment. A non-production environment is any system, hardware or software combination that is not the production environment. During software development, testing and maintenance there may be non-production environments serving different purposes to include unit, component, integration, system, user acceptance, performance and configuration testing. Each may individually or in combination [47] utilise separate environments with different hardware configuration, components, software versions or data[48]. An operational support environment used for the reproduction of live faults is a non-production environment configured as closely as possible to the production environment to enable the resolution of faults without impacting the production environment in the process.Organisations and third parties have been and are continuing to use copies of production data in non-production environments [49]. The conflict with the DPA is highlighted [50] and the lack of security explained by vendors of de-identification products and services [4, 5, 9]. Other software vendors see the benefits of using production data to improve the testing process[49, 51].The opposing views and conflicting advice may lead to confusion over whether using production data in a non-production system is acceptable,leading to difficulties selecting which, if any, of the many products and services will be of benefit in securing the data.

The following section looks at existing guidelines and advice to collect the salient points to be addressed.

2.3. Analysis of Advice, Guidelines and Frameworks

This section providesan analysis from different viewpoints, and identifiespoints required for a practical solution considering guidelines and advice, using vendor specific and vendor independent sources.From a process point of view, data classification is required to establish controls on systems.Data identification [6, 10] is the first step because unknown data cannot be adequately classified [11]. It may be necessary, initially, to limit the scope of data classification, as dealing with

Information and Computer Security

the data for an entire organisation may become an overwhelming task. A pilot exercise in a particular geographic or business area may be undertaken to give direction. The classification process also includes understanding how legal requirements and industry standards are applied to different business areas.Initially it may appear straightforward to align business processes directly with legal requirements or industry standards, but this may lead to an unworkablemodel. The provision of adult social care services provides an example of a business area being affected by a number of regulations and standards; Caldicott principles apply to the adult social care data, PCI DSS covers the payments made by credit card, and all other personal datais subject to DPA.The business reasons for the collection and retention of data at an organisation level should be scrutinised. Sensitive data which is not required for business purposes should not be collected; minimising unnecessary data collection will lead to less data requiring protection [10].Data classification at an organisation level requires data discovery at a system level [6], to check that only the expected data is held in the system. Any anomalies must be addressed.The scope moves from an organisational level to a system level, to the analysis of the non-production system requirements. Understanding the use of the non-production system is a key factor inachieving a balance of security and usability. Different instances of a system will have different requirements [6, 52]. For example, a live support system may require a full data set with particularly sensitive data de-identified, but an interface development system could fulfil requirements by using a subset of de-identified data or a generated test data set. If the level of de-identification is light, then further safeguarding practices such as access control and audit procedures are required. When the non-production system requirements have been assessed, the specific techniques for de-identification can be investigated.

Data treatment requirements refer to available techniques[52]and products appropriate to the system architecture. Capgemini [53]and Securosis[54] principles attempt to balance usability and security. De-identification should be managed, non-reversible and repeatable, resulting in data that is representative of the source data, with integrity maintained. To meet these principles, redaction techniques such as nulling out and gibberish generation would only suit free text such as memos and notes. IBM [7] looks at data privacy concerns related to analytics and ‘big data’ and considers older scripted techniques to be insufficient because they are not robust or scalable.Delphix[8]proposes using Data as a Service (DaaS). DaaSis seen as a way for organisations to hand over the risk and responsibility [55, 56]of data privacy to a third party who will identify data provenance and apply anonymization techniques appropriately. This approach was highlighted by ODCA as a challenging contractual issue [11].In the cases where light or no de-identification takes place, processes to determine supplementary security and documentary evidence for audit purposes are required.

The analysis of advice, guidelines and frameworks to find a practical solution to selecting and implementing a de-identification solution has highlighted the importance of the following areas:

- Legal and regulatory standards applicable
- Business data identification and classification
- Data within applications
- Requirements of the non-production application users
- Data treatment and protection as a compromise between security and usability
- Management of the process

In summary, there appears to be no shortage of definitions for what constitutes sensitive information for both general use and specific business sectors. The human element within information security is as important as technical factors, highlighting the need for efficient and appropriate business processes, and the treatment of sensitive data being an integral ethical consideration rather than a tick-

Information and Computer Security

box exercise. Techniques are available for the de-identification of data, with a range of products and vendor independent guidelines for the de-identification of sensitive data in non-production environments. There appear to be practical reasons why de-identification is not used extensively within the software delivery lifecycle. The next section expands the areas highlighted and proposes a framework to address them.

3. Proposed Framework for Minimising Data Leakage from Non-Production Systems

A framework is defined as "An essential or underlying structure; a provisional design, an outline; a conceptual scheme or system." [57].The proposed framework for minimising data leakage from non-production systemsoutlines a method for gathering knowledge to effectively manage the use of sensitive data outside of the organisation's production environment. The framework intends to offer practical and useful guidance to identify an efficient way for an organisation to treat and protect the sensitive data within its non-production environment.

Initially, the framework overview is developed to identify how non-production environments can be affected by legal, organisational and system structures, as well as human factors. It enables the business to consider how areas such as change management may affect the requirements of non-production environments. The effects of human factors on unauthorised data exfiltration also require consideration [45]. The framework is not a set of rules to be followed. The use of specific rules would only benefit a limited number of organisations or require a complex set of rules. The framework enables the organisation to consider its own particular position and tailor the processes specifically to its needs.

The first part of this section looks at the reasons for using the framework in the context of the conceptual business model.

The second part describes the six phases of the framework which are:

1) Know the legal and regulatory: standards showing demonstrable knowledge relating to the business of the organisation, and its use and storage of data.
2) Know the business data: describing the importance to the organisation of data as an asset, and how sensitive business data should be identified and protected.
3) Know the system: discovering data within business systems to confirm or improve the organisation's knowledge of its actual business information held.
4) Know the Environment: identifying the purpose of non-production environments in an attempt to make the best compromise between usability, security and cost.
5) Data treatment and protection: defining a procedure for non-production environments using the plan-do-check-act cycle.
6) Demonstrate knowledge: identify and implement the practices needed at each stage of the framework, and bring the information together to provide evidence.

3.1. Conceptual Business Model

Business systems may be described as a combination of data, application software and underlying technical infrastructure. They support the administration of the business, and facilitate its commercial purpose. Systems include support systems such as payroll, human resource, accounting, inventory, catalogues as well as ecommerce, health and education. A conceptual model, based on the NIST Enterprise Architecture Model [58], depicts the integration of legal and regulatory requirements; human interactions; business requirements; business systems and non-production environments in Figure 1 : Conceptual Business Model showing the Interaction of Elements Affecting the

Information and Computer Security

Organisationbelow. At all levels there is human interaction, whether from members of the public completing a weekly shop, call centre staff assisting on-line banking customers, accountants producing statutory accounts or engineers monitoring the technical infrastructure.

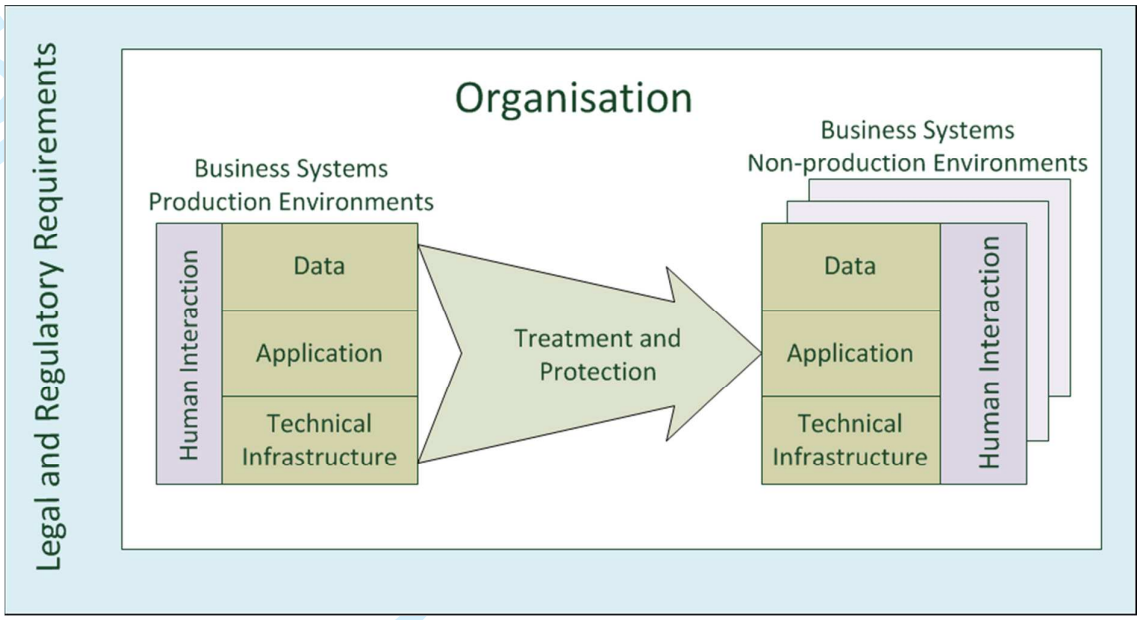


Figure 1 : Conceptual Business Model showing the Interaction of Elements Affecting the Organisation

The model shows how non-production environments may be perceived as separate from the main business systems, but how they remain subject to the same legal and regulatory requirements. Data, applications, technical infrastructure and human interactions with the non-production systems may be the same as, or different from, those associated with the production business systems.

There may be multiple non-production systems relating to each business system which are used for different stages of the software development lifecycle (SDLC). The differing needs may require different configurations of data, application, infrastructure and human interaction. For example, a production support environment, used to replicate and resolve live issues, can be argued to require an exact replica of the production environment. Taking this requirement literally, it would require the same data, application software and infrastructure. However, the human element may be different, with support analysts having access to the support environment,and would not expect to have access to the live environment. In the case of public facing systems, public users would be excluded from the non-production environment. An environment for offshore development would require representative data rather than a copy of the live data set.

These examples show the need for a practical aid to decision making and to identify factors influencing the requirements of non-production system users and to guide the decision makers through a process of providing non-production environments with an adequate security [59].

3.2. Description of the Framework

The proposed framework is depicted in Figure 2and describes a process flow from the legal and regulatory requirements to data treatment and protection; gained through understanding the organisation's business, the production system, the purpose and requirements of the non-production environment. Each stage builds on earlier stages of the process, but also provides appropriate

feedback. The knowledge of the stages should be demonstrated through governance, compliance, audit, policy and procedure.

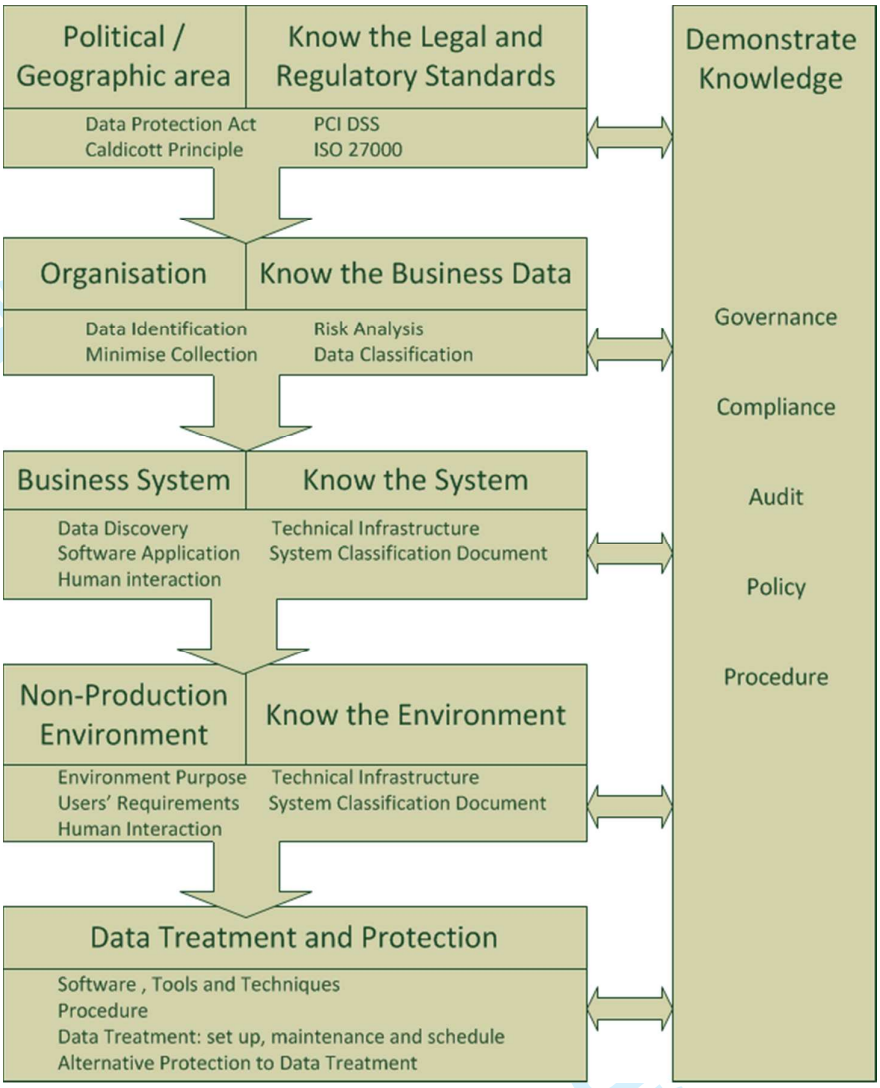


Figure 2: Framework for Minimising Data Leakage from Non-Production Systems

3.2.1.Stage 1 -Legal and Regulatory Standard

Demonstrable knowledge of legal and regulatory standards relating to the use and storage of data is important.The primary source of UK data protection legislation isthe Data Protection Act (1998) [23], which is upheld by an independent public body, the Information Commissioner's Office. Regulations also apply to different business areas such as Caldicott principles for health and social care data [31], PCI DSS for payment card data [32], and regulations which are applicable to data sharing across geographical areas, such as Privacy Shield; the EU – US framework for data transfer[37]. In addition there is intellectual property, business confidential information and data sharing for the prevention of fraud to consider.

3.2.2.Stage 2 -Know the Business Data

It is important for an organisation to appreciate laws and regulations affecting its business and to use organisation level risk assessments to consider threats to data security from insiders and business partners as well as outsiders [45]. The business's information management function would be expected

to cover the business aspects of this area. Organisations may also publish details of information management, governance or policy relating to specific privacy aspects. Knowledge of the business data is the identification of data using business terminology, knowing what information is held, for what purpose, its location and how it is classified[60]. Identification of sensitive data is important, as it is only possible to effectively protect what is known [10]. Collection and retention of business data should be considered during the cataloguing process. Identify the data required for business purposes and challenge data that is collected with no sound business reason. Minimising the collection and retention of data leads to a lower risk of data leakage, because data that is not collected and not held cannot be breached. NIST [10]recommends regular scheduled reviews of data retention policy and conformity to that policy. Plan the removal of unnecessary sensitive information. Using the example of social care provision payments; consider whether it is necessary to collect the National Insurance number of the payee when invoicing for social care support of a dependent[61].

Consider the retention schedule of data that is collected.

- Legal requirements for retention of data e.g. for HMRC purposes, Electronic Communications Data Retention Directive[62].
- Industry practice requirements for the retention of data, for example PCI DSS.
- Education/social care records for children with special needs.
- Customer led choices for convenience, e.g. retention of credit card details for future payments.
- Describe retention schedule for other data. Consider how to communicate the retention schedule to data controllers and subjects.

Document the organisational level business data, develop a data retention policy and schedule.Using the organisational level business data document, and the applicable legal and regulatory standards classify the data in terms of its type. Consider which classification scheme to use, UK Government uses ‘official’, ‘secret’ and ‘top secret’[63], Photopoulos[64] identifies a three popular classifications of ‘confidential’, ‘internal’ and ‘public’. NIST[10] recommends using ‘high’, ‘medium’ and ‘low’ to indicate the level of harm that may result from leaked data. Whichever method is considered to be best suited to the organisation’s business should be used to produce an organisation level data classification document, describing the data held at an organisational level along with its sensitivity rating and the applicable law or regulation. Including the reasoning behind the classification of sensitive data maybe useful for future compliance checks or when dealing with regulatory or legal changes.

3.2.3.Stage 3 -Know the System

The organisation, in the course of its business, will use a number of business systems to administer business tasks. The systems may be hosted internally, cloud based, outsourced or a combination. Whichever provision is used, the organisation should have knowledge of or control over the protection of the data within, whether that is complete responsibility for systems hosted internally, shared responsibility for cloud security or assurance from third party suppliers. Organisations may outsource systems for both cost saving and risk management reasons, some with greater satisfaction than others[65]. Knowing the system is about the data that the business actually holds within its systems. Data discovery is a method of identifying both structured and unstructured data residing within systems. Differences between expected data and data actually held should be analysed and, if necessary, fed back to the earlier stage.Each system may hold data relating to that which has already been classified at the organisational level, allowing the data in the systems to be catalogued against the business use. It is possible that data in systems is not known of or understood at an organisational level. For example, sensitive data such as National Insurance number may be routinely collected and retained for operational purposes that are not recognised as a business need. In these cases the

Information and Computer Security

differences may be highlighted and a decision made between the business expectations and the organisational requirements. A holistic approach to data identification and classification is recommended because data at a system level may not be considered sensitive, but becomes sensitive when combined with other data from another system.This stage in the process may lead to the creation of a system classification document containing details of data within a system using the organisational sensitivity classification from the 'Know the Business' stage.

3.2.4.Stage 4 -Know the Environment

Business systems may need functional and technical support, development, user acceptance testing, production support or offshore developments. Environments in addition to the production environment may be required. The use and contents of the environment may be within the control of the business area, or within the control of another business area such as IT, or an external organisation. The additional environment may not necessarily be an equivalent of the production environment. The non-production environment may be affected by different factors; it is important to understand its purpose and the needs of the users. In Figure 1 : Conceptual Business Model showing the Interaction of Elements Affecting the Organisationthe software application, technical infrastructure and human interactions in the non-production environment may differ from the production system, necessitating an investigation of how the data should differ from the production system to maintain the required level of protection [48] for the data.The environment users' requirements should to be compared with the data security needs to form a basis for the contents of the non-production system, and how any copies of production data are treated. This is an area where cost, ease of use and data security conflict because of the differences between the perceptions of need and the actual need. Effort may be required to alter the perception of need by guidance, training or example.

Developers may have a perceived need to use production data to perform realistic tests early in the lifecycle. This could appear reasonable, in an attempt to reduce errors later in the lifecycle. 'Test first programming' within the agile methodology XP (extreme programming) advocates designing the test before coding to ensure the coding meets the test[66], but does not require the use of production data. Therefore, whether following traditional or agile methodologies, the emphasis ought to be on the design of suitable tests, which may be automated and repeated, rather than using production data as the test base. This leans towards using generated data sets based on sound data analysis rather than copies of production environments for code or unit testing. Integration and interface testing relies on suitable test cases passing from one module or system to another. Different standards being applied to different modules or systems raises the possibility that modules or systems working independently but not necessarily together. Again, the quality of test cases is important. There is some justification for using a de-identified subset of production data, but generated data sets designed to flow from one system or module to the next would be useful, especially if those tests included suitable boundary cases. Testing to check for faults and failures is just as important as pushing production data through to make sure it does not fail. User acceptance data must be realistic and demonstrate that the system functions satisfactorily. The justification of using a full copy of production data may be made for the reasons that the users performing the tests already have access to the production data. It is important to consider how and where any test scripts and results are stored, as this highlights one of the leakage points of sensitive data. The preferred solution may be to use a partially de-identified data set to ensure that the most sensitive data is not propagated across file systems. Knowledge of the system would be required to identify the more sensitive areas.

Operational support environments, perhaps, have the greatest requirement for using a complete copy of production data, because of the need to replicate issues that have occurred in a production environment. It should be noted that, especially in complex systems, it is not always possible to

Information and Computer Security

replicate a fault on a cloned copy of production data as faults may be caused by a convergence of factors other than data, such as memory, traffic volume or conflict with other system usage. Where a copy of production data is to be used, then it may still be prudent to de-identify some of the most sensitive data. The judgement of the data to be used in the non-production environment could be the responsibility of the data owner, but to be decided in conjunction with those responsible for non-production environment users to give the best compromise in the conflicting triangle of security, usability and cost [67]. Some understanding of de-identification would lead to better decisions, for example knowing that a set of de-identified data can be recreated on multiple environments, so that results from one environment can be compared with results from another environment, thus changing the possible perception that faults cannot be tracked through environments.

If it is considered that the best option for the business is to allow developers to unit test on copies of production data, because there are sufficient other safeguards in place, the framework should reflect this decision, and describe the methods used to secure the sensitive data. It is advisable to document the agreed format of each non production environment, because requirements change. It is quite possible that an agreement has been made, for example, to use a copy of a production database and this may be acceptable to all parties because it is held onsite with sufficient access controls and no web availability. But a move to a multi-tenant cloud environment would change the access controls, the technical infrastructure and the human interaction. These changes would require the re-evaluation of the data security needs.The output of this process is the environment classification document which describes the system, and the requirements for the data in this environment and its expected use. If required, a further detail of degree of de-identification may also be applied.

3.2.5.Stage 5 -Data Treatment and Protection

The process of data treatment and protection describes how the data in the non-production environment should be created and treated, based on the non-production environment users' requirements and the categorisation of the production system data sensitivity.The balance between usability, perceived usability and data security must be considered. NIST [59] defines adequate security as appropriate to the risk and harm from the loss or misuse of data. Neither data treatment nor protection can be considered an easy option. There is no simple 'one-size-fits-all' solution.Data treatment and protection may be seen as a spectrum; ranging from generated data requiring no additional protection, to 'as live' data which may require 'as live' protection. When agreement on the level of data treatment and protection has been reached, the process of defining the method of treatment may begin. Generated data would need no further treatment and additional protection. Ensure that generated data contains no real data.Where no treatment is to take place, appropriate protection should be put in place and documented. The protection could include limiting access by revoking user rights to all except specific users.

If a level of de-identification has been agreed, the method must be decided upon, tested and implemented along with additional protection appropriate to the level of treatment. There are many de-identification techniques, tools and software available. This section is not intended to be allied to any particular tool or software. There is no 'magic button'. Whatever method of de-identification is used will require an iterative cycle of analysis, set up and testing before the data treatment is procedure complete. Plan-Do-Check-Act (PDCA)[68] may be used to describe that process (see Figure 3: Plan Do Check Act.

Information and Computer Security

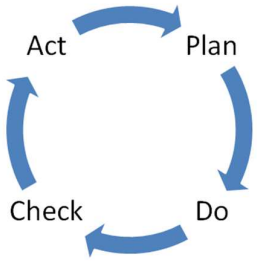


Figure 3: Plan Do Check Act

- Plan: Identify and analyse the problem.
  The non-production environmentrequirements have been documented during the 'Know the Environment' stage using the system classification document as a basis. This document may still be challenged to ensure it will both meet the security needs and the needs of the users. De-identification should be kept to a minimum; there is no benefit in de-identifying data which is neither sensitive nor a constituent part of sensitive data. The resulting de-identified data should be representative of the production data, and maintain referential integrity.The plan stage should identify tools and techniques that will best meet the requirements of the environment whilst maintaining the appropriate level of security. If considering the purchase of software or services, it may be prudent to plan a number of non-production environments on a number of business systems to ensure the software or service meets the requirements for all.
- Do: Develop and test potential solutions.
  This may involve using subsets of data, generating a series of potential solutions and performing a pilot exercise of the best solution.
- Check: Measure the effectiveness of the proposed solution, and improve as necessary.
  Check that de-identification provides the expected results. Combine the results of de-identification with the analysis of the test requirements. Ensure referential integrity is maintained. Look at unstructured data to make sure sensitive data does not remain in place. If the de-identified data is suitable to be passed on to the environment users then it is possible to progress to the Act stage, otherwise the Do and Check phases may be performed until a satisfactory outcome is found.
- Act: Implement the de-identification routine, including automation for maximum efficiency.
  The de-identification process should be repeatable.

As the quantity or the sensitivity of data increases in the non-production environment, the need for additional security from the application software, infrastructure or control over human factors also increases.

3.2.6.Stage 6 - Demonstrate Knowledge

The framework so far has dealt with the process flow from identification of legal aspects through to defining a data treatment and protection method. The process to demonstrate knowledge is included to provide assurances, which may otherwise be lacking, at each stage of the process. NIST[59] describes the purpose of assurance as giving confidence that intended security controls are effective. Evans[46] identifies cases where security controls are implemented, but there appears to be a lack of confirmation that those controls have been applied correctly or as intended. This stage of the framework is included toexplicitly extract that assurance from each of the other five stagesin the process. Effort must be made to ensure that the checks are done for the right reasons, rather than becoming another 'tick list', this may be done by defining acceptable practice and best practice for each stage in the framework as illustrated in Table 1: Examples of Acceptable and Best Practice. It is recognised that each organisation will have different requirements, so the examples are given as

Information and Computer Security

advice. NIST [10]also considers incident response to be an important control, because a breach of sensitive data from a non-production environment is still a breach.

Table 1: Examples of Acceptable and Best Practice

Ideally, the framework would be used in the order described. It is acknowledged that circumstances are not always ideal. The benefit of this framework is that it is not prescriptive, it allows the position to be assessed, and the knowledge required to be assessed before moving on. For example, if a requirement arose to send a database to a third party for testing application changes, this would be identified within the 'Know the Environment'process, assessing the non-production environment purpose and requirements.

4. Evaluation of the Framework

A set of guided interviews provided an initial review of the framework by professional people who have access to potentially sensitive data within the SDLC especially in non-production environments.The three objectives of the guided interview were to review the proposed framework and determine any benefit to the interviewee's organisation; identify any shortcomings of the framework and its application; and discuss requirements for a more detailed description of the framework.A qualitative research method was used because it seeks interviewee's opinion[69]. The guided interview included a verbal explanation of the framework documentto verify the interviewee understands the document andgive an opportunity to address any lack of clarity. A discussion of the interviewee's current use of sensitive data and non-production systems followed, to allow the interviewee to consider data use in their role as well as from a more abstract "what if" viewpoint. The framework and its stages were discussed in relation to the interviewee's role covering: how legal, business and technical knowledgeare addressed; improvements to their work; conflicts between the framework and their role; and improvements and additions to the framework.

All three interviewees had access to a variety of sensitive data, including but not limited to personally identifiable information of living and deceased persons, sensitive personal data, sensitive personal and business financial information, human resources records, and details leading to single and multi-agency prosecutions. The interviewees' knowledge and roles cover two broad areas:

- In depth expert knowledge of 'legal and regulatory standards', elements of 'business data' and 'business system', but with a little IT knowledge.
- A little informal knowledge of 'legal and regulatory standards' and 'business data'. In depth 'system' and 'environment' knowledge with some understanding of de-identification techniques.

The interviewees considered that the framework explained the importance of legal, business and technical knowledge. In reviewing the framework, it highlighted gaps in theirknowledgeand application of organisational standards.The process used to maintain systems was explained sufficiently.The framework assisted in improving their work by making a case for using some de-identification techniques; for example, replacing email addresses to reduce the risk of accidentally sending test documents to real users; and changing bank details so that test payment and direct debit files would not result in erroneous payments or debits if transmitted in error.Raising awareness of access restriction showed that there was an alternative to de-identification, and that such a decision should be made deliberately rather than by default.It also highlighted the amount of sensitive data in unstructured formats, such as test plans, sample reports, documentation and issue resolution notes. The explanation of non-production environments being required for different purposes showed that environments are not necessarily interchangeable. Points raised were that the requirements for development and test environments need to be considered. For example production support requires a

Information and Computer Security

copy of production data so that data fixes issued by the supplier may be tested before deployment to production. User acceptance testing (UAT) access could be 'as production' making sure that users have no more access in a non-production environment than in the production environment. The training environment does not need production data. Roles and responsibilities should be carefully considered. Functional and technical roles needs consideration, for example a functional consultant will require application system administrator access to perform and test configuration set up, but should not have access to a copy of production data. Understanding the purpose and use of an environment enabled the data and protection to be made appropriate to the use and audience.

Conflicts between the framework and the interviewee's role were identified. The perception of productivityidentified as a prominent conflict. Using generated test data was perceived by the interviewees as less productive initially, because of the additional resources required to generate data, scenarios and test cases. However when consideration was given to the use of automated testtools productivity was expected to improve. Testing aggregated data reportsusingde-identified or generated data setswas discussed, with concern expressed that this may allow individuals to be identified when the reports were run on production data. The use of commercial products was considered to adversely affect the stages of 'know the system', 'know the environment' and 'data treatment'because it was thought that there was a greater knowledge of in-house systems, whereas support for commercial systems tended to be purchased. Commercial systems also came with warranty and support restrictions. Some suppliers do notprovide support when unauthorised data changes are made, possibly leading to a limited choice of de-identification software.Improvements and additions to the framework were identified as providing a clearer explanation of the cataloguing process in 'know the business data'. A set of example documentation such as 'system classification document' would be beneficial. More consideration of the differences between commercial and in-house systems was required along with the impact of cloud and XaaS (anything as a service).

~~The overall impression was that the guidelines could be used by an organisation of any size using non-production environments.~~The overall impression of the reviewers was that the guidelines could be used by an organisation of any size using non-production environments. This impression should be moderated by the understanding that increasing scale and complexity of larger organisations could affect the use of the model. The structure of organisations which have become complex through takeover or merger is expected to be more difficult to assess than those which have been restructured to support growth. The diversity of business interests may also be a factor in complexity with multiple business interests being more challenging to support than limited diversification. One approach to treat the complexity may be to apply the framework initially to subdivisions of the organisation. Another may be to assess and rank the information sensitivity during 'know the system' to prioritise the importance for 'know the environment'.The framework explained the importance of the process clearly, and could be understood by someone with an interest in the protection of sensitive data, but little technical IT knowledge.It related well to both the role and the organisation.

5. Summary and Conclusion

This paper has researched the background for, and produced a framework for an organisation to manage sensitive data in its non-production environments.

Numerous de-identification tools and software, with guidance for usage, are available. In addition, vendor independent advice also exists. Human factors affect data security, with several high profile data breaches being attributed to human error. Exposing sensitive data in non-production environments increases the risk of data breach, simply by the additional availability

Information and Computer Security

The conceptual business model describes the interaction between factors affecting the organisation. Non-production business system environments may be perceived as separate, but are still subject to the same legal and organisational constraints as the production environments even though the system structure may differ. The requirements for environments used at different stages of SDLC are discussed. The framework diagram shows a process flow, with each stage being described in more detail. Factors affecting each stage of the process are explained; with the assurance stage interacting with each of the five stages and having examples of acceptable and best practice. The framework document was reviewed using guided interviews to determine possible benefit to an organisation, and uncover any shortcomings of the framework with a view to addressing them, and to identify further requirements. Generally, the framework was seen as useful, but requires expansion to consider any constraints required by commercial software vendors, and further research in to the effects of cloud services and other 'as a Service' provision.

The paper shows that there is some conflict between security and perceived usability, which may be addressed by challenging the perceptions of usability or identifying the compromise required.Non-production environments need not be the sole responsibility of the IT section, they should be of interest to the business area that is responsible for the data held. The business and IT should cooperate when identifying the use and protection of data, including that held in non-production environments. Ideally this will be done earlier in the acquisition cycle to be included as standard working practice rather than an afterthought.

The framework would benefit from proof of concept including a worked case study in the form of a table top exercise to consider the actual application of the framework. Te development of templates as examples of acceptable and best practice may be of benefit. An extension of the framework to consider if and how software vendors and 'as a service' could constrain the proposed stages may be of benefit.

References

[1] Shabtai, A. et al., *A Survey of Data Leakage Detection and Prevention Solutions* (2012), Springer Article

[2] United States of America Department of Homeland Security (2012) *Privacy Recommendations on the Use of Live Data in Research, Testing, or Training* Available from http://www.dhs.gov/sites/default/files/publications/Privacy_Recommendations_on_the_Use_of_Live_Data.pdf[accessed 23/08/2015]

[3] Akoka, J., Comyn-Wattiau, I., Du Mouza, C., Fadili, H., Lammari, N., Metais, E. and Cherfi, S.S.S., 2014. A Semantic Approach for Semi-Automatic Detection of Sensitive Data. *Information Resources Management Journal (IRMJ)*, *27*(4), pp.23-44

[4] Informatica (2011) *Best Practices for Ensuring Data Privacy in Production and Nonproduction Systems*[Online] Available from: https://www.informatica.com/downloads/6993_Data_Privacy_BestPractices_wp.pdf[Accessed 09/03/2016]

[5] Oracle Corporation (2013) *Oracle Test Data Management Pack*[Online] Available from: http://www.oracle.com/technetwork/oem/pdf/511875.pdf[Accessed 24/11/2015]

[6] Oracle Corporation (2013) *Data Masking Best Practice* [Online] Available from: http://www.oracle.com/us/products/database/data-masking-best-practices-161213.pdf[Accessed 09/03/2016]

[7] IBM (2013) *Data Masking Everywhere*[Online] Available from: https://tdwi.org/~/media/7C7559D0B69741018DAD7446142F2EAA.PDF. [Accessed 09/03/2016]

Information and Computer Security

[8] Delphix (2015) *Reinventing Data Masking*[Online] Available from: http://www.delphix.com/wp-content/uploads/2015/08/delphix-wp-reinventing-data-masking.pdf

[9] Howard, V. (2013) *Using copies of 'live' data in development and testing?*[Online] Available from: https://www.grid-tools.com/using-copies-of-live-data-in-development-and-testing/[Accessed 24/11/2015]

[10] NIST (2010) *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII) Special Publication 800-122*[Online] Available from: http://csrc.nist.gov/publications/nistpubs/800-122/sp800-122.pdf[Accessed 09/03/2016]

[11] Open Data CenterAlliance(2013) *Data Security Framework Rev 1.0*[Online] Available from: http://www.opendatacenteralliance.org/docs/Data_Security_Framework_Rev1.0.pdf[Accessed 09/03/2016]

[12] Duncan, G. Stokes, L. (2009), *Data masking for disclosure limitation*. WIREs Comp Stat, 1: 83–92. doi: 10.1002/wics.3

[13] Wilson, R.L., Rosen. P.A., (2003) Protecting Data through 'Perturbation' Techniques: The Impact on Knowledge Discovery in Databases *Journal of Database Management*, 14(2 Apr-June 2003), pp14-26

[14] Raghunathan, B., (2013) *The Complete Book of Data Anonymization: From Planning to Implementation*. CRC Press.

[15] Cormode, G. and Srivastava, D., (2009) Anonymized data: generation, models, usage. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data* (pp. 1015-1018).ACM.

[16] Gomatam, S., Karr, A.F. and Sanil, A.P.,(2005) Data swapping as a decision problem. *Journal of Official Statistics*, *21*(4), p.635.

[17] Poore, R. S., (2012) Format Preserving Encryption, In: Tipton, H.F., Nozaki, M.K. *Information Security Management Handbook* pp 289-293

[18] Machanavajjhala et al(2008)Privacy: Theory meets Practice on the MapIn: *IEEE 24th International Conference on Data Engineering* Cancun April 2008 pp277–286

[19] Kern,M., (2013). Anonymity: A Formalization of Privacy-l-Diversity. In *Proceeding zum Seminar Future Internet (FI), Innovative Internet Technologien und Mobilkommunikation(IITM) und Autonomous Communication Networks (ACN)* (Vol. 49).

[20] OECD (2008) OECD *Glossary of Statistical Terms*

[21] Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Nordholt, E.S., Spicer, K. and De Wolf, P.P (2012) *Statistical Disclosure Control Glossa*ry pp244-259 Wiley

[22] Sweeney, L. (2002) k-anonymity: a model for protecting privacy, *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, Vol. 10, (No. 5) pp. 557–570.

[23] *Data Protection Act 1998* (c. 29)

[24] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ

[25] PCI (2015) *Payment Application Data Security Standard v3.1*[Online] Available from: https://www.pcisecuritystandards.org/documents/PA-DSS_v3-1.pdf[Accessed 09/03/2016]

[26] ICO (n.d.) *Data protection principles*[Online] Available from: https://ico.org.uk/for-organisations/guide-to-data-protection/data-protection-principles/[Accessed 09/03/2016]

[27] ICO (n.d.) *Key definitions of the Data Protection Act*[Online] Available from: https://ico.org.uk/for-organisations/guide-to-data-protection/key-definitions/[Accessed 09/03/2016]

[28] ICO (n.d.) *Information security (Principle 7)*[Online] Available from: https://ico.org.uk/for-organisations/guide-to-data-protection/principle-7-security/[Accessed 09/03/2016]

Information and Computer Security

[29] ICO (2012) *What is personal data? – A quick reference guide*[Online] Available from: https://ico.org.uk/media/for-organisations/documents/1549/determining_what_is_personal_data_quick_reference_guide.pdf[Accessed 09/03/2016]

[30] ICO (2012) *Determining what is personal data*[Online] Available from: https://ico.org.uk/media/for-organisations/documents/1554/determining-what-is-personal-data.pdf[Accessed 09/03/2016]

[31] Great Britain Department of Health (2013) *The Information Governance Review*

[32] PCI Security Standards Council (n.d.) *PCI Security* [Online] Available from: https://www.pcisecuritystandards.org/security_standards[Accessed 09/03/2016]

[33] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/59

[34] ICO (2016) *The Guide to Data Protection*[Online] Available from: https://ico.org.uk/media/for-organisations/guide-to-data-protection-2-3.pdf[Accessed 09/03/2016]

[35] ICO (n.d.) *Sending personal data outside the European Economic Area (Principle 8)* [Online] Available from:https://ico.org.uk/for-organisations/guide-to-data-protection/principle-8-international/[Accessed 09/03/2016]

[36] European Commission *(n.d.) Commission decisions on the adequacy of the protection of personal data in third countries* Available *from* http://ec.europa.eu/justice/data-protection/international-transfers/adequacy/index_en.htm[Accessed 01/05/2016]

[37] European Commission (2016) Press Release IP/16/216 *EU Commission and United States agree on new framework for transatlantic data flows: EU-US Privacy Shield*

[38] Great Britain Department of Health (1997) *Report on the Review of Patient-Identifiable Information (Caldicott Report)*

[39] PCI (2014) *PIN Security Requirements* [Online] Available from:https://www.pcisecuritystandards.org/documents/PCI_PIN_Security_Requirements_v2__Dec2014_b.pdf[Accessed 09/03/2016]

[40] European Commission (2012) *MEMO/12/41Data protection reform: Frequently asked questions*

[41] European Commission (2016) *Reform of EU Data Protection Rules*[Online] Available from: http://ec.europa.eu/justice/data-protection/reform/index_en.htm[Accessed 07/05/2016]

[42] ICO (2016) *Preparing for the General Data Protection Regulation (GDPR) 12 steps to take now.* Available from https://ico.org.uk/media/for-organisations/documents/1624219/preparing-for-the-gdpr-12-steps.pdf [Accessed 30/05/2016]

[43] Great Britain, Department for Business, Innovation and Skills (2015) *2015 Information Security Breaches Survey* Ref: BIS/15/302

[44] Hurran, C. (2014) Cyber Insiders : A Board Issue*Cyber Security Review*, Summer 2014 , pp63-66.

[45] Silowash, G.J., Cappelli, D.M., Moore, A.P., Trzeciak, R.F., Shimeall, T. and Flynn, L., (2012) *Common sense guide to mitigating insider threats*.4th Edition

[46] Evans, M., Maglaras, L.A., He, Y. and Janicke, H., (2016) *Human Behaviour as an aspect of Cyber Security Assurance*. arXiv preprint arXiv:1601.03921.

[47] Lazić, L. Popovic, S. and Mastorakis, N.(2009), A simultaneous application of combinatorial testing and virtualization as a method for software testing. *WSEAS Transactions on Information Science and Applications*, (Issue 11, Volume 6, November 2009), pp.1802-1813.

[48] Graham, D., Van Veenendaal, E. and Evans, I.,(2008). *Foundations of software testing: ISTQB certification*. Cengage Learning EMEA.

Information and Computer Security

[49] Kim, M., Zimmermann, T., DeLine, R. and Begel, A., (2016), May. The emerging role of data scientists on software development teams.In *Proceedings of the 38th International Conference on Software Engineering* (pp. 96-107).ACM.

[50] Howard, P. (2006) *How your IT department is breaking data protection law*s [Online] The Register. Available from http://www.theregister.co.uk/2006/03/14/unknown_data_protection_breach/ [Accessed 24/11/2015]

[51] Graham, D. and Fewster, M., (2012). *Experiences of test automation: case studies of software test automation*. Addison-Wesley Professional.

[52] Net2000 Ltd., (2010) *Data Masking: What You Need to Know - What You Really Need To Know Before You Begin*[Online] Available from: http://www.datamasker.com/DataMasking_WhatYouNeedToKnow.pdf[Accessed 09/03/2016]

[53] Capgemini (2010) *Five Key Components for Data Masking*[Online] Available from: https://www.uk.capgemini.com/resource-file-access/resource/pdf/Enterprise_Data_Masking.pdf[Accessed 09/03/2016]

[54] Securosis (2012) *Understanding and Selecting Data Masking Solutions: Creating Secure and Useful Data*[Online] Available from: https://securosis.com/assets/library/reports/UnderstandingDataMasking_WhitepaperV2.pdf[Accessed 09/03/2016]

[55] Pringle, T (2014) *Data-as-a-service: the Next Step in the As-a-service Journey*[Online] Available from: http://www.oracle.com/us/solutions/cloud/analyst-report-ovum-daas-2245256.pdf, [Accessed 09/03/2016]

[56] Demirkan, H. and Delen, D., 2013. Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. *Decision Support Systems*, *55*(1), pp.412-421.

[57] *Oxford English Dictionary* (2011) 3rd Ed. Oxford: Oxford University Press

[58] NIST (1989) *Information Management Directions: The Information ChallengeSpecial Publication 500-167*[Online] Available from: http://www.nist.gov/customcf/get_pdf.cfm?pub_id=900419[Accessed 04/03/2016]

[59] NIST (2013) *Glossary of Key information Terms*, NISTIR-7298rev2

[60] Hutchinson, W. and Warren, M.,(2001) *Information Warfare: Corporate attack and defence in a digital world*. Elsevier.

[61] Shaurette, K. AND Schleppenbach, T.J., (2012). a "Zero Trust" model for Security. *Information Security Management Handbook*, *6*, p.175.

[62] Akrivopoulou, C. ed., (2010). *Personal Data Privacy and Protection in a Surveillance Era: Technologies and Practices: Technologies and Practices*. IGI Global.

[63] Great Britain Cabinet Office (2014) *Government Security Classifications*

[64] Photopoulos, C., (2011). *Managing catastrophic loss of sensitive data: a guide for IT and security professionals*. Syngress.

[65] Ward, J., Peppard , J., (2016) *The Strategic Management of Information Systems: Building a Digital Strategy.* Wiley

[66] Griffiths, Dawn, Griffiths, David, (2016) *The Agile Sketchpad: Understanding Agile's Core Concepts and Methods*[DVD] O'Reilly Media

[67] Sherwood, J.,Clark, A., Lynas, D. (2015) *Enterprise Security Architecture: A Business-Driven Approach.* CRC Press.

[68] Deming, W.E., Edwards, D.W. (1982) *Quality, Productivity and Competitive Position* (vol 183) Cambridge MA , Massachusetts Institute of Technology, Centre for Advanced Engineering Study.

[69] Walliman, N.(2010) *Research Methods: The Basics.* Routledge.

Information and Computer Security