Towards Automated Information Security Governance

Ariane Trammell[1], Benjamin Gehring[1], Marco Isele[2], Yvo Spielmann[3] and Valentin Zahnd[4]
[1]*Zurich University of Applied Sciences ZHAW, Winterthur, Switzerland*
[2]*Stadt Winterthur, Winterthur, Switzerland*
[3]*StepStone Group, Zurich, Switzerland*
[4]*Secuteer GmbH, Zurich, Switzerland*

Keywords: Security Management, Security Controls, Governance Risk and Compliance (GRC), Automation.

Abstract: Securing a company is not an easy task. Many organizations such as NIST, CIS, or ISO offer frameworks that offer comprehensive security measures. However, those frameworks are generally large and require expert knowledge to be tailored to a given organization. Since such experts are rare, we propose an automated solution that selects security controls and prioritizes them according to an organizations need. We performed initial steps towards the implementation of the proposed solution by evaluating how Natural Language Processing can be used to select security controls that are relevant for the assets of a company and by showing that we can prioritize the selected controls based on the current threat landscape. We expect the proposed solution to be a major benefit for all organizations that intend to improve their security posture but are limited in specialized personnel.

1 INTRODUCTION

According to the report on "SMEs and cybercrime" by the European Commission (Commission, 2022), 28% of SMEs have experienced cyber crime in the last 12 months and 70% are concerned about risks associated with cyber security such as hacking, phishing, or malware. There are numerous opportunities for these companies to improve their protection against cyberattacks. In order to achieve systematic improvement, various standards and best practices are available, such as the ISO 27001 standard (ISO, 2022), the NIST Cybersecurity Framework (NIST, 2018), the CIS Critical Security Controls (CIS, 2021) or the BSI IT-Grundschutz-Kompendium (BSI, 2023). These standards all offer good opportunities to sustainably improve a company's IT-security and are also partly integrated into so-called GRC tools (Governance, Risk and Compliance) such as (servicenow, 2023) or (sai360, 2023), which are intended to provide an overview of the security level achieved. However, the standards have a very large number (in some cases more than a thousand) of requirements that are called security controls. For this reason, IT security specialists are required for the implementation, who prioritize the security controls for an organization and also specify how these are to be implemented exactly. However, as can be seen from numerous media reports, it is precisely these security specialists who are currently lacking. This primarily affects small and medium-sized organizations, which have to secure their IT without dedicated security specialists. For them, there are now also some approaches that specifically address SMEs. For example (NIST, 2023) offers a collection of references for small organisations. Some are very broad in the form of a small fact sheet (such as (CISA, 2018)) while others try to customize their advice to the organization with a couple of high level questions (such as (FCC, 2023)). However, those offers still require significant security knowledge to implement the suggested measures as they are not tailored to an individual company or current threat landscape. Additionally, they are not supported by general GRC tools, which mandates a higher involvement of IT security professionals.

This lack of suitable solutions for SMEs combined with the lack of IT-Security professionals implies that often ad hoc security solutions are implemented rather than a standard is followed. The level of security achieved with this method depends heavily on the knowledge and initiative of individual employees. This means that the management level has no

Trammell, A., Gehring, B., Isele, M., Spielmann, Y. and Zahnd, V.
Towards Automated Information Security Governance.
DOI: 10.5220/0012357500003648

In *Proceedings of the 10th International Conference on Information Systems Security and Privacy (ICISSP 2024)*, pages 120-127
ISBN: 978-989-758-683-5; ISSN: 2184-4356

overall view of the state of IT security in their company and can only control it insufficiently. In order to advance the current state of the art, we propose a system that automatically recommends customized security controls to an organization. The focus here is on small and medium-sized organizations that currently cannot afford or do not want dedicated security specialists.

The reminder of this paper is organized as follows. In Section 2 we discuss how our approach differs from related work. In Section 3 we introduce our overall system architecture. In the next two Sections we look into two building blocks of the overall architecture namely security control selection (Section 4) and prioritisation (Section 5). For both building blocks we show implementation possibilities and initial evaluation results. In Section 6 we conclude our paper with a review of the achieved results and the identification of further work.

2 RELATED WORK

In this related work section we look at research projects that work towards an automated selection or prioritization of security controls. CIAM (Llansó, 2012) is an approach in which security controls are given weights in various categories by IT security specialists. Based on these weights, the controls are then prioritized. This inadequacy is taken into account in (La, 2023), where the weighting of the controls is made on the basis of real attacks. Here, the focus is on the technical security controls and not on those of processes. It is also not possible to adapt the scope of the controls to the size or sector of the corresponding organization. Thus controls might not be applicable if the affected infrastructure is not used, or important controls might be missing for an organization that uses specialized infrastructure.

Other approaches such as (Barnard and von Solms, 2000), (Yevseyeva et al., 2016) and (Neubauer et al., 2008) try to find the relevant controls for a project or IT-asset from a standard such as ISO or NIST. In all approaches, a lot of security expertise is still needed, for example, to capture security requirements, set weights, or weigh possible solutions against each other. Our system, on the other hand, is designed to suggest and prioritize relevant security controls without the need for IT-security experts. For this purpose, information from various sources is combined to automatically create a digital twin of the organization. This data is then used to select the controls that are relevant for an organization and in a second step those controls are prioritized.

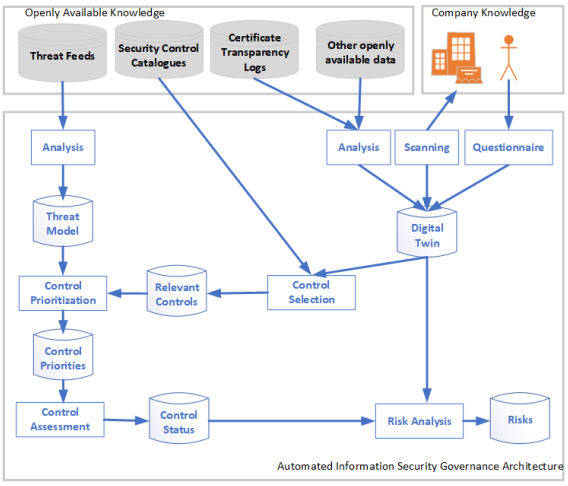


Figure 1: Overall System Architecture.

3 OVERALL SYSTEM ARCHITECTURE

To build a tool that supports SMEs with the selection and prioritization of appropriate security controls, we propose the architecture shown in Figure 1.

The basic idea is to combine openly available data and company internal knowledge to automatically select and prioritize security controls.

Security Control Selection. In order to select controls that are relevant for an organization, as much as possible has to be learned about this organization. To this end openly available external information is combined with company specific knowledge to build a digital twin of that organization. The information analyzed might range from certificate transparency logs (Transparency, 2023) to gain an insight into the web presence of a company over questionnaires to network scans of the organization where applicable. The control selection process uses the information gathered in the digital twin to select relevant controls from established security control catalogs such as provided by NIST, CIS, BSI, etc.

Security Control Prioritization. The prioritization of security controls is based on the relevant security control selection and information specifying the current threat landscape. This information could be obtained by a human assessment or automated by analysing threat feeds.

Risk Assessment. In order to be able to assess the overall risk related to cyber security of an organization, the implementation status of the security controls needs to be assessed. This assessment should be done as far as possibly automatically, however, this is not for all controls possible. For example, it is

comparatively easy to automatically assess whether all client devices are patched, but it is difficult to assess whether an appropriate business continuity plan is in place. The final risk assessment combines the implementation assessment of the individual security controls with the data available in the digital twin.

A dashboard containing the questionnaire, the prioritized controls, a control assessment possibility as well as the resulting risks allows the responsible personnel to plan security measures systematically and to see the effects in the resulting risks.

In the next two sections we look into the control selection and control prioritization in more detail.

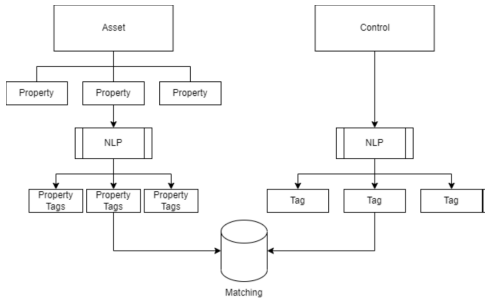
4 CONTROL SELECTION

One approach for the selection of appropriate controls is to make a mapping between IT-assets such as hardware or software components and security controls. This mapping has the advantage that the compliance of every IT-asset can be assessed individually. For many controls this is beneficial, since a control might only be relevant to a subset of the IT-assets and from that subset it might only be implemented by another subset. For example, a control that asks for a timely patching of devices might be implemented for client devices but not for IoT devices.

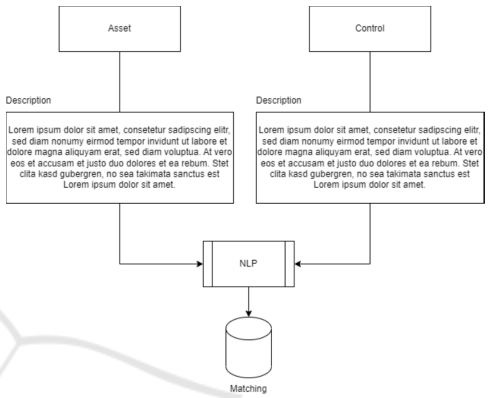
We explored *natural language processing (NLP)* as a mechanism to perform the matching of IT-assets to security controls. Since both IT-assets and security controls are described by natural language, NLP is an adequate mechanism to perform this mapping. We explored two different approaches, a *tag-based approach* and a *description-based approach*.

Tag-Based Approach. In the tag-based approach, NLP is used to generate tags from the control and asset descriptions. If a tag of a control matches a tag from an asset the control is applicable for that asset. The process to obtain tags differs slightly for controls and for assets, as assets are typically described by a set of properties but no long verbal description whereas controls typically have a longer verbal description. The method is shown in Figure 2a.

Description-Based Approach. The description-based approach directly compares the description of an asset with the description of a control. As assets generally doe not offer a long description, this description has to be obtained in an additional step, such as looking up the corresponding product description in the Internet. The matching is performed by NLP to extract the similarities between descriptions of assets and the controls. If a particular similarity value is reached, a security control is applicable for the given asset. This approach is shown in Figure 2b.



(a) Tag based matching.



(b) Description based matching.

Figure 2: NLP Based Matching Methods.

While the tag-based approach uses keywords to match controls and assets, the description-based approach uses the descriptions themselves. Both approaches have in common that they rely on an asset inventory that contains all IT-assets relevant for an organization. If such an inventory is not available, it would have to be compiled as a first step. In order to automate this process, techniques such as network scanning or accessing information from software management tools could be used.

We implemented and evaluated both approaches with the security controls provided by the NIST Framework (NIST, 2018) and the asset inventory of our organization.

4.1 Implementation of Tag-Based Approach

The tag-based approach follows the following steps: a) keyword extraction, b) clustering, and c) topic modeling. Each of those steps is defined in the following paragraphs.

Keyword Extraction. The goal of the keyword extraction phase is to find keywords that describe a

control and its usability well. In our implementation, the keyword extraction for controls is based on the control title, the control text as well as the discussion of each control. The keyword extraction for the IT-assets is based on the title as well as the properties of each asset.

We analyzed five different libraries for the keyword extraction. These are KeyBERT (KeyBERT, 2023), RAKE (rake nltk, 2023), spaCy (spacy, 2023), Textrazor (textrazor, 2023), and Yake (yake, 2023). Our evaluation showed that in our scenario Yake performs best overall in the extraction of keywords.

Clustering. The second step in the analysis process is clustering. The idea here is to combine similar keywords into a cluster to be able to create tags from them. As input, the cluster function receives all keywords and the references to the associated security controls. The KMeans algorithm was selected for clustering. This is because it gives good results, performs well, and is one of the most used algorithms for clustering. Furthermore, with KMeans, it is also possible to cluster words or sentences. The cluster size was set to 20% for the (control) tags and also 20% for the property tags. In addition, the tolerance value for cluster cleaning was set relatively high at 600% for the keywords and 800% for the controls in order to eliminate only the outliers and not catch too many of the somewhat larger clusters

Topic Modeling. With the help of topic modeling, a descriptive term should be found for the clusters. In the case of topic modeling, the LDA (Blei et al., 2003) algorithm can be used, which searches for a suitable topic for a given text.

4.2 Implementation of Description-Based Approach

In addition to the tag-based approach we evaluated the description based approach shown in Figure 2b. We used the well established TF-IDF (Rajaraman et al., 2014) value to analyze the descriptions of the controls and the IT-assets. Using the TfidfVectorizer from the Sklearn (scikit, 2023) library, a vector is created which contains the TF-IDF value for each word of a description. If two descriptions have the same words, the vectors have the same value at this point and are therefore similar to a certain degree. In order to receive a unique value for the similarity, the Cosine Similarity between the vectors is computed. The resulting value is used to determine whether a control matches a given asset.

4.3 Evaluation

We evaluated the implemented algorithms on the NIST security controls and a selection of 200 IT-assets from our organizations IT-asset inventory.

The following section presents the results of the implementation tests. It is divided into two sections, where on the one hand, the results with the tag-based approach are presented and, on the other hand, the results with the description-based approach.

Tag-Based Matching. Table 1 shows general statistics obtained with the tag-based approach described in Section 4.1. A total of 323 out of 1235 unique controls were recognized and assigned to the captured assets. Some of the controls have been assigned to multiple assets. Looking at the total number of matches, the number of controls matched increases to 5599. This means that, on average, each control that was matched was assigned about 17 times, and an asset has an average of 28 controls assigned to it. Of the 200 assets, 178 have received at least one control, which means no controls were assigned to 22 assets.

Table 1: Mapping based on tags.

Description	Value
Total matched unique controls	323
Total matched controls	5599
Number of assets with controls	178
Number of assets without controls	22

A manual verification was also performed to check if the assigned controls matched the assets. Since the amount of data is tremendous, we only evaluated 10 randomly selected assets. Table 2 shows the results of this verification. A total of 413 assignments were verified. Of these 413 assignments, 79 were identified as incorrect. This means that the assigned control doesn't match the asset. Thus, 334 assignments were correct in this subset. Thus, the application executed about 80% of the assignments correctly in the analysis with the tag-based approach.

Description-Based Matching. Table 3 shows an overview of the data from the tests with the description-based approach.

A total of 370 unique controls were matched. This represents a share of 28% of the total of 1235 available controls. The total number of matches is 1981. On average, each control is assigned to 6 different assets. 159 assets have received at least one control. So 41 assets have not been assigned a single control.

We also performed the manual analysis shown in Table 4. For the description-based matching, 10 assets were selected and manually checked. As far as possible, the same assets were selected as for the tag-based matching in order to obtain a possibility of

Table 2: Manual Evaluation of Tag Based Mapping.

Asset	Total assignments	Incorrect assignments
Google Drive [Web]	111	26
Cisco AnyConnect VPN Client	29	3
Microsoft Teams	7	0
Microsoft OneDrive for Business 2016	120	32
RV-APP-T-202	11	2
Citrix ICA Client	27	0
Microsoft Local Administrator Password Solution 6	35	5
Microsoft Windows 10 Enterprise	10	0
Microsoft SharePoint Online [Web]	29	0
VideoLAN VLC Media Player	34	11
Total	413	79

comparison. If no control was assigned to an asset, it was replaced by another, random one. In total, 137 assignments were recorded. Of these, 22 were identified as incorrect. In total, 115 assignments were classified as correct, which corresponds to a share of 83%.

Table 3: Mapping based on description.

Description	Value
Total matched unique controls	370
Total matched controls	2488
Number of assets with controls	159
Number of assets without controls	41

Comparison of Tag-Based and Description-Based Approach. To assess the quality of the results, we compared the tag-based and the description-based approach. As shown in Table 1 and Table 3 many of the controls identified by the tag-based matching were not found with description-based matching. With a total of 2488 matches, the description-based matching generally only identified half as many controls as the tag-based approach. This circumstance is also reflected in the detailed analysis, where many controls that should be matched are not identified. However, if a suitable control is found, this assignment is usually correct, as Table 4 shows. On the other hand, tag-based matching was also unable to find some controls that would actually be relevant. For example, for the asset "Cisco AnyConnect VPN Client", the "Remote Access" control was only assigned by description-based matching, but not by tag-based matching. Both, description-based matching and tag-based matching,

Table 4: Manual Evaluation of Description Based Mapping.

Asset	Total assignments	Incorrect assignments
Google Drive [Web]	3	0
Cisco AnyConnect VPN Client	32	2
Google Gmail [Web]	60	10
Microsoft OneDrive for Business 2016	3	0
SRV-APP-T-202	5	0
Citrix ICA Client	8	0
Microsoft Local Administrator Password Solution 6	3	0
Microsoft Windows 10 Enterprise	2	0
VideoLAN VLC Media Player	19	7
Microsoft Windows Media Player 12	12	3
Total	137	22

can identify some correct controls that are not found by the other algorithm. Therefore, further analysis should be performed to identify the number of applicable controls that were not identified by either of the algorithms.

Additionally, both methods are limited by the fact that only controls that are directly relevant to IT-assets are matched. But the control libraries contain additional controls that are not applicable to IT-Assets. Examples for such controls are controls that mandate processes or concepts. Since those controls are not bound to an IT-assets, they would not be selected by a corresponding algorithm. Therefore, the asset based approach has to be combined with further approaches such as an approach where controls are mapped to enterprise roles such as *network administrator*, *HR*, or *CISO* or a questionnaire that covers important gaps.

4.4 Discussion of NLP Based Control Selection

We showed that NLP can be used for the identification of security controls that are relevant to IT-assets. In this section we highlight some of the challenges.

No Ground Truth. As we did not have a known, verified mapping of security controls to IT-assets we had to perform the validation manually. This validation is time consuming and error prone and to a certain degree subjective. A validated ground truth would allow for a better evaluation of the results.

Parameter Selection. The quality of the results heavily depends on the chosen algorithms and their

configurations. If a ground truth would be available, several different parameters could be tested against that ground truth and the best parameters could be chosen. Without that ground truth the parameter selection was performed manually.

Short Asset Description. While the security control descriptions contain an adequate amount of text to be used with NLP, this is not the case for the asset description in a traditional asset inventory. Here the asset is only described with a few key words. To reach a higher quality in the selection of appropriate security controls, a more detailed description, including the value the asset presents to the organization, would be desirable.

In addition to NLP mechanisms, we could also think of using different approaches for the selection of the relevant security controls. Some possibilities include: *Deep Learning.* This would require a data set containing assets that already have assigned security controls. This data could then be used to create a neural network that establishes a relationship between certain assets and the controls. However, a large amount of data is required for this model. Furthermore, it is questionable how large the influence of subjectivity and the variation of asset specifications is on the quality of the model. The work of (Bettaieb et al., 2020) goes in this direction. Based on historical data, they try to find the appropriate controls for a new system using machine learning. The researchers achieved a precision value of 93%. However, the assignments are based on security requirements, which must first be determined.

Linguistic Rules. Another approach is to focus more on linguistic rules, as is done in (Li, 2017). Here, an attempt is made to identify security requirements employing linguistic rules and machine learning. The security requirements are divided into different categories such as ”threat-based”, ”asset-based”, etc., whereby a different structure of the linguistic rules is defined for each category. This approach could be used to determine security requirements. These are then used to identify potentially suitable controls. Machine learning could also be used for this step, or the model could be extended with the linguistic rules.

Large Language Models. Instead of using basic NLP methods pre-trained large language models could be used. Those have the benefit that we would not have to perform the parameter selection but that those are already tuned to match a large set of text.

Role Based Approach Instead of Asset Based Approach. An approach focusing on roles in a company, such as *network administrator*, *HR*, *CISO*, etc. could be used in addition to the approach focusing on IT-assets. Enterprise roles are generally described in plain text and contain all the responsibilities of a given role. The mapping could be performed based on the role description and the control description. Unlike the IT-asset based mechanism, this approach is applicable to all controls, since each control has to have someone that is responsible for its implementation. It would also facilitate the overall security governance, since the responsible role is clearly defined. With a mapping of roles to actual personnel, it would be possible to add contact details of a responsible person to controls.

5 CONTROL PRIORITIZATION

While the last section focused on the selection of appropriate controls, this section focuses on the prioritization of controls. The basic assumption is that a company has only limited resources and cannot implement all relevant controls simultaneously. Therefore, the controls that bring the most benefit for the security of the organization should be implemented first. In this chapter we describe one possibility on how to perform this prioritization. For this study, we used the BSI IT-Grundschutzkompendium (BSI, 2023). The IT-Grundschutzkompendium is a collection of security controls that have a set of attributes such as a role that is responsible for that security control or elementary risks that it is mitigating. Those risks range from natural hazards over misconfigurations to attacks or loss of qualified staff. Each security control is assigned to one of three levels (basic, standard, enhanced) which describes for which protection needs the control is required.

The idea is to evaluate the current threat landscape and use this input to prioritize the security controls.

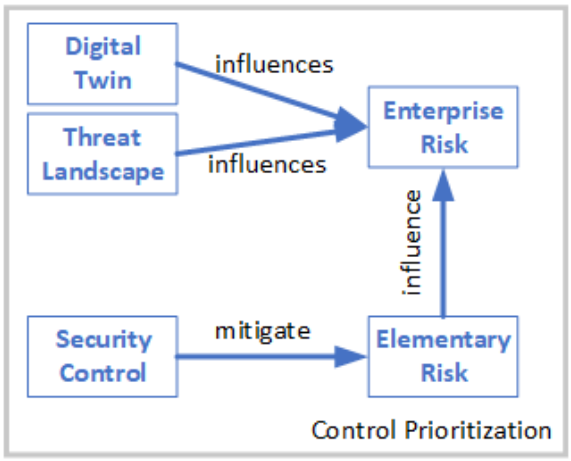


Figure 3: Control Prioritization Model.

As shown in Figure 3 on the one hand, the threat landscape influences the risk of a company and on the other hand the security controls mitigate the elementary risks. The missing link is between the elementary risks and the enterprise risks. We closed this link by combining elementary risks to enterprise risks. The enterprise risks can now be weighted either manually or automatically using a combination of threat feeds and the description available in the digital twin.

5.1 Risk Assessment

In this implementation we choose the manual risk assessment method. The responsible personnel assigns a risk score between 0 and 100 to each of the enterprise risks.

5.2 Control Prioritization

In order to calculate the priority of a control, its associated risk score is calculated. As shown in Figure 4, the risk score of a control is defined as the sum of the risk scores of all enterprise risks that it mitigates multiplied by the *importance* of that control.

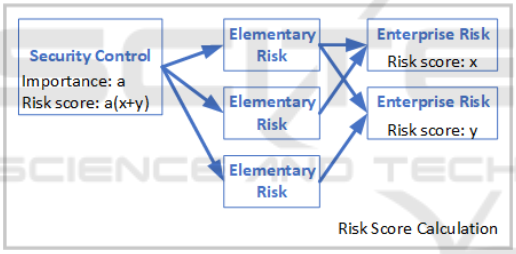


Figure 4: Risk Score Calculation.

The importance of a control is determined by its classification in the BSI IT-Grundschutzkompendium. The corresponding values are shown in Table 5.

Table 5: Value for the *importance* of a security control.

Description	Value
Basic	6
Standard	4
Enhanced	2

For example if a security control is associated with three enterprise risks which have risk score of 24, 42 and 51 and the control is classified as *standard*, the risk score is $\mathbf{4} * (\mathbf{24} + \mathbf{42} + \mathbf{51}) = \mathbf{468}$.

The risk score is then used to prioritize the security controls. The use of the *importance* ensures that more basic controls are always prioritized higher than more advanced controls and the sum of the enterprise risk scores ensures that controls that mitigate more likely risks are prioritized higher.

5.3 Risk Assessment

In order to complete the security management cycle, the current risks resulting from cyber threats should be assessed. To this end the implementation status of the controls need to be entered. In our basic implementation, a binary form (implemented/not implemented) is chosen that needs to be specified manually. More elaborate versions could either assess the implementation state of a control automatically or offer additional states such as partially implemented or implemented for a fraction of the IT-assets.

For each enterprise risk it the fraction of mitigating controls implemented over the totally assigned controls is calculated. This calculation could be enhanced by taking the weights (e.g., the *importance* of a control, respectively the risc score of an enterprise risk) in account.

Figure 5 shows the resulting dashboard.

Figure 5: Dashboard Security Control Prioritization.

6 CONCLUSION

We presented an architecture for a tool that could support small and medium sized organization with their security maturity. The architecture builds on openly available data such as established security control catalogues as well as open source threat information for risk assessment. In order to select the security controls that are relevant for a given organisation as much information as possible is collected by combining openly available data as well as company internal data such as questionnaires, network scans or existing assets inventories.

In the remainder of the paper we evaluated methods for two aspects of a corresponding tool: **control selection** and **control prioritization**. We showed that we could automate the selection of controls by NLP mechanisms. However, future work should analyse whether the selected controls are complete or whether some controls were missing, and the number of erroneously selected controls should be further reduced. Additionally, mechanisms should be developed that are applicable to the controls that are not covered by the IT-asset based approach.

For the control prioritization we showed that controls can be prioritized based on the currently observed threats. In our implementation we used a manual threat assessment, this work could be continued to automate the treat assessment when appropriate threat intelligence information is available.

In order to obtain a tool that covers all aspects of the proposed architecture, the two elements control selection and control prioritization need to be integrated and additional topics need to be addressed. This includes the development of a digital twin of an organization that is more specific than an asset inventory, and the development of possibilities that automatically assess whether a control is implemented.

We expect that a tool implementing all building blocks of the architecture would provide a significant step forward in supporting small and medium sized organization with their efforts towards securing their organization.

REFERENCES

Barnard, L. and von Solms, R. (2000). A formalized approach to the effective selection and evaluation of information security controls. *Computers and Security*, 19(2):185–194.

Bettaieb, S., Shin, S., Sabetzadeh, M., Briand, L., Garceau, M., and Meyers, A. (2020). Using machine learning to assist with the selection of security controls during security assessment. *Empirical Software Engineering*, 25.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3(null):993–1022.

BSI (2023). Bsi it-grundschutz kompendium. https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/IT-Grundschutz/IT-Grundschutz-Kompendium/it-grundschutz-kompendium_node.html. Accessed: 2023-10-19.

CIS (2021). Cis critical security controls. https://www.cisecurity.org/controls. Accessed: 2023-10-19.

CISA (2018). Cybersecurity resources road map - a guide for critical infrastructure - small and midsize businesses. https://www.cisa.gov/sites/default/files/c3vp/smb/DHS-SMB-Road-Map.pdf. Accessed: 2023-10-19.

Commission, E. (2022). Flash eurobarometer 496: Smes and cybercrime. https://europa.eu/eurobarometer/surveys/detail/2280. Accessed: 2023-10-19.

FCC (2023). Cyberplanner. https://www.fcc.gov/cyberplanner. Accessed: 2023-10-19.

ISO (2022). Iso/iec 27001:2022 - information security, cybersecurity and privacy protection. https://www.iso.org/standard/27001. Accessed: 2023-10-19.

KeyBERT (2023). Keybert. https://maartengr.github.io/KeyBERT/index.html. Accessed: 2023-10-19.

La, S. (2023). Prioritizing cybersecurity controls based on the coverage of attack techniques and attack probabilities. Technical report, ETHZ.

Li, T. (2017). Identifying security requirements based on linguistic analysis and machine learning. In *2017 24th Asia-Pacific Software Engineering Conference (APSEC)*, pages 388–397.

Llansó, T. (2012). Ciam: A data-driven approach for selecting and prioritizing security controls. In *2012 IEEE International Systems Conference SysCon 2012*, pages 1–8.

Neubauer, T., Ekelhart, A., and Fenz, S. (2008). Interactive selection of iso 27001 controls under multiple objectives. In Jajodia, S., Samarati, P., and Cimato, S., editors, *Proceedings of The Ifip Tc 11 23rd International Information Security Conference*, pages 477–492, Boston, MA. Springer US.

NIST (2018). Nist cyber security framework. https://www.nist.gov/cyberframework. Accessed: 2023-10-19.

NIST (2023). Small business cybersecurity corner. https://www.nist.gov/itl/smallbusinesscyber. Accessed: 2023-10-19.

Rajaraman, A., Leskovec, J., and Ullman, J. (2014). *Mining of Massive Datasets*.

rake nltk (2023). rake-nltk. https://csurfer.github.io/rake-nltk/_build/html/index.html. Accessed: 2023-10-19.

sai360 (2023). sai360 - an integrated approach to governance, risk and compliance. https://www.sai360.com/solutions/integrated-grc. Accessed: 2023-10-19.

scikit (2023). scikit. https://scikit-learn.org/stable/. Accessed: 2023-10-19.

servicenow (2023). servicenow - governance, risk and compliance. https://www.servicenow.com/products/governance-risk-and-compliance.html. Accessed: 2023-10-19.

spacy (2023). spacy. https://spacy.io/. Accessed: 2023-10-19.

textrazor (2023). textrazor. https://www.textrazor.com/. Accessed: 2023-10-19.

Transparency, C. (2023). Certificate trasparecy. https://certificate.transparency.dev/howctworks/. Accessed: 2023-10-19.

yake (2023). yake. http://yake.inesctec.pt/. Accessed: 2023-10-19.

Yevseyeva, I., Fernandes, V. B., van Moorsel, A., Janicke, H., and Emmerich, M. (2016). Two-stage security controls selection. *Procedia Computer Science*, 100:971–978. CENTERIS/ProjMAN / HCist 2016.