HOW TO INCREASE THE INVENTORY EFFICIENCY IN INFORMATION SECURITY RISK AND COMPLIANCE MANAGEMENT

Complete Research

Fenz, Stefan, Vienna University of Technology and SBA Research, stefan.fenz@tuwien.ac.at
Heurix, Johannes, XYLEM Technologies, Vienna, Austria, heurix@xylem-technologies.com
Neubauer, Thomas, Vienna University of Technology, thomas.neubauer@tuwien.ac.at

Abstract

The inventory process, i.e. the assessment of assets and implemented countermeasures, consumes a significant amount of time in the risk and compliance management process. Assets and countermeasures have to be identified and classified in terms of confidentiality, integrity and availability requirements. Depending on the organization's size this process may include thousands of assets and countermeasures. This paper presents a novel inventory approach for assets and already implemented technical, physical, and organizational countermeasures (based on tools for network device mapping, software inventory, asset management, etc.). To efficiently assess implemented organizational countermeasures (policies, guidelines, etc.) we developed a keyword- and rule-based approach which automatically identifies existing policies in the ISO 27002 control context. The method and its implementation support middle and large organizations at efficiently assessing assets and implemented countermeasures by highly automating the inventory process. The method is not bound to any organization type or industry sector.

Keywords: Information security, risk management, inventory

1 Introduction

Regardless of the approach used, the success of information security risk and compliance management activities highly depends on an accurate inventory of (i) assets which should be protected, and (ii) assets which can be seen as control implementations (e.g., countermeasures). Independent of the involved roles (management, technicians, external consultants, etc.), exact asset inventory data is required by these roles to successfully implement the risk management approach. In the context of information security (and this paper) assets, which should be protected, are limited to hardware, software (applications), and data. Control implementations are grouped into physical (locks, windows, etc.), technical (firewalls, malware scanners, etc.), and organizational (policies, guidelines, etc.) measures.

In order to provide methods and tools to increase the inventory efficiency in information security risk and compliance management this paper (i) reviews existing approaches and related work in Section 2, (ii) presents a novel inventory methodology in Section 3, (iii) presents our policy identification and inventory

method in Section 4, and (iv) describes the evaluation results of the developed methods in Section 5. The implementation of the method and the evaluation is conducted in the ISO 27002 context.

2 Existing approaches and related work

This section reviews existing inventory approaches, tools, and related work regarding asset and control implementation inventory in the context of risk and compliance management approaches.

2.1 Tangible Assets

In this paper a tangible asset is defined as an IT component which is of value to the organization. In this sense, tangible assets are grouped in hardware, software, and data.

2.1.1 Hardware

Hardware is a common abstraction layer in risk management as it allows to assess and rate physical items, such as clients, servers, storage devices, smart phones, printers, routers, and laptops. In today's highly interconnected world most of these devices are connected at least to the internal organization network which enables us to use network device mapping techniques to speed up the inventory of hardware assets.

Available network device mapping tools include nMap[1], Lumeta IPsonar[2], Adrem NetCrunch[3], Spice-Works Network Management[4], and Solarwinds Network Topology Mapper[5]. These tools utilize protocols such as SNMP (Case et al., 1989) and support (i) network discovery (network address spaces, routing devices, etc.), (ii) host discovery, and (iii) service discovery (identification of web services, wireless access points, network applications, etc.).

Besides network-based tools, several approaches to support the inventory via RFID (radio-frequency identification)-based methods exist (cf. (Werb and Lanzl, 1998), (Calio, Wyskida, and Frissora, 2011), (Meng et al., 2008), (Mardiasmo et al., 2008), (McKelvin Jr, Williams, and Berry, 2005)). Those approaches are based on RFID tags which are placed on the hardware devices and long-distance RFID readers to identify the tagged hardware assets.

(Nelson et al., 2013) developed an approach for automated asset tracking in data centers by using a vision-based robot with an ability to assess indicator LEDs (light-emitting diode) in servers and storage arrays. Unlike RFID-based methods, this method does not require the tagging of hardware and thereby saves time and costs in the inventory process.

2.1.2 Software

Creating and maintaining a software inventory is a prerequisite for treating software items as corporate assets (Ben-Menachem and Marliss, 2004). Many tools, such as Microsoft Inventory and Assessment Tools[6] or Matrix42 IT Inventory Management[7], are available to support the organization at the software inventory process. LOGINventory[8] is a network inventory tool which collects software and hardware information of all SNMP (Simple Network Management Protocol) capable devices. Due to the integration into Microsoft Management Console and the existing APIs (application programming interface) it works

[1] http://nmap.org/
[2] http://www.lumeta.com/product/ipsonar.html
[3] http://www.adremsoft.com/netcrunch/
[4] http://www.spiceworks.com/app/network-management/
[5] http://www.solarwinds.com/network-topology-mapper.aspx
[6] http://www.microsoft.com/sam/en/ca/invtools.aspx
[7] http://www.matrix42.de/produkte/modules/compliance/compliance/it-inventory-management/ueberblick/
[8] http://www.loginventory.de/

agentless. Another tool to gather required data for asset management and inventory is iET[9]. It works over TCP/IP and makes allows capturing the information from specific divisions by defining the IP (internet protocol) address range. The collected data is stored to a central database from where current data can be analyzed and compared with historical values. Viewing the software configurations remotely, tracking the changes and comparing the results are just some of the included features. A number of tools are equipped with license management often by using a kind of traffic light system to control and monitor the utilization. SMARQ[10] as another example is a real time inventory and asset management tool, which provides in addition to the automated tasks manual supplementing for information not recognized by the tools. The output of these tools can be used as a valuable input for the risk and compliance management process.

2.1.3 Data

The inventory of data is a sub-task of digital asset management. Without a digital asset management system, these valuable assets mostly with no defined structures or identifiers, would get lost. What kind of data is seen as a digital asset is specific to the organization. Valuable data might be customer data, construction plans, audio files, video files, e-mails, digital photographs, or source code. One problem with data inventory is that data across applications and packages is inconsistent regarding to the names and definitions. The aim is to inventory the data people actually need and categorize the output. A clear solution is to develop a catalog of key corporate data and a data dictionary. Depending on the organization-specific definition of digital assets, the inventory process can be supported by simple file searching/indexing tools which run on the organization's hardware components and provide details regarding the type, size, location, and content of the digital assets found. Tools like NogaLogic[11] help to construct a data inventory efficiently by identifying and analyzing the assets.

2.2 Control implementations

In risk and compliance management it is of utmost importance to assess existing control implementations to correctly calculate the risk and to avoid that control implementations are done twice. In information security the control implementation inventory is based on the output of the asset inventory described in the previous subsections. In the following we describe how to filter organizational, technical, and physical control implementations from the generic asset inventory results.

2.2.1 Organizational control implementations

Organizational measures include everything which regulates the processes and activities within an organization. Policies, guidelines, rules, and directives are amongst others organizational control implementations. In most cases these measures are implemented as human-readable text which has to be accepted by everybody in the target group (e.g., development team in the case of secure coding guidelines). Often, the physical document is signed by the recipient and hosted for future look-ups in electronic form at a central access point (e.g., the Intranet or file share of the organization). Search and indexing tools can be used in combination with a predefined set of patterns and key words to automatically identify potential electronically stored policies, guidelines, rules, etc. Section 4 shows our novel approach for automatically identifying potential organizational control implementations within the organization.

2.2.2 Physical control implementations

Physical measures, such as locks, security windows, or guards, include everything which can be utilized to physically protect the assets of the organization. The inventory of physical control implementations is

[9] http://www.iet-solutions.com/de/startseite/
[10] http://www.smarq.com/index.php
[11] http://www.nogacom.com/index.aspx

based on the hardware inventory results and is extended by inventory activities which go beyond the IT components (e.g., human guards or security doors).

2.2.3 Technical control implementations

Technical (logical) measures include everything which cannot be classified as organizational or physical control implementations. Examples are: access control facilities, identification and authorization, IDS (intrusion detection system), smoke detection systems, UPS (uninterruptible power supply), and surveillance cameras. Therefore, the results of the hardware and software inventory may be used to identify technical control implementations.

3 A structured risk and compliance management inventory method

This section provides a structured approach for the inventory process in the context of risk and compliance management activities. Figure 1 shows the generic model for the asset and control implementation inventory described in the following paragraphs.

1. Asset inventory
   a) Hardware: network device mapping tools can be used to assess hardware which is connected to the internal network. Devices which are not connected to the network can be assessed on the basis of existing inventory lists or by questioning the IT staff of the organization.
   b) Software: software inventory tools or vendor-specific infrastructure components, such as an internal Microsoft Windows Update Server, can be used to efficiently assess the software which is installed on the identified hardware components. Special software which is crucial to the company's success (e.g., special production support systems) are normally not recognized by these tools and require a manual identification by the process or business owners.
   c) Data: the kind of data which should be assessed has to be defined by the process or business owners. Based on their specification search and indexing tools can be used to automatically identify data which is of value to the organization. Besides the automated identification it is highly recommended to explicitly approach business and process owners regarding further data which is crucial to the organization.
2. Control implementation inventory
   a) Technical control implementations: parts of the software and hardware inventory results can be used for assessing existing technical control implementations. Suitable keywords (examples): intrusion detection/prevention system, firewall, uninterruptable power supply, identification and authorization system, access control, and encryption software.
   b) Physical control implementations: although a small part of the physical control implementations can be gathered from the hardware inventory results, the main part has to be assessed manually from the appropriate stakeholders (e.g., facility management, guards, security companies, or IT staff).
   c) Organizational control implementations: implemented policies, guidelines, rules, and directives have to be manually assessed by management, project management, or IT staff. The next section describes our novel method for automating the inventory of organizational control implementations.

4 Automated Policy Identification

The goal of the developed method is to increase the efficiency of the inventory process by using a combination of key words and search/indexing tools to automatically identify potential policies, guidelines, rules, or directives which are stored on digital repositories within the organization. Based on the information

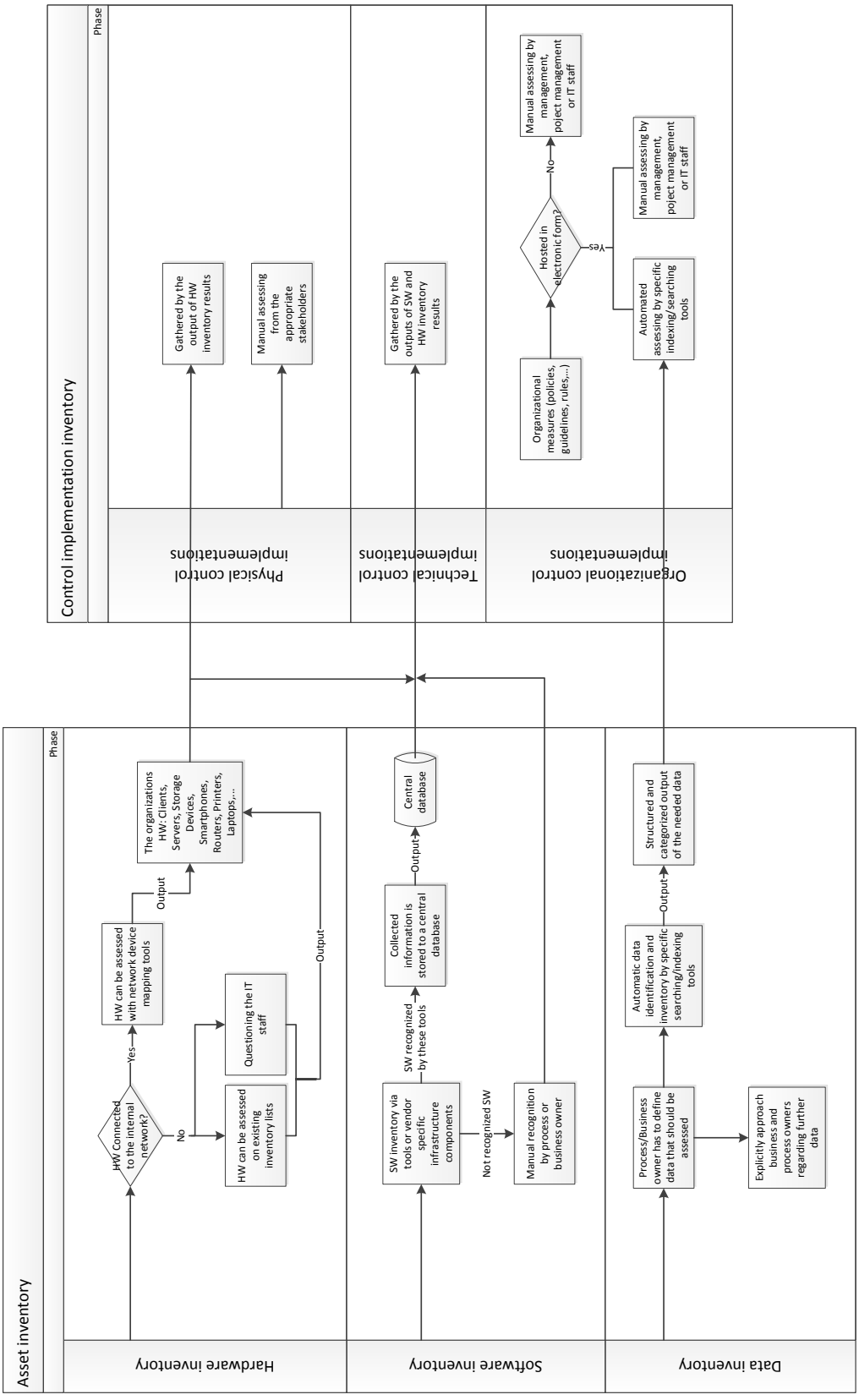

Figure 1. Generic asset and control implementation inventory

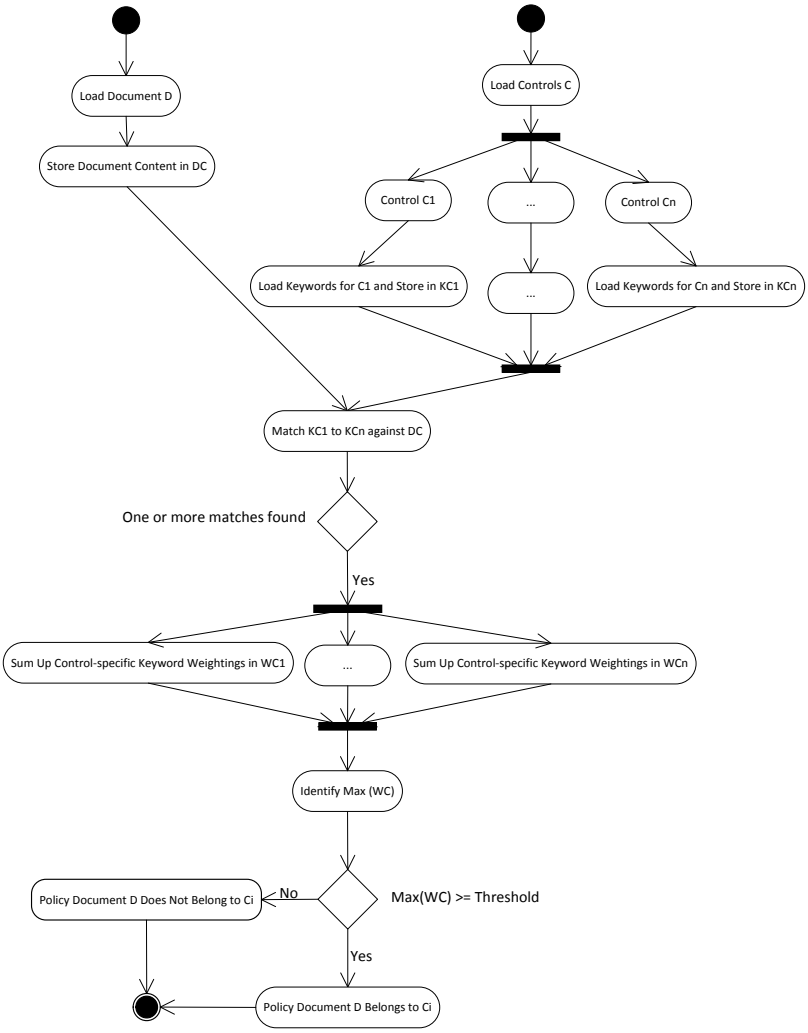

Figure 2. Policy identification process

security standard ISO (International Organization for Standardization) 27001/27002 and its individual controls we developed a set of control-specific key-words/patterns to support the automated detection of organizational control implementations.

While Figure 2 shows the fundamental structure of the proposed method for identifying policies, Table 1 explains the variables which are used in this calculation schema. In the following paragraph the steps from the input to the output are explained in more detail and the calculation schema is mathematically illustrated by considering the given inputs:

1. Load document D from repository
2. Store document content (i.e., its words) in variable DC
3. Load control set C
4. For each control i in control set C
   a) Load keyword and keyword sets from database and store in KCi
   b) Match content of KCi against content of DC
   c) If one or more matches are found

Variable	Description
D	Potential policy document
DC	Content of the document (i.e. its words)
C	Set of ISO 27002 controls
Ci	ISO 27002 control i
KCi	Keywords and keyword sets for ISO 27002 control i
WCi	Sum of keyword weightings for ISO 27002 control i

Table 1. Policy identification model: variables and descriptions

i. Sum-up control-specific keyword weights in WCi (k is the number of keywords)

$$WCi = \sum_{z=1}^{k} [Weight(KCi_z)]$$

d) Identify the maximum value in WC

$$Max(WC) = Max\{WC1, WC2, ..., WCn\}$$

e) If Max(WC) exceeds the defined threshold

i. Policy document D can be assigned to control Ci

4.1 Implementation

This subsection shows in detail the implementation of the method, how the program and especially the user interface works and what the expected results are. The Automated Policy Identification tool can be executed in the command line whereby the executable program is a .jar file. Different parameters used within the program are:

- -k or -keyword [XLS FILE PATH]: specifies the keyword file name and path. An Excel 97-2003 file is required.
- -i or -pdf [PDF FILE PATH]: the file path of the PDF policy repository (all sub directories are included).
- -o or -output [OUTPUT FILE PATH]: an optional parameter which specifies the file path of the output file. Depending on the other parameters the results are either all controls matching with a single document or the controls with the maximum weight regarding to more documents. The output file format is .csv.
- -r or -report [OUTPUT FILE PATH]: works only with single documents. Compared to -o it shows the single keywords matching with the document.

An example for a single document, for example for the "Information Security Policy" from Ruskwig[12] can be constructed like this:

```
java -jar AutomatedPolicyIdentification.jar -k c:\keywords.xls -i
c:\Documents\Information_Security_Policy.pdf -o c:\single_result.csv
```

To get results for multi documents stored in a folder "Documents" the following command is used:

[12] `http://www.ruskwig.com/iso27001/iso_27002_policies.htm`

java -jar AutomatedPolicyIdentification.jar -k c:\keywords.xls -d c:\Documents -o c:\multi_result.csv

With the use of the command for single documents and the output parameter -o the results for the "Security Policy Document" would be structured as follows:

Number of the Control\Control Name\Control description\Weight

In the output file the matching controls are ranked concerning to the weights, which means that the controls at the top fits best with the document and the last ones at least. The output with the parameter -r would be presented as follows:

Number of the Control\Control Name\Implementation guidance\Keywords\Keywords Marked \Weight

In this case, the results are ranked according to the number of controls. Starting with the lowest number where at least one keyword fits with the document. The main difference is that all the keywords fitting with the document are listed.

5 Evaluation

In this section the evaluation results of the method presented in Section 4 are described by the use of a fictive organization. We assume that for our fictive organization the ISO 27002 controls are of interest and used as inputs in the calculation schema.

Every ISO 27002 control has a short description and an implementation guidance. In Control 5.1.1, for example, the definition of an information security policy, how it is structured, as well as responsibilities for the management and references, e.g., to more detailed security policies or procedures concerning this control are determined.

For every control we defined keywords and keyword groups. Concerning Control 5.1.1, which is used as a main example in this section, some keywords and their individual weightings are (amongst others):

- information security policy (100)
- security policy document (100)
- control objectives (85)
- risk management (75)
- risk assessment (75)
- security policy (100)
- management review (75)
- implement ISO 27002 (75)
- business continuity management (60)
- information security definition (90)
- information security scope (90)
- security policy statement (90)
- explanation of security policy (90)
- communicate to all (30)
- policy document (55)

Further keywords for the controls 7.1.3, 8.1.3, 10.5.1, 10.8.5, 11.3.3, 11.4.1, 11.4.6, 12.6.1, 13.1.1 and 14.1.2 are shown in detail in Table 2. Each keyword has a weight (from 0 to 100) specific to its control context. Higher values mean that these keywords are more specific to the control than keywords with lower values. In the above example *information security policy* is a keyword that fits nearly exactly with Control 5.1.1 information security policy document and therefore has a higher range than *management review* or *business continuity management* for example.

For testing we used 19 publicly available policy templates which were downloaded from SANS[13] and Ruskwig[14].

To match existing policies with a special document, for example, with the "Information Security Policy" from Ruskwig, the words have to be identified and stored in a specific structure. Figure 1 in Section 4 shows that the next step is to match the keywords with the words of the document. Keywords which appear in our document are stored with their individual weighting. Depending on the parameters, explained in Section 4.1 the program lists the keywords matching with the document in a determined output file. To name just a few examples, fitting keywords for the document "Information Security Policy" are among others *Security Policy*, *Risk Assessment* and *implement ISO 27002*. Generally there exists a large list of keywords and on average there is normally at least one control with 10 to 30 keywords fitting with the policy document. The aim is to find the control with the greatest number of keywords fitting with the policy.

After storing the matched keywords of one control, their individual weightings can be added and with the help of a threshold (defined by domain experts based on prior test runs), the affinity can be determined. Results greater than the threshold match most likely the policies to this document. For all the smaller results the contrary is true.

To evaluate the results of this example the control ranked highest is 5.1.1 "Information security policy document" with a weight of 4000. Other possible controls according to the program are 13.1.1 "Reporting information security policy" with the weight 3500 or for example 5.1.2 "Review of the information security policy" which achieves a weight of 2300. In the output file all controls, where at least one keyword matches with the words from the policy, are listed.

With regards to the other policies from Ruskwig and SANS, the results with the highest weight for each policy, and so the control which fits best, are listed in Table 3. Our tool not only identified policy documents, it also assigned them to the most relevant ISO 27002 controls. A manual review of the result set has shown that more than 80% of the policies were assigned to the correct ISO 27002 controls.

6 Conclusion and outlook

In this paper, we have presented a generic and structured inventory approach for hardware, software and data assets in the context of information security risk and compliance management activities. Based on this generic approach we developed a method for automatically identifying policy documents in digital repositories and assigning them automatically to the most fitting ISO 27002 control. The evaluation has shown that the developed key word set can be used to achieve a classification result quality which reduces the effort at the inventory process.

Within the evaluation phase we identified the following shortcomings of the developed approach: (i) non-english policy documents were not identified because of the English keywords, (ii) some policy documents where not kept on the central file servers and were therefore not found by our approach (some policies are communicated only via e-mail or are stored on the Intranet portal), (iii) collaborative editing of the keywords was a cumbersome process because participants had to maintain the keywords in a central spreadsheet, (iv) the current spreadsheet approach does not support multilingualism or automated synonym identification (e.g., via Wordnet), and (v) within the spreadsheet there was no possibility to

[13] http://www.sans.org/security-resources/policies/

[14] http://www.ruskwig.com/iso27001/iso_27002_policies.htm

7.1.3 - acceptable use of assets	use of assets	acceptable use	e-mail information	e-mail confidentiality	e-mail use
	use of internet	internet usage	risks with internet usage	mobile devices	bluetooth devices
	inappropriate use	scan e-mails	monitor	phishing	spam
8.1.3 - terms and conditions of employment	ethic	ethics policy	establish a culture of openness	trust	effective ethics
	protect	tolerate	wrongdoing	damaging	knowingly or unknowingly
	ethical practices	unethical behavior	terms and conditions	responsibilities	code of conduct
10.5.1 - information back-up	information backup	backing up information	recover information	recover application systems	backup procedure
	data repository	storage medium	backup storage	data center	backup status
	configuration of the backup	backup level	backup copies	rotation scheme	media failure
10.8.5 - business information systems	voice mail	voice mail	personal communication devices	PCD	sharing of business information
	bluetooth	loss	sensitive information	personal use	safety
	wireless device	voice mail box	theft	personal PCD	interconnecting facilities
11.3.3 - clear desk and clear screen policy	clean desk policy	clean screen policy	confidential information	disclose	locked drawers
	logged off	terminals	keyboard locking mechanism	public information	classified
	password token	malicious entity	protect	mail points	pin code function
11.4.1 - policy on use of network services	network services	security assessment	network	allowed access	network connection
	web application	identify potential weakness	application release	web application security	network service access
	management controls	connection agreement	connection	use of	management procedures
11.4.6 - network connection control	dial-in	dial-in access	dial-in connection	dial-in access policy	network access rights
	connection	corporate network	access control policy	restrictions	network
	using a dial-in connection	connect	use by authorized personnel	assets	scanned
12.6.1 - control of technical vulnerabilities	vulnerability management	system hardening	infrastructure hardening	service release	security patch
	types of patches	patch identification	change management	vulnerability canning	asset tracking
	anti virus process	lab testing	configure firewall	testing	evaluation
13.1.1 - reporting information security events	information security incident	incident response	escalation procedure	incident report	loss of service
	loss of data	unwanted disruption	reporting procedure	how to report	immediately
	malicious incident	accidental incident	security breaches	non compliance	information security event
14.1.2 - business continuity planning framework	information security aspects	human errors	disaster recovery	risk assessment	business continuity
	disaster plan	equipment replacement plan	risk team	critical information	risk assessment process
	business continuity strategy	affect confidentiality	planing process	immediate actions	security risk

Table 2. Example keywords for the used controls

Policy	Control	Control Name	Weight
Email AUP	7.1.3	Acceptable use of assets	3000
Information Backups	10.5.1	Information back-up	2300
Information Security Policy	5.1.1	Information security policy document	4000
Infrastructure Hardening	12.6.1	Control of technical vulnerabilities	2700
Internet AUP	7.1.3	Acceptable use of assets	3100
Reporting Information Security Incidents	13.1.1	Reporting information security events	4600
Secure Extranet AUP	7.1.3	Acceptable use of assets	2400
Technical Vulnerability Patch Management	12.6.1	Control of technical vulnerability	3700
Working In A Foreign Country	7.1.3	Acceptable use of assets	1300
Acquisition Assessment Policy	7.1.3	Acceptable use of assets	1500
Bluetooth Security Policy	7.1.3	Acceptable use of assets	1700
Dial-in Access Policy	11.4.6	Network connection control	1700
Equipment Disposal Policy	9.2.6	Secure disposal or re-use of equipment	1300
Ethics Policy	8.1.3	Terms of conditions of employment	1400
Information Sensitivity Policy	11.3.3	Clear desk and clear screen policy	2800
Internal Lab Security Policy	7.1.3	Acceptable use of assets	2500
Personal Communication Device	10.8.5	Business information systems	1800
Risk Assessment Policy	14.1.2	Business continuity and risk assessment	2000
Web App Security Assessment Policy	11.4.1	Policy on use of network services	2200

Table 3. Result set for all policy documents from Ruskwig and SANS

identify which changes were made by certain users and there was no possibility to discuss changes or suggested changes within the user group (e.g., via comments or a message board).

Further research will address the limitations of the current approach as follows:

Moving from spreadsheets to ontologies We will replace the initial spreadsheet approach with an ontology-based approach to solve the language and the synonym problem. Ontologies allow us (i) to define concepts independent of the used language and (ii) to connect the concepts to further ontologies/taxonomies such as Wordnet to automatically identify suitable synonyms.

Using collaborative ontology editing tools After developing the ontology we will use collaborative ontology editing tools, such as WebProtege, to enable users to extend and modify the keyword set in a collaborative way. WebProtege also provides the possibility of a concept-specific online discussion. I.e., users can discuss the extension and modification of the ontology (e.g., adding new keywords) in detail within the tool.

Support further digital repositories Besides the integration of file servers we will implement an approach which monitors also e-mail repositories and Intranet resources for policy documents. This will be achieved by subscribing relevant e-mail mailing lists and by directly accessing available Intranet resources.

References

Ben-Menachem, M. and G. S. Marliss (2004). "Inventorying information technology systems: supporting the paradigm of change." *Software, IEEE* 21 (5), 34–43.

Calio, B., D. Wyskida, and M. Frissora (2011). "Integrating RFID technology to improve IT asset management controls, playing an integral part in datacenter relocation." In: *Emerging Technologies for a Smarter World (CEWIT), 2011 8th International Conference & Expo on*. IEEE, pp. 1–6.

Case, J., M. Fedor, M. Schoffstall, and C. Davin (1989). *A simple network management protocol (SNMP)*. Network Information Center, SRI International.

Mardiasmo, D., S. Tywoniak, K. Brown, and K. Burgess (2008). "Asset management and governance—An analysis of fleet management process issues in an asset-intensive organization." In: *Infrastructure Systems and Services: Building Networks for a Brighter Future (INFRA), 2008 First International Conference on*. IEEE, pp. 1–6.

McKelvin Jr, M. L., M. L. Williams, and N. M. Berry (2005). "Integrated radio frequency identification and wireless sensor network architecture for automated inventory management and tracking applications." In: *Proceedings of the 2005 Conference on Diversity in Computing*. ACM, pp. 44–47.

Meng, S., W. Chen, G. Liu, S. Wang, and L. Wenyin (2008). "An asset management system based on RFID, WebGIS and SMS." In: *Proceedings of the 2nd international conference on Ubiquitous information management and communication*. ACM, pp. 82–86.

Nelson, J. C., J. Connell, C. Isci, and J. Lenchner (2013). "Data center asset tracking using a mobile robot." In: *Proceedings of the ACM SIGMETRICS/international conference on Measurement and modeling of computer systems*. SIGMETRICS '13. Pittsburgh, PA, USA: ACM, pp. 339–340. ISBN: 978-1-4503-1900-3. DOI: 10.1145/2465529.2466584. URL: http://doi.acm.org/10.1145/2465529.2466584.

Werb, J. and C. Lanzl (1998). "Designing a positioning system for finding things and people indoors." *Spectrum, IEEE* 35 (9), 71–78.