Comparable and Repeatable Information Security Level Evaluation

Mari Seeba[1]

[1]University of Tartu, Narva mnt 18, Tartu, 51009, Estonia

Abstract

To safeguard citizens' digital lifestyles and the functioning of societal systems, countries enact regulations (e.g., GDPR, NIS2) mandating cybersecurity measures in organisations to improve security. We must repeatedly evaluate the improvement rate in organisations and collect the data for a state-level overview to measure the improvement rate over time. There are developed instruments to assess or measure security, but they lack of best practices for evaluating compliance in a way that considers environmental changes while ensuring consistent security evaluation over time and across organisations (e.g., benchmarking) simultaneously. This PhD project introduction paper introduces the artifact - a framework for security level evaluation (F4SLE) in organisations based on chosen baseline standards with the method to update the instrument content and its user stories, utilising the design science research method. The F4SLE is used in piloting experiments by 70 organisations in Estonia and South Moravia (a district of the Czech Republic) to validate the framework and its user stories. The final results are a work in progress.

Keywords

security standards, security evaluation, framework for security level evaluation, security assessment, organisation

1. Introduction

Our contemporary way of life relies heavily on digital solutions, making us susceptible to availability, integrity, and confidentiality vulnerabilities. These threats elude the awareness of ordinary users. Consequently, legal regulations, such as NIS2 [1] and GDPR [2] of the European Union, aim to safeguard data subjects, fortify our digitised lifestyle, and place the onus on organisations to implement security controls.

Based on regulations in the European Union, member states should establish their security requirements and standards for organisations. However, there are no direct rules for choosing the standard. Some security management system standards like ISO27002 [3] NIST CSF [4] guide organisations in ensuring adequate cyber resilience for digitalised systems and networks. Some member states have developed their approaches like the full framework with national standards and security controls catalogue in German BSI IT Grundshutz Kompendium [5]; Estonian Information Security framework (E-ITS) includes security measures or Latvian regulation of procedures for Ensuring Compliance of Information and Communication Technology Systems with Minimum Security Requirements [6].

Evaluating the achievement of security goals must be broad and cover the entire organisation and its processes. Mohebbi et al. [7] propose two resilience strategies: implementing robust security countermeasures in layers and preparing the whole organisation for unexpected risks through incident management training and exercising. Therefore, a comprehensive implementation and evaluation of essential security countermeasures is imperative for the organisation, its partners, and clients to formulate a robust cyber resilience strategy. Organisations can confirm compliance with security requirements through various methods, including third-party audits, internal monitoring, and post-analysis of incidents. These methods are either resource-intensive or create insurmountable assumptions but still do not reflect the comprehensive overview. Also, these methods do not provide aggregated views of the information security situation to the state-level policy-makers. However, such methods are unsuitable for an organisation planning to implement an information security management system.

CAiSE 2024 Doctoral Consortium

mari.seeba@ut.ee (M. Seeba)

0000-0002-9066-2467 (M. Seeba)

At the same time, state-level policymakers also need evaluations of the situation and the organisation's input. In addition to reports based on technical scans and incidents (e.g. CERT-EU Threat Landscape report [8]), information about the security situation at the national level is collected during organisation surveys, statistical studies and interviews (e.g., [9, 10, 11, 12]). The result is that organisations are asked similar questions from different institutions but do not get feedback or immediately applicable benefits for their situation with such studies.

While different security assessment and evaluation instruments are available [13, 14, 15], there is a gap for a standardised, low-entrance-barrier security evaluation framework, which is regularly updated, and the evaluation results are comparable over time and by sectoral considering changes in the cybersecurity field, but also support policymakers with collected data.

This PhD project focuses on a solution to improve the information security level of organisations across the entire country, e.g., Estonia. The project is motivated by the need to develop and implement national information security standard for the organisations of Estonia. Therefore, Estonian cases and examples are used. The validation is conducted simultaneously in Estonia and the South Moravia district of the Czech Republic to generalise the artifact usage outside Estonia.

This paper's structure covers the PhD project problem statement in Section 2, related work in Section 3, the introduction of the use of design science research method and artifacts in Section 4, validity threats of the developed artifact in Section 5 and the project's conclusion in Section 6.

2. Problem Statement

If governments have identified a security management standard for their country's organisations (E-ITS [16] in Estonia), evaluating the effectiveness of these standards or baselines necessitates an ongoing evaluation of the security level. A single evaluation does not show the changes in security level and adaptation to the dynamic nature of security landscapes. We must reflect on these changes in security-level evaluation instruments as security methods and threat landscapes evolve. It is essential to avoid creating a false sense of security by evaluating outdated security measures that may no longer provide security. We also need horizontal evaluation for benchmarking with other organisations/sectors and country-level data collection. Hence, there is a need for a security-level evaluation instrument that is repeatable, adaptive, and updatable, ensuring the generation of consistent and comparable results over time. In addition, it is necessary to understand the context of the security evaluation and the associated user stories required by regulation NIS2 [1].

Based on that, here are the research question (RQ) and its subquestions (SRQ) for this PhD project: **RQ:** *How to evaluate the change in information security level in and across organisations?* And its subquestions: **SRQ1:** What are the requirements to develop information security management standards for public sector organisations on the national level? **SRQ2:** How to evaluate an organisation's information security level? **SRQ3:** How can the security evaluation instrument attributes be updated so that evaluation results remain consistent and comparable over time? **SRQ4:** What are the user stories for security-level evaluation instruments usage and results' interpretation?

3. Related Work

In the past decade, systematic literature review articles on security evaluation instruments have mentioned more than 100 instruments [15, 13, 17, 14]. The main disadvantage of evaluation instruments is their short lifespans, which, for example, are directly related to the lifespan of research projects. Leszczyna [13] in his review addresses the gaps like missing real-world application and evaluation methods of instruments, missing real supporting tools, insufficient documentation for implementation, undefined target users, assessment incompleteness, and the issue of no updating methods being provided for continuous maintenance. Spruit et al. [18] introduce its instrument but highlight concerns and uncertainties about the instrument model's long-term viability and applicability in the face of evolving information technology developments, organisational diversity, and sector-specific challenges. Also, De

Bruin [19] provides specific suggestions for creating maturity models in security but does not manage updating and maintenance issues.

The NIS2 directive [1], which member states in EU must transpose by October 2024, requires member states to be able to ensure, and therefore evaluate, the status and progress of organisations' security implementation. Countries act independently to create security evaluation instruments (some examples: Finnish [20], Italian [21], Greek [22], Spanish[23], Austria [24], Czech [25]). The separate creation of instruments is motivated by a security risk or compliance with national legislation or standards; for example, in Estonia, the Public Information Act restricts the release of detailed security data from the institution or the expectation to support a national security framework. In the case of the instruments of the countries given as examples, we see that surveys are updated based on which superior reports are put together annually [25]. The renewal methods are not clearly described for other instruments[22], or the instrument has already been used unchanged for more than three years [21], or there is no workable method to collect data in the centre [20, 21, 24, 23].

Data on the security situation can also be collected at the country level. For example, Global Cybersecurity Index[26] or newly created European Cybersecurity Index EU-CSI [27] and NCSI [28] metrics are used as security evaluation instruments. The uniqueness of the NCSI [28] methodology compared to other metrics is its open raw data as a collection of links. These links allows dig for details of the security of the assessed country (e.g. standards, cyber security strategies, regulations). However, when assessing organisations' security, these metrics rely on secondary sources, such as the presence of security certificates or other measures of organisational security. National statistics offices or national audit institutions also collect data on security by specific issues (e.g., [10, 9, 11]). However, it does not provide a complete overview of the implementation of risk management measures. The disadvantage of all general-level security evaluation instruments from the organisation's point of view is the lack of (immediate) feedback on the organisation's information security situation and the option to plan improvements.

There is one root source in the social sciences – Oppenheim's work [29], which, according to Google Scholar, has been cited more than 10,000 times in the creation of assessment instruments or surveys. This study supports the creation of all kinds of metrics in all areas. However, it does not provide domain-specific recommendations for the design of underlying source criteria, risks, timeframes, and source requirements or for instrument content validation in our security context.

In addition, the security evaluation instruments with their data collection methodology [15, 13, 17, 14] are designed for the organisation's internal use (mainly based on individual assessment), which does not provide an opportunity to create a benchmark or collect data centrally. Legislation (e.g., [1]) has also changed significantly during the last years, requiring central security evaluation overviews and, therefore, data collection, data interpretation options, and relevant user stories for collected data.

4. Research Approach and Contributions

The PhD project for increasing the level of information security in organisations uses the Design Science Research Method (DSRM) by [30], going through DSRM stages (identification of the problem, setting goals, design and development of the artifact, demonstration and evaluation of the result) (see Fig. 1). These stages are passed iteratively when answering each sub-research question (SRQ1, SRQ2, SRQ3, SRQ4), while iterations also occur within each SRQ's answering. The stages of the work according to the research questions are, in turn, aligned with the DSR method in Figure 1. The main artifact of the whole project is the framework for security level evaluation of the organisation (F4SLE) and the method to update the security evaluation level instrument (MUSE) for updating it as an output of SRQ2 and SRQ3.

Security management system standard choosing criteria (SRQ1). To respond to SRQ1, we elicited the requirements for national standards based on European Union Member States Cybersecurity strategies and implementation plans using the National Cybersecurity Index (NCSI) [28] database. Then,

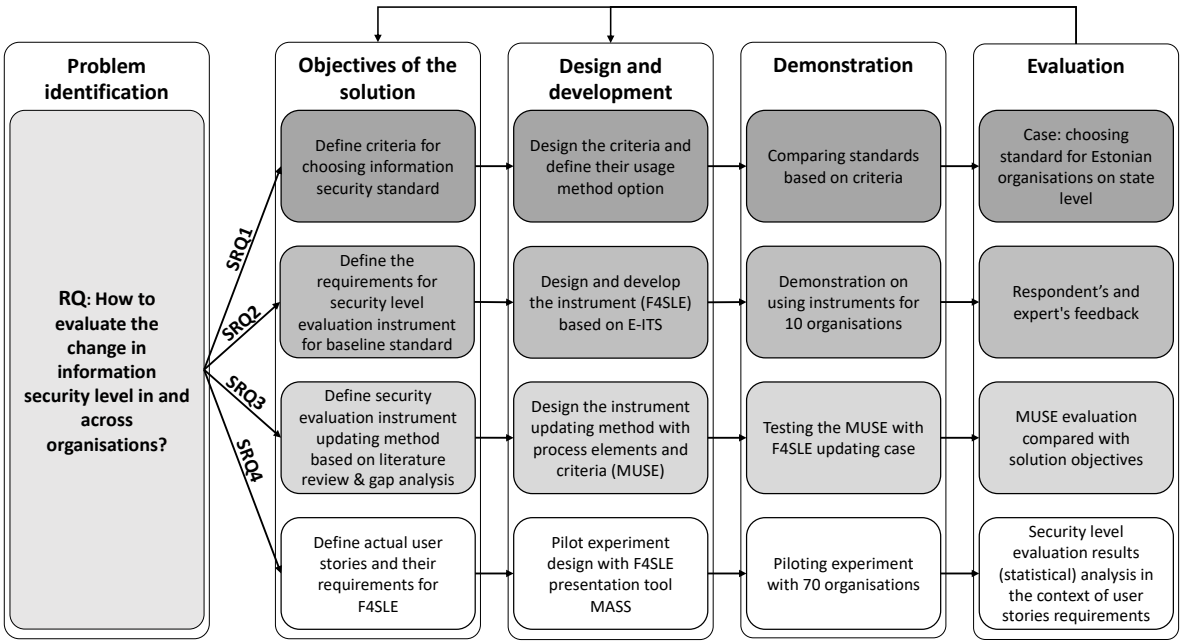

Figure 1: PhD project overview using DSR method based on [30]

we categorised elicited requirements into 3 modules and 15 characteristics to compare the standards and comply with national cybersecurity strategy requirements. For validation, we used the Estonian case - to find the most suitable standard or framework approach based on Estonian requirements. Results [31] showed there wasn't a perfect fit, but with some adoptions, the most suitable compliance with Estonian requirements was the German BSI IT-Grundschutz Kompendium [5]. (Estonian Information Security Framework E-ITS is developed and enforced at 2022 based the German BSI IT-Grundschutz Kompendium [5]. E-ITS has been significantly shortened, simplified, and adapted to the conditions of Estonia compared to German Kompendium.)

Framework for security level evaluation (SRQ2). Using the E-ITS [16] catalogue, we developed the framework with ten security dimensions, four maturity-like levels and around 200 security attributes divided into dimensions and levels. The framework's content attributes in dimensions are divided into process-based (five dimensions) and asset-based (the other five dimensions) to ensure a complete overview of the organisation's security situation. Ten organisations tested the F4SLE. Based on the respondents' and experts' evaluation and feedback, we updated and improved the framework during three iterations [32]. The artifact F4SLE [33] satisfied four solution objectives:

1. Framework should cover a wide area of security-related topics.
2. Framework should produce quantifiable and comparable results.
3. Framework should be quick and easy to implement and understand.
4. Framework should be aligned with a security standard.

A method for updating security level evaluation instruments (SRQ3). We designed the general updating process based on the literature review and defined the needed security-specific elements (roles, inputs, activities and outputs). The designing process supplemented the method with descriptions of the elements and each activity's criteria. The **m**ethod to **u**pdate the **s**ecurity **e**valuation level instrument (MUSE) process covers determined rule-based updating, reference standard mapping, expert review, pre-testing and validating activities. Each activity input, related roles and responsibilities have passing criteria. During the demo, we validated each activity element of the MUSE using several validation strategies described by [34, 35]. A predefined security baseline standard was used to check

the construct validity. For content validity, the MUSE required full compliance with a standard other than the predefined baseline standard to fix the similar scope in the updated version. The face validity strategy was realised using the newest cybersecurity threat landscape report suggestions for attributes as actuality checks in cybersecurity. For the criterion validity approach, we used pre-testing evaluation data of real respondents and a statistically significant positive correlation relationship (ρ closer to +1.0) between both F4SLE instrument versions [36, 37] pre-testing them simultaneously. If the primary goal of F4SLE was to create an opportunity to compare results between different organisations, the introduction of the renewal method allows for comparing the dynamics of results over time in the case the attributes change (security domain changes). This is ensured by a sufficient level of aggregation and updating attributes following the validation criteria.

User stories of F4SLE (SRQ4). To validate the F4SLE artifact in the context of the PhD project, we conducted a pilot experiment to evaluate the information security levels of 70 organisations simultaneously in both Estonia and South Moravia. The validation aims to prove F4SLEs generalisation options outside Estonia and its legislation area. Statistical analysis of cross-over organisations is a work in progress, as well as identifying security level evaluation requirements from regulations and assessing obtained statistical results against these requirements. Subsequently, the user stories (e.g., auditing tool, dashboard of situational awareness, monitoring, input to CSIRTs) of F4SLE can be evaluated as a facilitator for meeting regulatory requirements. Still, we already developed the web-based tool MASS [38] to display F4SLE attributes and provide an organisation's results for improved user convenience for respondents. MASS's design issues and requirements are detailed in the master's thesis [39]. MASS facilitated the centralised collection of organisation responses. For result interpretation, MASS provided immediate output to the respondents, evaluating their risk levels and the current status of security levels across ten dimensions defined in F4SLE [40]. Organisations could also instantly compare their results with a benchmark prepared from the outcomes of other organisations.

5. Threats to Validity

We designed F4SLE based on the E-ITS baseline standard [16] and assumed that E-ITS is developed with due diligence and considering all security aspects when defining the security dimensions and levels. These dimensions and levels are used for generalisation to reduce the number of security attributes and allow comparability in time. To reduce the E-ITS-related validity risk, we defined the content validity criteria to map each attribute with other ISMS standard clauses (see SQR3). We also included the environment where E-ITS is unknown in the piloting experiment.

F4SLE does not evaluate individual security countermeasures (attributes) but generalises the results into dimensions and levels. This creates a situation where, when the security situation changes, it is impossible to observe what is happening with a single attribute on a generalisation level, which can cause the complexity of interpreting the results if the nature of the dimensions is unclear. To reduce the interpretation risk, creating a formal, conceptual model of F4SLE in the PhD thesis is necessary, considering the elements accompanying the instrument updating method.

Currently, we have accepted the interpretation risk in the context of Estonia because organisations in Estonia are using E-ITS and interpretation skills are improving. Outside Estonia, the F4SLE presentation layer tool supports users with examples of interpretation. Data collected at the centre are interpreted only by specialists of the baseline E-ITS standard.

Survey-like studies can cause respondent risks (e.g., fatigue, inattention, falsification). A bigger reference group is needed to validate the F4SLE's benchmark as an acceptable benchmark. This will be out of this PhD project's scope but is planned as further work to repeat the data collection and find automatisation options for answering the attributes.

6. Concluding Remarks

The primary objective of this doctoral project was to find an option to evaluate the change in security level in and across organisations. To achieve this, we initially identified criteria for selecting a standard applicable to the state and its organisations [31]. Subsequently, we developed the F4SLE (information security level evaluation framework) [33] and the MUSE method [36] for updating the instrument's content to ensure sustained comparability to satisfy the evaluating the change of security level. Finally, we employed the updated F4SLE to evaluate the information security level of 70 actual organisations; work in progress should validate its effectiveness and potential user stories.

Each intermediate outcome of this PhD project stands independently applicable. For instance, the criteria employed in selecting standards for countries can be extrapolated to organisational contexts; the requirements of the security evaluation instrument can guide the development of instruments based on its own standard, and the MUSE method is expressly designed to update the content of survey-based security level evaluation instruments. The validated user stories furnish valuable feedback for creating security evaluation instruments.

7. Acknowledgments

The supervisor of this PhD thesis project is Prof. Raimundas Matulevičius.

This research is supported by the European Union under Grant Agreement No. 101087529. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

[1] European Parlament, Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive), 2022. URL: https://eur-lex.europa.eu/eli/dir/2022/2555.

[2] European Commission, 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) (Text with EEA relevance), 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[3] ISO/IEC 27002:2022 Information security, cybersecurity and privacy protection — Information security controls, International Organization for Standardization, 2022.

[4] NIST, The NIST Cyber Security Framework 2.0, 2024. URL: https://www.nist.gov/cyberframework.

[5] German Federal Office for Information Security (BSI), BSI IT-Grundschutz Kompendium, 2020. URL: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/Kompendium/IT_Grundschutz_Kompendium_Edition2020.html.

[6] Latvian Cabinet, Regulation No. 442, Procedures for Ensuring Compliance of Information and Communication Technology Systems with Minimum Security Requirements, 2015.

[7] S. Mohebbi, Q. Zhang, E. Christian Wells, T. Zhao, H. Nguyen, M. Li, N. Abdel-Mottaleb, S. Uddin, Q. Lu, M. J. Wakhungu, Z. Wu, Y. Zhang, A. Tuladhar, X. Ou, Cyber-physical-social interdependencies and organizational resilience: A review of water, transportation, and cyber infrastructure systems and processes, Sustainable Cities and Society 62 (2020) 102327. doi:10.1016/j.scs.2020.102327.

[8] CERT-EU, Threat Landscape Report 2023, 2024. URL: https://cert.europa.eu/publications/threat-intelligence/tlr2023/, accessed: 2024-03-05.

[9] Statistics Estonia, Community Survey on ICT usage and e-commerce in Enterprises, 2023. URL:

https://www.stat.ee/en/find-statistics/methodology-and-quality/esms-metadata/20505, accessed: 2023-12-14.
[10] National Audit Office of Estonia, Guaranteeing security and preservation of critical state databases of estonia, 2017. URL: https://www.riigikontroll.ee/DesktopModules/DigiDetail/FileDownloader.aspx?FileId=14218&AuditId=2462, accessed: 2024-03-05.
[11] The State Audit Office of the Republic of Latvia, Can we rely on the access to information systems and the receipt of e-services?, https://www.lrvk.gov.lv/en/getrevisionfile/29525-5Aio6j7MwYsuSG4nKlzFVmCMG0JZircA.pdf, 2022.
[12] NIS Cooperation Group, Cybersecurity and resiliency of Europe's communications infrastructures and networks, Technical Report, European Commission, 2024. URL: https://digital-strategy.ec.europa.eu/en/library/report-cybersecurity-and-resiliency-eu-communications-infrastructures-and-networks, accessed: 2024-03-05.
[13] R. Leszczyna, Review of cybersecurity assessment methods: Applicability perspective, Computers & Security 108 (2021) 102376. doi:10.1016/j.cose.2021.102376.
[14] N. T. Le, D. B. Hoang, Can Maturity Models Support Cyber Security?, in: 2016 IEEE 35th International Performance Computing and Communications Conference (IPCCC), 2016, pp. 1–7. doi:10.1109/PCCC.2016.7820663.
[15] M. Khaleghi, M. R. Aref, M. Rasti, Comprehensive Comparison of Security Measurement Models, Journal of Applied Security Research (2022) 1–69. doi:10.1080/19361610.2021.1981089.
[16] RIA (Estonian Information System Authority), E-ITS. Portal of Estonian Information Security Standard, 2022. URL: https://eits.ria.ee/.
[17] A. M. Rea-Guaman, I. D. Sánchez-García, T. S. Feliu, J. A. Calvo-Manzano, Maturity models in cybersecurity: A systematic review, in: 2017 12th Iberian Conference on Information Systems and Technologies (CISTI), 2017, pp. 1–6. doi:10.23919/CISTI.2017.7975865.
[18] M. Spruit, M. Röling, ISFAM: the Information Security Focus Area Maturity Model, in: Proceedings of the European Conference on Information Systems (ECIS) 2014, 2014. URL: http://aisel.aisnet.org/ecis2014/proceedings/track14/6.
[19] T. de Bruin, R. Freeze, U. Kulkarni, M. Rosemann, Understanding the Main Phases of Developing a Maturity Assessment Model, Australasian Conference on Information Systems (2005).
[20] Finnish Transport and Communication Agency National Cyber Security Centre, Cybermeter, https://www.kyberturvallisuuskeskus.fi/fi/palvelumme/tilannekuva-ja-verkostojohtaminen/kybermittari, 2024. Accessed: 2024-02-13.
[21] CINI, Framework Nazionale per la Cyber Security e la Data Protection , 2024. URL: https://www.cybersecurityframework.it/en, accessed: 2024-04-27.
[22] Hellenic Ministry of Digital Governance Government department, Cybersecurity self assessment tool, 2021. URL: https://mindigital.gr/wp-content/uploads/2022/03/cybersecurity-self-assessment.xlsm, accessed: 2024-04-27.
[23] The Spanish National Cybersecurity Institute, Herramienta de Autodiagnóstico , 2021. URL: https://adl.incibe.es/#, accessed: 2024-05-27.
[24] Bundeskanzleramt, Österreichisches Informationssicherheitshandbuch, https://www.sicherheitshandbuch.gv.at/, 2024. Accessed: 2024-06-13.
[25] National Cyber and Information Security Agency of the Czech Republic, 2022 Report on the State of Cybersecurity in the Czech Republic, https://nukib.gov.cz/download/publications_en/2022_Report_on_the_State_of_Cybersecurity_in_the_Czech_Republic.pdf, 2023. Accessed: 2023-11-08.
[26] International Telecommunications Index, Global Cybersecurity Index, https://www.itu.int/en/ITU-D/Cybersecurity/Pages/global-cybersecurity-index.aspx, 2020. Accessed: 2024-05-20.
[27] European Union Agency for Cybersecurity, EU Cybersecurity Index, https://www.enisa.europa.eu/topics/cybersecurity-policy/nis-directive-new/eu-cybersecurity-index, 2024. Accessed: 2024-05-20.
[28] e-Governance Academy (eGA), NCSI National Cyber Security Index, 2021. URL: https://ncsi.ega.ee.
[29] A. N. Oppenheim, Questionnaire design, interviewing and attitude measurement, Continuum,

2001.
[30] K. Peffers, T. Tuunanen, M. A. Rothenberger, S. Chatterjee, A Design Science Research Methodology for Information Systems Research, Journal of Management Information Systems 24 (2007) 45–77. doi:10.2753/MIS0742-1222240302.
[31] M. Seeba, R. Matulevičius, I. Toom, Development of the Information Security Management System Standard for Public Sector Organisations in Estonia, in: Business Information Systems, 2021, pp. 355–366. doi:10.52825/bis.v1i.43.
[32] M. Seeba, Estonian Information Security Standard (E-ITS) based security level evaluation instrument, Technical Report, University of Tartu Institute of Computer Science, 2021. doi:10.23673/re-298.
[33] M. Seeba, S. Mäses, R. Matulevičius, Method for Evaluating Information Security Level in Organisations, in: R. Guizzardi, J. Ralyté, X. Franch (Eds.), Research Challenges in Information Science, Springer International Publishing, Cham, 2022, pp. 644–652. doi:10.1007/978-3-031-05760-1_39.
[34] H. Taherdoost, Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research, SSRN (2016). doi:10.2139/ssrn.3205040.
[35] C. L. Kimberlin, A. G. Winterstein, Validity and reliability of measurement instruments used in research, American journal of health-system pharmacy 65 (2008) 2276–2284. doi:10.2146/ajhp070364.
[36] M. Seeba, A. amefon Obot Affia, S. Mäses, R. Matulevičius, Create your own MUSE: A method for updating security level evaluation instruments, Computer Standards & Interfaces 87 (2024) 103776. doi:10.1016/j.csi.2023.103776.
[37] M. Seeba, Framework for Security Level Evaluation (F4SLE) E-ITS based ver 2021-1, 2022. doi:10.23673/re-372.
[38] MASS tool to present F4SLE, 2023. URL: https://mass.cloud.ut.ee/massui/#/.
[39] M. P. Murumaa, Designing a Security Sensitive Self-assessment Framework, Technical Report, University of Tartu, 2023. URL: https://comserv.cs.ut.ee/ati_thesis/datasheet.php?id=77886&language=en.
[40] M. Seeba, T. Oja, M. P. Murumaa, V. Stupka, Security Level Evaluation with F4SLE, in: Proceedings of the 18th International Conference on Availability, Reliability and Security, ARES '23, Association for Computing Machinery, New York, NY, USA, 2023. doi:10.1145/3600160.3605045.