

GEIGER



Deliverable D1.2

Architecture

Point of Contact	Samuel Fricker
Institution	FHNW
E-mail	samuel.fricker@fhnw.ch
Phone	+41 79 196 9629

Project Acronym	GEIGER
Project Title	GEIGER Cybersecurity Counter
Grant Agreement No.	883588
Topic	H2020-SU-DS03
Project start date	June 2020
Dissemination level	Public
Due date	M12
Date of delivery	M12
Lead partner	ATOS
Contributing partners	FHNW, KSP, KPMG, UU, MI
Authors	José Francisco Ruiz (ATOS), José Javier de Vicente (ATOS), Max van Haastrecht (UU), Injy Sarhan (UU), Rolan Kab (KPMG), Samuel Fricker (FHNW), Martin Gwerder (FHNW), Louis Baumgartner (FHNW), Marco Spruit (UU), Wissam Mallouli (MI), Bettina Schneider (FHNW)
Reviewers	KPMG, FHNW, MI

This document contains information that is treated as confidential and proprietary by the GEIGER Consortium. Neither this document nor the information contained herein shall be used, duplicated, or communicated by any means to any third party, in whole or in parts, except with prior written consent of the GEIGER Consortium.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 883588 (GEIGER). The opinions expressed and arguments employed herein do not necessarily reflect the official views of the funding body.

Revision History

Version	Date	Author	Comment
1.0	05/04/2021	José F. Ruiz and José Javier de Vicente (ATOS)	Initial version
1.1	09/04/2021	Max van Haastrecht and Injy Sarhan (UU)	GEIGER Indicator – First draft
1.2	11/04/2021	Rolan Kab (KPMG)	KPMG GEIGER bot
1.3	14/04/2021	Max van Haastrecht and Injy Sarhan (UU)	Edits to section 3.1.6.1. Also affects references and table numbering.
1.4	16/04/2021	Samuel Fricker, Martin Gwerder, Louis Baumgartner (FHNW)	Edits to Sections 1.3.1 CYSEC, 3.2 Toolbox, and 3.1.4.2 (use case-based scenarios of using the GEIGER Indicator)
1.5	21/04/2021	José F. Ruiz and José Javier de Vicente (ATOS)	GEIGER Cloud
1.6	22/04/2021	Rolan Kab (KPMG)	GEIGER Conversation module and proxy module
1.7	23/04/2021	José F. Ruiz and José Javier de Vicente (ATOS)	First general review and comments
1.7.1	27/04/2021	Samuel Fricker, Martin Gwerder (FHNW)	Section 3.2.2 Toolbox Software Design
1.7.2	30/04/2021	Samuel Fricker, Louis Baumgartner (FHNW)	Section 3.2.1 Toolbox User Experience and Interface Design
1.8	30/04/2021	Max van Haastrecht, Injy Sarhan, and Marco Spruit (UU)	Edits to Section 3.1.4 and Appendix C.
1.9	05/05/2021	José F. Ruiz and José Javier de Vicente (ATOS)	Review and comments
2.0	07/05/2021	Wissam Mallouli (MI)	Edit sections 1.3.12, 1.3.13 and 1.3.14 and comments
2.1	27/05/2021	Samuel Fricker, Louis Baumgartner (FHNW)	Edits to section 1.3
2.2	28/05/2021	José F. Ruiz and José Javier de Vicente (ATOS)	Review and comments
2.3	30/05/2021	Wissam Mallouli (MI)	Review and comments
2.4	31/05/2021	Rolan Kab (KPMG)	Review and comments
2.5	31/05/2021	Samuel Fricker, Bettina Schneider (FHNW)	Finalisation for submission

Contents

Abbreviations	viii
List of Tables	ix
List of Figures	x
Summary	1
1. Introduction	2
1.1 Requirements and needs of GEIGER	2
1.2 Process and methodology	2
1.3 Cybersecurity Tools	3
1.3.1 CYSEC Mobile Learning	3
1.3.1.1 Adaptation of Cybersecurity Coach to Mobile Learning Application	3
1.3.1.2 User Experience: Structured Lessons	3
1.3.1.3 User Experience: Unstructured Dialogue	6
1.3.1.4 Data Model	7
1.3.1.5 CYSEC Architecture and Toolbox Interoperability	8
1.3.2 Risk Assessment Engine (RAE)	9
1.3.3 Information Sharing Platform (ISP)	10
1.3.4 Conversation Module	12
1.3.4.1 Conversation Module Sequence Diagrams	13
1.3.4.2 Conversation Module Package Diagrams	14
1.3.5 KPMG Proxy Module	14
1.3.6 Fraud Detection	14
1.3.7 Document Harvesting	15
1.3.8 Employee Virtual Assistant	15
1.3.9 Kaspersky Interactive Protection Simulation (KIPS)	15
1.3.10 KMS-SDK	15
1.3.11 CyberSafety Management Games (CSMG)	15
1.3.12 Montimage IDS (MIDS)	16
1.3.13 Cyber-Range	16
2. Architecture	17
2.1 Overview	17
2.2 Architecture of GEIGER	17
2.3 GEIGER Scenarios	20
2.3.1 Local Only	21
2.3.2 GEIGER Cloud (Cloud + local)	21
2.3.3 GEIGER Cloud Next (Cloud + local + external)	22
2.4 Functionality	22
2.4.1 Information of the Apps Running Locally	22
2.4.2 Storing Data of Results of Apps in the Cloud	23

2.4.3	Requesting data of results of apps	24
2.4.4	Information of the MSE (only local)	24
2.4.5	Storing Cloud-Relevant Data of the Company for Further Analysis	25
2.4.6	Requesting Cloud-Relevant Data of Company for Calculation of Indicator	26
2.4.7	Storing Information of CTI	27
2.4.8	Requesting Information of CTI for the Indicator	28
2.4.9	Storing Information of Infrastructure Apps in the Cloud	29
2.4.10	Sending Information of Infrastructure Apps to the GEIGER Indicator	30
2.4.11	Storing Information of External Apps in the Cloud	31
2.4.12	Sending Information of the External Apps to the GEIGER Indicator	32
2.4.13	Generation and Storing Personal Information of the Employee	32
2.4.14	Use of the Personal Information of the Employee	33
2.4.15	Storing Refined Personal Information Data on the Cloud for Further Analysis	34
2.4.16	Storing (Relevant) Information from MSEs	35
2.4.17	Requesting Refined Data for GEIGER Indicator	36
2.4.18	Storing Relevant Information for CTI Coming From MSEs	37
2.4.19	Requesting Refined Data of Company for Calculation of Indicator	38
2.5	Data Exchange and Communications	38
2.5.1	Communication Adapters	38
2.5.2	APIs for Online and Local Database Access	39
2.5.2.1	Local Storage API	39
2.5.2.2	Cloud Storage API	40
2.5.3	Data Exchange Protocol	40
2.6	Roles	42
2.6.1	Technical Engineer	42
2.6.2	Cybersecurity Trainer	42
2.6.3	Organization	42
2.6.4	CERT / CSIRT	42
2.6.5	Cybersecurity Defender	43
2.6.6	End-User	43
2.7	Use Cases	44
2.7.1	Technical Engineer	44
2.7.2	Cybersecurity Trainer	44
2.7.3	Organization	45
2.7.4	CERTs and CSIRTs	45
2.7.5	End-User on PC and Internet	46
2.7.6	End-User on PC without Internet	47
2.7.7	End-User with Mobile Device	48
3.	GEIGER Components	49
3.1	GEIGER Toolbox: GEIGER Indicator	49
3.1.1	GEIGER Indicator Concept	49
3.1.2	GEIGER Indicator Data Model	51
3.1.3	GEIGER Indicator Data	52

3.1.3.1	MSE Profiles	52
3.1.3.2	GEIGER Indicator Threats	53
3.1.3.3	GEIGER Indicator Tool Metrics	54
3.1.3.4	GEIGER Indicator Global Recommendations	55
3.1.4	GEIGER Indicator Algorithm	56
3.1.4.1	Incorporating Hierarchy	57
3.1.4.2	Algorithm Aggregation	59
3.1.5	GEIGER Indicator Output	61
3.1.5.1	Initial Output	61
3.1.5.2	Output Update	62
3.1.6	GEIGER Indicator Sustainability	62
3.1.6.1	Impact Determination	62
3.1.6.2	Concept Adaptability	64
3.1.6.3	Concept Resilience	64
3.2	GEIGER Toolbox	66
3.2.1	Toolbox User Experience and Interface Design	66
3.2.1.1	User Journey and Personas	66
3.2.1.2	Wireframe of the Toolbox User Interface	67
3.2.1.3	Operationalisation of Quality Requirements	78
3.2.2	Toolbox Software Design	79
3.2.2.1	Distributed Database Design	81
3.2.2.2	Plugin Mechanism	82
3.2.2.3	Scoring and Knowledge Overview	82
3.2.2.4	Cross-Platform Environment	83
3.2.2.5	Internalization and Localization	83
3.3	GEIGER Cloud	84
3.3.1	GEIGER Cloud Overview	84
3.3.2	GEIGER Cloud Function and Objectives	84
3.3.3	GEIGER Cloud Design and Architecture	84
3.3.3.1	GEIGER Cloud Core	85
3.3.3.2	GEIGER Cyber Threat Intelligence Sharing	85
3.3.3.3	GEIGER Infrastructure Applications	86
3.3.4	GEIGER Cloud Data Exchange	86
3.3.5	KPMG GEIGER Conversation module	86
3.3.5.1	Local storage under the GEIGER Toolbox Core	87
3.3.5.2	Conversation module GEIGER Cloud data storage	88
3.3.6	KPMG ISP with Information Sharing	88
3.4	GEIGER Cyber Threat Intelligence (CTI) Sharing	88
3.4.1	GEIGER Information Sharing	88
3.4.2	KPMG Conversation Module	88
3.4.3	KPMG Proxy Module	89
4.	Conclusions and Future Work	90

5. Annexes	91
Appendix A: Definitions	91
Appendix B: Threat mapping	91
Appendix C: Detailed recommendation information	94
Appendix D: Algorithm variables	97
Appendix E: Data model entities	98
6. References	101
6.1 GEIGER Indicator	101

Abbreviations

AHP	Analytic Hierarchy Process
AI	Artificial Intelligence
CERT	Computer Emergency Response Team
CIS	Centre for Internet Security
CISO	Chief Information Security Officer
CS	Cybersecurity
CSIRT	Computer Security Incident Response Team
CSC	Critical Security Controls
CSD	Certified Security Defender
CSF	Cybersecurity Framework
CTI	Cyber Threat Intelligence
DB	Database
GUI	Graphical User Interface
HMI	Human-Machine Interface
ISAC	Information Sharing and Analysis Center
ISP	Information Sharing Platform
MCDM	Multiple-Criteria Decision Making
MISP	Malware Information Sharing Platform
ML	Machine Learning
MSE	Micro or Small Enterprise
NCSC	National Cyber Security Center
OS	Operating System
OSINT	Open-source intelligence
PaaS	Platform as a Service
PI	Personal Information
RAE	Risk Assessment Engine
SDK	Software Development Kit
SLA	Service Level Agreement
TLP	Traffic Light Protocol
UR	User Requirement
UUID	Universally Unique Identifiers

List of Tables

Table 1: A selection of user requirements for the GEIGER solution.	49
Table 2: The three MSE profile attributes category, sector, and country.	52
Table 3: The GEIGER indicator threats. We indicate the definition of each threat and the source for this definition.	53
Table 4: Statistics on the collected metrics for the GEIGER indicator threats.	54
Table 5: Mapping of the NIST CSF functions to CIS Critical Security Controls (CSCS).	63
Table 6: Controls to ensure the adaptability and resilience of the GEIGER indicator concept.	65
Table 7: Personas supported by the Toolbox.	67
Table 8: Requirements-refining user stories for personalized assessment.	69
Table 9: Requirements-refining user stories for guidance and help.	70
Table 10: Requirements-refining user stories for pairing of devices and employees.	71
Table 11: Requirements-refining user stories for the handling of events.	73
Table 12: Requirements-refining user stories for management of tools.	75
Table 13: Requirements-refining user stories for navigation.	76
Table 14: Requirements-refining user stories for toolbox settings.	77
Table 15: Implementation of quality requirements by the user interfaces of the GEIGER toolbox.	78
Table 16: Definitions employed in the GEIGER indicator development.	91
Table 17: The guiding principles for mapping threats from the ENISA threat taxonomy (ENISA, 2016b) to the GEIGER setting.	91
Table 18: Mapping of threats from the ENISA Top 15 threats (ENISA, 2020) and the detailed ENISA threat taxonomy (ENISA, 2016b) to the GEIGER indicator threat concepts.	92
Table 19: Sources used to construct our global recommendation list.	94
Table 20: The global recommendation list employed in the GEIGER indicator solution.	94
Table 21: The global GEIGER Indicator recommendations are mapped to security control categories, where each recommendation maps to at least one security control category.	96
Table 22: GEIGER indicator algorithm variable definitions.	97
Table 23: Data Model Entities and Attributes.	98

List of Figures

Figure 1: Example of a structured lesson.	6
Figure 2: Example of discussion that is part of the unstructured dialogue.	7
Figure 3: Data model for the specification of awareness, instructional, and learning sequences that can be delivered by CYSEC.	8
Figure 4 – Architecture of the Information Sharing Platform	11
Figure 5 – Conversation Module active interaction - Sequence Diagram	13
Figure 6 - Conversation Module proactive interaction - Sequence Diagram	13
Figure 7 - Conversation module – Use Case Diagram	14
Figure 8 - GEIGER Architecture	18
Figure 9 - GEIGER Toolbox architecture	19
Figure 10 - GEIGER Cloud architecture	20
Figure 11 - GEIGER scenarios	21
Figure 12 - Information of apps running locally flow	22
Figure 13 - Storing data of results of apps in the Cloud flow	23
Figure 14 - Requesting data of results of apps flow	24
Figure 15 - Information of the MSE (local) flow	24
Figure 16 - Storage of company data in the cloud flow	25
Figure 17 - Requesting cloud data of the MSE for risk score calculation flow	26
Figure 18 - Storing information of CTI flow	27
Figure 19 - Requesting CTI information for indicator flow	28
Figure 20 - Storing data of infrastructure apps in the cloud flow	29
Figure 21 - Sending information of infrastructure apps to the indicator flow	30
Figure 22 - Storing information of external apps in the cloud flow	31
Figure 23 - Sending information of external apps to indicator flow	32
Figure 24 - Generation and storing employee PI flow	32
Figure 25 - Storing refined PI data on the cloud flow	34
Figure 26 - Storing information from MSEs flow	35
Figure 27 - Requesting data for indicator flow	36
Figure 28 - Storing relevant information of MSEs flow	37
Figure 29 - Requesting refined data of MSE for indicator flow	38
Figure 30 - Detail of the GEIGER Controller in the GEIGER architecture	39
Figure 31 - Detail of the GEIGER Cloud adapter in the GEIGER architecture	39
Figure 32 - Technical engineer use case	44
Figure 33 - Cybersecurity trainer use case	44
Figure 34 - Organization use case	45

Figure 35 - CERT use case	45
Figure 36 - End user on PC + internet use case	46
Figure 37 - End user on PC without internet use case	47
Figure 38 - End-user with mobile device use case	48
Figure 39: The view on cyber-systems that serves as the basis for the GEIGER indicator algorithm.	50
Figure 40 - GEIGER Indicator Data model	51
Figure 41 - An example MSE hierarchy.	58
Figure 42 - GEIGER Indicator aggregation	61
Figure 43: Summary of User Journey (details: see user journey specified in D1.1).	66
Figure 44: Wireframes for personalised assessment.	68
Figure 45: Wireframes for guidance and help.	70
Figure 46: Wireframes for pairing of devices.	71
Figure 47: Wireframes for pairing of employees.	71
Figure 48: Wireframes for the handling of events.	73
Figure 49: Wireframes for management of tools.	75
Figure 50: Wireframe for navigation.	76
Figure 51: Wireframes for toolbox settings.	77
Figure 52: Toolbox architecture (middle, red) and external components (left, yellow: GEIGER Cloud; right, blue: tool integrated as a plugin).	80
Figure 53: Distributed database consisting of local storages on the MSE's devices and connected with a shared end-to-end encrypted channel.	80
Figure 54: Hierarchical structure of the data.	81
Figure 55 - Diagram of the GEIGER Cloud	85
Figure 56 - MSE relations and data information	87

Summary

The document defines and describes the architecture proposal for the GEIGER cybersecurity platform. It covers all the components of the GEIGER solution including the most important blocks: GEIGER Indicator, GEIGER Toolbox, GEIGER ISAC and GEIGER Cloud. In addition, it explains all the subcomponents as well as the tools that are used or integrated in GEIGER. These tools can be classified in three main categories:

- GEIGER tools,
- Cloud infrastructure apps,
- External tools.

Given that different MSEs and business can benefit from trusting on GEIGER, the platform architecture has considered various scenarios or use cases, such as

- a local only mode,
- a GEIGER Cloud and local scenario,
- a Cloud, local and external scenario.

This variety of cases can help get a real idea on the flexibility of the proposed solution.

Regarding data, for the purpose of clarifying how it is managed in GEIGER, information flows have also been described in the document, covering issues such as who and how to request information, who is responsible for providing requested data, or the path data are expected to traverse to the final destination. This can help the reader to understand GEIGER data management.

Finally, both GEIGER roles and use cases are described within the present document. It is necessary to know how the various actors involved can interact with GEIGER and what their actions in the platform can be.

1. Introduction

1.1 Requirements and needs of GEIGER

Nowadays cybersecurity is acquiring vital importance in every aspect of our lives. One of the most critical fields where cybersecurity is becoming more relevant is the business environment – no matter what industry the MSEs develop their activity.

GEIGER is born from a necessity: to provide cybersecurity and make it available and affordable for medium, small and microenterprises around Europe while, at the same time, focusing on helping these organisations to achieve a higher level of security in their day-to-day operations.

The nature of the GEIGER solution entails the necessity of being a user-centric platform where the end-user can definitely be aware of the various risks of the environment. For this purpose, the GEIGER Indicator shall provide updated information on the level of risk. As per this feature, GEIGER requires updated or even (if possible) real-time information that can help determining the risk level accurately. Besides, GEIGER is called a *complete solution*, meaning that it will be possible to have it installed on more than one device owned by the MSE. The possibility of carrying GEIGER within reach empowers the end-user and makes the platform much more versatile and useful.

Moreover, GEIGER possesses a useful degree of adaptability. Being a cybersecurity solution, there exists an implicit requirement for adaptation, evolution, and change, as well as for answering to new threats effectively. This requirement is satisfied with the addition of new tools, which can engage with GEIGER, provide new information, and enhance possibilities depending on future needs.

Finally, it is important to outline that GEIGER has a double approach:

- i. The GEIGER Cloud part, which is focused on providing real-time information with the power and possibilities that the cloud environment offers nowadays.
- ii. The GEIGER Toolbox, whose purpose is to make GEIGER available for any device and relies on a local storage and approach.

Both platforms are interconnected, making GEIGER a solution that improves the cybersecurity level of MSEs.

1.2 Process and methodology

The idea behind the GEIGER project is to **increase cybersecurity knowledge and awareness** of the owners and employees of MSEs. To achieve success in the project it is necessary to couple different tools from various partners and take advantage of both the Cloud and the on-premise features of the solution.

Since the very beginning of the GEIGER project, the aim was to empower MSEs and make users more conscious of the various threats, which could endanger their business. For that purpose, GEIGER perfectly mixes the Cloud and the on-premise possibilities, empowering the user too freely choose how to access the platform.

In the GEIGER platform, **suggestions, ideas, and efforts from various partners have been integrated** continuously to enrich the product. The various prototypes have been refined with the help and contribution of the partners. It has been necessary to carefully address the following questions:

- i. What information needs to be stored amongst all information available (data from tools, provided information by CERTs and CSIRTs, data gathered from the business, etc.)?
- ii. Where to store that information: on the cloud, locally or both?
- iii. How should the GEIGER Toolbox and the GEIGER Cloud interact and under what premises? Which APIs are needed to ease this communication?
- iv. How should the risk score of the GEIGER Indicator be calculated? What threats and categories should be considered? How should these be weighted?

- v. How should the different tools of the partners integrate with the GEIGER platform? What amount of data are they going to share with GEIGER?
- vi. What is the best way to make information available to the user? What would be the simplest way to achieve this?

Work package meetings have helped the partners to determine and clarify the aspects mentioned. Furthermore, **keeping a continuous track of the work** performed is part of the methodology followed in the GEIGER project. These continuous follow-ups enable the early detection of any misalignment, and allow addressing any problem that may arise. In addition, regular meetings helped to be involved and aware of the work that was being accomplished by the other partners, and they were a good opportunity to voice doubts, problems, warnings or concerns.

Work has been scheduled and shared among partners in advance. Anticipation means possibility of fixing any problem and help task leaders to best accommodate the work to be done to the abilities of the partners engaged in the task. The objective was to accomplish the best possible fit amongst abilities, capacities, and activities to be done.

1.3 Cybersecurity Tools

1.3.1 CYSEC Mobile Learning

1.3.1.1 Adaptation of Cybersecurity Coach to Mobile Learning Application

CYSEC is a cybersecurity awareness tool developed by FHNW in the SMESEC project. It is a coaching platform for users, which was originally designed to be accessed via browser by means of a web User Interface. CYSEC provides the human end-user with the ability to consume sequences for becoming aware of issues and options for addressing these issues, for stepwise following instructions to achieve a desired outcome, and for learning about cybersecurity topics at the levels of knowing and understanding.

CYSEC is being ported from a PC-based platform to mobile platforms, where it becomes CYSEC Mobile Learning. The tool adapted to address the MSE end-user's learning and guidance needs. These needs are triggered when the GEIGER Indicator offers recommendations for securing the company with sensor and tool installations, configuration instructions, and learning experiences. The new challenge being addressed is to allow even novice ICT users to understand the instructions they receive and get help if guidance is needed in the application of the instructions.

GEIGER mobile learning includes two modes of knowledge delivery: short, structured lessons delivered in the form of a micro-teaching, and an unstructured dialogue between the learners and Security Defenders¹. The structured lessons aim at raising awareness and letting the learners understand a recommended cybersecurity topic. The unstructured dialogue aims at enabling the learners to be able to apply the cybersecurity topic in their own specific context and devices.

Such hybrid delivery of knowledge has been pursued by initiatives in domains other than cybersecurity, for example in the learning to code by SoloLearn², Stackoverflow³, or Instagram's help instructions. It offers highly personalised support of the learners while allowing the delivery of knowledge to scale beyond what would be possible with a simple teacher-learner relationship.

1.3.1.2 User Experience: Structured Lessons

A lesson consists of various informational slides. All lessons are at installation already integrated inside the application and unlocked through the GEIGER Toolbox. Cybersecurity experts and educators design these

¹ See 2.6.5 for a definition of this term

² <https://www.sololearn.com>

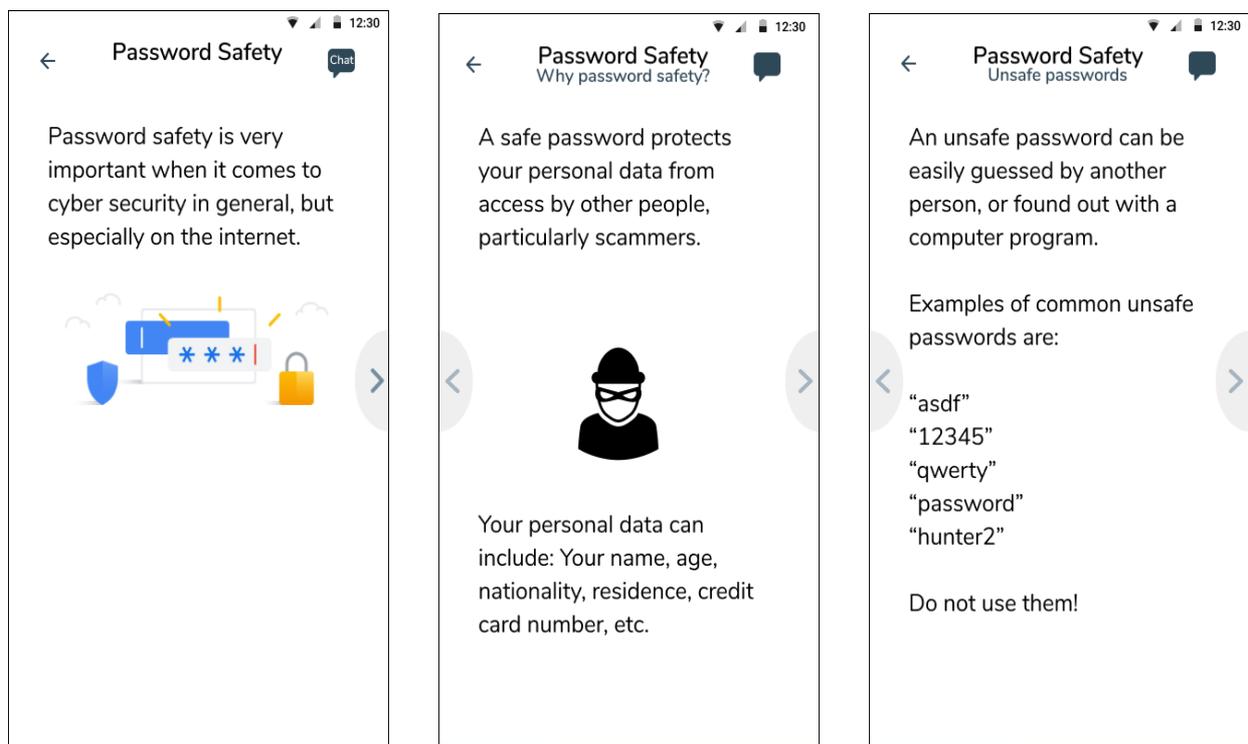
³ <https://www.stackoverflow.com/>

lessons. To avoid overwhelming the user with information, lessons are designed to be compact overall, with the visuals providing small pieces of information or instructions each.

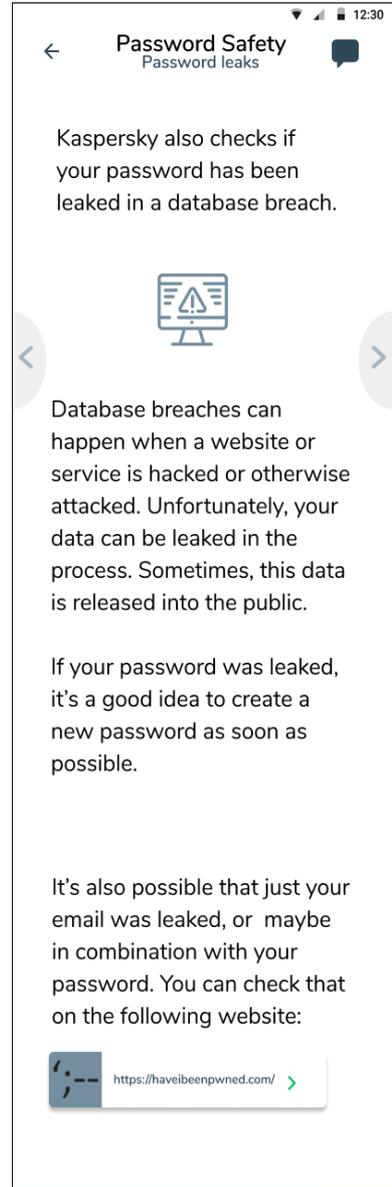
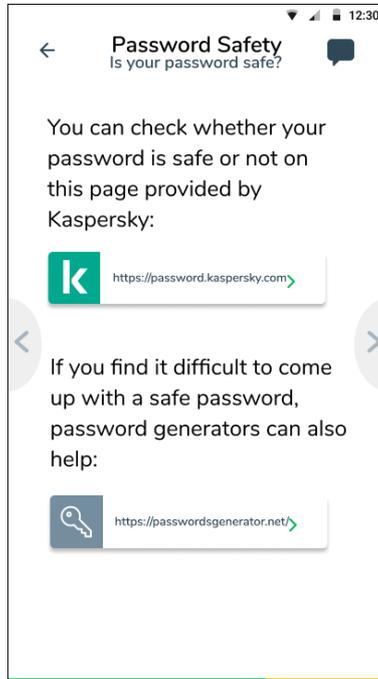
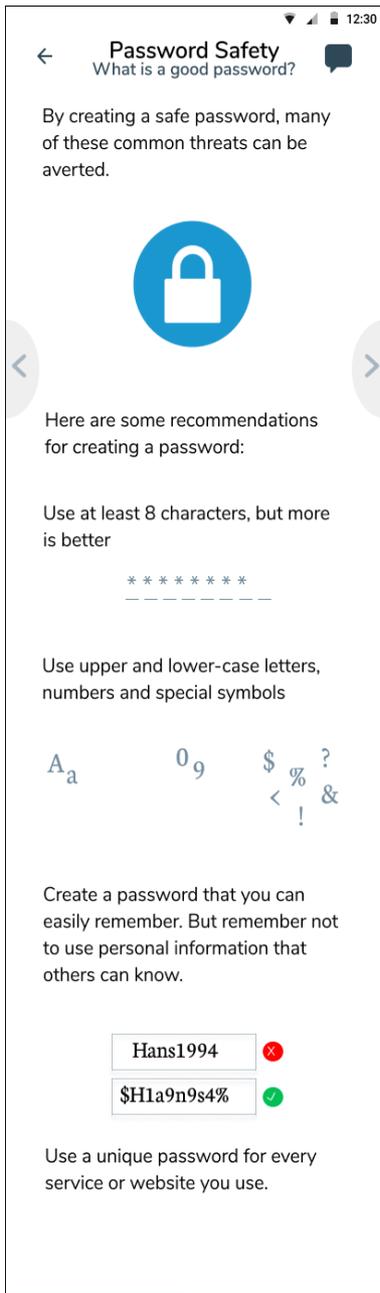
Like Instagram's help instructions, the user can swipe through the visuals from left to right. In addition, there are controls on the left and right side of the screen for offering a book reading-like interaction style. To give a user a sense on how far the learner has progressed through the lesson, a progress bar is displayed at the bottom of the screen, giving an approximation for how long the remaining part of the lesson will continue. Specific slide numbers are not displayed as lessons are supposed to be rather short and numbers could discourage the user from engaging.

At the end of each lesson, a learning-score is awarded. This form of gamification serves two main purposes. It displays the user's overall participation with the app and can be used as a symbol of status amongst his/her co-workers, motivating others through competition to get a higher score and interact with the application. It shows the experience a user has in the communication part of the app.

Figure 1 shows sample screens for a lesson on password safety are shown⁴:



⁴ The use of the Kaspersky password checker is for illustration only. Other services may be referred to instead, e.g., based on the capabilities of the tools integrated as plugins into the GEIGER toolbox.



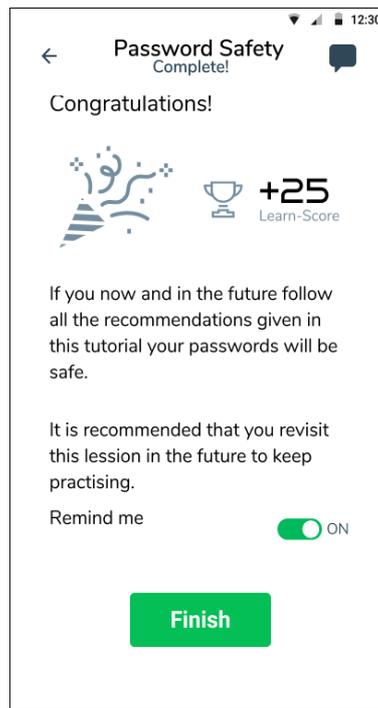


Figure 1: Example of a structured lesson.

1.3.1.3 User Experience: Unstructured Dialogue

CYSEC Mobile Learning includes a discussion section for each learning sequence. The discussion can be accessed at the top right, where users can ask questions and discuss with each other. The discussion section is an overlay on top of the slide content. It can be closed by tapping the “x” element at the top left.

A discussion consists of message threads. A message is visualised with the user’s picture, their score indicating overall experience, their pseudonym, and the message. Underneath each message, the date on which it was published as well as an option to reply to a specific comment is given.

On writing the first comment, a user is kindly reminded that personal information must be avoided to be communicated and that communication on the platform be done in a respectful manner.

Figure 2 shows a sample discussion.

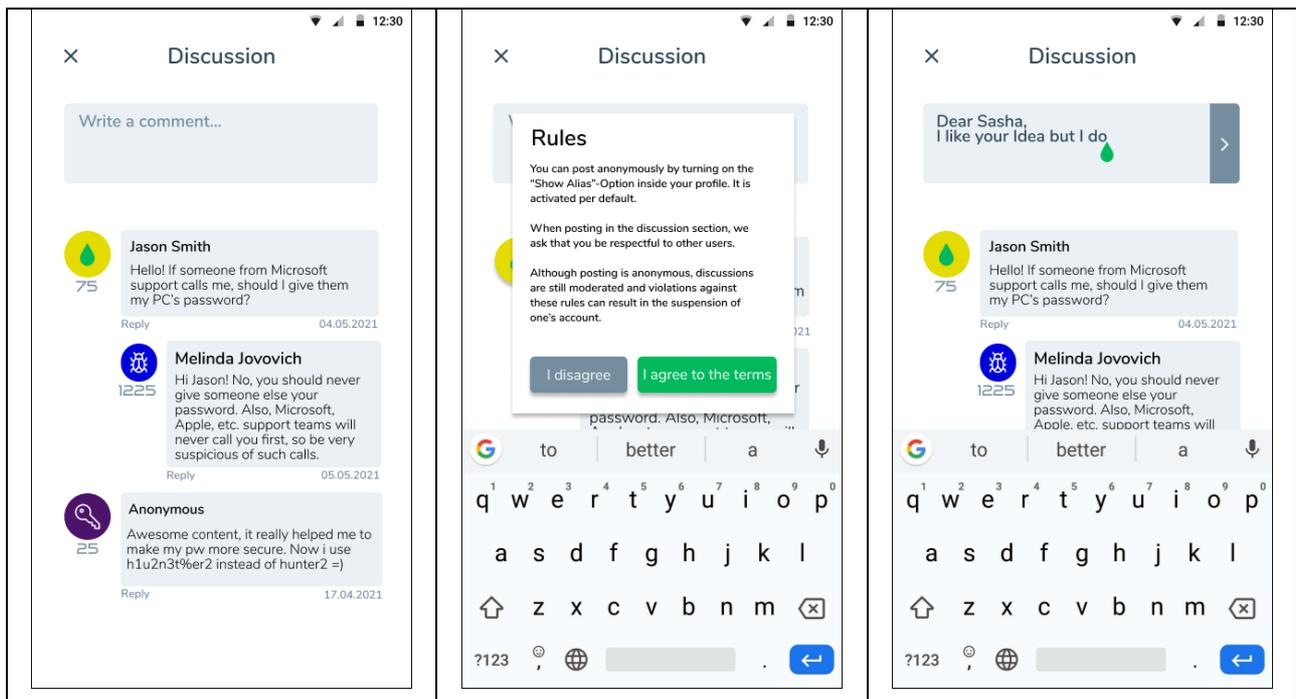


Figure 2: Example of discussion that is part of the unstructured dialogue.

1.3.1.4 Data Model

The data is stored in a hierarchical format following the following structure:

```
learning interface = { modules }
```

```
module = title, description, { dialogue }
```

(* a module reflects a theme with several topics requiring a learning dialogue *)

```
dialogue = ( block ) feedback
```

(* a dialogue can be triggered for a specific recommendation or within a module (local URL) *)

(* a dialogue leads to a specific learning achievement that can be recorded in xAPI *)

```
block =
```

```
question,
```

```
[what explanation, why explanation, how walkthrough],
```

```
reflections
```

```
question = multiple choice
```

(* may provide in-situ feedback *)

```
explanation = text, [video]
```

```
walkthrough (device or app type) = { screenshot enhanced with comments | interactive component like link tester or cybergage }
```

```
reflections = { messages from (explicit message author | anonymous) for (self | designated users | paired users | named user groups) }
```

(* get tips for tools that we would not be allowed to advertise *)

(* share screenshots from previously undocumented devices or apps *)

(* share own examples *)

(* clarify questions *)

```

(* offer feedback to work results *)
(* stimulate reflection with lessons-learned *)
user group = [ my coachees | curators ]
feedback = summary of achievement, opportunity to record in GEIGER profile, set
reminder for repetition
(* offer a motivating recognition for the user *)
(* allow the user to share the recognition with others with a link allowing
them to do the same: designated users | paired users | device's share function
*)

```

Figure 3: Data model for the specification of awareness, instructional, and learning sequences that can be delivered by CYSEC.

1.3.1.5 CYSEC Architecture and Toolbox Interoperability

CYSEC follows a recourse-efficient client/server architecture. On the client side, the lessons are executed and presented to the end-user, and the learning progress is reported to the toolbox. A lesson may be triggered by the GEIGER Toolbox. On the server side, definitions of the structured lessons are stores and a download of these lessons offered to the client side.

The GEIGER Toolbox offers an xAPI⁵ statement storage for reporting the user's achieved learning experiences and reasoning about the user's knowledge. The storage is organized as follows in the Toolbox data structure:

```

- :- | ---Users
    | | ---user1
    | | | ---xapi_experiences
    | | | | ---my_xapi_verb_object_statement
    | | ---Global
    | | | ---xapi
    | | | | ---verbs
    | | | | | ---my_verb
    | | | | ---objects
    | | | | | ---my_object

```

The xAPI Statement Storage consists of two parts, a user-specific part and a global part. The xAPI statement is extended with context information.

- The user-specific part captures the user's learning achievements with xAPI statements. The statements are recorded without stating the actor; the actor is always represented by the node to which the xAPI statement is added.
- The global part defines the verbs and objects that may be used to compose an xAPI statement. The content of the global part is defined in the GEIGER Cloud with the GEIGER selection of xAPI verbs⁶ and extended with GEIGER-specific verbs. When the toolbox is initialised or synchronising with the GEIGER Cloud, these definitions are retrieved from the GEIGER Cloud.
- The context information includes a timestamp of when the corresponding learning objective has been achieved, how many points have been awarded, and when the learning achievement will be depreciated.

The GEIGER Toolbox allows a tool like CYSEC to be called from within a GEIGER Indicator-generated *Recommendation*. For this aim to be achieved, the *Action* specified in a *Recommendation* must adhere to the following structure:

⁵ <https://github.com/adlnet/xAPI-Spec/blob/master/xAPI.md>

⁶ <http://xapi.vocab.pub/verbs/index.html>

- protocol: "geiger://"
- plugin: unique identifier of an integrated tool
- xapi statement: verb/object

Example of an *Action* as part of a *Recommendation* specification:

```
...
Action=["geiger://mobile_learning/configure/ios_backup"],
...
```

To build a valid URL, the verb and object must be specified by representing the spaces with underscores "_". When reporting learning achievements, these underscores should be replaced with spaces again.

The xAPI Statement Storage can be written with the GEIGER communication API as follows:

```
Node n = LocalAPI.getStorage().get(":Device:User:user1:xapi_experiences");
// edit node with xapi update
localAPI.getStorage().update(n);
```

The xAPI Statement Storage allows any tool or new component to retrieve information about the user's knowledge and reason about it. The GEIGER Toolbox Storage supports simple queries but no joins.

For example, a new component written to determine the user's learning level according to the GEIGER educational reference model could interact with the xAPI Statement Storage as follows:

```
Node n = LocalAPI.getStorage().get(":Device:User:user1:xapi_experiences");
// check the xAPI statements against the set of learning objectives to be
// achieved to qualify for the level 1...
// if qualified, update the following in the Node n:
|   |--xapi_experiences
|   |--{ "verb": { "id": "http://adlnet.gov/expapi/verbs/achieved" },
|       "object": { "id": ":Global:xapi:objects:L1" } }
localAPI.getStorage().update(n);
```

1.3.2 Risk Assessment Engine (RAE)

RAE is the cybersecurity tool for risk assessment developed by ATOS. Its purpose is to assess the cyber risk of the company by means of executing a risk-model based algorithm. RAE also suggests several possible mitigation measures, which will be proposed differently and in accordance to the results of the assessment.

The tool firstly gathers information of the company, that is, a profile, by means of presenting a questionnaire to the user. After weighting the information, RAE engine considers how it can affect the security aspects (confidentiality, integrity, and availability). The software relies also on some vulnerability scanners, which can help gathering extra information of vulnerabilities detected in the system. RAE also considers some algorithms based on risk pattern modelling.

RAE is made up of the following components:

- **Indicator value generator:** this component is responsible for translating the data gathered (information about the company profile, events and alarms of the monitoring module and vulnerabilities found) into indicators, i.e., information, which the models can understand. Data will be stored in a data warehouse where it is accessible for the rest of the components.
- **Triggering detector:** it notices any change in the inputs for the models, which mean that rules will be triggered or not depending on the scenario.
- **Instantiators** (DEXi model and R model): they are in charge of creating new instances of the models. For that purpose, they use indicators as inputs.
- **Executors** (DEXi model rules and R model rules): as the name suggests, they execute the rules to assess the model. Results are sent to the aggregator component:

- DEXi model rules executor, which runs qualitative risk assessment based on the DEXi model file which has been previously defined.
 - R-model rules executor, which runs quantitative risk analysis by means of R scripts.
- **Aggregator:** adds and integrates results and produces the risk assessment of the system.
 - **Data warehouse:** it is the storage component of RAE and it consists of two database backend services: a relational database and a document-based store. All the data necessary for the risk calculation is kept here, such as profile information of the organization, configuration parameters defined by the user, risk catalogues, indications and reports, any event gathered by the sensors and findings of the vulnerability scanners.

RAE works in real-time once it is launched but it does perform a risk assessment based on one of the following scenarios:

- i. At the user's will. For example, a vulnerability scan is launched or the questionnaire with information of the company is filled.
- ii. In a 'semiautomatic way', that is when:
 - There is a change on the indicators managed by RAE. For example, the user decides to change some of the answer(s) provided in the questionnaire.
 - The user changes the model to work with.
 - After a vulnerability scan is completed, the value of some indicators varies.
 - Events or alarms monitored change.

With regard to security, RAE employs HTTP REST API as a common interface for communications: HTTP REST calls (GET, POST, PUT) are the way the software can send and receive information, while JSON is the format of the objects in transit. Any request is secured with HTTPS and the OAuth2 authentication standard.

It is worth mentioning the Decision Support System (DSS). This module performs some important tasks such as displaying to the user the results of the RAE. It also provides the report to the user, in which several mitigation measures are normally included (depending on each case) as well as some analytical features, which can clarify the results provided by RAE.

Finally, regarding its integration with the GEIGER platform, RAE can assist on privacy assessments. The tool is being actively enhanced with the aim of:

- a) Be smoothly integrated in the GEIGER application environment
- b) Provide information concerning privacy risks in a friendly way for the end-user.

1.3.3 Information Sharing Platform (ISP)

Nowadays, the Malware Information Sharing Platform (MISP) standard helps obtaining information related to security incidents. However, it does not manage and control what information is sent through the network.

The Information Sharing Platform is a framework developed by ATOS, which main objective is to allow a quick and safe exchange of sensitive information. It can also compile and process OSINT information from local systems. The basis for ISP are the MISP instances and the framework adds an extra security layer to achieve a higher control over data exchange. The objectives of ISP include restricting access to threat and intelligence information to authorised personnel only, the creation of policies, easing management of information, and the secure sharing of cyber threat information.

The architecture of the tool is the client-server approach, while the server should be running on Linux OS and the client needs to be authenticated prior to the access with the help of a centralized system. Both client and server have been developed to be as lightweight as possible. On a component perspective, the Information Sharing Platform has been written with Python language, while Mongo DB is the storage system and Flask is

the minimalist Python framework for the creation of the web application. Figure 4 shows the architecture of the platform:

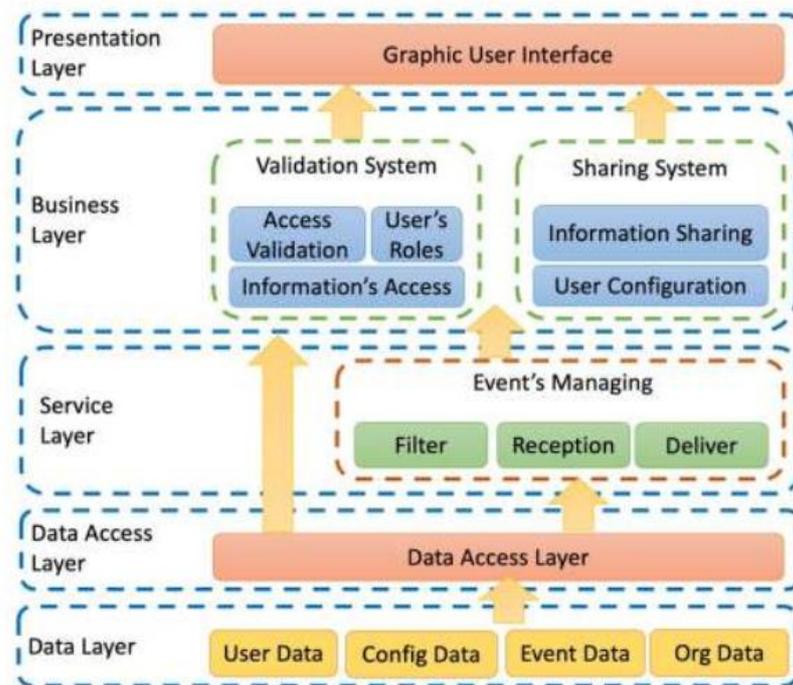


Figure 4 – Architecture of the Information Sharing Platform

- The presentation layer contains the Graphic User Interface (GUI), whose main functions are both displaying information and allow for the user to introduce any needed data. This dashboard relies on a Bootstrap validation that increases accuracy of information provided by the user.
- The business layer is responsible of the logic within the platform. It manages information but also shares it accordingly, that is, once the role and permissions of the user have been verified, the business layer will provide information requested by him/her. It is important to outline that data are encrypted while in transit, so communications are safer.
- The purpose of the service layer is to filter, deliver, and receive events. To accomplish this objective, it needs to employ the API provided by the connected MISP instance.
- Finally, data layers are designed to contain the knowledge base where all information is stored. This includes user, event, configuration as well as any other data, which the organization labels as relevant.

We are extending this tool in GEIGER by allowing the filtering of specific information for MSEs and two specific functionalities:

- Extraction of specific information and mapping to the needs of GEIGER,
- Creation of MISP files to share with other CERTs.

We are working closely with CERT-RO to create a specific structure of the information to be shared, so it can then be extended for using with other additional CERTs we want to connect with. The extraction of the information is done by identifying the elements of a MISP file that are necessary for the GEIGER Indicator or user experience, facilitating the process and usage of this information. We will also create an API that will facilitate the use of these functionalities by any tool or component of GEIGER (or to be added in the future) so we can further extend this functionality and facilitate its integration with other solutions.

1.3.4 Conversation Module

This is a modular Chatbot platform aimed at interacting with the employees of the MSEs. This cybersecurity awareness solution has been developed by KPMG and provides administration abilities to manage the platform.

The motivation behind the creation of the Conversational Module is to create an automatic and self-sustain interaction with the client (MSE's end user), in order to understand recurring security issues, predict future threats, detect patterns of security exposing actions, and transfer the results through the KPMG-Proxy Module to the GEIGER Cloud. This information will be used, as described previously, by the GEIGER Information Sharing Platform (ISAC) for generation of MISPs files. This information will also be used for the analysis and calculation of the MSE-specific threat score, which is part of the GEIGER Indicator.

The Conversation Module is divided into two main sections:

- i. Active interaction: Terminology describing a situation where the end user (MSE's end user) initializes the conversation with the chat bot to understand and share the threats they are under.
- ii. Proactive interaction: Describing a scenario when the Conversation Module itself initializes the conversation in front of the client (MSE's end user) to warn and guide them about the relevant threats.

Storing and transmitting data:

As for the information that it passes on from the Conversation Module:

- After the client has completed the process in front of the chat bot, we transfer the collected information to the GEIGER Cloud so it can then be used by the GEIGER ISAC via the KPMG-Proxy module for future analyzations, calculations and generation of MISPs files for sharing with CERTs, CSIRTs or any other entity.
- Some of the information is stored in the GEIGER Cloud Data Storage for further analytics and analysis, processed and used by the GEIGER Indicator.
- The private MSE's users and devices information is stored locally under the GDPR policy.

As for the information that passes to the Conversation Module:

- KMS-SDK⁷ sends to the Conversation Module regular updated information regarding current threats that are relevant to the client, and as a result, the proactive interaction begins.

Triggers for invoking the Conversation Module

- i. Active interaction: The MSE's end users invoke the chat bot whenever they want to start a new conversation. When the client tries to communicate with the chat bot a new session begins, which will continue until the end of the conversation process, or alternatively when the page session ends.
- ii. Proactive interaction: The Conversation Module maintains a listener mechanism that listens to the KMS-SDK, which feeds it with the new updates regarding the current threats that are relevant to the specific client. Whenever the Conversation Module gets updated by the KMS-SDK, a new session begins, and the Conversation Module initiates a new conversation with the client (MSE's end user).

The Conversation Module is made up of a client platform connected to backend servers deployed on the cloud. The Conversation Module requires an internet connection to perform:

- The client is expected to access via a web browser.
- Updates are provided based on the internet connection.
- The server has been deployed on cloud, and PaaS is provided as the chosen architecture.

⁷ See section 1.3.10 for more details on this component.

1.3.4.1 Conversation Module Sequence Diagrams

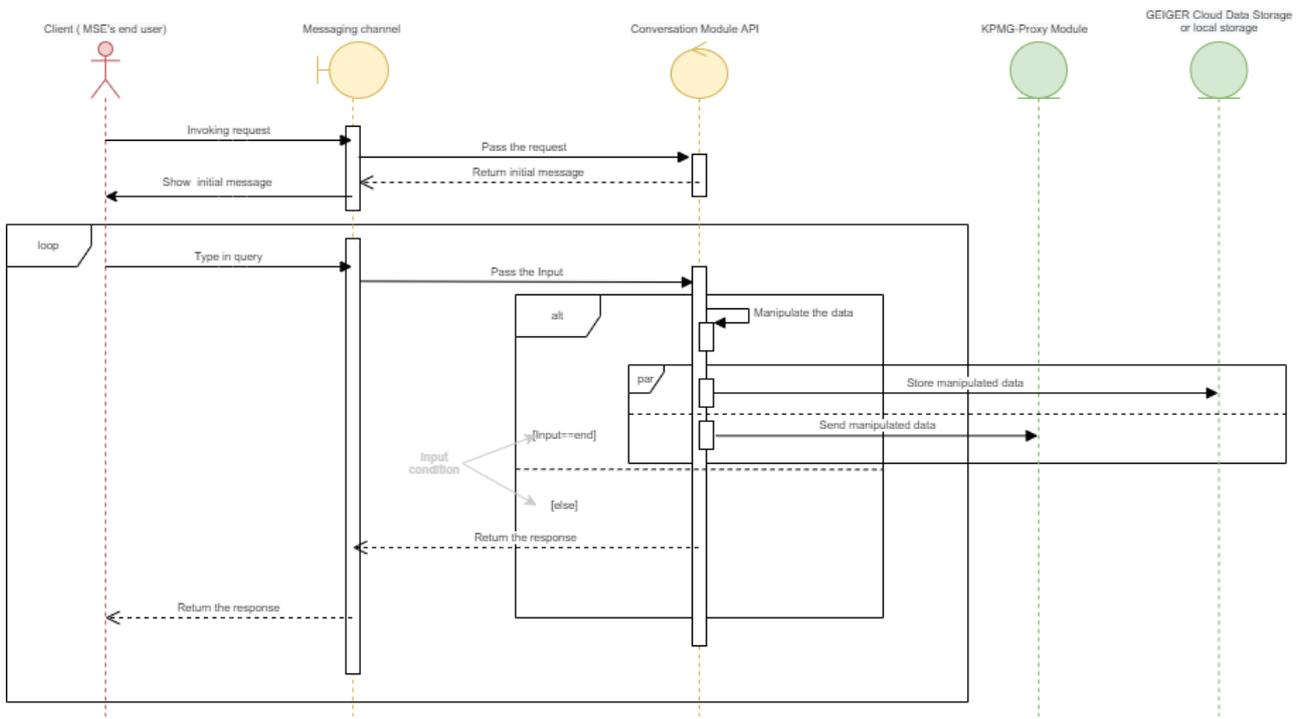


Figure 5 – Conversation Module active interaction - Sequence Diagram

When the MSE’s end user initiates a conversation in front of the chat bot, a request is sent to the Conversation Module API. A new session begins and a response containing an initial message is sent to the client. When both parties reach the end of the call, the session ends, and the conversation content is processed and transmitted simultaneously to both KPMG-Proxy Module and to the Cloud Data Storage for future analysis.

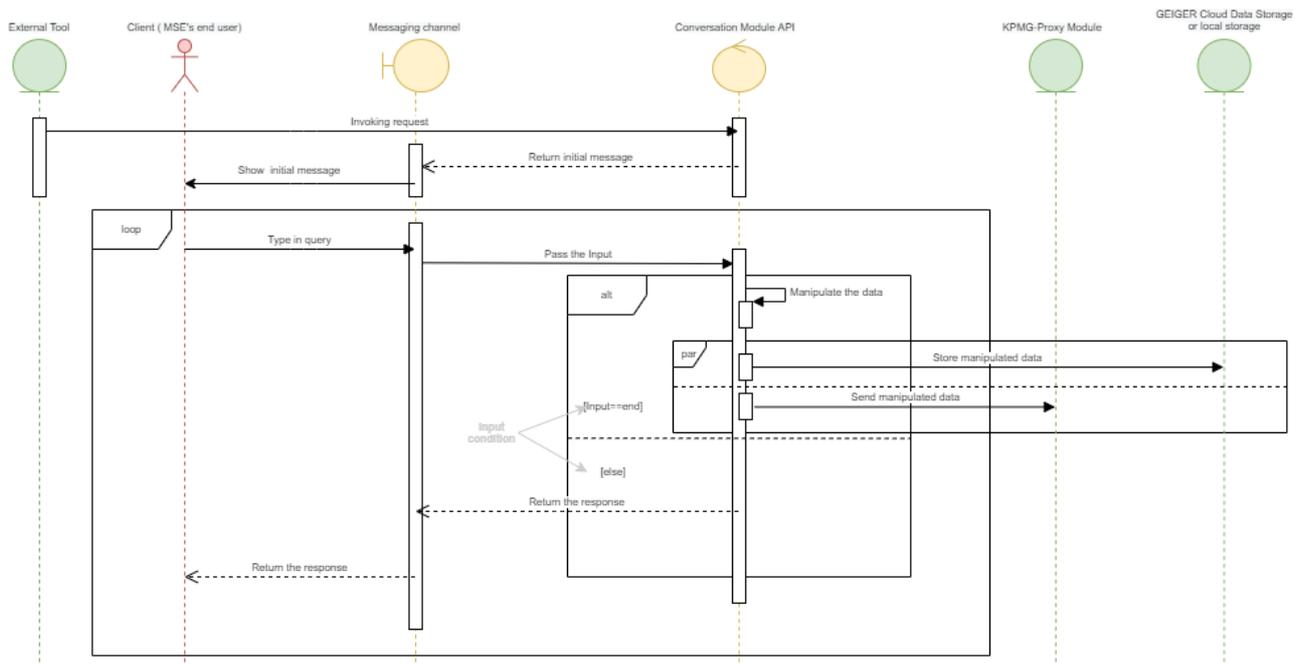


Figure 6 - Conversation Module proactive interaction - Sequence Diagram

When the KSP-SDK sends an alert notification to the Conversation Module, it initiates a conversation in front of the client. A request is sent directly to the Conversation Module API, a new session begins and a response containing an initial message is sent to the client. When both parties reach the end of the call, the session ends, and the conversation content is processed and transmitted simultaneously to both KPMG-Proxy Module and to the Cloud Data Storage for further analysis.

1.3.4.2 Conversation Module Package Diagrams

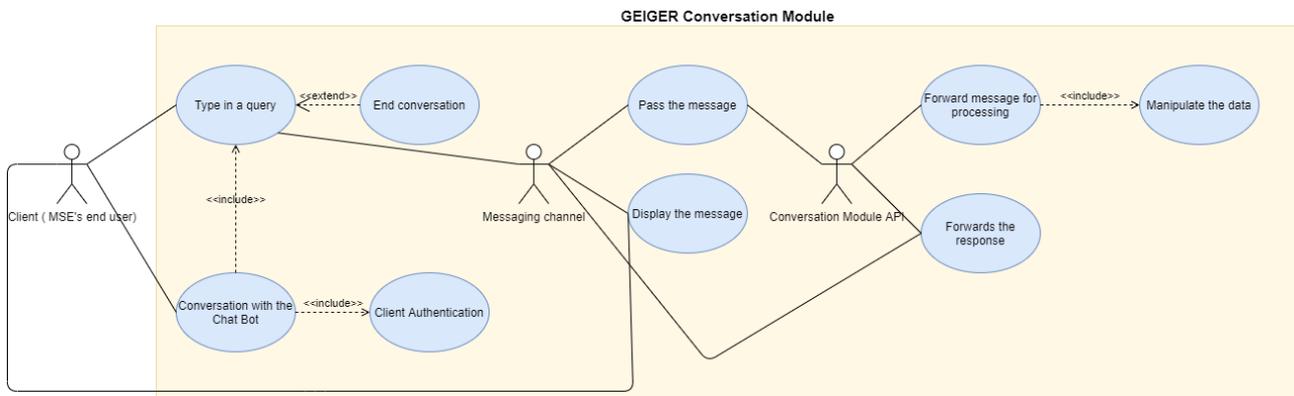


Figure 7 - Conversation module – Use Case Diagram

When a conversation session begins, the client (MSE's end user) has to type a query in front of the chat bot. The conversation continues until the client ends the call, or alternatively when the Conversation Module reaches a saturation point where it has enough information regarding the call. On each interaction, the query is sent to the Messaging Channel, where the query is processed and its content passed on the Conversation Module API where the next answer is processed and sent back to the client. At the end of the call, the final data set is manipulated and delivered to the KPMG-Proxy Module.

1.3.5 KPMG Proxy Module

This module is designed as a proxy between the parties enabling manipulation, extraction, and processing of data, while ensuring format consistency and efficiency improvement between the producer and consumer parties. It serves as a file transfer system for exchanging, sharing or transmitting the manipulated data between different services over a network or internet connection.

The KPMG-Proxy Module has two main functionalities:

- I. A process designed for manipulating, processing, and extracting data from the source system and making it compatible with the destination system before writing into it.
- II. Acting as a communication proxy channel for sharing, transmitting, or transferring the manipulated data between the systems.

1.3.6 Fraud Detection

It is a real-time solution developed by KPMG, whose purpose is to detect illegal operations or frauds such as transaction anomalies, money laundering and questionable or prohibited relationships between employees and clients of financial institutions.

Fraud detection is offered as a standalone application, which can be deployed either on the cloud or on premise. Its architecture follows the client/server model: the client can access via a web browser, and the server is deployed on the cloud, making use of PaaS. The server requires some logic to work and detect fraud, i.e., it relies on some Artificial Intelligence (AI) infrastructure, which can recognize patterns of illegal or suspicious activity.

1.3.7 Document Harvesting

Document harvesting is the proposed solution of KPMG to gather information from written sources of the MSEs and identify frauds and risks. It is based on both machine learning (ML) and Artificial Intelligence (AI) to learn from the MSEs' documents and entities within them to extract reliably and automatically information from large sets of documents.

The solution either can work on cloud or on premise, depending on the MSE's needs and it has been developed following the client/server approach, where the server is deployed on the cloud and offered as PaaS. It relies on AI to recognize traces of activities and detect fraud. On his behalf, client is required to access via web browser.

1.3.8 Employee Virtual Assistant

This is a solution oriented to call centre departments for call orchestration and call transcription analysis developed by KPMG. The software is modular and combines call centre administration abilities to manage call diagnostic in real time. Amongst its capabilities, it allows call automated transfer, first response and remediation, analytics in real time and call diagnosis based on transcriptions.

Employee Virtual Assistant has been built as a backend and frontend solution, which can be run either on the cloud or on premise. While the server requires some containerization for the deployment, the user just needs a web browser to access. Call management is performed by means of APIs and call analytics is achieved with the help of AI tools, which can sometimes be intense in terms of resource consumption. On the other hand, call orchestration is a bot process meant to be lightweight as far as processing capabilities are involved.

1.3.9 Kaspersky Interactive Protection Simulation (KIPS)

KIPS is a software developed by KSP whose purpose is to provide experiential training to users, including prevention and mitigation of threats, and training with controls. The tool can be classified as a learning software provided as a service, which requires little configuration to be employed. It follows a client/server architecture where the server is deployed on the cloud and the user (client) can access via web browser. KIPS tool requires an active internet connection to work and employs JSON format as a means of sending scores to third parties.

1.3.10 KMS-SDK

Kaspersky Mobile Security Software Development Kit or KMS-SDK is a software development kit (SDK) to be integrated in a mobile app to help preventing and detecting cyber threats. This tool is focused on the mobile device protection field either in Android or iOS. Although the SDK does not need internet connection to work, signatures will not be updated if not connected periodically to the internet.

KMS-SDK provides data protection as well as several others privacy mechanisms and employs AES-256 for data encryption.

1.3.11 CyberSafety Management Games (CSMG)

CSMG is a training tool developed by Kaspersky whose aim is to help users increasing their cybersecurity knowledge and awareness. The software helps the user to learn and acquire best practices in usual workday situations in an environment where cyber threats are constantly evolving. CSMG is a learning game, which can check user's knowledge in the security field.

CSMG has been designed following a traditional client-server architecture but, while the server is deployed on the cloud, client, i.e., user, can access by means of a web browser so an internet connection is required. CSMG can also share the scores with third parties so it employs JSON standard for that purpose.

1.3.12 Montimage IDS (MIDS)

As the name suggests, the MIDS⁸ is a network monitoring solution, which can passively analyse network traffic to detect potential threats and anomalies. The tool can also engage with other security tools, given its capacity to correlate data coming from external sources. Although there is a standard or default configuration, user can adjust some parameters to fit the end user needs so it allows a certain flexibility.

Montimage IDS is based on a client/server architecture. It can be deployed on the end user device but in this case, it is mandatory for the user to run Linux OS or it can be deployed in a standalone server (e.g., in the cloud) receiving mirrored traffic from the end user network or device. Despite being a monitoring system, the components have been built to be lightweight enough so it will not exhaust resources and processing computing on the client.

1.3.13 Cyber-Range

The Montimage Cyber-range is a cyber game and training tool, which can help the user learning about phishing attacks and other cyber threats. The application takes care of the data of the user because it includes some privacy mechanisms. It seeks to increase the cybersecurity knowledge and awareness of the user and focuses on some the most important attacks companies can suffer. The Montimage Cyber-range is offered as a mobile app with a client/server architecture available both on Android an iOS platforms and is being developed with lightweight components.

⁸ http://montimage.com/products/MMT_DPI.html

2. Architecture

2.1 Overview

GEIGER has been conceived as a platform with the purpose of helping users to increase their cybersecurity knowledge and awareness as well as to know, at a glance, the level of risk the business is exposed. Considering these objectives, the high-level architecture shows two clearly different environments:

- a) The Cloud, where data are available anywhere and everywhere. Tools can feed the GEIGER Cloud at any moment, so there are no limitations.
- b) The GEIGER Toolbox or a local environment, where processing is performed. The toolbox is designed to provide information to the device where GEIGER is installed.

Nowadays not only big companies but also – and specially – MSEs are a target for cybercriminals. MSEs have fewer resources and, therefore, protection against sophisticated threats such as zero-days, ransomware attacks, phishing, SQL injection or, even, crypto jacking, not to mention viruses, trojans, and worms is not a piece of cake. However, GEIGER, helps fighting and mitigating the effects of these threats mainly in two ways:

- By having more updated information of threats and attacks provided by CERTs and CSIRTs. This information can help users be more conscious of what the risks are and how attackers typically try to delude users.
- With the help of both GEIGER Infrastructure and GEIGER External Tools: some of them can effectively help preventing attackers and threats achieving their target and, therefore, getting a higher level of protection for the MSE.

The proposed architecture mixes effectively those mentioned resources to actively help both preventing and fighting cyber threats for MSEs.

2.2 Architecture of GEIGER

The GEIGER architecture is represented in Figure 8:

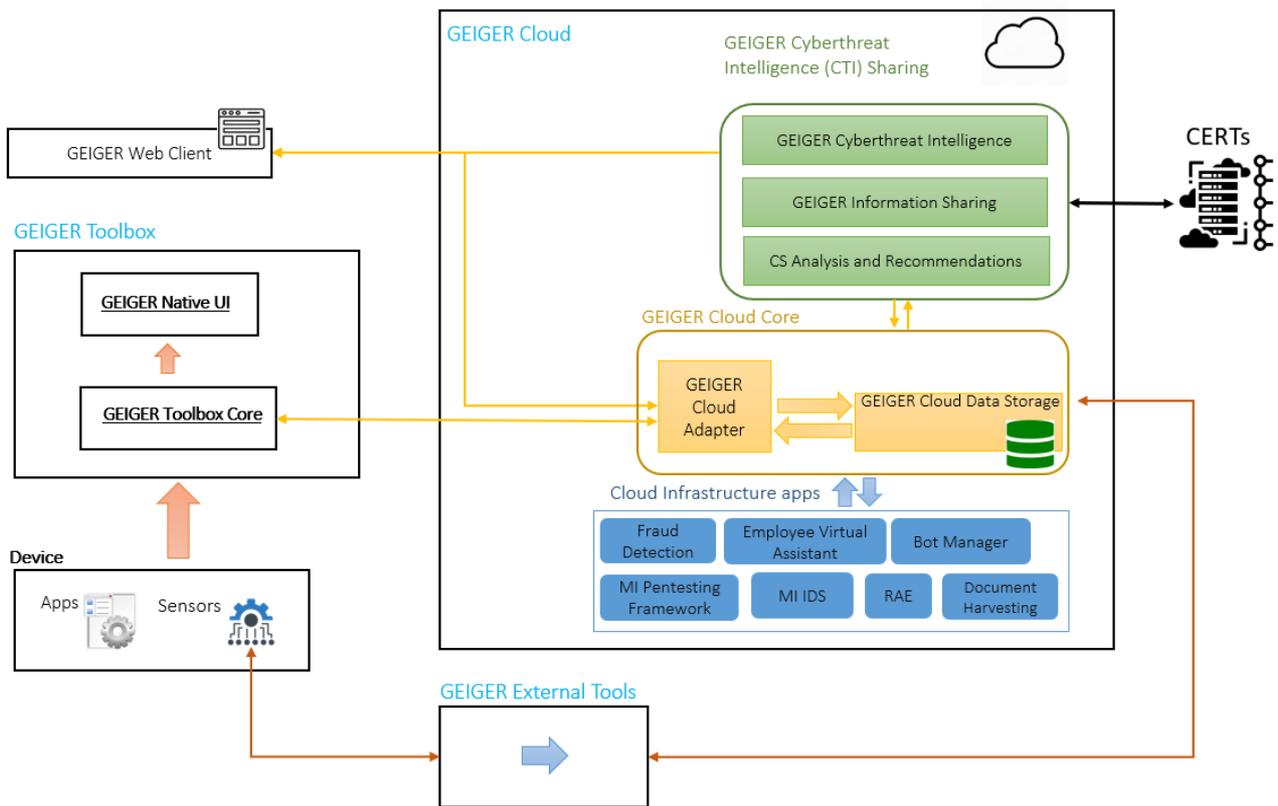


Figure 8 - GEIGER Architecture

The GEIGER platform is based on the interaction of several components:

- i. The **GEIGER Toolbox**, the place where risk score calculations are performed, includes:
 - The **GEIGER Logic and analysis** sub-component is embedded into the GEIGER Toolbox and will be responsible for executing the algorithm which will estimate the level of risk.
 - The **native User Interface (UI)** is a point of contact with the end-user. It has been designed in a modern, friendly, and user-oriented way to ease interaction. It is also a way for end-users to introduce some information on the GEIGER platform.
 - **Sensors** of external applications are deployed in the **GEIGER Device**. These sensors gather information from the device and send it to their private server, given that these applications follow a client/server architecture. Destination of these data is the GEIGER Cloud Core.

GEIGER Toolbox

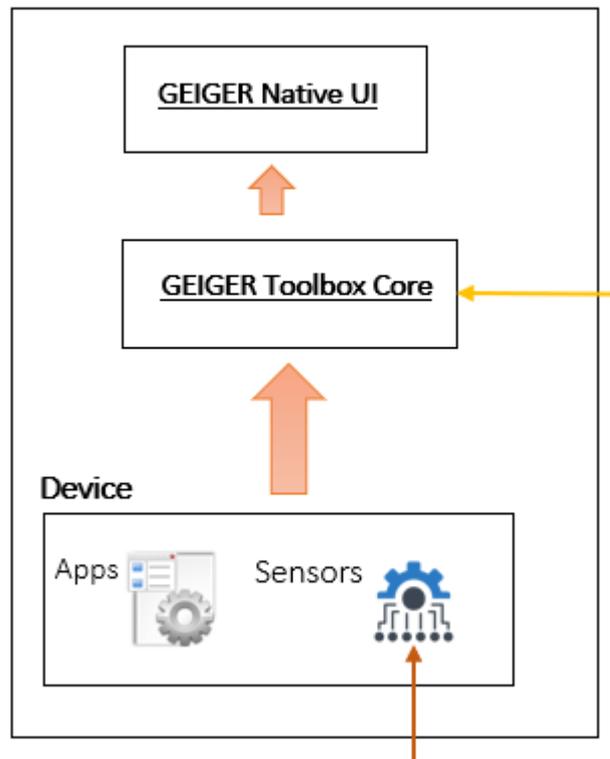


Figure 9 - GEIGER Toolbox architecture

- ii. The **GEIGER External Tools** include all those proprietary tools, which will **not be deployed** into the GEIGER platform. Their respective private server will receive information sent from the sensor deployed on the device. Then, the server is responsible for providing data to the GEIGER platform by means of directly reaching the GEIGER Cloud Core. The Cloud Data storage will be the place where keep all this data. A high-level of data protection and authentication is integrated in this communication.
- iii. The **GEIGER Cloud** is a complex component made up of:
 - The **GEIGER Cyber Threat Intelligence (CTI) Sharing**: as part of the GEIGER Cloud, it performs some functions such as:
 - gathering information from CERTs and CSIRTs,
 - providing data collected to be stored on the GEIGER Cloud Core,
 - offering security recommendations by means of the Cybersecurity (CS) Analysis and recommendations subcomponent.
 - The **GEIGER Cloud Core**: it hosts the cloud storage, which is the place where data gathered from CERTs, CSIRTs, external and internal (or integrated) GEIGER tools is stored. This cloud repository can provide information to the toolbox when requested by means of the GEIGER Cloud Adapter, which is a bridge for connecting the cloud part with the toolbox.
 - The **GEIGER Cloud infrastructure apps**: this includes all applications expected to be integrated into the GEIGER platform, i.e., deployed on the GEIGER servers. Cloud infrastructure apps provide information to the GEIGER Cloud, which is necessary for performing the risk level calculation.
- iv. CERTs and CSIRTs. These external entities interact with GEIGER and provide:
 - Updates on threats, vulnerabilities,

- information of recent attacks and security events,
- any other data which, may be relevant to perform the risk score calculation.

The point of contact with the GEIGER platform is the CTI Sharing component.

Additionally, GEIGER will provide data to the CERTs in the way of MISP files with cyber threat information of the MSEs.

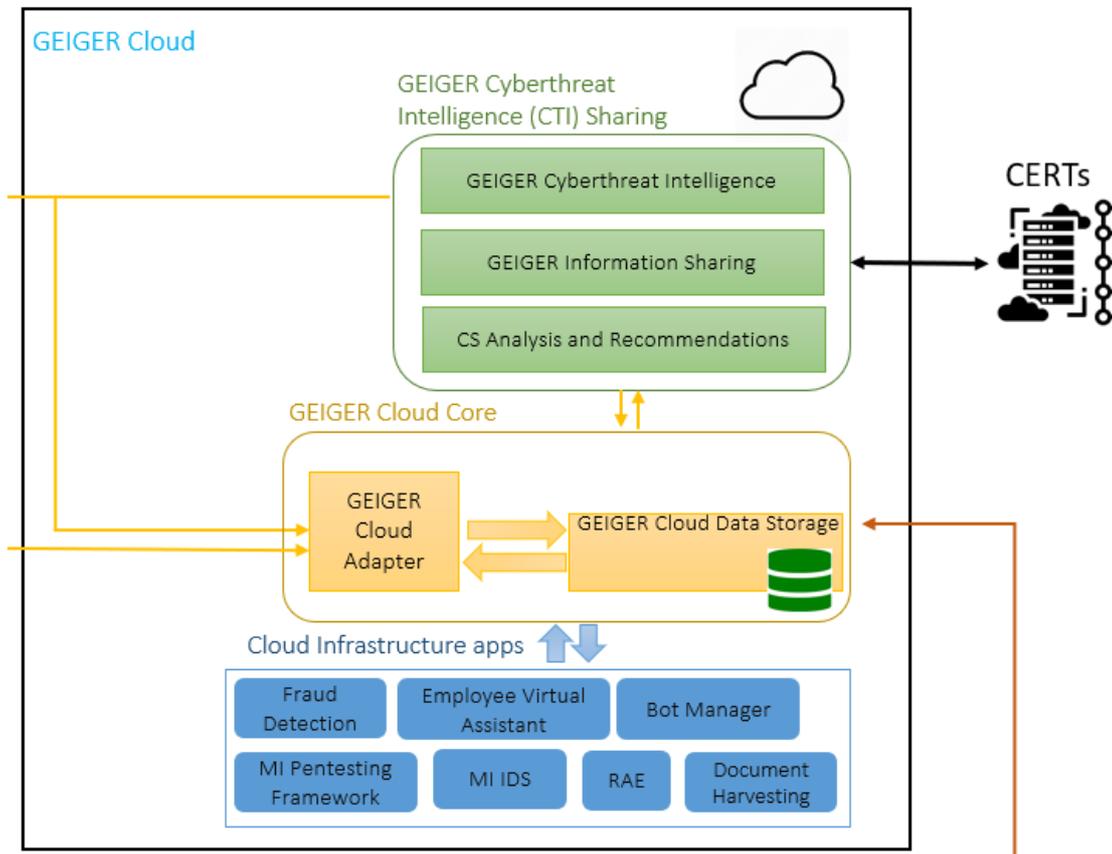


Figure 10 - GEIGER Cloud architecture

- v. **GEIGER Web client:** this interface is a place where the end-user can interact with the GEIGER platform. Although it provides quite interesting possibilities such as user registration, its functionality is reduced when compared with the possibilities of the native UI. Its main functions are related to:
- User registration,
 - provide access to GEIGER information,
 - show notifications to the end-user.

2.3 GEIGER Scenarios

The GEIGER architecture can be defined according to three different scenarios:

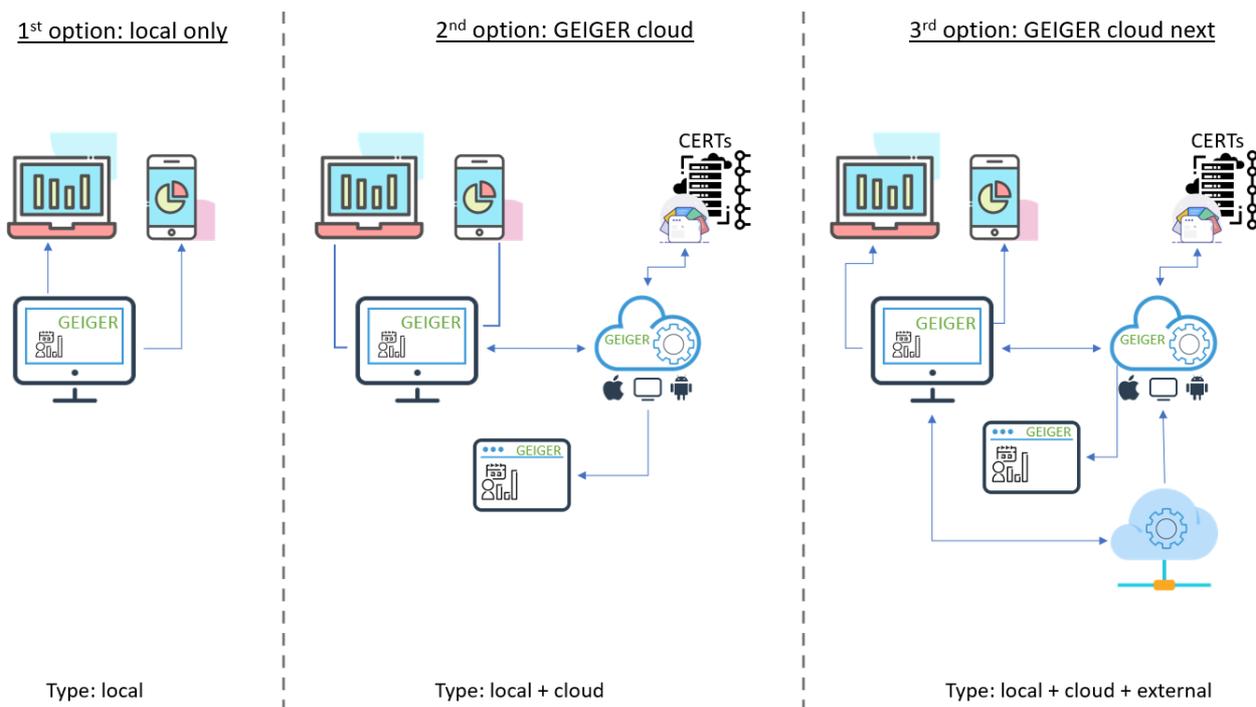


Figure 11 - GEIGER scenarios

2.3.1 Local Only

In this scenario, the end-user employs GEIGER without assistance of the GEIGER Cloud component. It is defined by the following points:

- ✓ All the processing and calculations of the risk score are to be performed on the GEIGER Toolbox, more specifically, on the GEIGER Logic and analysis sub-component.
- ✓ Information needed for the risk score calculation comes from the data stored locally.
- ✓ With the course of time, **reliability on the calculation decreases** if the data stored is not refreshed, that is, updated with GEIGER Cloud information.
- ✓ Risk score calculation is passed to the GEIGER Indicator to be displayed to the end-user.

The local only scenario is feasible to be employed with different devices such as tablets, laptops, or mobile devices. It is important to emphasize the limited functionality of the local-only mode, given that GEIGER is based on updated information.

2.3.2 GEIGER Cloud (Cloud + local)

The GEIGER Cloud scenario engages both the local and the Cloud components of GEIGER. This approach mixes effectively features from both components to make GEIGER more flexible and updated compared with the local only scenario. It is based on the following premises:

- ✓ Processing and calculations of the risk score are performed on the GEIGER Logic and analysis sub-component, part of the GEIGER Toolbox component.
- ✓ CERTs and CSIRTs are an active part of this scenario. Therefore, they are responsible for providing **updated information of threats and vulnerabilities** to the GEIGER platform. Data is stored in the GEIGER Cloud and requested by the GEIGER Toolbox for the risk score calculations. One of the strengths of GEIGER lies on the fact that information is periodically updated as new threats and vulnerabilities are being discovered daily.

- ✓ In addition, there are some **tools integrated into the GEIGER platform that provide information**. With the advantage of the deployment of these applications into the GEIGER servers, data will be provided continuously or by demand.
- ✓ The GEIGER **Indicator displays the level of risk to the end-user**. Besides, the GEIGER Cloud can provide updated data when requested.
- ✓ There is also a friendly, user-oriented UI where the user can interact with the platform.

2.3.3 GEIGER Cloud Next (Cloud + local + external)

This scenario renders a slight variation to the GEIGER Cloud (Cloud + local). The difference comes from the consideration of external apps, which will not be deployed into the GEIGER servers. The most important points of this architecture would be:

- ✓ Risk score calculations are performed by the GEIGER Logic and analysis sub-component.
- ✓ Information for the risk score calculations can be obtained from **different sources**:
 - CERTs and CSIRTs provide information of threats and vulnerabilities.
 - Applications deployed in the GEIGER servers gather and store data into the Cloud storage.
 - Furthermore, external applications provide information to be stored into the Cloud storage. These apps have their sensors deployed on the GEIGER Toolbox for the purpose of gathering data. Communication is established in a typical client/server approach between the sensor of the app and its private server. Then the server is the one responsible for sending the information to GEIGER.
- ✓ Graphical web interface is also available on this scenario, as well as the GEIGER Indicator for the purpose of displaying the information to the end-user.
- ✓ The introduction of external applications provides more flexibility to GEIGER, given that they gather data directly with their sensors.

2.4 Functionality

Following we describe the different information flows we have identified in the GEIGER architecture.

2.4.1 Information of the Apps Running Locally

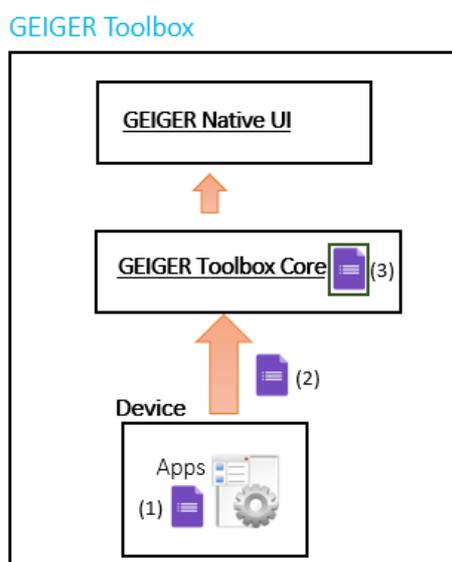


Figure 12 - Information of apps running locally flow

In this flow, the apps generate data locally, i.e., in the GEIGER Toolbox. This information is sent to the GEIGER Toolbox Core, where it will be stored in the GEIGER local data storage. Information from the local data storage will also be used afterwards by the risk score algorithm to calculate the level of risk.

2.4.2 Storing Data of Results of Apps in the Cloud

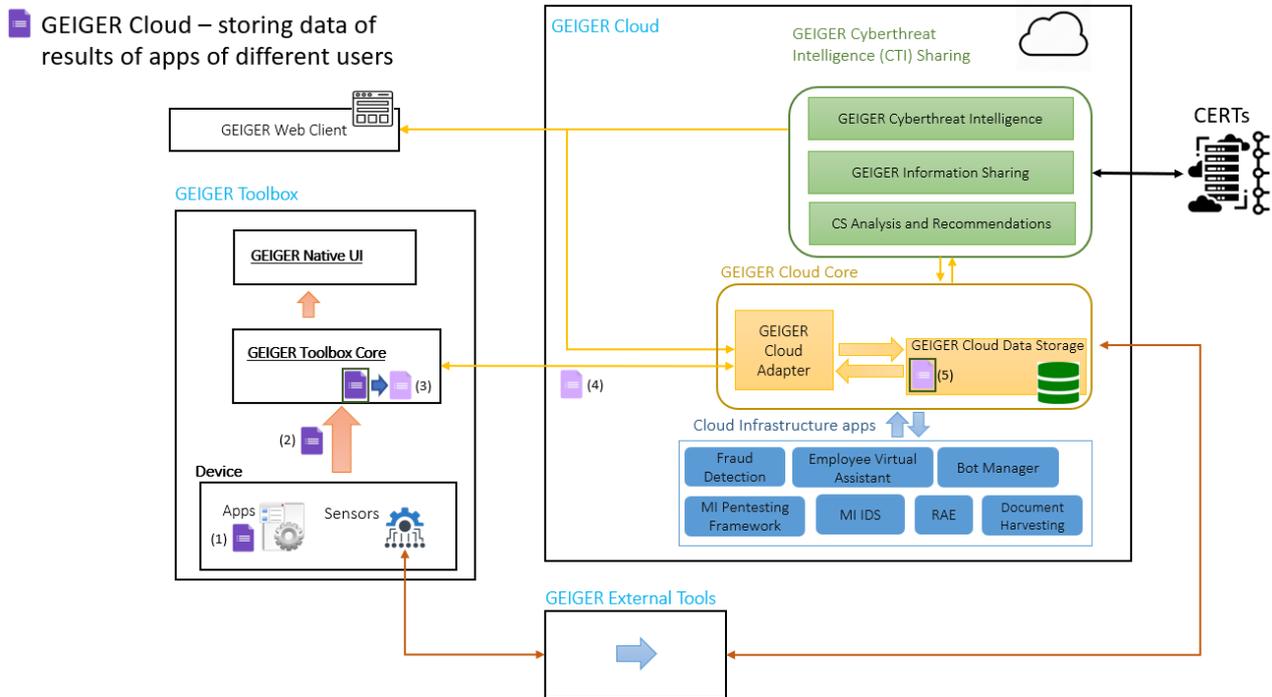


Figure 13 - Storing data of results of apps in the Cloud flow

This flow shows how data obtained from calculations in the apps are stored in the GEIGER Cloud:

- i. Applications produce some data as a result of their operation. This information is generated in the device subcomponent of the GEIGER Toolbox.
- ii. Data are sent to the GEIGER Toolbox Core, where this information is analysed, processed and stored in the local data knowledge base.
- iii. By means of both the GEIGER Controller and GEIGER Cloud Adapter, data pass through the GEIGER Toolbox into the GEIGER Cloud.
- iv. Finally, data reach the GEIGER Cloud data storage, where it will be stored in the cloud data knowledge base.

2.4.3 Requesting data of results of apps

GEIGER Cloud – requesting data of results of apps of different users

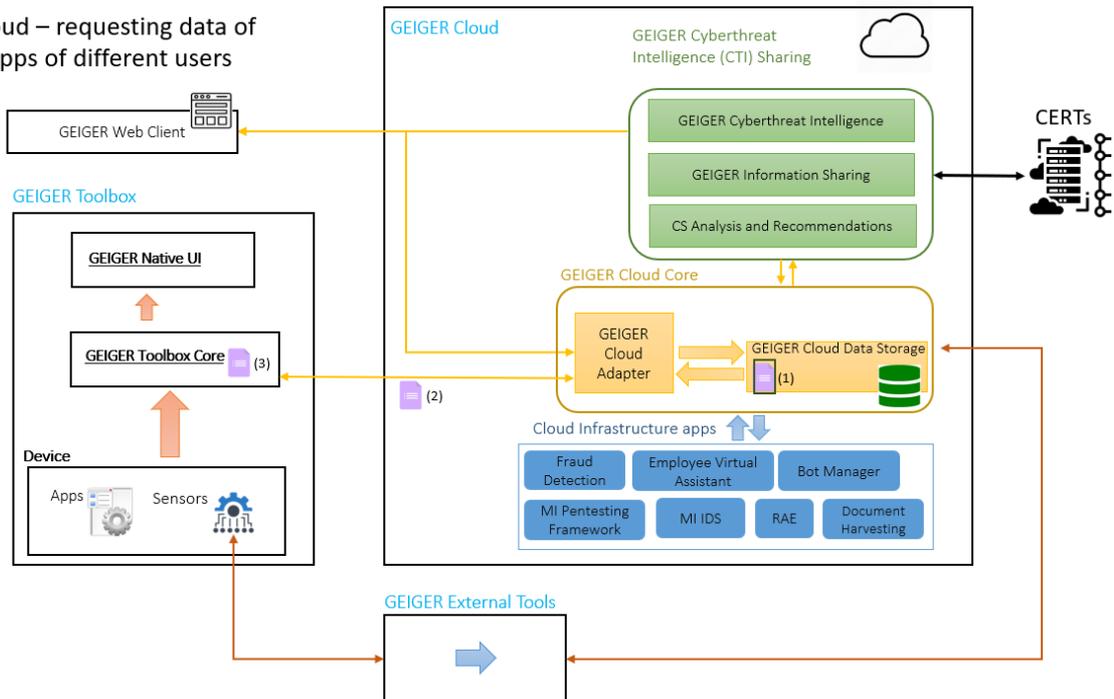


Figure 14 - Requesting data of results of apps flow

This information flow shows the way data stored in the GEIGER Cloud are gathered:

- i. GEIGER Cloud Data storage possesses both data provided by Infrastructure apps and External Tools. This information is needed to perform the risk score calculations in the GEIGER Toolbox.
- ii. Information passes both the GEIGER Cloud Adapter and the GEIGER Controller to be passed to the GEIGER Toolbox.
- iii. Finally, information reaches the GEIGER Toolbox Core where it can be employed to estimate the level of risk.

2.4.4 Information of the MSE (only local)

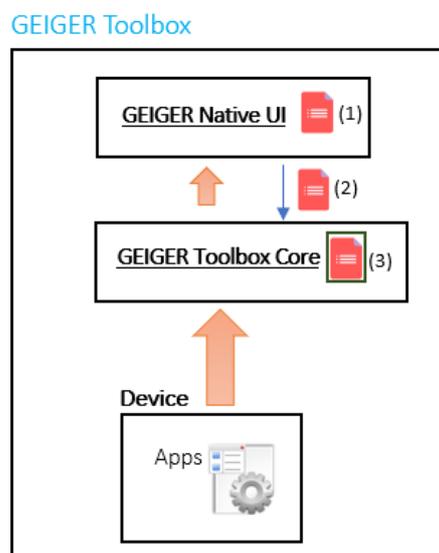


Figure 15 - Information of the MSE (local) flow

Information of the MSE should include various registers such as industry or data about the kind of environment where the MSE operates, organizational and geographical information, data from devices such as operating system, updates, employee assigned, etc. This data flow should follow the steps indicated:

- i. The end-user introduces information of the MSE with the help of the GEIGER Native UI.
- ii. Data are stored in the GEIGER Toolbox Core, where it is sent to the local storage for future use.

2.4.5 Storing Cloud-Relevant Data of the Company for Further Analysis

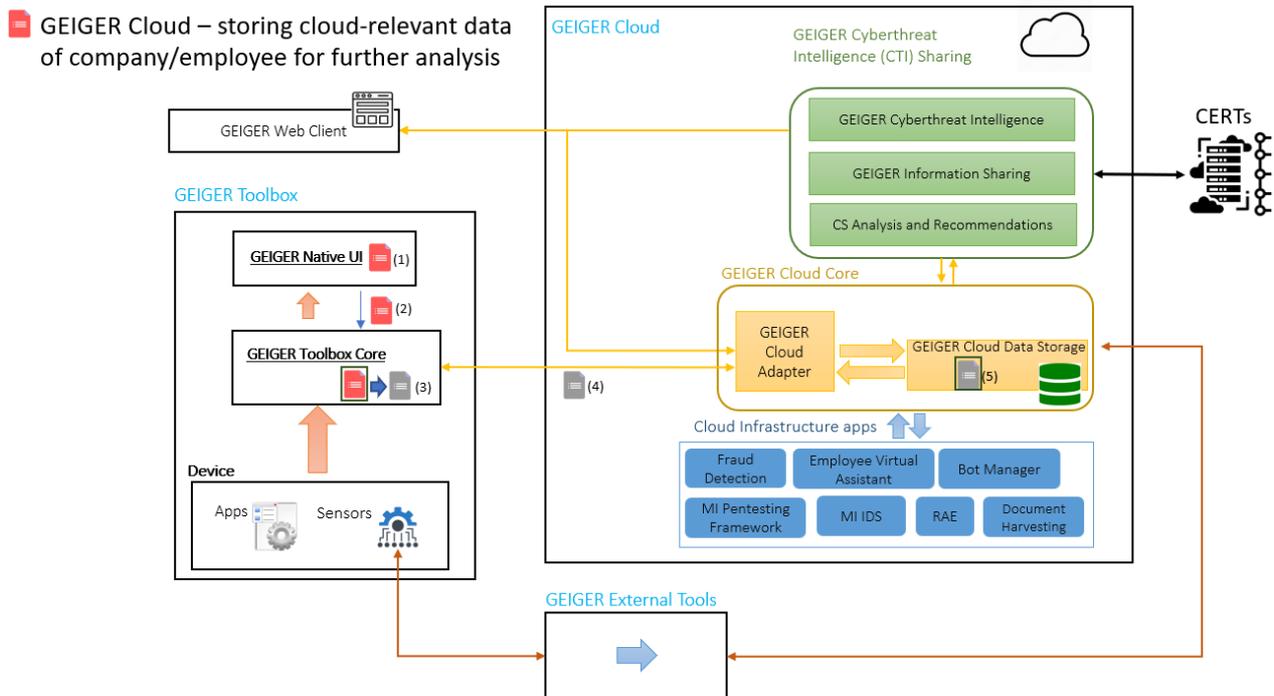


Figure 16 - Storage of company data in the cloud flow

As it has been described, the company owns its private information including data about their business environment, organizational and geographical information as well as data from the company's devices. The **MSE is responsible for indicating what of this private information will be also stored in the cloud (if any)**, a process described in this information flow:

- i. The end-user introduces information of the company into the system by means of the GEIGER Native UI.
- ii. Data are passed and stored locally in the local storage of the GEIGER Toolbox Core.
- iii. Then, data are sent to the GEIGER Cloud with the help of the GEIGER Controller and Cloud adapter.
- iv. Finally, the information reaches the GEIGER Cloud Core where it is stored in the GEIGER Cloud data storage.

2.4.6 Requesting Cloud-Relevant Data of Company for Calculation of Indicator

GEIGER Cloud – requesting cloud-relevant data of company for calculation of indicator

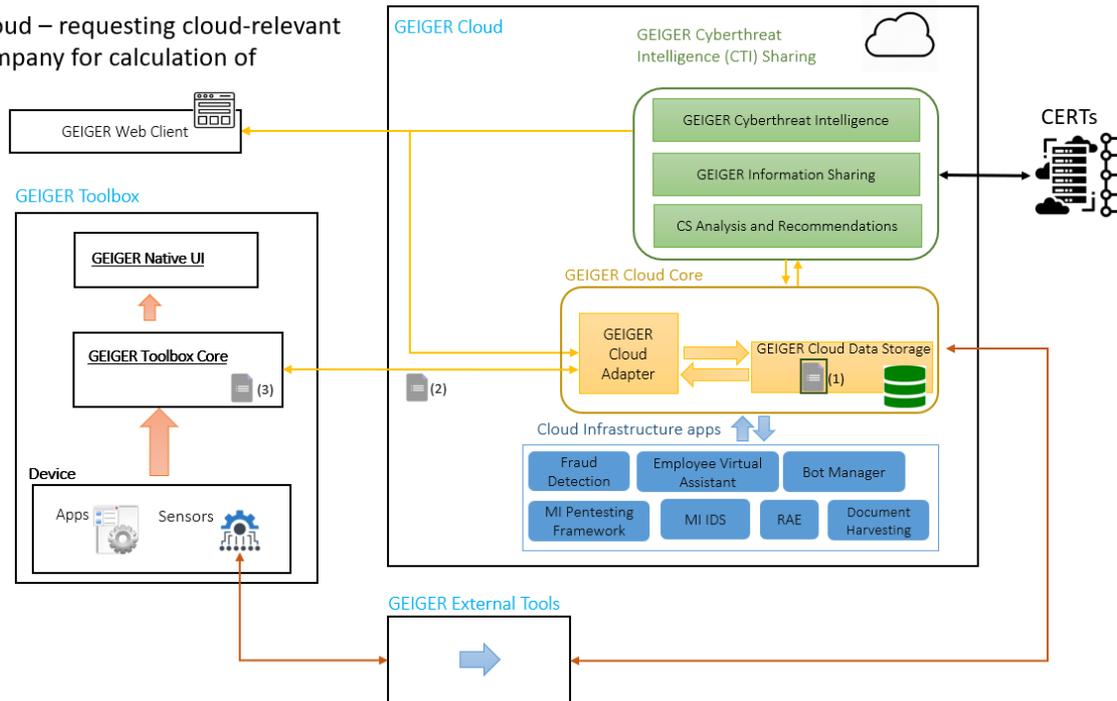


Figure 17 - Requesting cloud data of the MSE for risk score calculation flow

This scenario shows the way data stored in the GEIGER Cloud are requested by the GEIGER Toolbox with the purpose of helping estimate the level of risk:

- The Cloud stores updated information needed to perform the GEIGER risk score. This information is stored into the Cloud data storage.
- Data cross the adapters (Cloud adapter and GEIGER controller) and is passed to the GEIGER Toolbox Core.
- Once data are in the GEIGER Toolbox Core, it can be used by the algorithm in charge of calculating the risk score.

2.4.7 Storing Information of CTI

GEIGER Cloud – storing information of CTI

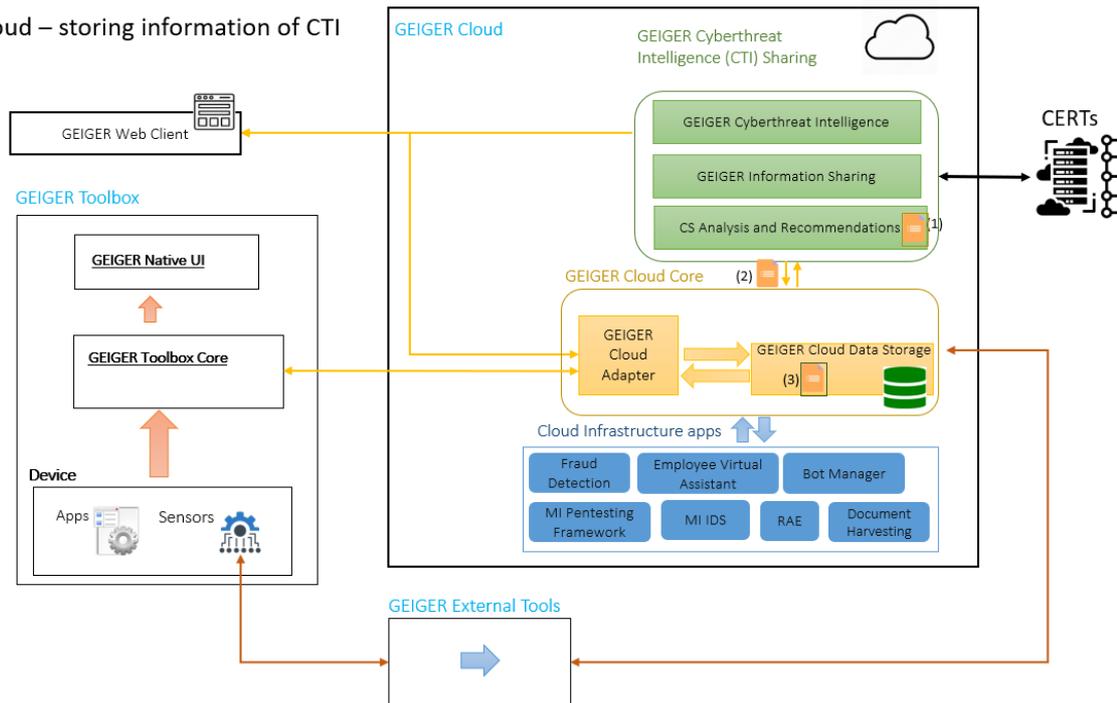


Figure 18 - Storing information of CTI flow

Cyber Threat Intelligence (CTI) Sharing component of the GEIGER platform is responsible for receiving and managing data, which is originated in the CERTs and CSIRTs outside the GEIGER environment. Given that there can sometimes be an overwhelming amount of information about new threats and vulnerabilities, the GEIGER platform is designed to store only which is necessary, that is, the essential pieces of data to perform the risk calculation.

Regarding the flow of information, this one describes how data coming from CTI are stored:

- Data are originated in CERTs and CSIRTs. This should include updated information of threats, vulnerabilities, attacks, etc. This information is sent to the GEIGER CTI Sharing.
- Information is then generated in the CS Analysis and recommendations sub-component of the GEIGER CTI Sharing component.
- Information is sent to the GEIGER Cloud Core.
- Finally, data are expected to be stored in the cloud, more specifically, in the GEIGER Cloud data storage, which is the online repository of the GEIGER platform.

2.4.8 Requesting Information of CTI for the Indicator

- GEIGER Cloud – requesting information of CTI for the indicator

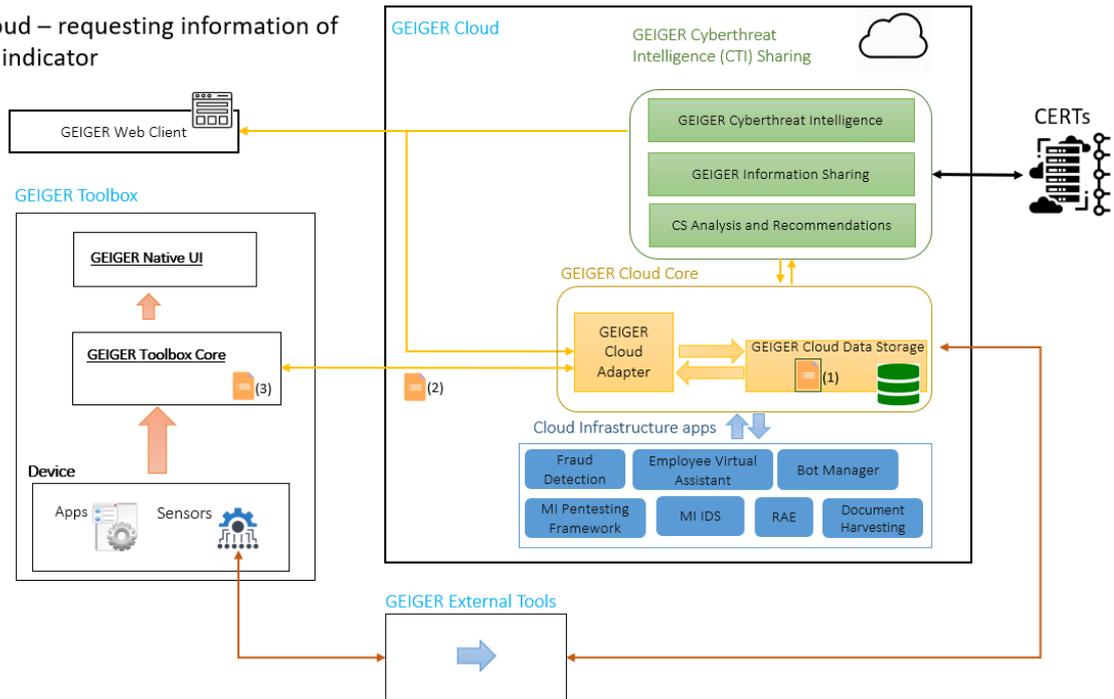


Figure 19 - Requesting CTI information for indicator flow

This information flow complements what has been described in the information flow *Storing information of CTI* (section 2.4.7) and describes how the information provided by the CTI is requested by the GEIGER Toolbox.

- As it has been previously described, updated information (of threats, vulnerabilities, attacks...) provided by CERTs and CSIRTs is kept into the Cloud data storage.
- The algorithm performing the risk score calculation needs these data when performing an estimation of the level of risk. Therefore, information traverses both the Cloud adapter and the GEIGER Controller to be routed and delivered to the GEIGER Toolbox Core
- Finally, data arriving the GEIGER Toolbox Core can be employed by the GEIGER logic and analysis component to calculate the risk level.

2.4.9 Storing Information of Infrastructure Apps in the Cloud

GEIGER Cloud – storing information of apps in the GEIGER Cloud

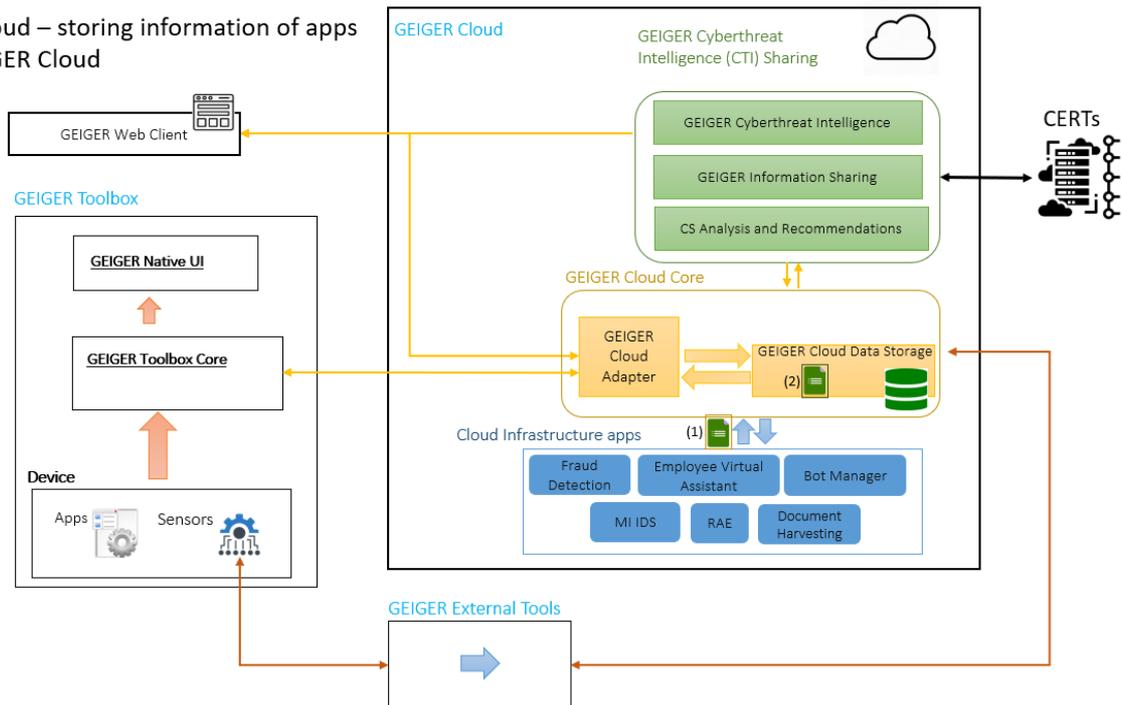


Figure 20 - Storing data of infrastructure apps in the cloud flow

GEIGER Cloud Infrastructure apps include all those apps, which are expected to be deployed in the GEIGER servers. The information flow is the following:

- i. Infrastructure apps produce some data as a result of their operation.
- ii. Data are sent to the GEIGER Cloud Core.
- iii. Finally, the information is stored in the GEIGER Cloud Data storage subcomponent of the GEIGER Cloud Core.

2.4.10 Sending Information of Infrastructure Apps to the GEIGER Indicator

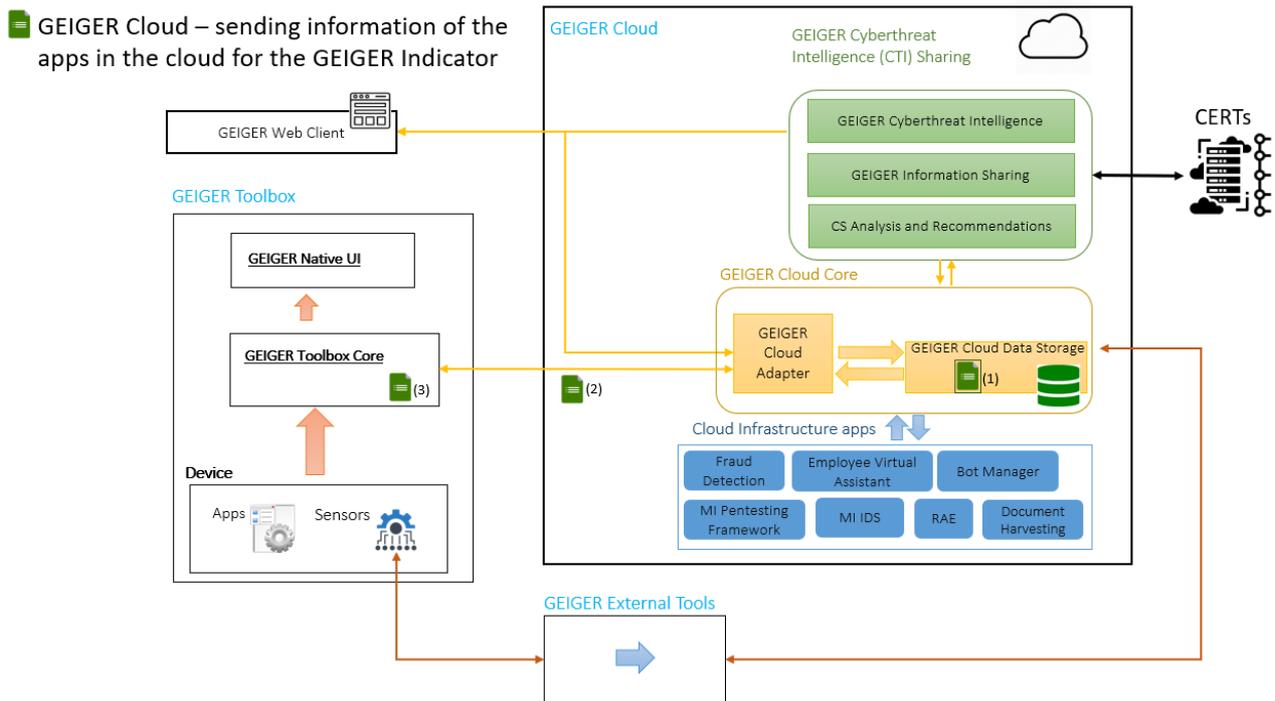


Figure 21 - Sending information of infrastructure apps to the indicator flow

As it has been described, Cloud infrastructure apps are the ones deployed in the GEIGER server, i.e., integrated into the GEIGER platform. These applications generate information, which is stored in the online repository (the Cloud data storage) of the GEIGER Cloud core:

- i. Information stored in the cloud is requested for the calculation of the level of risk.
- ii. Data traverse both the GEIGER Cloud adapter and the GEIGER controller.
- iii. Finally, information reaches the GEIGER Toolbox Core where it can be used by the algorithm responsible of the risk score calculation.

2.4.11 Storing Information of External Apps in the Cloud

GEIGER Cloud – storing information of external apps in the GEIGER Cloud

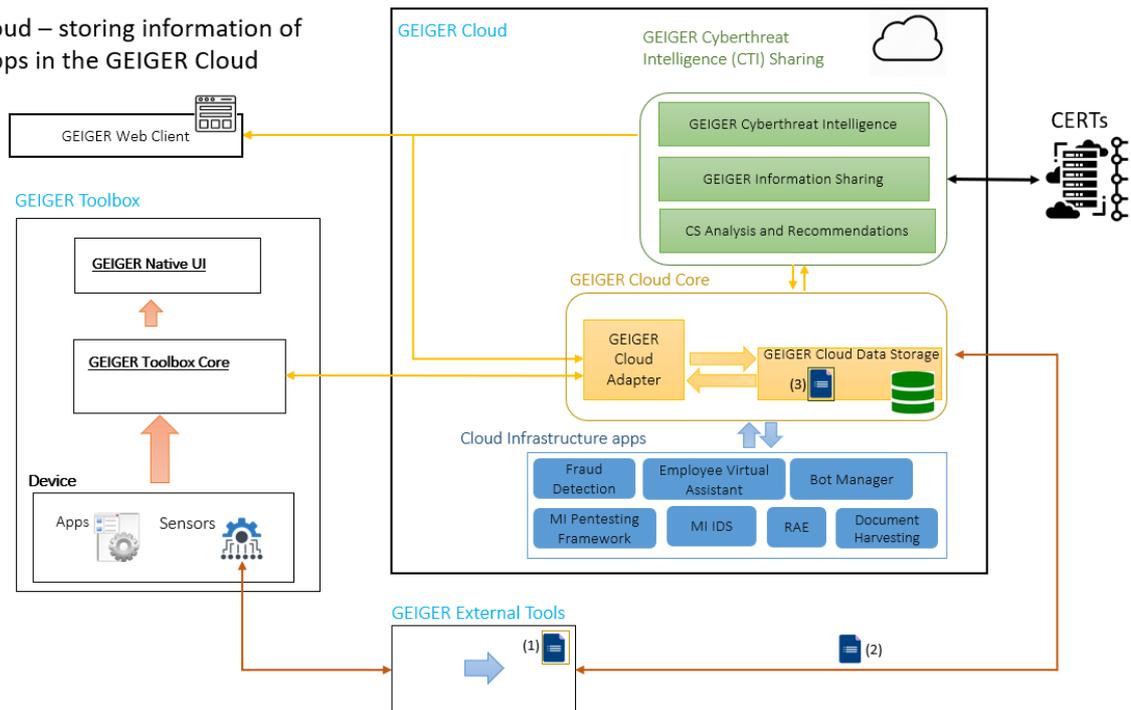


Figure 22 - Storing information of external apps in the cloud flow

As a logical counterpoint to the Cloud infrastructure apps, GEIGER External tools will not be deployed into the GEIGER platform but will gather data from their sensors, located in the device, and sent it to their private servers:

- i. GEIGER External Tools sensors are the ones sending information to their respective servers.
- ii. Then, data will be sent to the GEIGER Cloud data storage of the GEIGER Cloud to be stored. Once in the cloud storage it can be requested by the Toolbox when necessary.

2.4.12 Sending Information of the External Apps to the GEIGER Indicator

GEIGER Cloud – sending information of the external apps to the GEIGER Indicator

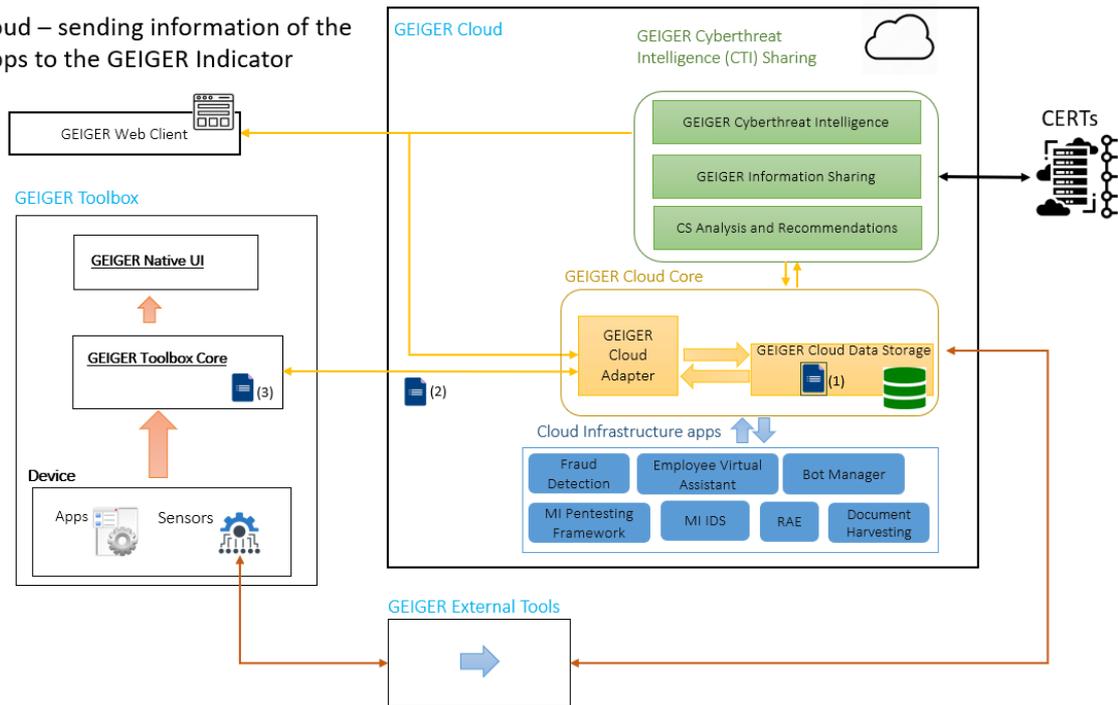


Figure 23 - Sending information of external apps to indicator flow

GEIGER External Tools produce data, which are employed for the risk score calculation. This information produced in External Tools is stored in the GEIGER Cloud storage, as it has been described:

- i. Data produced by External tools is stored in the GEIGER Cloud data storage.
- ii. This information is sent by means of the Cloud Adapter and GEIGER Controller to the Toolbox Core.
- iii. Finally, data are passed to the risk score calculation algorithm.

2.4.13 Generation and Storing Personal Information of the Employee

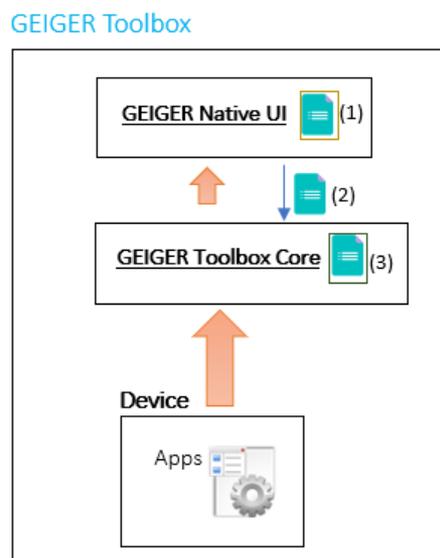
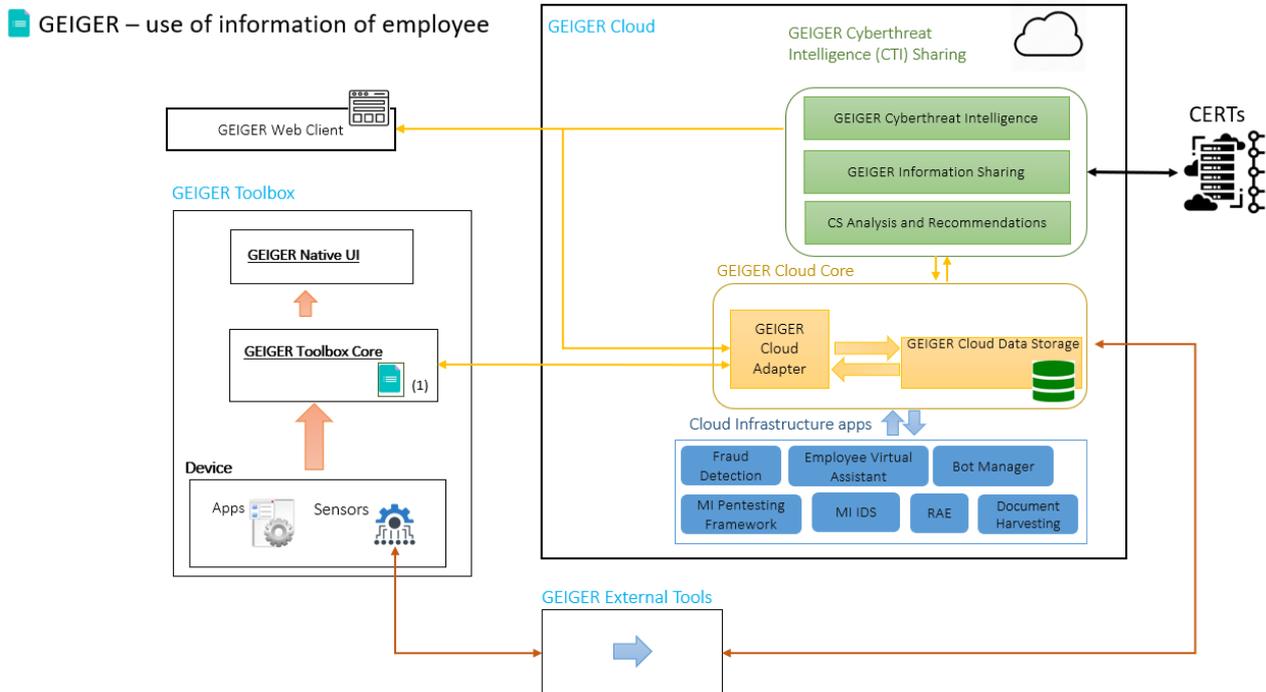


Figure 24 - Generation and storing employee PI flow

Personal information (PI) of the employee follows the following flow:

- i. The end-user employs the GEIGER Native UI to introduce information in the GEIGER platform. Therefore, GEIGER Native UI is the point where information is generated.
- ii. Data are passed to the Toolbox Core.
- iii. Finally, the information is stored locally in the local storage of GEIGER, where it will remain until requested.

2.4.14 Use of the Personal Information of the Employee



The end-user generates Personal Information (PI) of the employees of the MSE with the help of the GEIGER Native UI as it has been previously described:

- i. Then, personal information of the employee is stored locally in the local storage of the GEIGER Toolbox Core where it is available until it is claimed.

2.4.15 Storing Refined Personal Information Data on the Cloud for Further Analysis

GEIGER Cloud – storing refined data of company/employee for further analysis

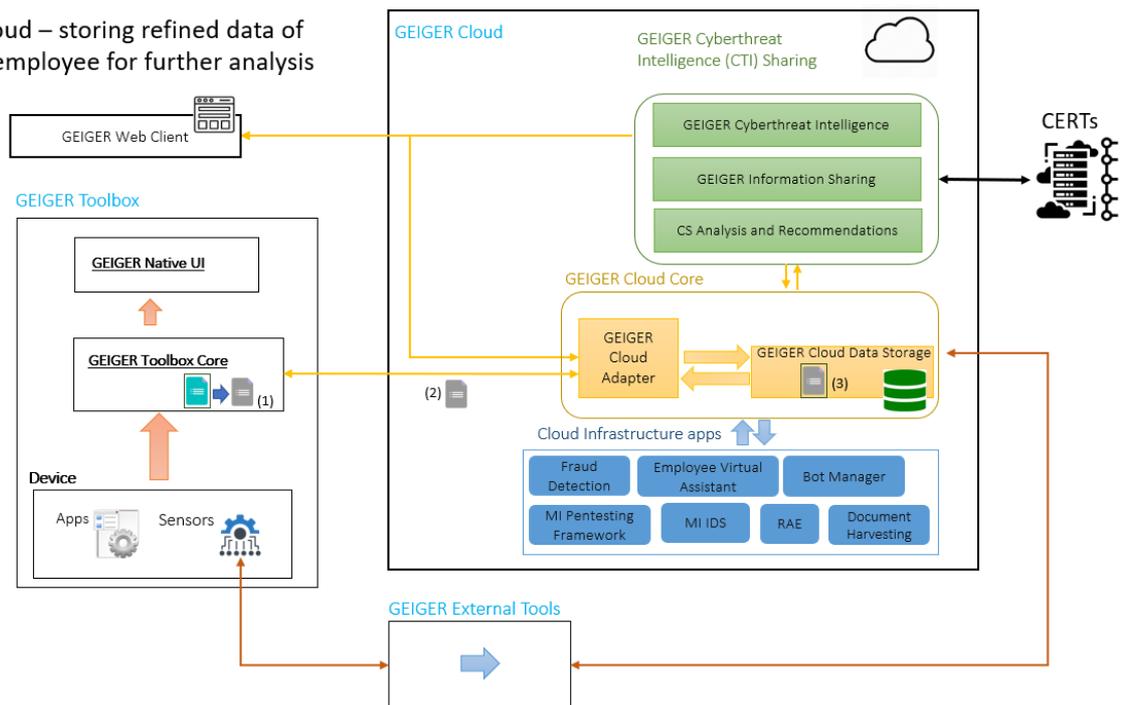


Figure 25 - Storing refined PI data on the cloud flow

The end-user introduces personal information (PI) of the employees and the MSE. Data are stored in the local storage of the GEIGER Toolbox Core. In this use case, information shall reach the GEIGER Cloud:

- Data are stored in local data knowledge base of the GEIGER Toolbox Core.
- Then the information passes through the GEIGER Controller and the GEIGER Cloud Adapters, to reach the GEIGER Cloud Core.
- Finally, data are stored on the GEIGER Cloud Data Storage.

2.4.16 Storing (Relevant) Information from MSEs

GEIGER Cloud – storing relevant information for CTI coming from MEs

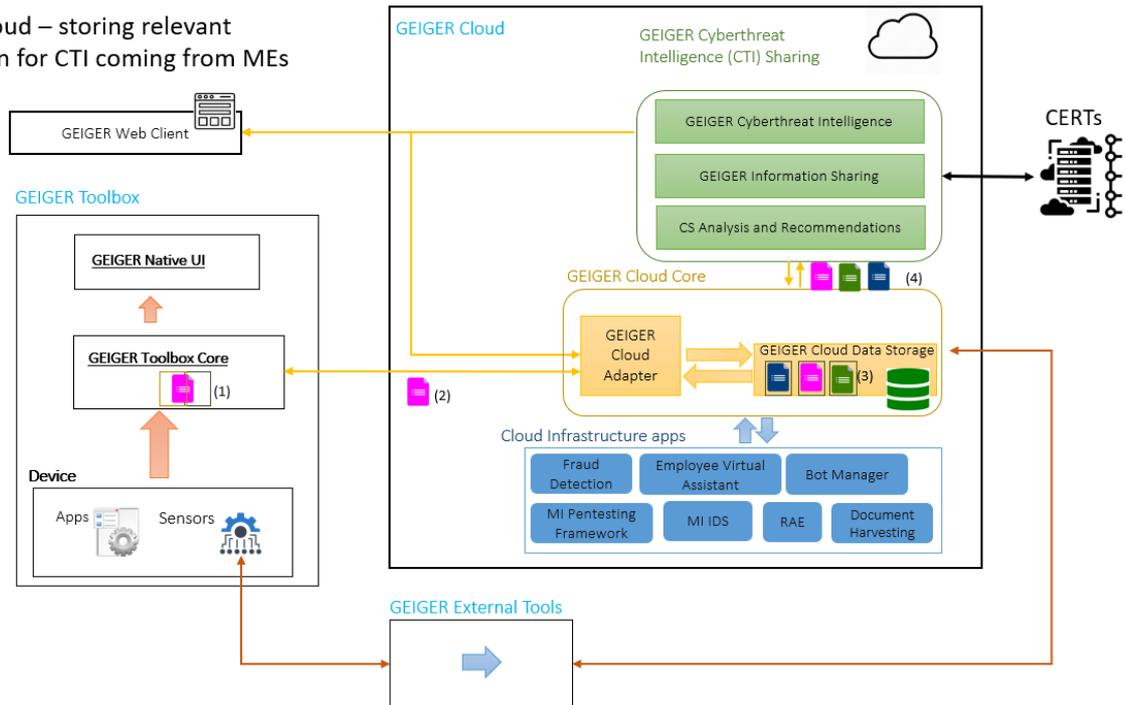


Figure 26 - Storing information from MSEs flow

GEIGER Cloud also stores data, which can be used by the Cyber Threat Intelligence Sharing component. To be more precise this is data from MSEs stored in the Toolbox Core:

- i. Information of the MSEs is generated locally, i.e., in the GEIGER Toolbox.
- ii. Data flow from the Toolbox Core to the GEIGER Cloud with the help of both adapters (GEIGER Controller and Cloud adapter)
- iii. The information is stored in the GEIGER Cloud data storage. Online storage retains various kinds of data including:
 - o private information from MSEs,
 - o data collected and sent by cloud infrastructure applications, i.e., information provided by tools deployed in the GEIGER servers,
 - o any information provided by GEIGER External Tools,
 - o any of these kinds of data can be shared with the GEIGER CTI Sharing component if necessary.
- iv. Finally, the information can be shared with the GEIGER CTI Sharing component if necessary.

2.4.18 Storing Relevant Information for CTI Coming From MSEs

GEIGER Cloud – storing relevant information for CTI coming from MSEs

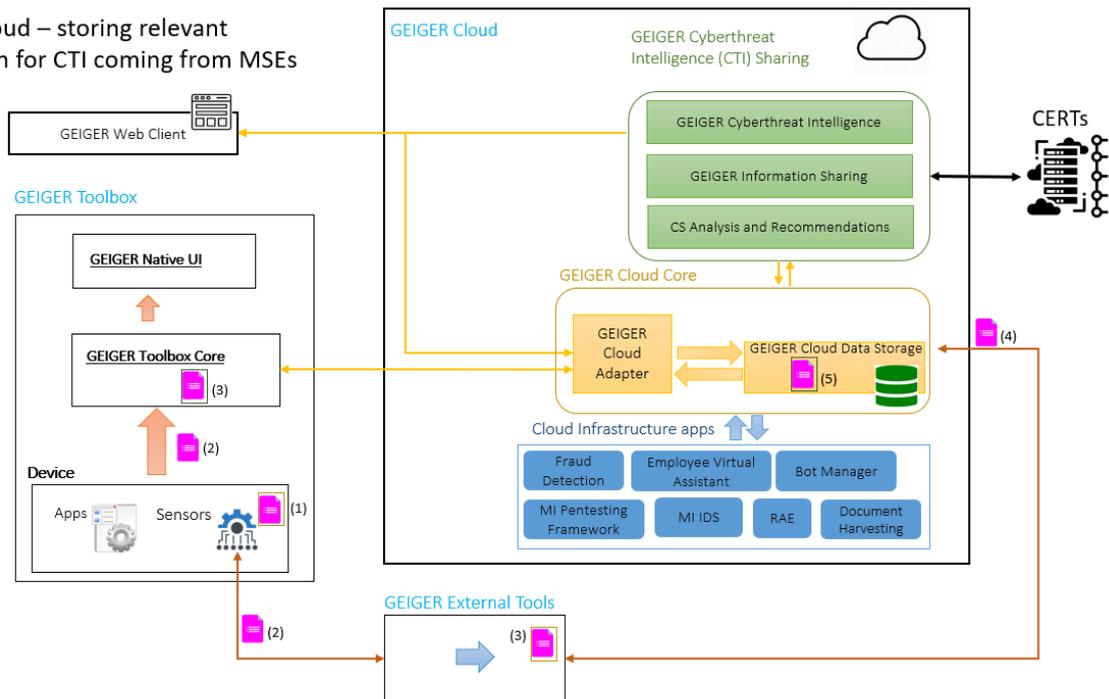


Figure 28 - Storing relevant information of MSEs flow

The flow indicates how relevant information of MSEs is managed in the GEIGER platform:

- i. Application sensors in the MSE gather information, which is necessary for the CTI.
- ii. Data will follow two paths:
 1. One flow indicates that data are sent to the GEIGER Toolbox Core where the information is stored in the local database storage of the GEIGER Toolbox.
 2. The second flow involves GEIGER External Tools: data gathered from their sensors are sent to the GEIGER Cloud data storage to be stored in the cloud database storage.
- iii. Finally, information can be shared with the GEIGER CTI Sharing component.

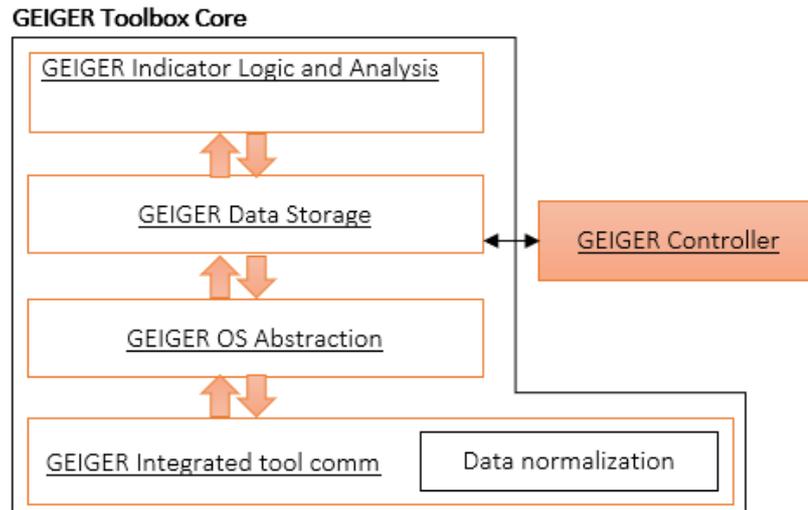


Figure 30 - Detail of the GEIGER Controller in the GEIGER architecture

- ii. GEIGER Cloud adapter: this sub-component of the GEIGER platform is placed in the GEIGER Cloud Core and manages all traffic targeting the online storage coming from the GEIGER Toolbox. As per communications with the GEIGER Toolbox are frequent, any outgoing traffic shall be routed by the Cloud adapter too.

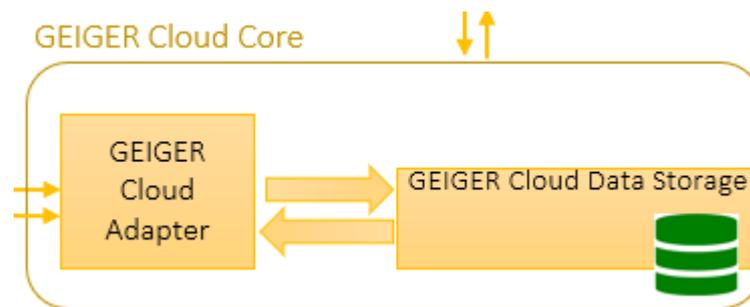


Figure 31 - Detail of the GEIGER Cloud adapter in the GEIGER architecture

2.5.2 APIs for Online and Local Database Access

APIs are the vehicle of a standard communication. They help by easing data exchange while reducing the amount of work to be done to send and request information. GEIGER depends on two APIs for the database access: Local storage API and Cloud storage API

2.5.2.1 Local Storage API

Local storage API or ‘GEIGER Controller’ provides the following features:

- It can secure communications in a non-intrusive way: it signs (Diffie-Hellman encryption) the information flow.
- It grants access to the information and registers plugins.
- It can also notify or even ‘wake up’ plugins. In addition, it can also gather information of the current plugins available.
- Control can be passed to the GEIGER Toolbox when needed.
- It can flag data as outdated as well as remove a plugin if necessary.

2.5.2.2 Cloud Storage API

Cloud storage API also known as 'GEIGER Cloud Adapter' acts as a point of contact with the online knowledge storage. It requests information from other APIs such as the GEIGER Controller and allows for data retrieval from the online database storage. Information can come from various sources, such as:

- cybersecurity-related data about threats, attacks, vulnerabilities, etc. provided by the CTI sharing component,
- requests of information to be sent to the local storage,
- requests on information to be stored in the cloud, such as MSE authorized data,
- data meant to be passed to the CERTs, which may come from the GEIGER Toolbox,
- data provided by Cloud Infrastructure apps to be stored online,
- information gathered by GEIGER External Tools to be stored in the online database.

2.5.3 Data Exchange Protocol

Data to be exchanged on the GEIGER platform follows an extension of the Traffic Light Protocol (TLP) code in terms of labelling:

- TLP:BLACK: information must not get out of the device.
- TLP:RED means that is not for disclosure. In terms of GEIGER, this information can be shared across MSE's devices.
- TLP:AMBER: information disclosure is restricted or limited to be shared amongst nodes of the MSE while, at the same time, could be used in the GEIGER Cloud to be refined.
- TLP:GREEN: information is authorised to circulate amongst MSE's devices as well as GEIGER Cloud entities.
- TLP:WHITE: information can be shared with no restrictions within the GEIGER platform.

Data can be adapted to different kinds of sub-nodes, including the following ones:

- **Users** → gathers information of users:

- First name.
- Last name.

User nodes should be named according to Universally Unique Identifiers (UUIDs) to prevent collisions.

- **Devices** → this sub-node compiles all information from each device:

- Name (recommended UUID to avoid collisions) and type of device.
- OS and OS version of device.
- Owner of the device, that is, user UUID.

- **Enterprise** → the sub-node includes all the information about the MSE:

- Name and profile of the company.
- Sector of operation.
- A list of assets separated by commas (",").
- Location of the company.

- **Enterprise:users** → it is used to store information of the employees of a company. The sub-node incorporates:

- Name and last name of the employee.
- Role of the user.
- Knowledge level of the user (ranking from 0 to 4).
- **Global: threats** → it contains a list of threats:
 - Name of the threat (UUID recommended).
 - Description of the threat.
- **Global:recommendations** → a list of recommendations:
 - Short description (30 characters maximum) and type of recommendation.
 - Long description.
 - Action.
 - List of threat UUIDs related to the recommendation separated by commas.
 - User role associated with the recommendation.
 - Knowledge level required.
 - Financial cost required (Boolean).
 - Device type and OS required.
 - How to implement the recommendation, that is, a list of steps delimited by commas.
 - Assets required to perform the recommendation.
- **Global:userRole** → a list of user roles:
 - Role name (UUID suggested).
- **Global:securityDefenders** → information of security defenders of GEIGER:
 - First and last name.
 - Company.
 - Contact information, including both telephone number and email.
- **Global:securityDefendersOrganizations** → UUID of the organizations to which security Defenders belong to.
- **Global:location** → a list containing all jurisdictions managed in the GEIGER platform.
- **Global:assets** → contains a list of assets of every MSE where GEIGER is used.
- **Global:profiles** → a list of nodes containing:
 - Name and description of the profile.
 - Key-value pairs describing the threat and its weight within the profile.
- **Global:cert** → information of CERTs available in GEIGER:
 - Name and location of CERT.
- **Keys** → includes information of keys stored:
 - Path to the key.
 - Encoded key.
- **Local** → information related to local data storage:
 - Current user and device.

- Cloud account and cloud identity linked to the public key.
- EnterpriseKey needed to obtain keys.

2.6 Roles

After analysing the list of requirements described in D1.1 and close discussion with use case partners and technical providers, we have identified the following roles and profiles for the GEIGER platform:

2.6.1 Technical Engineer

The technical engineer is the person responsible for providing support to GEIGER, that is, they create and deliver technical cybersecurity solutions, which are to be integrated into the platform. Depending on the situation, this role should help identifying problems and develop strategies while gathering information to finally contribute with the better solution for GEIGER. There are various objectives to meet with the solutions proposed such as helping end-users to achieve a better cybersecurity knowledge, fixing problems with the platform or even providing new functionality to enhance GEIGER capabilities upon demand.

2.6.2 Cybersecurity Trainer

GEIGER has been designed as a complete cybersecurity platform for MSEs. This means that should include the possibility of acquiring cybersecurity knowledge. Cybersecurity trainers are responsible for providing training to GEIGER end-users. Their responsibilities include creating courses and delivering cybersecurity materials that are integrated in the GEIGER training platform. By means of security education, owners and employees of the MSEs can achieve a higher understanding of all surrounding threats which can endanger their business and become more proficient in the cybersecurity field.

2.6.3 Organization

The Organization has a prominent role because is the one which manages the GEIGER platform. This role is in charge of providing GEIGER to end-users, that is, make GEIGER accessible as a platform to MSEs personnel, both owners and employees. Besides, an organization should accomplish several other tasks including:

- ✓ Hosting the cybersecurity tools as a service.
- ✓ Integrating new cybersecurity tools in the platform: GEIGER is constantly evolving, which means that new tools can be added to adapt to threats and vulnerabilities discovered. That flexibility makes GEIGER ready to adapt to the new challenges proposed by the cybersecurity environment.
- ✓ Making training courses accessible and available for end users.

2.6.4 CERT / CSIRT

A CERT is a response team able to deal with incidents related to information technologies. The CERT is in charge of developing measures (preventive, reactive) to mitigate the effect of an event. A CSIRT is a team of professionals whose main mission is to provide support when there is a security incident. Apart from the actions derived from an incident CERTs and CSIRTs help by spreading good security practices.

As far as GEIGER is involved, both the CERT and CSIRT is expected to provide information, that is, data about cyber threats, security incidents, vulnerabilities discovered and so on. They will be engaged with the GEIGER Cloud component. Due to the amount of data to be produced by a CERT or CSIRT, the GEIGER platform should only gather and store data needed to perform the risk score calculation, discarding all those data that might be considered less useful to feed the GEIGER Indicator.

2.6.5 Cybersecurity Defender

Within the GEIGER Ecosystem there are specific profiles for 'GEIGER Security Defenders'. They will be trained in various cybersecurity aspects and can pass an assessment to become a 'Certified Security Defender'. Their functions can be to:

- i. Guide the GEIGER installation at the MSE.
- ii. Monitor the running of the GEIGER application in an MSE.
- iii. Train employees to improve the cybersecurity knowledge/skills and to handle GEIGER correctly.
- iv. Assist the user on the implementation on cybersecurity recommendations.

Security Defenders can be internal (e.g., designated employees) or external (e.g., pertinent service providers) to the MSE. Particularly for internal Security Defenders GEIGER training schemes are assumed to work for IT-lay-persons, too.

2.6.6 End-User

The end-users are both the MSE owners and employees, i.e., people who should use GEIGER in the day-to-day basis. GEIGER has been conceived as a cybersecurity platform to help business deal with cybersecurity. It is important to outline that GEIGER is flexible and does not require any kind of knowledge or proficiency of the user as far as cybersecurity is involved.

An end-user can employ GEIGER for different purposes such as the following:

- To protect their business and, at the same time be more aware of the threats and vulnerabilities that could impact their organization. One of the objectives of GEIGER would be to make the end-user conscious of how serious these threats and risks are.
- To realise how well or bad their organization is performing in terms of cybersecurity: what is the level of risk, what measures could be adopted and whether the ones placed are effective or not.
- To increase their cybersecurity knowledge by receiving training.
- To be alerted and know more about new threats and vulnerabilities that could endanger the business. This should include to discover what measures and controls could possibly mitigate the effects of risks.
- To ask GEIGER about cybersecurity recommendations.
- To perform awareness activities.
- To be protected against fraud.

2.7 Use Cases

2.7.1 Technical Engineer

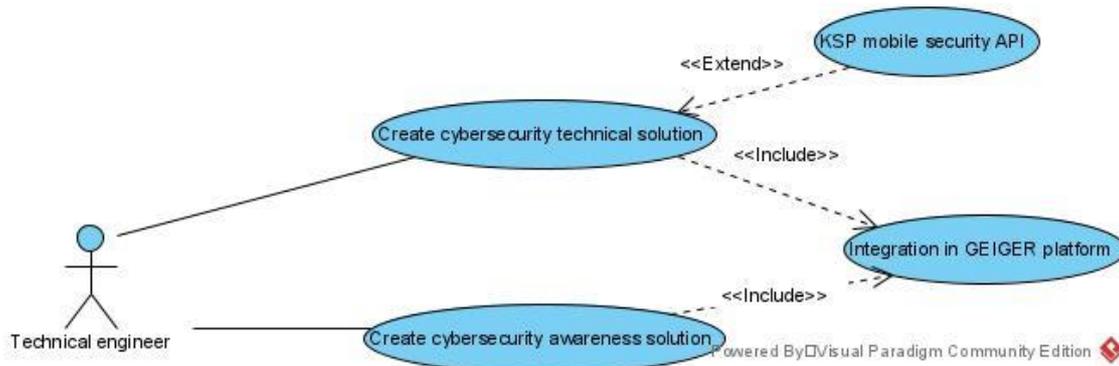


Figure 32 - Technical engineer use case

The technical engineer is expected to provide support to GEIGER. Therefore, technical engineers are expected to:

- Create cybersecurity technical solutions: those should include various scenarios where they can help identifying problems with the platform, development of strategies and the process of gathering information to propose a solution that really fix the problem.
- Encourage awareness to end-users and help them achieve a higher degree of cybersecurity knowledge.

No matter what kind of solution is provided, it is expected to be integrated in the platform.

2.7.2 Cybersecurity Trainer



Figure 33 - Cybersecurity trainer use case

GEIGER might be also considered a learning platform. The platform offers cybersecurity training to the end-user. Cybersecurity trainer is the role responsible for delivering knowledge to end-users. Duties of the trainer should include:

- Creation of cybersecurity training.
- Provide training to the end-users (this is achieved with the integration of the training into the GEIGER platform).

An implicit requirement for the cybersecurity trainer will be to prove proficiency in the cybersecurity field.

2.7.3 Organization

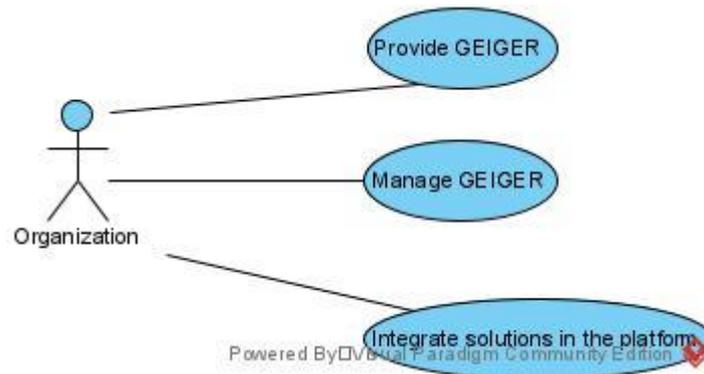


Figure 34 - Organization use case

The Organization is the entity, which manages GEIGER. The responsibilities of the organization are:

- Provide GEIGER platform to the end users: the organization should make sure all end-users in the MSE can access GEIGER.
- Management of the GEIGER platform: this is a high-level task which includes:
 - Installing new tools or updating the current ones.
 - Changing setup and making appropriate adjustments.
 - Hosting the cybersecurity tools as a service.
 - Make training courses available to end-users.
- Integrate solutions: the organization will add any solution to the platform. One example would be the cybersecurity training courses.

2.7.4 CERTs and CSIRTs

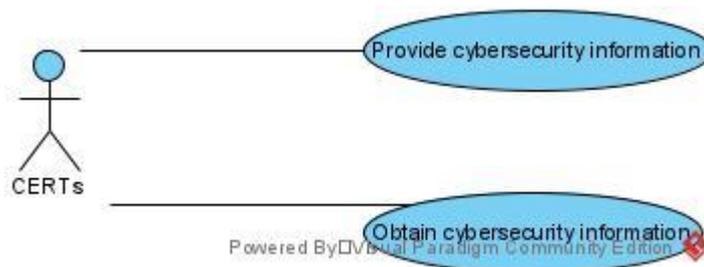


Figure 35 - CERT use case

A CERT deals with incidents in IT. Main objectives of the CERT regarding the GEIGER platform are:

- Gather cybersecurity information, including:
 - New or updated threats.
 - Recent vulnerabilities discovered.
 - Data of security incidents that may be of interest.
 - Any other information, which could be considered useful to protect the MSE.
- Provide data gathered in the previous point to GEIGER platform, where it will be stored and employed in the calculation of the level of risk. It is important to note that not all information sent by CERTs will be stored, that means only information needed for the calculation of the level of risk.

2.7.5 End-User on PC and Internet

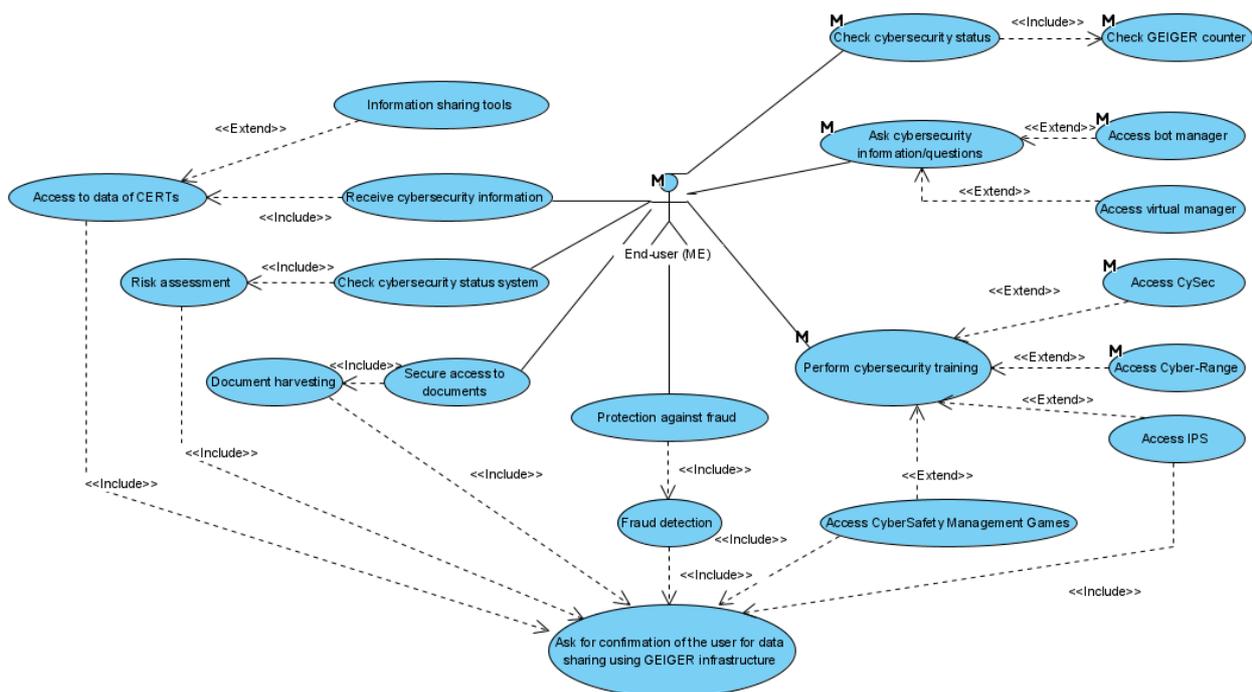


Figure 36 - End user on PC + internet use case

GEIGER aims to be a useful tool for end-users, that is, MSEs owners and employees. Users will employ GEIGER in many ways, including the following:

- Receive cybersecurity training: this is achieved in two ways:
 - By accessing and completing training courses, end-users can increase their knowledge in cybersecurity. In addition, they will be more conscious of how serious cyber threats can be.
 - By asking questions to the GEIGER platform, the end-user can quickly find answers and manage cybersecurity concepts better.
- Check risk score and status of the system:
 - The GEIGER Indicator gives the user an immediate estimation of the level of risk. It is possible to check GEIGER Indicator from any device.
 - When an internet connection is available, the end-user is updated with the most recent vulnerabilities and threats provided by CERTs.
 - The user can check whether there is any risk for the MSE as well as prioritize and understand the severity of them.
- With the help of document harvesting tool, which is part of GEIGER, the end-user can securely check any information or access documents.
- Be protected against fraud: GEIGER is a real-time cybersecurity platform, which helps the end-user fighting fraud and alerts when there is a suspicious activity.

2.7.6 End-User on PC without Internet

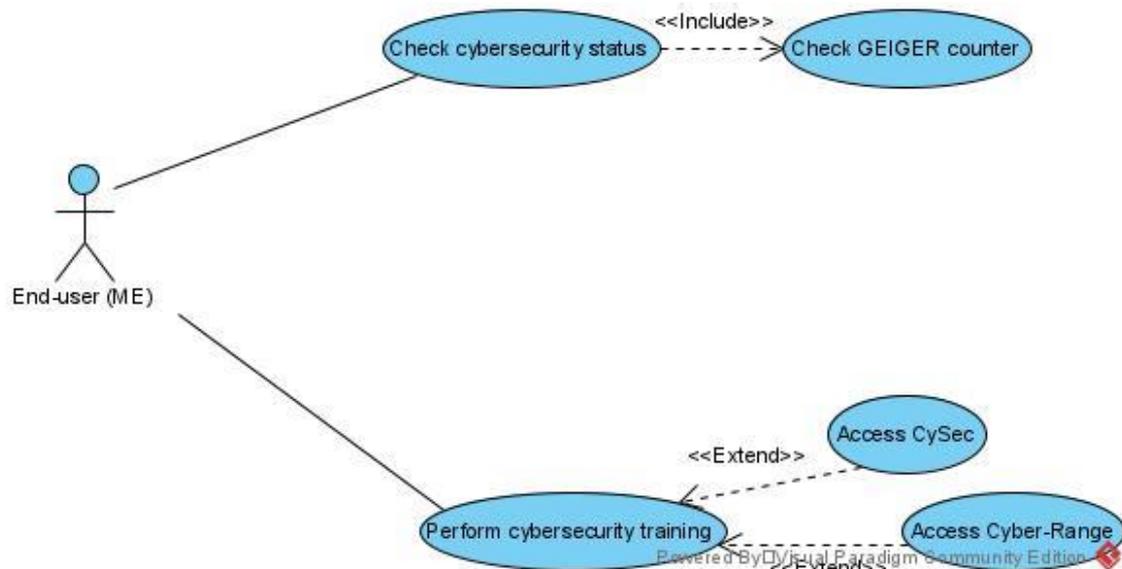


Figure 37 - End user on PC without internet use case

End-users can access GEIGER even if they are not connected to the internet: GEIGER is a connection-oriented platform, but there is no problem to use it offline. However, it is important to note that functionality available to the end-user will be slightly reduced compared to the possibilities when online.

In such case, GEIGER will make use of local data stored. In this scenario, users will be able to:

- Check GEIGER Indicator, which may be updated with information stored locally. It is important to outline that, in terms of cybersecurity, information changes quickly so data provided by local storage has a short lifetime. As a result, GEIGER Indicator value will be less reliable as time goes by.
- Perform cybersecurity training with the help of those GEIGER tools which do not require an active internet connection.

2.7.7 End-User with Mobile Device

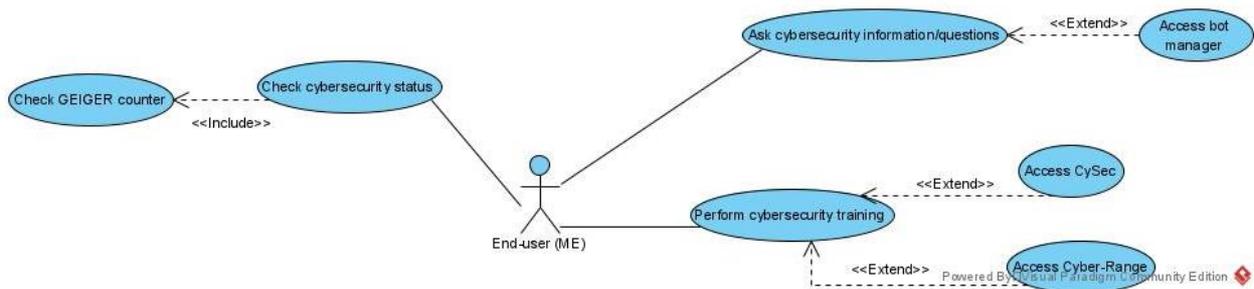


Figure 38 - End-user with mobile device use case

One of the advantages of GEIGER is the possibility of employ the platform with mobile devices such as tablets or mobile phones. The purpose is to make GEIGER available to end-users in any situation, even when they are not in front of their computer. Features accessible when in a mobile device include:

- Check the risk level by means of the GEIGER Indicator and know immediately which is the level of risk. According to what has been described in the PC use cases, it would be advisable to have an active internet connection so data can be properly updated.
- Perform cybersecurity training, which includes:
 - The possibility of asking and solving any cybersecurity questions at the very moment.
 - Access to some of the GEIGER training tools to accomplish any training provided by the platform.

3. GEIGER Components

3.1 GEIGER Toolbox: GEIGER Indicator

3.1.1 GEIGER Indicator Concept

The GEIGER Indicator solution will allow users to calculate their GEIGER score, a measure of the cybersecurity risk they and their MSE are facing. Based on the characteristics of an MSE and the results of the GEIGER Indicator score calculation, users will receive recommendations for actions to mitigate cybersecurity risk.

We used the user requirements that were collected as a part of the work for the requirements deliverable of GEIGER as input for formulating the GEIGER Indicator concept. We can see from the user requirements in Table 1, that our solution should be easy to use and not too intrusive (**UR1, UR6**). Additionally, it should consider both technical and organisational (or social) characteristics of MSEs (**UR2**). The GEIGER Indicator concept will incorporate expert knowledge to facilitate easy use without requiring too much data. The organisational and social sides of cybersecurity are covered by the Education Framework linked to the GEIGER solution. A crucial part of our solution is to align the GEIGER Indicator model with the GEIGER Education Framework, to be able to fit cybersecurity recommendations (i.e., countermeasures) to the knowledge level of users. This conforms to best practice as suggested by, among others, Padayachee (2012) and Shojaifar et al. (2020).

Table 1: A selection of user requirements for the GEIGER solution.

Use case country	Use case	User requirement
Romania	MSE start-ups	UR1. The solution should be intuitive and usable.
		UR2. The solution should assess both technical system properties and good cybersecurity practice.
Switzerland	MSEs	UR3. The solution should provide knowledge on how to secure an MSE.
		UR4. The solution should facilitate improvement of cybersecurity status to the point that the MSE is considered secure.
The Netherlands	Accountants of MSEs	UR5. The solution should be clearly linked to existing rules and regulations, providing more comfort and assurance in using the solution.
		UR6. The solution should facilitate simple – preferably automated – data collection.

The need for guidance, comfort, and assurance (**UR3, UR4, UR5**), is not extensively considered in existing literature. All these aspects relate to trust. Regardless of how well we incorporate behavioural theories, such as Protection Motivation Theory (PMT) and Self-Determination Theory (SDT, Menard et al., 2017), people will not use the GEIGER solution if they do not trust it.

To promote trust in the GEIGER solution, we have chosen to partner with governmental institutions such as CERTs and National Cyber Security Centers (NCSCs) in creating our cybersecurity risk assessment model. This allows us to incorporate the rules and regulations mentioned in **UR5** and actively partner with the governments they originate from. We also explicitly incorporate the legal and compliance part of the cybersecurity picture in our threat mapping.

Our indicator concept builds on existing views of cybersecurity measurement, such as the view presented in Casola et al. (2020). Their research was motivated by another EU Horizon 2020 project: MUSA. The MUSA project – short for Multi-Cloud Secure Applications – aims “to support the security-intelligent lifecycle management of distributed applications.” In this sense, the MUSA setting, like many other cybersecurity settings, merits a more complex model than our MSE scenario. Recall that our users prefer an intuitive and

usable solution (**UR1**). Figure 39 presents an adaptation of the view presented in Casola et al. (2020) that is better suited to our MSE setting.

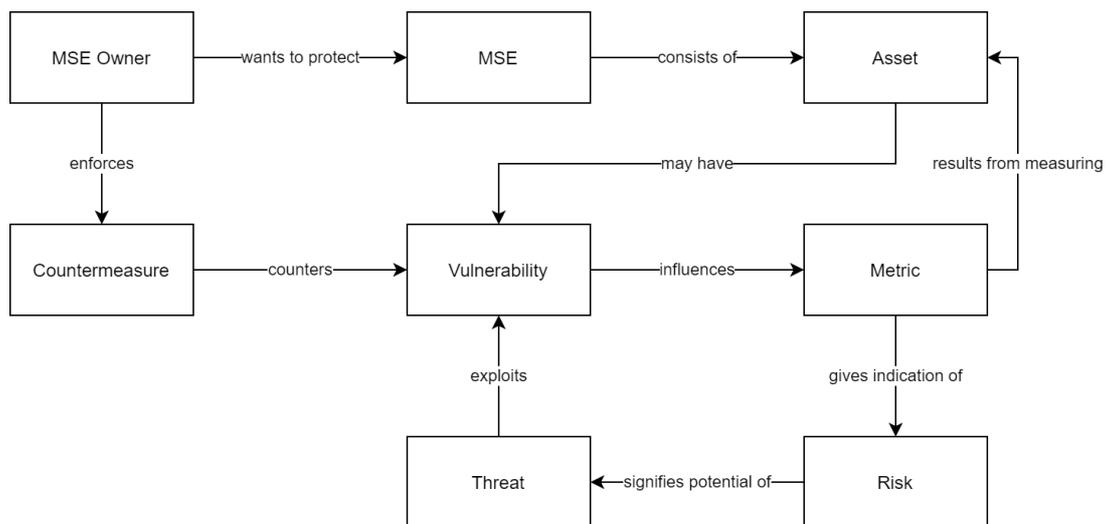


Figure 39: The view on cyber-systems that serves as the basis for the GEIGER indicator algorithm.

As Figure 39 shows, the GEIGER Indicator solution aims to indicate the cybersecurity risk of an MSE, by measuring the (security-related) properties of the assets in an MSE. The metrics that result from these measurements are influenced by vulnerabilities, which can be exploited by threats and countered by countermeasures (sometimes referred to as controls). In this sense, the GEIGER Indicator concept aligns with the threat-vulnerability-control paradigm commonly employed in the security field (Pfleeger et al., 2015). Table 5 in Appendix A defines the security terminology we use for the GEIGER Indicator concept.

Many security measurement solutions in a socio-technical setting exclude the real-life threat environment (Gollmann et al., 2015). This is interesting, since threats are a big part of getting MSEs to understand why they should act. Solutions that simply list a set of vulnerabilities or focus areas, skip the step of motivating the user, which is essential in GEIGER. Therefore, our approach starts from threats. Common threat concepts are generally understood by MSEs, as shown in a recent survey by the Australian Cyber Security Centre among small businesses (ACSC, 2020). Of the respondents, 83% could understand and explain *malware* and 79% could understand and explain *phishing*.

3.1.2 GEIGER Indicator Data Model

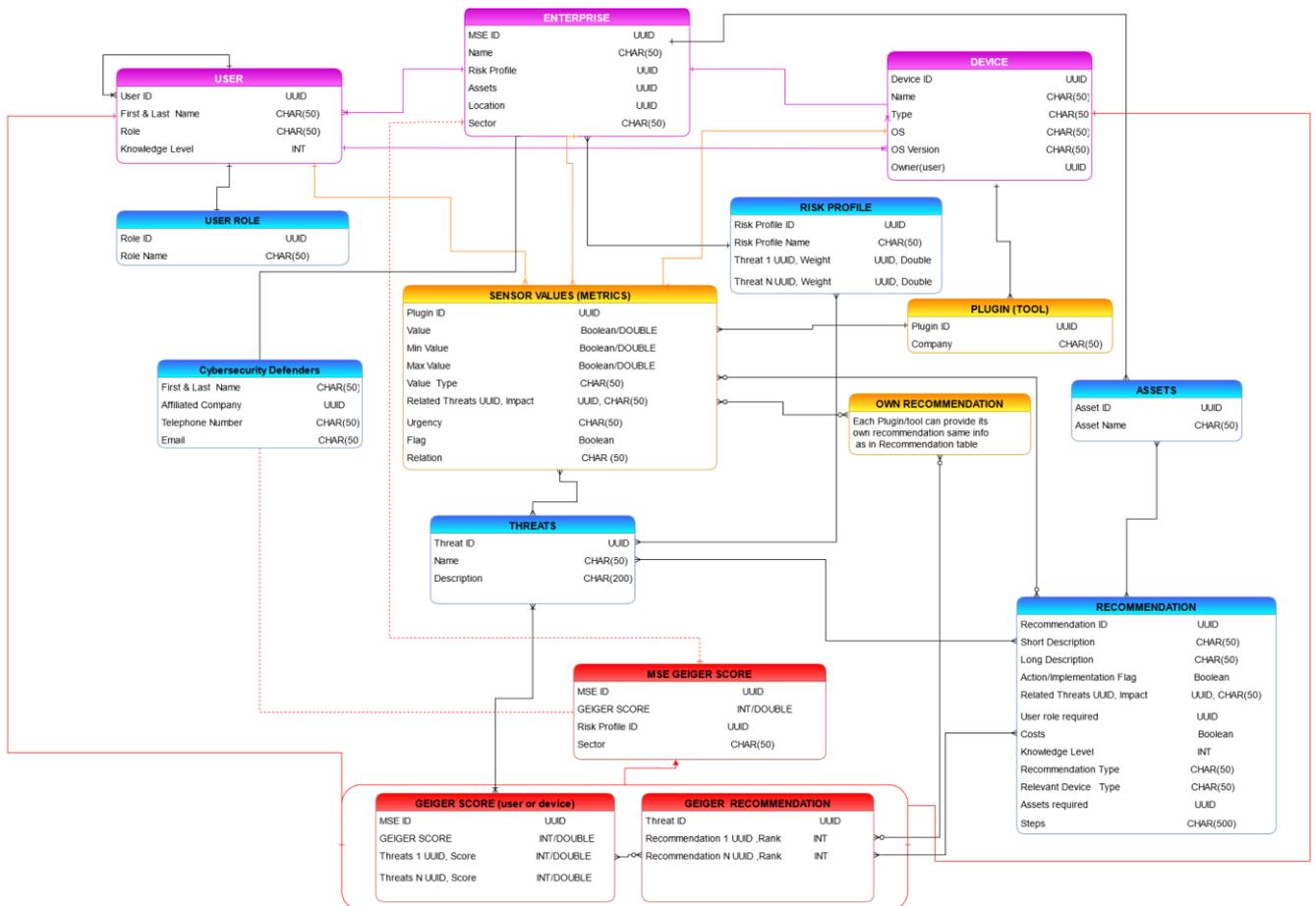


Figure 40 - GEIGER Indicator Data model

Figure 40 depicts the GEIGER Indicator score data model. The data model is in line with the local data storage concept. Each MSE has three main components:

- The Enterprise entity contains all the information concerning the MSE. One of its properties is the risk profile classification, as each risk profile has different threat impacts.
- The User entity encloses information regarding the employees of an MSE including their current knowledge level. Furthermore, the owner of the MSE (main) user must be identified.
- The Device entity includes data on all devices associated with an MSE. Each device is owned by a user.

Additionally, each tool (plugin) installed on a device will store its information on its metrics as indicated in Sensor Values (Metrics) entity. The 'flag' attribute is a Boolean attribute that represents the relationship between the metric and the score, if it is set to '1,' it indicates that the metric score contributes positively to the overall score. Moreover, each metric value must relate to either: User, Device or Enterprise to reflect the metric score in the corresponding GEIGER score, as explained in Section 3.1.4.

The Recommendation entity represents the global set of the provided recommendations discussed in Section 3.1.3.4 and listed in Appendix C. Each recommendation in the global set must relate to one or more threats with a specific impact 'high', 'medium' or 'low', in addition to more information such as the type of recommendation, if it is a user-related or a device-related recommendations and if any cost is associated with it. The implementation of such a recommendation is indicated by the user and the UUID of implemented recommendations are stored in either User entity if it is a user-related recommendation or Device entity if it is a device-related recommendation.

In addition to global recommendations, tools can provide their own recommendations, which is specific to a tool metric. In contrast to the global recommendation, the implementation of such recommendations is indicated by a change in the metric value of the tool that raised the recommendation. Nonetheless, certain information must be provided along with the recommendation description.

The GEIGER Indicator concept interacts heavily with the GEIGER Education Ecosystem. The recommendations of the GEIGER Indicator concept are assigned a required knowledge level, which connects them to the knowledge levels of the competence grid in the GEIGER Training Plan. The connectedness to the GEIGER Education Ecosystem warrants alignment between the data models of the GEIGER indicator concept and the programme for the education of security defenders and MSEs. Alignment was achieved through several alignment meetings and co-development of the indicator and education data models.

For more information on each entity and its associated attributes, we refer to Table 23 in Appendix E. It should be emphasized that the GEIGER Indicator concept adheres to the data minimization principle; no unnecessary data related to users is collected.

3.1.3 GEIGER Indicator Data

Besides the data provided by GEIGER tools in the form of metrics, the GEIGER Indicator solution also depends on general inputs that define the way in which the GEIGER Indicator algorithm operates. MSE profiles are used to provide MSEs insights tailored to their category, sector, and country. The GEIGER Indicator threat model defines the cybersecurity threats that are used in the GEIGER Indicator solution, and the relative risks associated with each for all MSE profiles considered. Lastly, a list of global recommendations allows us to ensure completeness in our advised countermeasures to MSEs.

3.1.3.1 MSE Profiles

It is well recognized that SMEs are a diverse set of organisations that cannot simply be considered as a single group (European Digital SME Alliance, 2020). To align with this view, we create MSE profiles, that allow us to dynamically tailor our risk estimations and suggested countermeasures to each MSE. The European Digital SME Alliance distinguishes four different SME categories: start-ups, digitally dependent SMEs, digitally based SMEs, and digital enablers (European Digital SME Alliance, 2020). Start-ups are generally considered to fall into either the digital enabler or digitally based categories. Since MSEs are smaller than SMEs and start-ups already fall under these other two categories, we choose to exclude the start-up category in our initial approach.

In addition to the MSE category, the MSE country and sector play an important role in how we should assess the cybersecurity of the MSE. We allow MSE sectors to follow the statistical classification of economic activities in the European community (Carré, 2008), also known as NACE. Country codes should follow the ISO 3166-1 alpha-2 two letter standard. Table 2 provides an overview of the three dimensions that together (can) constitute the profile of an MSE. In the current version of the GEIGER Indicator solution, only the MSE category is used to determine the profile. As more data is incorporated in the GEIGER solution, we will be able to additionally form profiles using the country and sector of an MSE.

Table 2: The three MSE profile attributes category, sector, and country.

MSE Profile Attribute	Possible Values
Category	<ul style="list-style-type: none"> • Digitally dependent MSE. • Digitally based MSE. • Digital enabler.
Sector	Sectors of the NACE (Carré, 2008) classification.
Country	Country codes from the ISO 3166-1 alpha-2 standard.

3.1.3.2 GEIGER Indicator Threats

Many threat taxonomies exist that use different definitions of the term ‘threat’ and often even have internal conflicts and contradictions. In the threat actions as specified within the VERIS framework (VERIS, 2017) we encounter general terms such as ‘Hacking’ and ‘Social’. Verizon uses VERIS in their Data Breach Investigation Report (DBIR, Bassett et al., 2020). In this report, Verizon distinguishes between cybersecurity incidents and data breaches.

ENISA, on the other hand, uses a detailed threat taxonomy (ENISA, 2016b). Their reports on the top 15 threats (ENISA, 2020) also provide a more granular view than the VERIS taxonomy. Interestingly, ENISA include ‘Data breach’ as a threat, rather than it being the result of a threat as in the Verizon DBIR. ENISA recognizes that ‘Data breach’, and the related ‘Identity theft’, are not regular threats. They state that “they are consequences of successful threats,” which is in line with the DBIR definition (ENISA, 2019).

There are more subtleties like these contained in the ENISA threat taxonomy. Nevertheless, ENISA is a highly influential organisation in our context. CERTs and NCSCs regularly use taxonomies like the ENISA Reference Incident Classification Taxonomy (ENISA, 2018) and the similar CERT-XLM taxonomy used in the MISP threat sharing platform (MISP, 2021). Therefore, we choose to use the ENISA Top 15 threats (ENISA, 2020) as the basis for our GEIGER indicator threats, and analyse relationships with additional taxonomies, to arrive at a pivot mapping like that of the ENISA Reference Incident Classification Taxonomy (ENISA, 2018).

Table 17 in Appendix B lists the guiding principles for mapping the ENISA top 15 threats (ENISA, 2020), as well as other threats contained in the overall ENISA taxonomy, to the GEIGER Indicator threats. It should be noted that our threat list is not a new taxonomy, but rather a selection and grouping of relevant threats from the existing ENISA threat taxonomy.

By applying these principles, and by including input from the Romanian CERT, Swiss NCSC, and Dutch Digital Trust Center, we obtained the GEIGER Indicator threats as specified in Table 3. Table 17 in Appendix B presents a mapping of ENISA threats to GEIGER Indicator threats, where we indicate the guiding principles behind our decisions.

Table 3: The GEIGER indicator threats. We indicate the definition of each threat and the source for this definition.

GEIGER indicator threat	Definition	Source
Malware	Short for malicious software. Malware is any program written with the intent to carry out harmful actions.	CyBOK (2019), NIST (2021)
Web-based threats	Web-based threats are an attractive method by which threat actors can delude victims using web systems and services as the threat vector.	ENISA (2020)
Phishing	Phishing is the fraudulent attempt to steal user data such as login credentials, credit card information, or even money using social engineering techniques.	ENISA (2020)
Web application threats	Threats to the security of web applications and services, often abusing misconfigurations, weaknesses, or vulnerabilities in the implementation of these applications.	ENISA (2020), OWASP (2020)
Spam	The abuse of electronic messaging systems to indiscriminately send unsolicited bulk messages. It is considered a cybersecurity threat when used as an attack vector to distribute or enable other threats.	ENISA (2020), NIST (2021)
Denial of service	The prevention of authorized access to resources or the delaying of time-critical operations. A denial-of-service condition is accomplished by flooding the targeted host or network with traffic until the target cannot respond or simply crashes.	CISA (2009), NIST (2021)

Data breach	A data breach is a type of cybersecurity incident in which information (or part of an information system) is accessed without the right authorisation, typically with malicious intent, leading to the potential loss or misuse of that information.	ENISA (2020)
Insider threats	The potential of an entity with authorised access to harm an information system or enterprise through destruction, disclosure, modification of data, and/or denial of service.	NIST (2021)
Botnets	A network of compromised computers controlled by the same cybercriminal. Operating in a peer-to-peer (P2P) mode or from a Command and Control (C2) centre, botnets are remotely controlled by a malicious actor to operate in a synchronised way.	CyBOK (2019), ENISA (2020)
Physical threats	Threats related to the tampering, damage, theft, and loss of physical assets.	ENISA (2020)
Ransomware	Ransomware is a type of malware that infects the computer systems of users and manipulates the infected system in a way that the victim cannot (partially or fully) use it and the data stored on it. The victim receives a blackmail note pressing them to pay a ransom to regain full access to systems and files.	ENISA (2016c), ENISA (2020)
External environment threats	Threats of financial, reputational, or legal damages due to non-compliance with regulations, standards, or other agreements with third parties (e.g., SLAs). Also includes the threats resulting from changing financial and economic circumstances and the actions (intended or unintended) of external stakeholders such as customers and suppliers.	Davis (2014), ENISA (2016b)

3.1.3.3 GEIGER Indicator Tool Metrics

This section includes tool metrics information necessary for the GEIGER indicator. Tool owners are asked to indicate all the metrics that can be provided by their tools. As each tool can provide more than one metric, data was collected on a metric-level rather than tool-level. The required metric data collected is shown in Figure 40 under 'Sensor Values (Metrics)' entity.

Each metric must relate to a specific cyber-system, either the device or the user. The value of the metric is reflected in the corresponding GEIGER score. In educational tools, metric values can contribute positively to the GEIGER score. This is indicated by a Boolean 'flag' variable that is set to '1' if the metric is contributing positively. Additionally, metric value type (Boolean, integer or double) must be specified along with the minimum and maximum values that can be achieved are necessary for calculation purposes as described in Section 3.1.4.

Each metric value obtained from a tool must relate to one or more threats with a specific impact. This information is provided by the tool owner, where the impact is either set as 'high', 'medium' or 'low'. The threats covered in the GEIGER indicator are discussed in detail in Section 3.1.3.2. It is important to highlight that 'urgency' attribute is only tied to Boolean metrics and to send notifications to user based on how critical the metric value is. Possible values are 'low': no notification, 'medium': notification after 1 week, 'high': notification after 1 day and 'critical': immediate notification.

After all the necessary information concerning the metric is collected, we can now use this data as an input to the GEIGER indicator. Statistics of the collected metrics are shown in Table 4, where we can clearly see that more metrics will need to be collected to facilitate a fully functioning GEIGER indicator algorithm.

Table 4: Statistics on the collected metrics for the GEIGER indicator threats.

	Description	Value
--	-------------	-------

Tools	Number of tools that are providing data for GEIGER indicator.	9
Metrics	Total number of metrics provided by all tools.	46
User related metrics *.	Total number of metrics that will be reflected in the user's score.	19
Device related metrics. *	Total number of metrics that will be reflected in the device's score.	27
Positive metrics.	Number of metrics that will affect the score positively.	14
GEIGER indicator threats * (number of metrics affecting each threat)	Malware	22
	Web-based threats	4
	Phishing	7
	Web application threats	0
	Spam	2
	Denial of service	0
	Data breach	15
	Insider threats	5
	Botnets	0
	Physical threats	0
	Ransomware	15
External environment threats	2	

3.1.3.4 GEIGER Indicator Global Recommendations

Figure 39 showed that MSE owners can enforce countermeasures to counter vulnerabilities and lower the cybersecurity risk they face. To ensure that the countermeasures we suggest within the GEIGER solution are complete and that they overlap as little as possible, we use a global recommendation (or countermeasure) list. This list was constructed by consulting numerous sources, as presented in Table 19 in Appendix C. From the sources we collected a long list of over 300 recommendations, which was then merged to form a list covering the unique recommendations made by the different sources. Table 20 of appendix C lists the global recommendations used in the GEIGER indicator solution.

To prove that our list of recommendations can be considered complete for the MSE situation, we must look towards research and standards. Yigit Ozkan and Spruit (2021) present a set of 17 security control categories for SMEs. By investigating three further security control guidelines for SMEs (ENISA, 2015; ENISA, 2017; FTC, 2018b), we unearthed one additional security control category relevant in the GEIGER setting: physical security.

By mapping the GEIGER indicator global recommendations to the security control categories, we were able to conclude that our recommendations cover all categories adequately, except for 'Human resources security.' Yigit Ozkan and Spruit (2021) include background checks and "pay attention to people you work

with and around” in the human resources security control category. Although a hiring and firing policy, which considers cybersecurity can certainly be relevant for MSEs, we do not see this as part of the role of GEIGER. Therefore, we feel justified in not covering this category of recommendations in the GEIGER indicator global recommendations. For more details on the global recommendations and the security control category mapping, see Appendix C.

3.1.4 GEIGER Indicator Algorithm

The GEIGER Indicator algorithm turns threat information, metric values, and countermeasure implementation indications into GEIGER scores. We formalise the GEIGER Indicator algorithm through a mathematical model. Central to this model is the cyber-system, as defined in Refsdal et al. (2015). All devices and employees within the MSE are cyber-systems. The enterprise itself is also a cyber-system. We let S denote the set of cyber-systems of an MSE, with a specific instance of a cyber-system being denoted by s .

We begin with the set T of all GEIGER threats. These are the threats as outlined in Table 3. We initially specify a set P of MSE profiles. Profiles will allow us to offer dynamic prioritisations of threats, based on MSE country, sector, and category. The Digital SME Alliance outlines four SME categories: digital enabler, digitally based, digitally dependent, and start-ups (European Digital SME Alliance, 2020). Start-ups are defined as a sub-category of digital enablers and digitally based SMEs. Since MSEs are smaller than SMEs and start-ups already fall under these other two categories, we choose to only use the first three categories for our MSE profiles. MSE sectors follow the statistical classification of economic activities in the European community (Carré, 2008). Country codes should follow the ISO 3166-1 alpha-2 two letter standard. Given the GEIGER EU domain, only European country codes will be used.

For each profile p in P , we determine the relative risk r_{pt} associated with the threat t in T . We should note at this stage that although the fact that the profiles in the set P allow for a dynamic algorithm tailored to the specific MSE situation, a vast amount of data must be collected before being able to create the $|P|=3$ (categories) * 50 (countries) * 21 (sectors) = 3150 different threat weightings. It is likely that the weightings, and thus the values of the individual risks r_{pt} , will be identical for certain profiles. Initially, we will only provide unique weightings for the three MSE categories digital enabler, digitally based, and digitally dependent. Once CERT/NCSC incident data is incorporated into the GEIGER solution, more unique profiles can be constructed.

We now have a way to prioritise the different threats t in T , using profiles p in P , where threats receive relative risk indications r_{pt} . To figure out how a particular cyber-system s in S scores on each threat, we need metrics measuring the security state of that cyber-system. Let M be the set of metrics. For each metric m and cyber-system (e.g., an MSE, an employee, a device) s , the normalized (to between 0 and 1) value of the metric is given by v_{ms} . We define a corresponding impact of a metric on a threat i_{mt} . Impacts have a value between 0 and 1, including both 0 and 1. In practice, an impact of 0 means a metric is completely unrelated to a threat, and the provider of the metric will leave this impact unspecified. If there is a relation, metrics can have one of three impact values: low (0.1), medium (0.5), and high (1.0). These specific values were chosen based on three principles:

1. The value of ‘high’ should numerically translate to the highest possible value, which is 1.0.
2. The value of ‘medium’ should translate to the middle of the impact spectrum, which is 0.5.
3. The value of ‘low’ should be higher than 0, since there is an impact, but should be significantly lower than the value of ‘medium’.

We allow cybersecurity metrics to relate both positively and negatively to cybersecurity risk. A metric such as ‘level of cybersecurity awareness’ relates negatively to cybersecurity risk since a higher value indicates less security risk. Theoretically, a single metric could even relate positively to the cybersecurity risk associated with one threat, while relating negatively to the cybersecurity risk of another threat. For now, we leave this option open in the algorithm, by specifying a threat-specific Boolean indicator δ_{mt} , which equals 1 when a metric relates positively to cybersecurity risk. For now, the threat-specific nature of this indicator is not accommodated in the data model and we have not encountered metrics that exhibit the mentioned behaviour.

When a threat t in T achieves a high score due to the metric values and metric impacts, we recommend the user to take countermeasures. Let C be the set of countermeasures. We define a similar impact variable i_{ct}

for each countermeasure c in C and threat t in T . Once more, impact is assumed 0 if no impact of a countermeasure on a threat is specified. Otherwise, the impact of a countermeasure on a threat is: low (0.1), medium (0.5), or high (1.0).

For metrics, we want to keep track of which have been calculated for a cyber-system s . Let λ_{ms} be a Boolean indicator variable, which equals 1 if metric m in M has been calculated for cyber-system s in S . Similarly, we let λ_{cs} indicate whether countermeasure c in C has been implemented for cyber-system s in S .

The threat-specific GEIGER score for a cyber-system s in S , which is (a part of) an MSE with profile p in P , for threat t in T , is then given by:

$$G_{spt} = 50 + 50 \times \frac{\sum_{m \in M} \delta_{mt} \lambda_{ms} i_{mt} v_{ms}}{\sum_{m \in M} \delta_{mt} \lambda_{ms} i_{mt}} - 25 \times \left(\frac{\sum_{m \in M} (1 - \delta_{mt}) \lambda_{ms} i_{mt} v_{ms}}{\sum_{m \in M} (1 - \delta_{mt}) \lambda_{ms} i_{mt}} + \frac{\sum_{c \in C} \lambda_{cs} i_{ct}}{\sum_{c \in C} i_{ct}} \right)$$

In words, the above equation can be captured as:

$$G_{spt} = 50 + 50 \times \text{normalized impact score increasing metrics} \\ - 25 \times (\text{normalized impact score decreasing metrics} \\ + \text{normalized impact implemented countermeasures})$$

We assume for simplicity in the above equations that at least one of each type of metric has been calculated and at least one countermeasure has been implemented, so that we do not divide by zero. In practice, when nothing has been calculated for a specific element, we set the value of this element to 0. This implies that the starting threat-specific GEIGER score is always 50. The score ranges from 0 to 100.

We choose to normalize the metric values by whether calculation took place, so that the threat-specific GEIGER score is immediately responsive. If we would normalize by all relevant metrics, the first results could have a very low impact on the score, giving the user a false impression of security.

The total GEIGER score for the cyber-system s in S , which is (a part of) an MSE with profile p in P is then given by:

$$G_{sp} = \frac{\sum_{t \in T} G_{spt} r_{pt}}{\sum_{t \in T} r_{pt}}$$

Note that this weighted average approach yields an overall cyber-system GEIGER score ranging between 0 and 100, since the threat-specific GEIGER scores also ranged between 0 and 100.

When we consider practical calculation of scores, different metric values will relate to different sub-systems of the MSE. Additionally, just because a countermeasure has been implemented by a particular employee (e.g., education on phishing), does not mean this implies a change of score for other employees. The employees and devices of an MSE are separate entities, that should be treated separately. Additionally, there will be metrics and countermeasures that apply to the enterprise. Hence, rather than calculating the MSE GEIGER score in the above manner, we will aggregate the scores of the MSE sub-systems, as explained below.

3.1.4.1 Incorporating Hierarchy

MSEs, like any business, are generally organised in a hierarchical structure. In the simplest case, this involves a single MSE owner. The next simplest case is a two-person company where the MSE owner manages a single employee. More complex hierarchies can involve intermediate managers who supervise groups of employees.

Due to privacy concerns, GEIGER users cannot view any personal security information of other employees within their MSE without their consent. This is true even for the MSE owner. Hence, security information sharing needs to be approved by the concerned subjects and consent can be withdrawn at any time.

In practice, a manager may want to see the GEIGER scores of the employees they manage. We facilitate this in the following manner. By default, the initial GEIGER user within the MSE is the main user. This may be the MSE owner, or an IT specialist within the MSE. As more employees start using the GEIGER application, they will need to be matched to their MSE.

Any employee can use GEIGER to calculate their GEIGER indicator scores. However, unless they are paired with a supervisor and are incorporated in the MSE hierarchy, their scores cannot be incorporated in the overall MSE score. The exception is the main GEIGER user of the MSE, who serves as the root node for the hierarchy and therefore does not require a supervisor.

Note that pairing itself does not give any rights to the supervisor, other than allowing the supervisor to request the supervised employee to share their data. The pairing process establishes a connection between toolbox instances. This allows a supervisor to request security information from the employees they supervise. An employee receives the request and can choose to consent to his or her current personal security information being shared with the supervisor. Even when a supervisor and an employee are paired, and the employee agrees to sharing, the GEIGER solution only shares aggregated security information.

We restrict each employee to having exactly one supervisor. If the employee accepts the request, the manager is placed above the employee in the GEIGER hierarchy. Our approach additionally requires that there are no supervision loops (i.e., employees supervising their supervisors), since this would create problems in score calculation.

Our approach has as advantage that sharing of (privacy sensitive) data is kept to an absolute minimum. Aggregate data of employees are only shared with their direct supervisor, and only when they provide permission to share. The GEIGER indicator algorithm uses the aggregate data of the supervisor themselves, along with the aggregate data of the employees they supervise, to recursively calculate aggregate scores and pass them up the hierarchy. When we reach the root node of the hierarchy, the main user of GEIGER within the MSE, we can calculate the overall MSE score.

Figure 41 shows an example of an MSE hierarchy. Each user can view their own employee and device scores per threat, given the profile p of the MSE. They can additionally view aggregated scores of employees they supervise and who have given consent to share data. The overall aggregate score of the main user in the MSE, Ana, represents the MSE GEIGER score. This is the only score that may be shared outside of the MSE.

This is just an example of a hierarchy. The principles described in this text will also apply to single-person companies, companies with a flat hierarchy under the sole supervision of the MSE owner, or companies in which the MSE owner establishes a person managing cybersecurity (also called a Chief Information Security Officer, CISO).

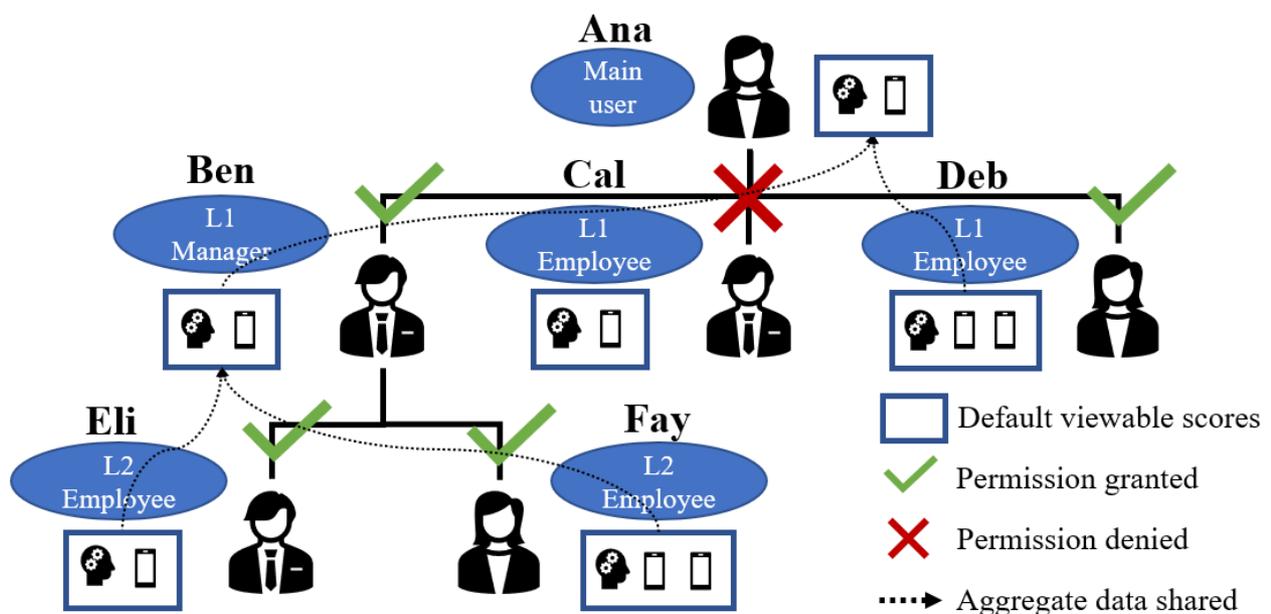


Figure 41 - An example MSE hierarchy.

In the example shown in Figure 41, Eli's aggregate score may be calculated by considering the score of the device owned by Eli and his user score. For Fay, the same process applies, only she includes two device scores

in her aggregate, since she owns two devices. For their manager, Ben, the aggregate score is calculated by considering his user and device scores, as well as the aggregate scores of Eli and Fay.

Similarly, for the MSE main user, Ana, the aggregated score can be calculated by considering her personally owned device and her user profile, as well as the aggregated scores of Ben and Deb. Cal's information is not incorporated in scoring since he has chosen not to share his information.

There is a risk that accounts may be paired in a circular loop. Such a loop would exist, for example, if Ben would pair his device by designating Eli to be his supervisor. Similarly, such a loop would exist if Ana would pair her device by designating Ben or Eli to be her supervisor. As this would make score calculation intractable, a loop detection algorithm should be implemented as a mitigation measure that prevents a pairing attempt if a loop would be created with it.

The scenario in Figure 41 also includes the denial of permission to share aggregate data. Ana supervises three employees, of which two have granted permission to view aggregated scores. Since Ana is still paired to Cal, she is aware that Cal is not sharing data. Ana can initiate a discussion with Cal if she feels his score should be incorporated and that they should jointly collaborate to secure the MSE. A different problem emerges when a user within the MSE is not even paired, as it would not be clear within the GEIGER application that this user is incorrectly omitted from the MSE hierarchy. If it turns out in user testing that there are scenarios where MSE users are not paired, we can take mitigating actions such as clearly communicating the need to pair all participating employees or incorporating automatic pairing suggestions when a user starts using GEIGER.

This example also illustrates how analysis may be performed in the organisation to diagnose the root-causes for a poor MSE GEIGER score. Ana can see from her GEIGER instance whether one of her personal scores are poor, or whether one or more of her directly supervised employees has a problematic aggregate score. Ana can also see which directly supervised employees have decided to not share their aggregate score. In both cases, Ana can initiate discussion about cybersecurity with the concerned person to identify the root causes together in a collaborative, discussion-oriented way. The same can be done by Ben, who supervises Eli and Fay. This stimulation of discussion offers an opportunity in the organisation to approach cybersecurity constructively and collaborate to establish a strong cybersecurity culture in the organisation.

3.1.4.2 Algorithm Aggregation

There are two main types of entities to deal with in MSE GEIGER score aggregation: the employees and the devices. They both play their own role in determining the cybersecurity posture of the MSE, and thus we choose to treat each equally in the total MSE score.

Individual employees and devices are scored according to the calculations described in the previous section. Our hierarchy permits the sharing of aggregate data from an employee to their supervisor. An employee e in S will have a user score pertaining to their cybersecurity knowledge and ability, as well as device scores for any device they are the owner of. We define the subset S_e of S , as the set of cyber-systems belonging to employee e in S . This includes the employee and the devices they own.

The more data we have on a particular cyber-system s in S , the more certain we are of the GEIGER score of that system. When we are more certain of a GEIGER score relating to system a in S than we are of system b in S , the GEIGER score of a should receive a higher weight. To incorporate this concept, we define the variable n_s for a system s in S to be the total number of metrics calculated for that system s . Using this n_s term in scoring, allows us to naturally ensure that we prioritise those systems for which we have more information in the scoring mechanism.

Although implementing recommendations influences scoring, we do not count implemented recommendations as contributing to the n_s term. The reason is that we assume that when a recommendation has not been indicated as implemented, that it has not been implemented. This means that becoming aware of a recommendation implementation does not add information, it simply alters the state of the recommendation. For metrics this is different, as we do not assume a default value.

Assuming for the moment that the employee does not supervise anyone, we define the aggregate score of this employee e , who is a part of an MSE with profile p in P , to be:

$$G_{ep}^{agg} = \frac{\sum_{s \in S_e} G_{sp} n_s}{\sum_{s \in S_e} n_s}$$

We assume here that for at least one system s in S_e , n_s is greater than zero. This aggregate score can be shared with a supervisor, so that the supervisor can incorporate the aggregate score in their own aggregate score. This implies that the above equation becomes more complex in the case that employee e in S supervises one or more other employees. Let E denote the set of employees, a subset of S . Then, we indicate that employee \hat{e} is directly supervised by employee e , by stating that \hat{e} is in the set E_e . The following formula corresponds to the complete aggregate score calculation:

$$G_{ep}^{agg} = \frac{\sum_{s \in S_e} G_{sp} n_s + \sum_{\hat{e} \in E_e} G_{\hat{e}p}^{agg} n_{\hat{e}}^{agg}}{\sum_{s \in S_e} n_s + \sum_{\hat{e} \in E_e} n_{\hat{e}}^{agg}}$$

Note that this formula equates to the earlier formula when $|E_e|=0$, the situation where an employee does not supervise any other employees. We define the aggregated n term for an employee e in E to be the total number of calculated metrics included in the aggregate GEIGER score of e :

$$n_e^{agg} = \sum_{s \in S_e} n_s + \sum_{\hat{e} \in E_e} n_{\hat{e}}^{agg}$$

To be able to perform this calculation, two items need to be shared by the supervisee \hat{e} with their supervisor e :

1. $G_{\hat{e}p}^{agg}$: The aggregate GEIGER score of \hat{e} .
2. $n_{\hat{e}}^{agg}$: The total number of calculated metrics used in the aggregate GEIGER score of \hat{e} .

We assume that only the main user has no supervisor. The aggregate score for the main user serves as a proxy for the MSE score, since all the scores of other employees and devices are incorporated in this aggregate score. Letting α in E be the main user, we obtain for an MSE with profile p in P :

$$G_p^{MSE} = G_{\alpha p}^{agg}$$

For a list of all variables used in the GEIGER indicator and their definitions, see Appendix D. Figure 42 shows an example of how hierarchical aggregation is used to recursively calculate scores, to eventually arrive at the MSE score. Eli has a user score of 60, which resulted from the calculation of 4 metrics. He additionally owns one device, with a GEIGER score of 20, resulting from the calculation of a single metric. Eli calculates his aggregate GEIGER score as:

$$G_{Eli}^{agg} = \frac{60 \times 4 + 20 \times 1}{4 + 1} = \frac{260}{5} = 52$$

Eli then passes his aggregate score of 52 on to Ben, along with the total number of metrics used to arrive at his score, 5. Fay follows a similar process, with the difference that she is the owner of two devices. Ben can then calculate his aggregate GEIGER score as:

$$G_{Ben}^{agg} = \frac{(10 \times 8 + 90 \times 2) + 52 \times 5 + 55 \times 6}{8 + 2 + 5 + 6} = \frac{850}{21} \approx 40.48$$

For the purposes of this demonstration, we round off this number to 40. Ana incorporates the aggregate scores of Ben and Deb in her calculations, as well as her own user and device scores, to arrive at her aggregate score of 35. Since Ana is the main user within the MSE, this is also the MSE GEIGER score.

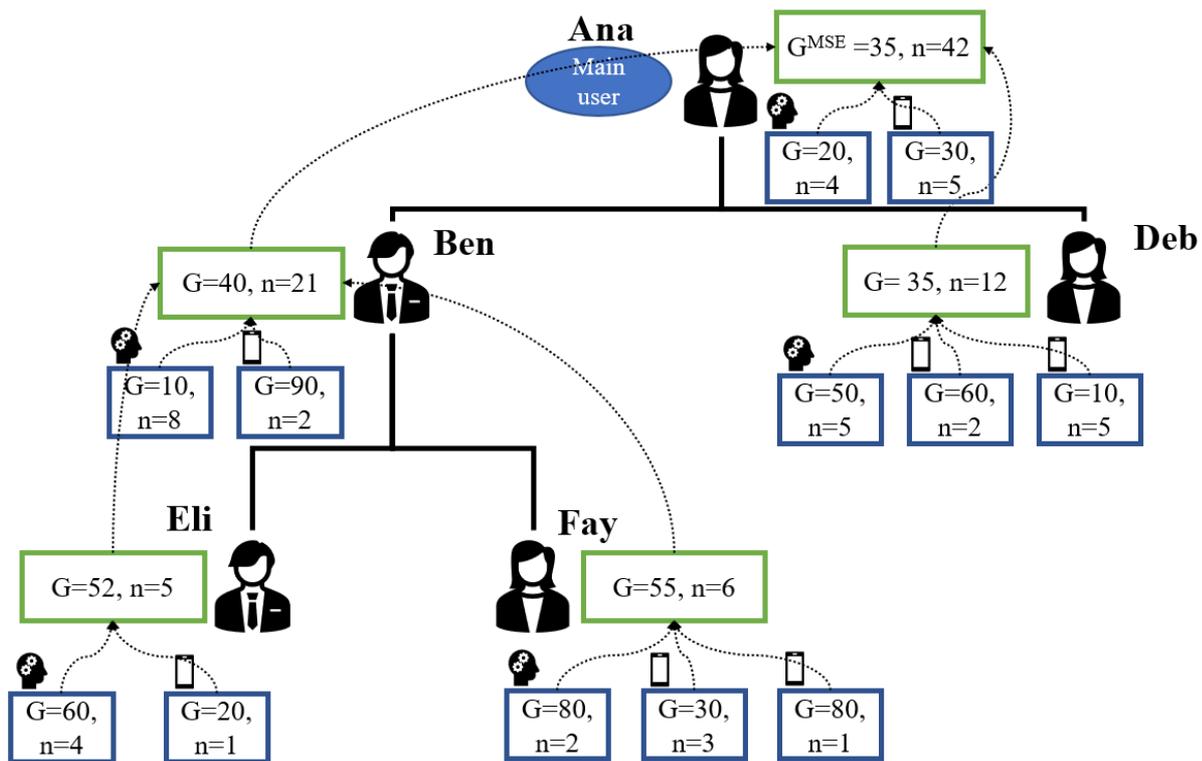


Figure 42 - GEIGER Indicator aggregation

3.1.5 GEIGER Indicator Output

This section focuses on the output of GEIGER Indicator, which is composed of: GEIGER score and GEIGER recommendations. In addition, we discuss the process involved in updating GEIGER scores.

3.1.5.1 Initial Output

As mentioned in the previous sections, three types of scores are generated: employee aggregate scores, user scores, and device scores.

1. GEIGER employee aggregate score: G_{ep}^{agg} . The score can be viewed by the user it applies to, but is not threat-specific.
2. GEIGER User Score: The current logged-in user can view their user-specific score G_{sp} , the user can also view the score of each threat affecting the user score separately G_{spt} .
3. GEIGER Device Score: The logged-in user can view their current device score G_{sp} , the user can also view the score of each threat affecting the device score separately G_{spt} .

Additionally, as explained in 3.1.4.2, a supervisor can view the aggregated score of the employees under their supervision if the employee consents to data sharing. In addition, aggregate threat-specific score of the current user and the current device in use is displayed to the user.

For the User and Device GEIGER scores a set of recommendations is provided for each threat. The provided recommendations are ranked based on their impact on the GEIGER score. It is worth nothing that only recommendations matching the role and the knowledge level of the current user are displayed.

3.1.5.2 Output Update

There are four scenarios that can result in an update in the GEIGER score:

- A change in an existing or a new sensor (metric) value.
- Implementation of a recommendation.
- Update in pairing or sharing information of employees. Updates in threat information obtained from CERTs.

In either scenario, GEIGER score is updated, and an updated set of recommendations is provided.

It is important to highlight that implementation of a recommendation is indicated by a user and is only tied with the global set of available recommendations, as soon as the recommendation is implemented the user will notice a change in the GEIGER score.

In the case of recommendations that are raised by tools, an implementation flag is not needed, as the recommendation will result in a change in the metric value of the tool that raised the recommendation. Hereby, this resembles the first scenario as a result the GEIGER score will updated.

3.1.6 GEIGER Indicator Sustainability

3.1.6.1 Impact Determination

The determination of impacts – either low, medium, or high – forms an important part of the GEIGER Indicator solution. The impact of each metric on each threat must be defined, as well as the impact of each countermeasure on each threat. Given the absence of internally available data in the GEIGER Indicator solution during the development phase, impacts must be determined with the help of cybersecurity standards and frameworks and human (expert) input.

For the initial determination of metric impacts on threats, the owners of the tools producing these metrics were requested to provide the impacts. This process was guided by the GEIGER Indicator development team, to ensure any misunderstandings were addressed adequately. Throughout this process it became clear that tool owners have an intricate understanding of the metrics resulting from their tools. They were able to link relevant threats to their metrics, but the decision process to arrive at impacts was experienced as challenging.

This led us to develop guidelines for impact determination for both the tool owners determining the impact of their metrics on threats, as well as for the GEIGER Indicator developers themselves, who have the task of determining the impact of countermeasures on threats.

The Framework Core Functions of the NIST Cybersecurity Framework (Barrett, 2018) and 20 Critical Security Controls (CSCs) of the Centre for Internet Security (CIS, 2019) help in this regard. The NIST Cybersecurity Framework (CSF) includes the following functions: identify, protect, detect, respond, and recover. These functions are directly related to the goals we aim to achieve with our countermeasures and are commonly used to aid cybersecurity decision making (OWASP, 2016; Dutta and Al-Shaer, 2019). Similarly, metrics can be mapped to the Framework Core Functions. For example, a metric resulting from awareness training for employees, will be mapped to the 'Awareness and training' category of the 'Protect' function, meaning it will receive a default impact of 'High'.

Simultaneously, the functions relate to progressive stages of cybersecurity incidents, from before the incident (identify and protect), to during the incident (detect and respond), to after the incident (recover). Given the decreasing likelihood of each stage occurring, we can assign impacts to metrics and countermeasures based on the function they correspond to. Using this logic, metrics and countermeasures

are assigned a high impact in the identify and protect phases, a medium impact in the detect and respond phases, and a low impact in the recovery phase.

The CIS security controls (CIS, 2019) are helpful in this regard, as they are listed in order of importance. One may expect then that more important controls receive a higher impact in the GEIGER indicator solution. Table 5 maps the NIST CSF Core Functions to the CIS CSCs. CIS performed a similar exercise for an earlier version of their controls and the core functions (Sager, 2015), which we translated to the current contents of both frameworks. Apart from a couple of minor changes in content and a changed ordering of the CIS controls, the frameworks are largely the same as in 2015.

We can see from Table 5 that the 'Identify' function is given the highest importance by far within the CIS CSCs. In general, our approach maps well to what the CIS controls indicate, except for the 'Protect' function, which could be assigned a 'Medium' impact based on Table 5. However, given the fact that 'Awareness and Training' and 'Data Security' are important in the MSE context, as witnessed in the GEIGER requirements collection phase, we feel justified in assigning a 'High' impact to the 'Protect' function.

Table 5: Mapping of the NIST CSF functions to CIS Critical Security Controls (CSCs).

NIST Cybersecurity Framework		CIS CSCs		GEIGER
Functions	Categories	Control #	Average #	Impact
Identify	Asset Management	1, 2	2.0	High
	Business Environment			
	Governance			
	Risk Assessment	3		
	Risk Management Strategy			
	Supply Chain Risk Management			
Protect	Identity Management and Access Control	4, 14, 15, 16	11.4	High
	Awareness and Training	17		
	Data Security	13		
	Information Protection Processes and Procedures	5, 9, 11, 18		
	Maintenance			
	Protective Technology	7, 8		
Detect	Anomalies and Events	6, 19	10.7	Medium
	Security Continuous Monitoring	3, 8, 16		
	Detection Processes	12		
Respond	Response Planning	19	12.0	Medium
	Communications			
	Analysis	6		
	Mitigation	3		
	Improvements	20		
Recover	Recovery Planning	10	15.0	Low
	Improvements	20		
	Communications			

The expert(s) deciding on specific impacts can evaluate for each metric or countermeasure individually whether it is warranted to deviate from our default impact determination. A possible reason for this could be that the metric does not easily map to one of the NIST CSF Functions. We already see this happening in Table 5, as for example control number 16 ("Account monitoring and control") is mapped to both the 'Protect' and 'Detect' functions.

Another cause for deviation can be that tool owners feel their metric relates strongly to one threat, but weakly to another. An example is a metric that measures whether a specific e-mail protection is in place. This maps to the NIST CSF 'Protective Technology' category, which is part of the 'Protect' function. Therefore, by default this metric should receive a 'High' impact for the threats it relates to. If a tool owner judges the metric to relate strongly to 'Phishing' but weakly to 'Malware', they should be able to let this be reflected in their impact determination.

To provide experts a structured manner to deviate from the default approach, we use a multiple-criteria decision making (MCDM) tool, such as the Analytic Hierarchy Process (AHP) tool developed by Goepel (2013). Using this approach allows us to offer security experts flexibility in their impact determinations, while maintaining a unified GEIGER Indicator solution. All the impact decisions are made with the support of the GEIGER Indicator development team.

3.1.6.2 Concept Adaptability

With the dynamic nature of the cyber threat landscape, any cybersecurity risk assessment tool should be adaptable (Evesti and Ovaska, 2013). This is especially true for the GEIGER Indicator solution, given the diverse set of MSEs that should be assessed. The importance of considering MSE characteristics to tailor cybersecurity solutions is well-recognised (Mijnhardt et al., 2016). Our indicator concept addresses the need for adaptability by creating MSE profiles that allow us to determine different threat risks for different types of MSEs. Additionally, we allow for the possibility of metrics and recommendations that are only applicable under certain conditions, such as whether an MSE possesses a particular asset. In this sense, the GEIGER indicator concept is adaptable and dynamic by nature.

Of course, the cyber threat landscape itself can change, as witnessed by the changes in the ENISA Top 15 Threats over the years (ENISA, 2020). In the context of machine learning and data mining, this is known as concept drift (Widmer and Kubat, 1996). The consequence of concept drift is that the threat risks originally determined for the GEIGER Indicator algorithm may no longer be valid. The GEIGER Indicator solution will solve this issue by using incident data collected from CERTs to automatically update the threat risks for all MSE profiles. The aggregated data over a period of one month will be used to update the threat risks. This new information will be given a relatively low weight, to not completely alter the threat risks every month. By using this data, we will in future be able to extend the MSE profiles beyond MSE categories (digitally dependent, digitally based, digital enabler), to also include the MSE sector and MSE country.

For the GEIGER solution, adaptation also entails adapting to changes in the GEIGER toolbox. New tools may be added, and old ones may be deleted. Tools may also alter the metrics and recommendations they provide to the GEIGER solution. Given that tool owners follow the instructions for metric and recommendation impact determination outlined in Section 3.1.5.1, their inputs to the GEIGER solution are automatically processed by the GEIGER Indicator solution. It is possible that changes in the GEIGER Toolbox result in a lack of coverage for specific GEIGER threats. When this is the case, the GEIGER indicator algorithm will detect this and notify the GEIGER curator through an event in the GEIGER Data storage. Ideally, the GEIGER curator identifies coverage issues beforehand when negotiating with tool owners.

When changes in the GEIGER Toolbox affect GEIGER Indicator scores, GEIGER users should also be informed. Once again, users are notified through an event in the data storage, which is collected and displayed by the GEIGER user interface. Given that these changes may be experienced as confusing by users, we will implore tool owners to only make changes to their inputs when necessary.

3.1.6.3 Concept Resilience

Besides being adaptable, the GEIGER Indicator solution should be resilient. By resilience we mean that the GEIGER Indicator solution should be able to anticipate, avoid, and adjust to shocks in the external environment (Ortiz-de-Mandojana and Bansal, 2016). We distinguish resilience from adaptability, to emphasize that resilience relates to controls infused in the GEIGER Indicator solution to ensure that it is robust to changes.

These changes could be both on the micro level (e.g., individual MSEs) and on the macro level (e.g., cyber threat landscape). Regarding individual MSEs or groups of MSEs, there may be issues with missing or dirty data (Kim, 2003). When we do not have enough data to calculate a threat score for an MSE, we can indicate this to the user. When there is enough data, but proportionally only a small amount of the possible information is supplied by the MSE, we can additionally indicate this to the user. Since the GEIGER indicator algorithm keeps track of which metrics have been calculated and which metrics can be calculated for an MSE, we can indicate to the user a degree of uncertainty for the GEIGER scores. This uncertainty will decrease as more information is provided by the MSE. In this way, the GEIGER Indicator algorithm is not only able to deal with missing data but can also communicate to the user that data is missing and what effect this has on the GEIGER scores.

Dirty data is more difficult to address, as the GEIGER Indicator algorithm has no way of detecting that data is dirty. An example of dirty data is incorrectly entered data (Kim, 2003). If data is collected through interaction with the user, and the user makes an error in this process, then dirty data enters the GEIGER Data storage.

We posit two solutions. Firstly, given the adaptable nature that GEIGER operates in, it is justifiable to periodically ask users to ratify the data they have entered. This not only offers a possibility to update the data in case of changes, but also a possibility to correct data that may have been entered incorrectly. We suggest that all manually entered data should be updated at least every month. Secondly, the GEIGER user interface should include controls to prevent incorrect data entry where possible. Instructions to the user should be clear and whenever data is being entered that will be used for the GEIGER Indicator, we should perform a confirmation check with the user before storing the data in the GEIGER Data storage.

Issues with metric data are not the only challenge. Changes in the cyber threat landscape may be so extensive that they change the set of threats that are relevant to MSEs. In this situation, the measures discussed in Section 3.1.5.2 to counter concept drift are insufficient, since they only apply when the set of threats remains the same. When changes in the cyber threat landscape go beyond concept drift, we must update the GEIGER indicator threats manually. We advise to evaluate the need for updating every three months, based on the input we have received from several of the CERTs involved in the GEIGER project.

We constructed the set of threats for the GEIGER Indicator solution with resilience in mind. By including 'external environment threats' and grouping threats together whenever possible, we hope to have ensured that our initial set of threats remains relevant for an extended period. Of our 12 GEIGER Indicator threats, 11 were already included in some form in the top ENISA threats of 2012 (ENISA, 2012); at the time of writing nearly a decade ago. The only threat not included in the 2012 top threats is the 'external environment threats' category. However, given the current trends in legislation such as GDPR and the increase in supply chain attacks, this threat can be expected to gain in relevance in coming years.

The issues and controls discussed in the above sections are presented in Table 6. We trust that the controls built into the GEIGER Indicator solution will ensure a sustainable future for the GEIGER Indicator within the GEIGER solution.

Table 6: Controls to ensure the adaptability and resilience of the GEIGER indicator concept.

Issue	GEIGER Indicator Control	Period
Concept drift cybersecurity threats	Update threat risks using aggregated CERT data.	1 Month
No threat coverage due to GEIGER Toolbox changes	Detect lacking coverage and notify GEIGER curator and GEIGER users.	Automatic
No threat coverage due to missing data	Notify user of lacking coverage.	Automatic
Lacking threat coverage due to missing data	Communicate uncertainty in GEIGER scores due to missing data to user.	Automatic
Incorrectly entered data	Periodically ask users to ratify data they have previously entered.	1 Month
Incorrectly entered data	Include controls in the GEIGER user interface, such as confirmation checks when entering GEIGER indicator data.	Automatic

Changed cyber threat landscape	Periodically evaluate the need for updating the GEIGER threat list. Update if necessary.	3 Months
Changed cyber threat landscape	Construct the list of GEIGER threats with resilience and potential future developments in mind.	Automatic

3.2 GEIGER Toolbox

3.2.1 Toolbox User Experience and Interface Design

3.2.1.1 User Journey and Personas

The toolbox is the main instrument that GEIGER offers to MSEs for assessing their security and receiving guidance and help for improving their security. In the overall user journey specified in D1.1, the toolbox is encouraged to be downloaded through a link available on the GEIGER homepage (D1.1 Figure 10 Steps 4-6) and is used for personalised assessment (Steps 7-12), guidance and help for doing improvements (Steps 13-14), and recognition for achieved improvements (Step 14-16). Figure 43 gives a summary overview of that user journey.

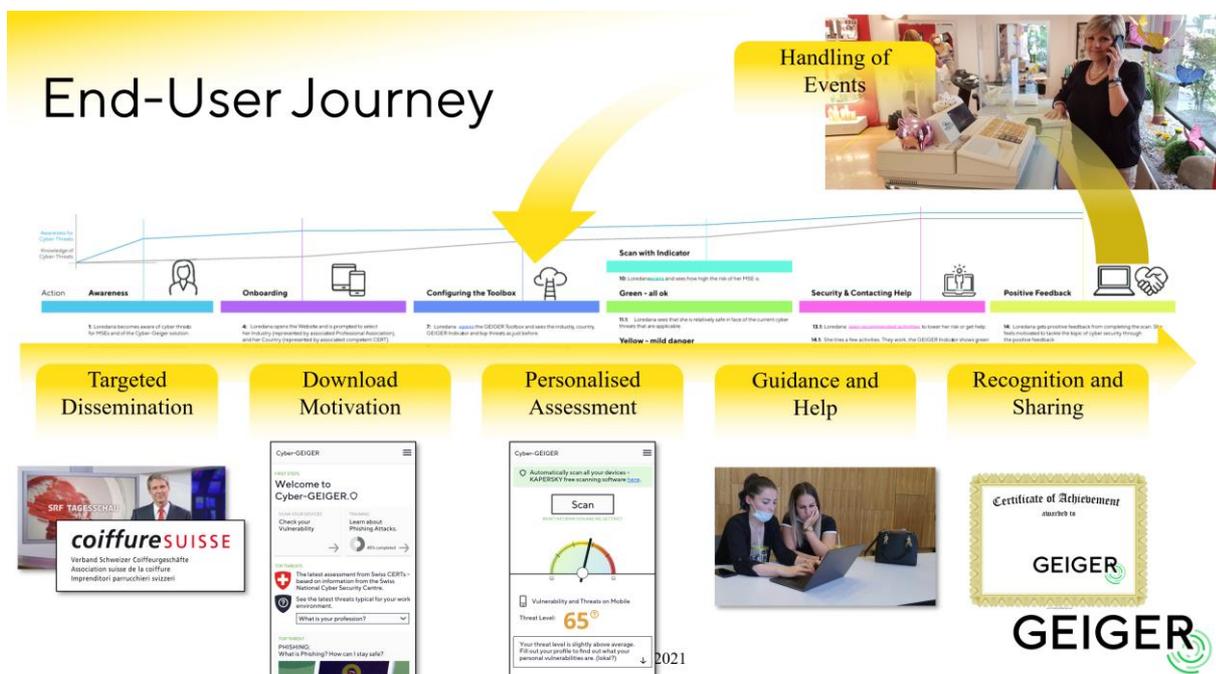


Figure 43: Summary of User Journey (details: see user journey specified in D1.1).

The GEIGER Toolbox is confronted with multiple categories of MSEs and multiple types of end-users within the MSEs. The categories of MSEs are those described in D1.1 Section 3.2.1 and cover digitally dependent MSEs, digitally based MSEs, digital enabler MSEs, and start-up MSEs. The sizes of the MSEs range from one to ten people. The MSE may have a Chief Information Security Officer (CISO) appointed, or that role is knowingly or unknowingly implemented by the owner of the MSE. The ICT knowledge level is diverse with a tendency of low awareness about cyberthreats and cybersecurity solutions that are not maintained over time. The ICT environment is diverse among MSEs, covering the mobile platforms Android and iOS and the PC-based platforms Windows, macOS, and Linux. The toolbox targets primarily the lowest-maturity categories, starting with the use case MSE Coiffeur Loredana and extending to the increasingly more mature use case MSEs. The toolbox helps them to improve, while considering all platforms to be supported.

Within the MSEs, the GEIGER toolbox differentiates multiple personas that have been drawn from the interaction with the use case MSEs. Table 7 gives an overview.

Table 7: Personas supported by the Toolbox.

Persona	Characterisation	Support by Toolbox
CEO	The owner of the MSE. Has a vital interest in the business outcome and survival of the MSE, which are affected by risk. In addition to being an employee, wants to understand the risk of the MSE and achieve protection with minimal effort or for free and by involving the employees. Decides about the reporting of incidents to authorities like CERTs.	Gives risk score aggregated from personal knowledge and devices and supervised employees. Shows how each supervised employee contributes to the score to allow initiation of discussions about how to reduce risk. Can submit incident to CERT.
Employee	Reports to a supervisor or CEO. Is accountable for cybersecurity of his/her devices and good cybersecurity and data protection behaviour. Needs to be motivated to contribute to security in the MSE. Is interested in privacy towards others, including the CEO, and secrecy of the work done for the company.	Gives risk score aggregated from personal knowledge and devices. Allows to push risk score to supervisor. Offers recommendation for risk reduction. Notifies about incidents and offers help for resolution. Offers privacy settings and control over tools running on devices.
Supervisor	In addition to being an employee, supervises employees. Is a role model for cybersecurity in the MSE. Is accountable for cybersecurity in his/her area of influence. Is interested in involving the supervised employees for the protection of the area.	Gives risk score aggregated from personal knowledge and devices and supervised employees. Shows how each supervised employee contributes to the score to allow initiation of discussion about how to reduce risk.
CISO	Supervisor with the responsibility of coordinating the cybersecurity of the MSE, including all employees. Reports protection results, risk of the MSE, and any incidents to the CEO.	Same as for supervisor
Current or Former Malicious Employee	Shows disappointment or other negative feelings about the MSE. Engages in espionage, weaponization, blackmailing, or sabotage to hurt the company.	Constrains visibility of security information to own knowledge and devices. Gives the CEO the ability of removing the employee in GEIGER from the company.

3.2.1.2 Wireframe of the Toolbox User Interface

The user interface of the GEIGER Toolbox has been designed by considering the functional and quality requirements stated in the deliverable D1.1, by consulting experts for cybersecurity in MSEs and by following a state-of-the-art design process⁹.

The following steps were pursued to define the content and structure of the toolbox user interface. These results were documented and validated with expert review in the form of a low-fidelity wireframe.

⁹ The toolbox UI designers based their method thinking on <https://www.netsolutions.com/insights/user-experience-design-process/> and <https://www.uxpin.com/studio/ui-design/what-is-a-wireframe-designing-your-ux-backbone/>.

1. Study in-depth the user journey and functional and quality requirements.
2. Search for existing solutions with similar features and identify commonly used designs.
3. Create a content inventory that identifies the screens and describes the elements of these screens.
4. Define the structure of the screens with a wireframe.

The aim of the wireframe was to show all the features and content based on the requirements and to build up a logical information architecture from this. The wireframes allowed to define the human-computer interaction design to a large extent. The later phase of high-fidelity design will describe how the wireframes are implemented.

The following summarises the UI design for each feature exposed by the user interface. Directly contributing the steps of the user journey are the features personalized assessment, guidance and help, and handling of events, including incident resolution and reporting. Features indirectly supporting the use of the toolbox are the navigation, pairing of devices and employees, management of tools, incident resolution and reporting, and toolbox settings.

Personalized Assessment

Personalized assessment includes the user interface that provides the human end-user with the ability to interact with the GEIGER Indicator. It offers the scan button to trigger the indicator algorithm and offers the human end-user's global score as well as the scores for the top-five threats the end-user is exposed to. It allows to drill down into a specific threat by showing the user-specific and device-specific scores and offering a list of recommendations ranked according to impact on risk reduction. All scores and recommendations are retrieved from the GEIGER Indicator. The user interface also provides the human end-user to compare his score with score statistics of similar companies, for him- or herself over time, and list the history of events that have contributed to changes in the score over time.

Figure 44 shows the toolbox wireframes.

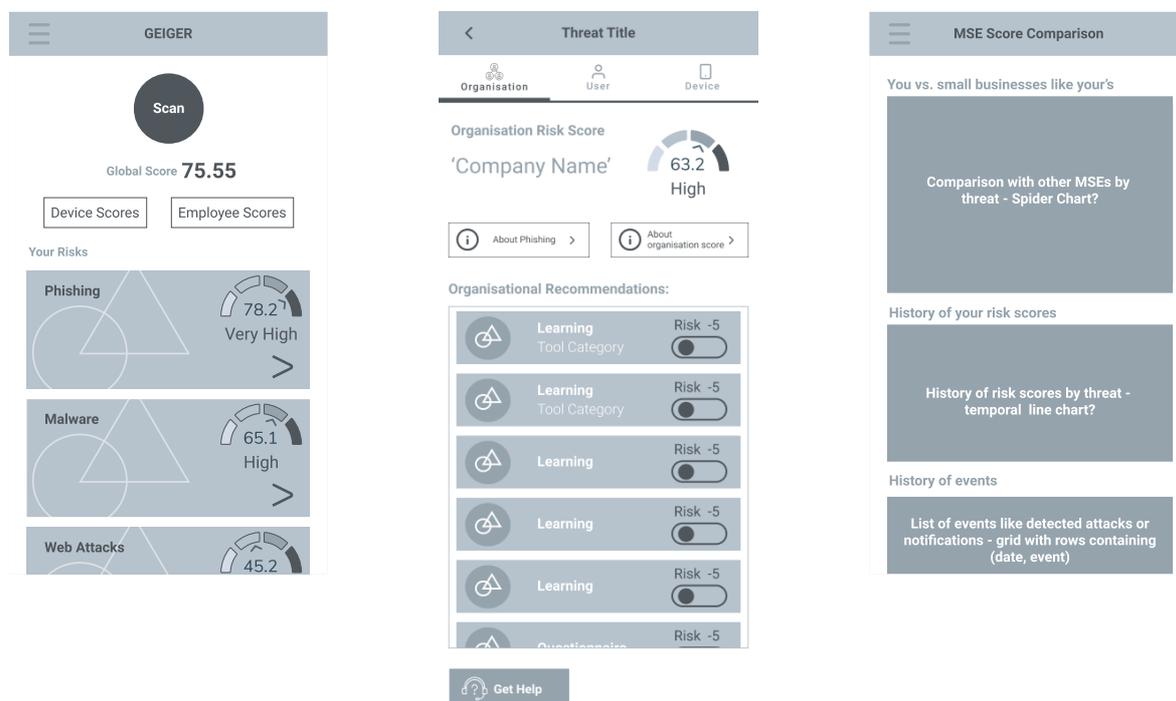


Figure 44: Wireframes for personalised assessment.

Table 8 specifies the user stories supported by the toolbox user interface and refining the Requirements stated in D1.1:

Table 8: Requirements-refining user stories for personalized assessment.

Feature	Requirements-refining User Stories ¹⁰	Comments
T.F02 MSE Profiling	T.F02.R01: TU01 The user can see the risk score. TU02 The user can see the risk scores of his/her top threats. TU03 The user can inspect his/her risk associated with a threat. TU04 The user can see the risk scores of paired devices ¹¹ . TU05 The user can see the risk scores of paired employees ¹² .	Figure 44 left. Figure 44 left. Figure 44 left -> mid. Figure 46-3 Figure 47-3
T.F02.2 Scanner	T.F02.R21: TU06 The user can trigger a scan of the user, device, and paired devices and employees ¹³ . T.F02.R22 is without user involvement.	Figure 44 left.
T.F03 GEIGER Indicator and Recommendations	T.F03.R01: TU07 The user can see the risk score for a given threat. T.F03.R02, T.F03.R04 (compliance recommendations): TU08 The user can see the top risk-reducing recommendations. T.F03.R03 is without user involvement. T.F03.R04 (comparison), T.F03.F05: TU09 The user can see the a spider chart for comparison with the MSE community. TU10 The user can see a line chart for understanding the evolution of the own score. TU11 The user can see a list of events that affected the evolution of the own score.	Figure 44 mid. Figure 44 mid. Figure 44 right. Figure 40 right. Figure 40 right.

Guidance and Help

For guidance and help, tool CYSEC is foreseen. CYSEC introduces the human end-user into cybersecurity topic recommended by the toolbox. Such introduction will explain what the topic is, why the topic is important for the end user, and guide the end user step-by-step through installations of tools, configuration of settings, documenting decisions (e.g., MSE policy), and offering information to employees and other stakeholders. CYSEC is integrated as a plugin into the toolbox.

In addition, guidance and help includes mini questionnaires allowing the end-user to provide feedback to the GEIGER Indicator, thus improving the indicator's accuracy.

¹⁰ Labelling scheme (example): The user story identifier (short identifier: TU01) is appended to the identifier of the respective requirement (full identifier: T.F02.R01.TU01).

¹¹ Paired devices will only share their risk scores and no details about their profile, thus limiting the risk of the GEIGER toolbox becoming a platform for attacks.

¹² Paired employees will only share their risk scores and no details about their profile, thus limiting the risk of the GEIGER toolbox becoming a platform for attacks. The respective risk score is only shown if the respective employee has consented to sharing it.

¹³ Employees that receive the scan request can decide whether they want to share their current risk score, thus fulfilling privacy needs.

In addition, the toolbox provides a directory of certified security defenders that are of relevance given the SME end-user’s geographic area and industry. The security defenders are a source of knowledge and help that otherwise would not be available to the MSE. A Security Defender is only listed if the defender gave consent for being listed.

Figure 45 shows the toolbox wireframes.

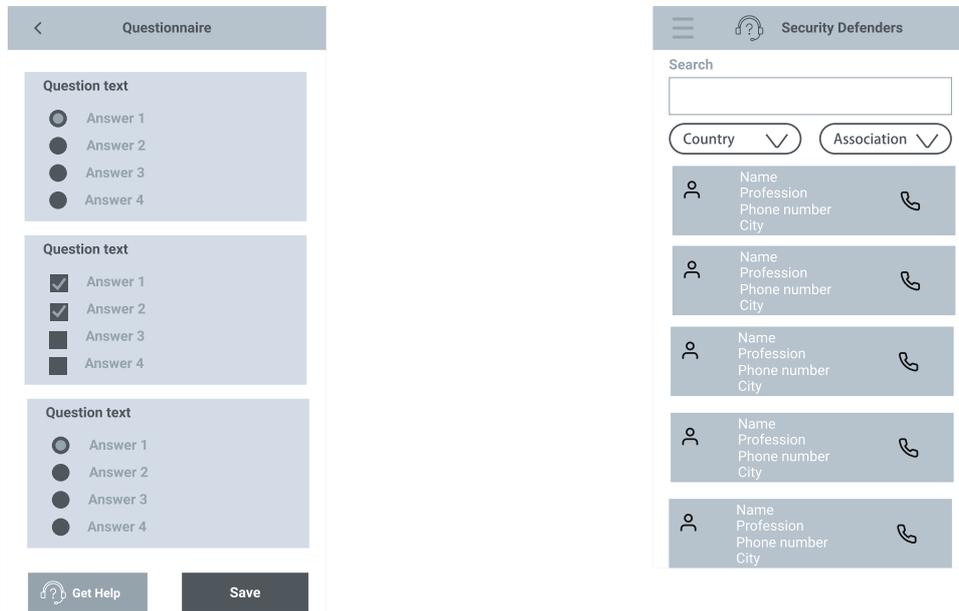


Figure 45: Wireframes for guidance and help.

Table 9 specifies the user stories supported by the toolbox user interface and refining the Requirements stated in D1.1:

Table 9: Requirements-refining user stories for guidance and help.

Feature	Requirements-refining User Stories	Comments
T.F02.1 Questionnaire	T.F02.R11, T.F02.R13, T.F02.R14: TU12 The user can answer a single choice question. TU13 The user can answer a multiple choice question.	Figure 45 left. Figure 45 left.
C.F05 Certified Security Defenders Directory	C.F05.R04: TU14 The user can browse security defenders. C.F05.R05: TU15 The user can filter security defenders by country. TU16 The user can filter security defenders by association. TU17 The user can filter security defenders by keyword.	Figure 45 right. Figure 45 right. Figure 45 right. Figure 45 right.

Pairing of Devices and Employees

The pairing of devices and employees includes the user interface that provides the human end-user with the ability to establish or remove a connection ‘pairing relationship’ between two toolbox instances. The pairing is established by scanning a QR code. The removal on a simple click on a button. The pairing with the QR code is intended to ensure proximity of the two devices, stimulating discussion if an employee is being paired.

Figure 46 and Figure 47 show the toolbox wireframes.

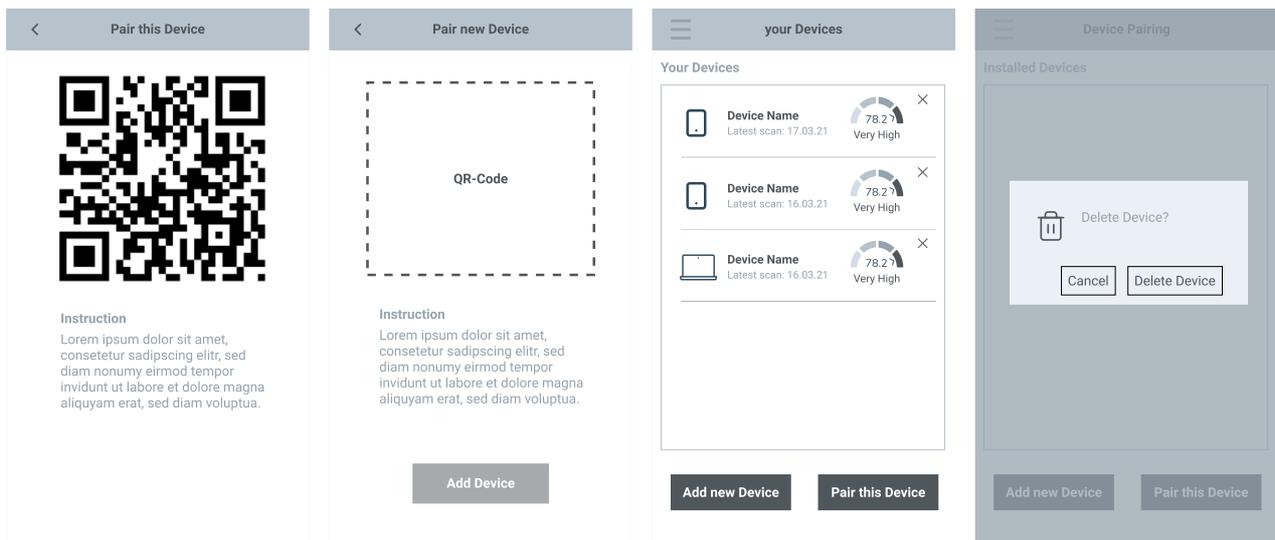


Figure 46: Wireframes for pairing of devices.

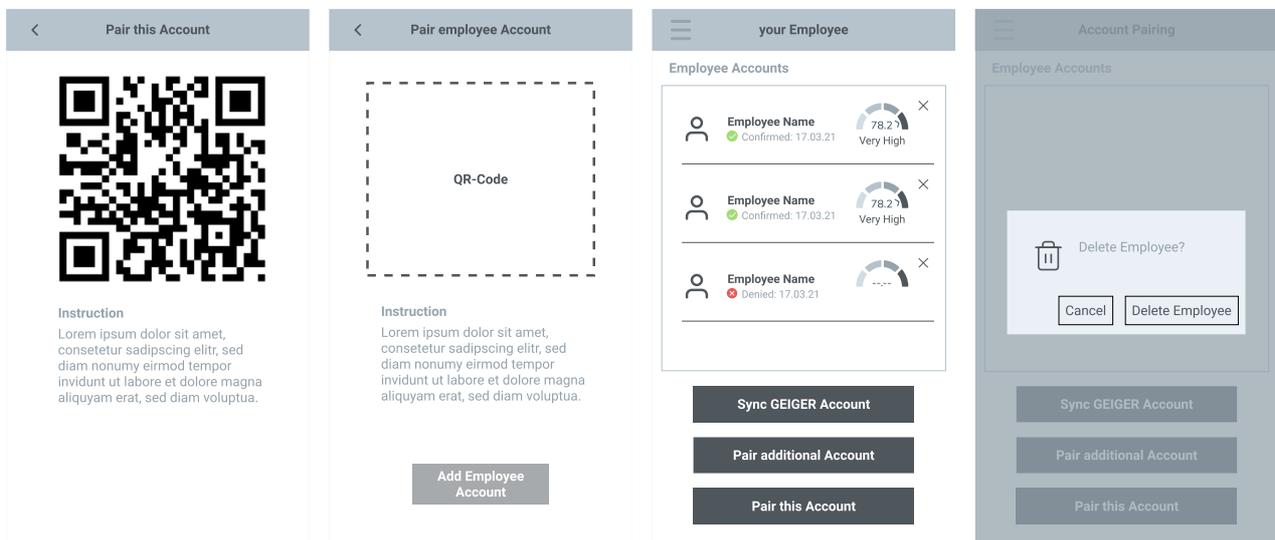


Figure 47: Wireframes for pairing of employees.

Table 10 specifies the user stories supported by the toolbox user interface and refining the requirements stated in D1.1:

Table 10: Requirements-refining user stories for pairing of devices and employees.

Feature	Requirements-refining User Stories	Comments
T.F01.2 Device Pairing	<p>T.F01.R11: TU18 The user can show a QR code for enabling the pairing of a device. Already paired devices will also be paired with the other toolbox.</p> <p>TU19 The user can scan a QR code of a device for establishing a pairing. The user’s name is replicated as the owner of the paired devices.</p> <p>TU20 The user can see the scores of the paired devices.</p> <p>TU21 The user can remove a paired device. The user’s name is removed from the paired device.</p>	<p>Figure 46-1.</p> <p>Figure 46-2.</p> <p>Figure 46-3.</p> <p>Figure 46-4.</p>

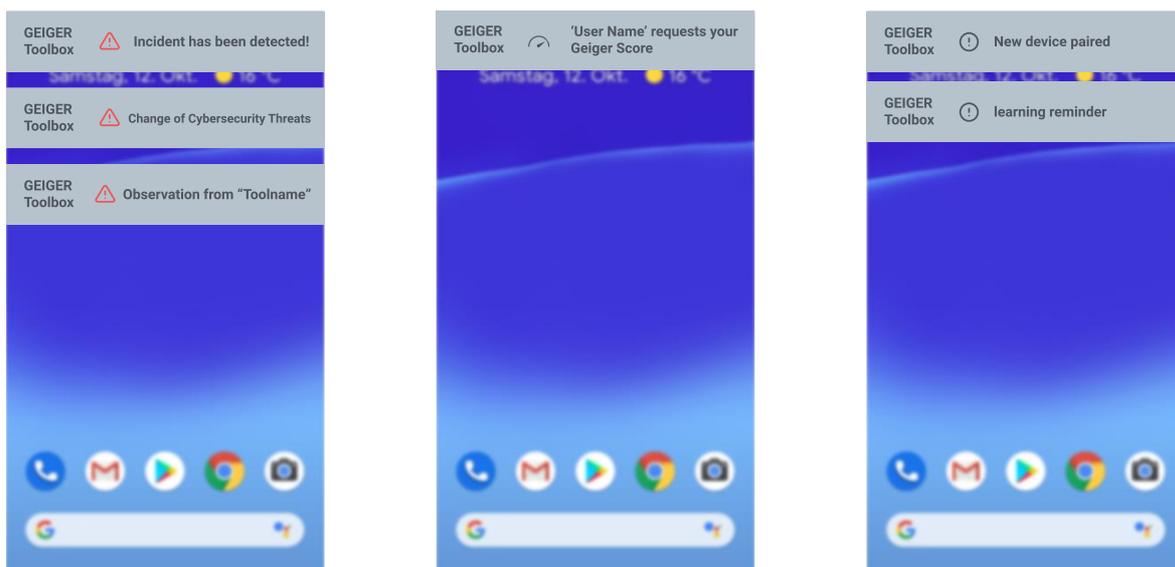
<p>T.F01.2 Cloud Account Pairing</p>	<p>T.F01.R21: TU22 The user can show a QR code for enabling the pairing with the cloud. TU23 The user can scan a QR code for establishing a pairing. The user’s name is replicated as the owner of the paired devices. TU24 The user can remove the pairing with the cloud. T.F01.R22 is without user interaction.</p>	<p>Figure 46-2. Figure 46-4.</p>
<p>T.F01.3 Employee Account Pairing</p>	<p>T.F01.R31: TU25 The user can show a QR code for enabling the pairing with a supervisor. TU26 The user can scan a QR code of an employee for establishing a pairing. TU27 The user can see the scores of the paired employees. TU28 The user can remove a paired employee. TU29 The user can remove a paired supervisor.</p>	<p>Figure 47-1. Figure 47-2. Figure 47-3. Figure 47-4. Figure 47-4.</p>

Step: Handling of Events

The handling of events includes push messages provided to the human end-user and warnings embedded in the toolbox user interface. The messages and warnings allow the human end-user to be aware of events like new threats communicated by the relevant CERT, changes in the prevalence and criticality of existing threats, incidents that the end-user has encountered, and updated scores shared by supervised employees. The end-user can react to the message or warning to get forwarded to the appropriate toolbox screen or integrated tool.

The resolution and reporting of incidents are handled by the KPMG chatbot tool, which is integrated as a plug-in into the toolbox. The detailed chatbot dialogue for guiding the resolution and reporting is specified in the respective section of this document.

Figure 48 shows the toolbox wireframes.



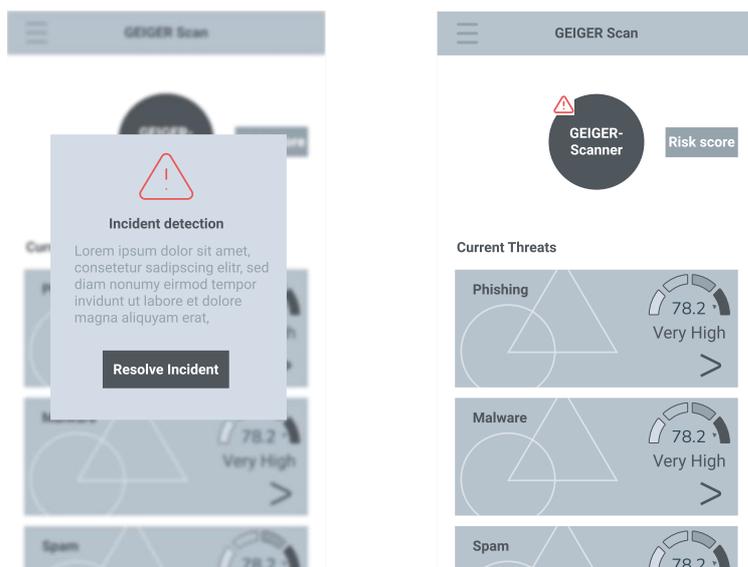


Figure 48: Wireframes for the handling of events.

Table 11 specifies the user stories supported by the toolbox user interface and refining the requirements stated in D1.1:

Table 11: Requirements-refining user stories for the handling of events.

Feature	Requirements-refining User Stories	Comments
T.F01.2 Device Pairing	T.F01.R11 (device pairing): TU30 The user can receive a push message indicating the successful pairing of a device. TU31 The user can receive a push message indicating the successful removal of a paired device.	Figure 48 top-right. Figure 48 top-right.
T.F01.2 Employee Account Pairing	T.F01.R31: TU32 The user can receive a push message indicating the successful pairing with an employee or supervisor. TU33 The user can receive a push message indicating the successful removal of a paired employee or supervisor. New capability: security information sharing TU34 The user can receive a push message indicating the supervisor's request of the user's updated score. The message forwards the user to the toolbox main screen. TU35 The user can receive a push message indicating an employee's updated score. The message forwards the user to the employees screen.	Figure 42 top-right. Figure 48 top-right. Figure 42 top-mid. Figure 44 left. Figure 47-3.
T.F02.2 Scanner	T.F02.R21: TU36 The user can receive a signal on the main screen that the user's score is outdated and a new scan needs to be performed.	Figure 42 bottom-mid.
T.F02.3 Education Reporting	T.F02.31 is handled within the CYSEC tool. T.F02.R32 is without user involvement. T.F02.R33: TU37 The user can receive a push message indicating an	

	<p>educational achievement. The message forwards the user to the toolbox main screen.</p> <p>New capability: learning reminder TU38 The user can receive a push message indicating the need to revisit a learning experience. The message forwards the user to the learning tool with the learning experience as a parameter.</p>	<p>Figure 44 left.</p> <p>Figure 48 top-right.</p>
T.F05.1 Incident Notification	<p>T.F05.R11 is without user involvement.</p> <p>T.F05.R12: TU39 The user can receive a push message indicating a tool observation. The message forwards the user to the main screen showing incident data. TU40 The user can receive a message describing an incident. The message forwards the user to the chatbot for incident resolution and reporting.</p>	<p>Figure 48 top-left.</p> <p>Figure 48 bottom-left.</p>
T.F07 Threat Updates	<p>T.F07.R01: TU41 The user can receive a push message indicating a change of threats. The message forwards the user to the toolbox main screen.</p> <p>T.F07.R02: TU42 The user can receive a push message indicating a change of data protection regulation. The message forwards the user to the toolbox main screen.</p> <p>T.F07.R03: TU43 The user can receive a push message indicating a change of recommendations. The message forwards the user to the toolbox main screen.</p>	<p>Figure 48 top-left.</p> <p>Figure 44 left.</p> <p>Figure 48 top-left.</p> <p>Figure 44 left.</p> <p>Figure 48 top-left.</p> <p>Figure 44 left.</p>

Management of Tools

The GEIGER toolbox is an open platform that intends to allow any third-party cybersecurity tool to be integrated as a plugin. The open character allows extensibility to the massive market of cybersecurity tools and technologies, hence creates innovation potential. The toolbox governs data exchange with the tool and offers the matching between a recommendation and an installed tool. The user interface provides the end-user with the ability to inspect installed tools, add tools, and remove tools.

Figure 49 shows the toolbox wireframes.

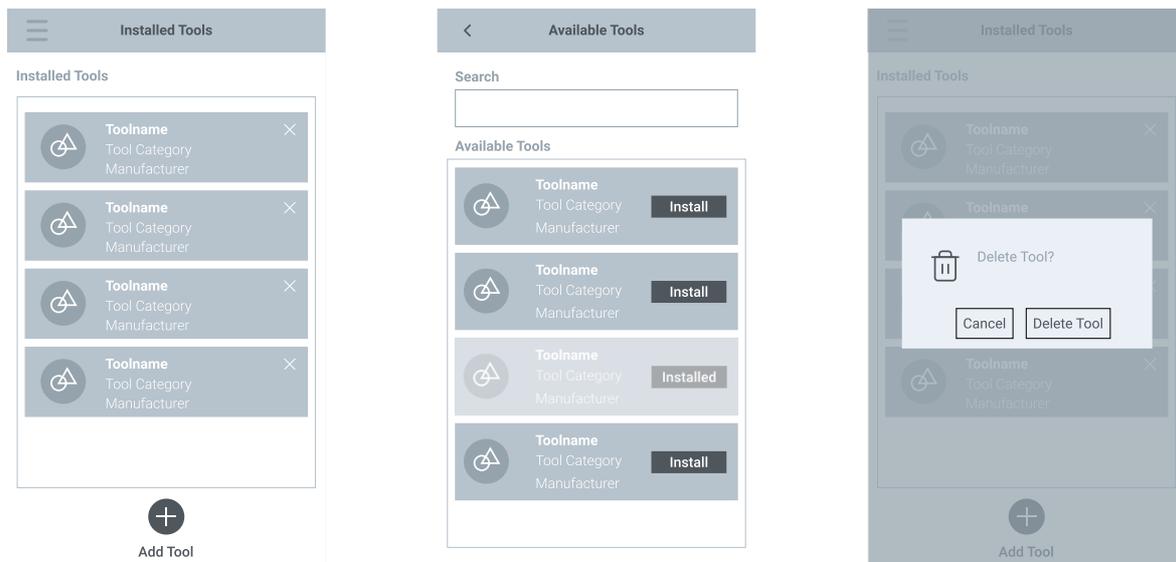


Figure 49: Wireframes for management of tools.

Table 12 specifies the user stories supported by the toolbox user interface and refining the requirements stated in D1.1:

Table 12: Requirements-refining user stories for management of tools.

Feature	Requirements-refining User Stories	Comments
T.F04.1 Cybersecurity Tool Installation	T.F04.R11: U29 The user can see the available tools. U30 The user can install a tool. T.F04.R12 is without user interaction. T.F04.R13: U31 The user can see the installed tools. U32 The user can uninstall and remove the tool.	Figure 49 mid. Figure 49 left. Figure 49 left. Figure 49 right.

Navigation

The toolbox offers a menu that allows the user to reach any of the views provided by the toolbox. The menu is indicated by a hamburger icon and allows to reach the view by a single click.

Figure 50 shows the toolbox wireframes.

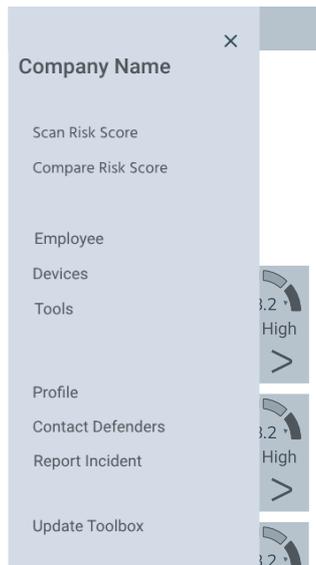


Figure 50: Wireframe for navigation.

Table 13 specifies the user stories supported by the toolbox user interface and refining the requirements stated in D1.1:

Table 13: Requirements-refining user stories for navigation.

Feature	Requirements-refining User Stories	Comments
T.QR06 Usability / Learnability	T.QR06.1: U33 The user can open the toolbox menu. U34 The user can reach a chosen view through the toolbox menu.	Figure 50 Figure 50

Toolbox Settings

The toolbox allows the human end-user to manage identities used for accessing data, specify context to which the toolbox adapts, and exert his/her rights defined in the GDPR. The settings affect both personal data and confidential data at the same time, thus implementing the quality feature T.QR08 Data Protection. The toolbox uses simple forms and intuitive controls to offer this functionality.

Figure 51 shows the toolbox wireframes.

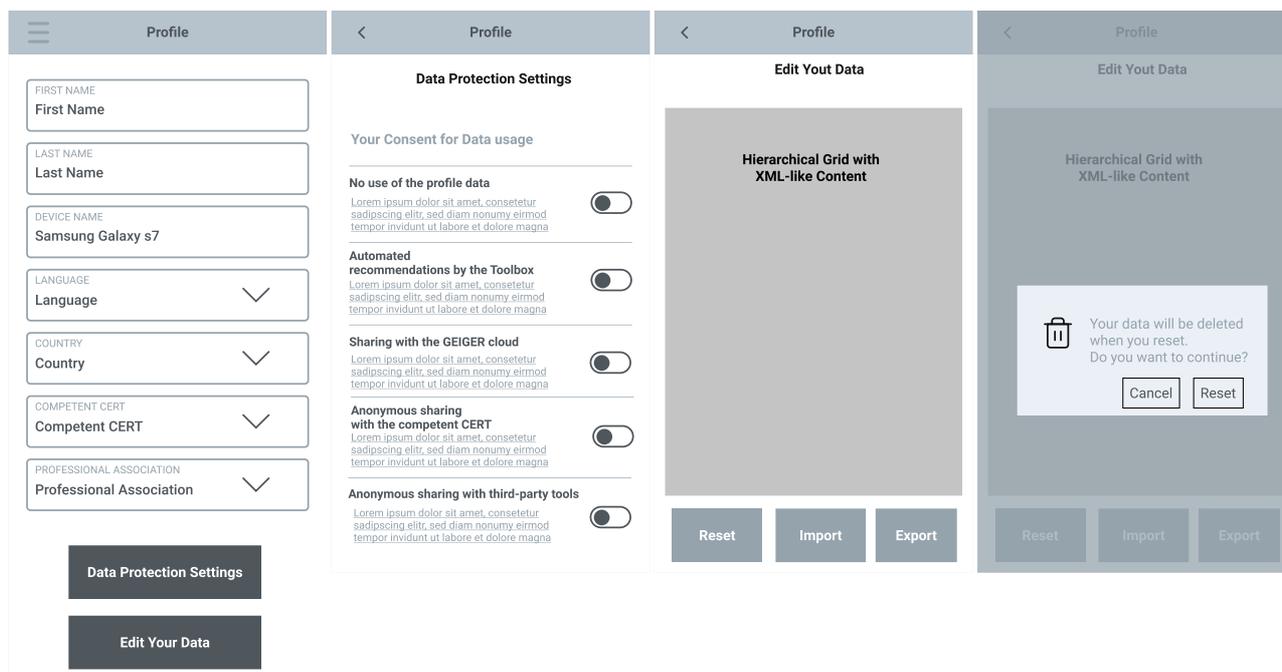


Figure 51: Wireframes for toolbox settings.

Table 14 specifies the user stories supported by the toolbox user interface and refining the Requirements stated in D1.1:

Table 14: Requirements-refining user stories for toolbox settings.

Feature	Requirements-refining User Stories	Comments
T.F06 Data Management	<p>T.F06.R01: U35 The user can export the data maintained by the toolbox in a human-friendly machine-readable format.</p> <p>T.F06.R02: U36 The user can edit the data maintained by the toolbox. U37 The user can reset the data maintained by the toolbox.</p> <p>T.F06.R03: U38 The user can import data that has been exported.</p>	<p>Figure 51-3.</p> <p>Figure 51-3. Figure 51-3 and 4.</p> <p>Figure 51-3.</p>
T.F06.1 Dynamic Consent	<p>T.F06.R11: U39 The user can prevent the use of his/her data (TLP:BLACK). U40 The user can agree to the automated processing of his/her data by the toolbox (TLP:RED). U41 The user can agree to the automated processing of his/her data by the GEIGER cloud (TLP:AMBER). U42 The user can agree to the automated processing of his/her pseudonymous data by the user’s chosen CERT (TLP:GREEN). U43 The user can agree to the automated processing of his/her pseudonymous data by installed tools (TLP:WHITE).</p> <p>T.F06.R12 is handled by the respective tool.</p>	<p>Figure 51-2.</p> <p>Figure 51-2.</p> <p>Figure 51-2.</p> <p>Figure 51-2.</p> <p>Figure 51-2</p>

Functionality Provided by Tools Integrated as Plugins

Not considered in the UI Wireframes of the toolbox are functionality provided by tools that are integrated as plug-ins into the GEIGER toolbox, including tools for questionnaires (T.F02.1), education reporting (T.F02.3), asset protection (T.F04), and incident reporting and resolution guidance (T.F05).

3.2.1.3 Operationalisation of Quality Requirements

For the personas described in Table 7, the user interface of the GEIGER Toolbox offers a user experience that satisfies the quality requirements stated in D1.1 as shown in Table 15.

Table 15: Implementation of quality requirements by the user interfaces of the GEIGER toolbox.

Quality Requirement	Operationalisation by the GEIGER toolbox user interface
<p>T.QR01 Suitability / Informative</p>	<p>T.QR01.1-a As its primary feature, the GEIGER Toolbox offers a personalised risk score. The score is aggregated of knowledge, devices, and any supervised employees the human end-user is responsible for. The score indicates the pragmatic criticality of undertaking protection action for the MSE, respectively the employee.</p> <p>T.QR01.1-b To allow diagnosis in terms of vulnerability identification within the MSE, the aggregated score is broken down into scores for individual threats and into scores for each device and any supervised employees.</p> <p>T.QR01.1-c To allow treatment in terms of risk reduction within the MSE, the score is justified with a threat-specific ranked list of protection recommendations for the end user and the device. The recommendations include mini questionnaires to declare the presence or absence of aspects critical for cybersecurity or data protection (such as a question “do you store private customer data on this device?”), configuration to be applied on the device or applications, sensor and protection tools to be installed, and learning objectives to be achieved. The presence or absence of the aspect, respectively of the protection is shown to the end-user.</p> <p>T.QR01.2-a The GEIGER Toolbox offers the option of seeing a short description of a recommendation.</p> <p>T.QR01.2-b The GEIGER Toolbox offers the option, thanks to the integrated tool CYSEC, for a structured walkthrough for understanding and implementing the protection. This option blends learning and acting for the human end-user.</p>
<p>T.QR06 Usability / Learnability</p>	<p>T.QR06.1-a The GEIGER Toolbox offers a simple form for the setting of pseudonyms for the user, device, and company</p> <p>T.QR06.1-b The GEIGER Toolbox offers a simple form for localisation settings, including the choice of language, competent CERT, and relevant professional association.</p> <p>T.QR06.1-c The GEIGER Toolbox offers a simple form to inspect data maintained in the toolbox instance, and to export, import, and reset that data.</p> <p>T.QR06.1-d The GEIGER Toolbox uses the widely established hamburger button icon for in-app navigation, thus building on visual cues recognised by smartphone users.</p> <p>T.QR06.2-a The GEIGER Toolbox adheres to the GEIGER style guide specified in the deliverable D5.1.</p>

<p>T.QR07 Different Languages</p>	<p>See T.QR06.1-b for settings.</p> <p>T.QR07.0-a The GEIGER Toolbox dynamically fetches texts and images from the back-end to be displayed on the user interface. The texts and images are those of the language chosen by the human end-user. If texts or images are not available in that language, the English variant is chosen.</p> <p>T.QR07.1-a The GEIGER Toolbox released at M12 supports English.</p> <p>T.QR07.2-a The GEIGER Toolbox released at M18 will support German. The translation will be performed by the Swiss use case.</p> <p>T.QR07.3-a The GEIGER Toolbox released at M18 will support Dutch. The translation will be performed by the Dutch use case.</p> <p>T.QR07.4-a The GEIGER Toolbox released as M18 will support Romanian. The translation will be performed by the Romanian use case.</p>
<p>T.QR08 Data Protection</p>	<p>T.QR08.1-a The GEIGER Toolbox offers a simple for privacy settings. The options include use of the data for automated processing in the toolbox, GEIGER Cloud, or third-party tools, and sharing of data with the chosen competent CERT. By default, the strictest and most limiting settings are applied, implying no use of data.</p> <p>T.QR08.1-b The GEIGER Toolbox offers the end user with the choice to relax privacy settings that are needed for a given function to be performed. The end-user may chose to inspect the concerned data, allow the one-time use of the data, or change the corresponding privacy setting.</p> <p>See T.QR06.1-c for transparency concerning the data stored in the GEIGER Toolbox.</p> <p>T.QR08.2-a T.QR08.1-a and T.QR08.1-b apply for secret data in the same way they apply for personal data.</p>

3.2.2 Toolbox Software Design

The GEIGER Toolbox is a three-layered architecture following the model-view-controller paradigm. The front-end interacts with the end-user through a native user interface that displays screens and data and forwards end-user commands to the interfaces provided by lower layers. The controller layer comprises the indicator, including calculating the indicator values and justification of the values with recommendations. A developer may extend the model layer with business logic that controls complex toolbox behaviour. The back-end offers core functionality and access to the local and distributed storage mechanisms. The back-end also provides communication with the GEIGER Cloud and with tools that are integrated as plugins.

Figure 52 gives an overview of the layers and components.

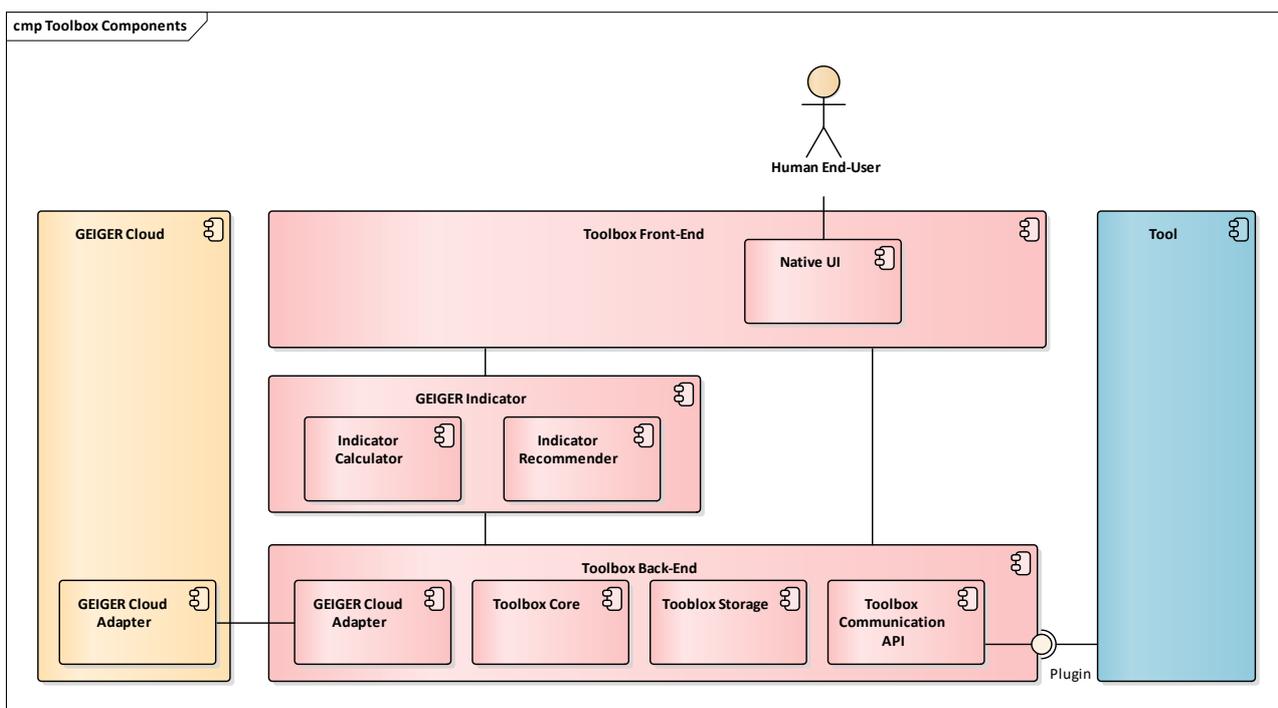


Figure 52: Toolbox architecture (middle, red) and external components (left, yellow: GEIGER Cloud; right, blue: tool integrated as a plugin).

The toolbox manages the MSE’s cybersecurity profile, offers cybersecurity risk assessment, and connects tools for sensing, protecting, and educating the MSE’s devices and employees and for detecting, handling, and reporting incidents. The architecture implements a cloud-edge approach to minimize cybersecurity and data protection risks: a custom-developed distributed secure database deployed on the devices of the MSE’s mobile cross-platform environment. Figure 53 illustrates.

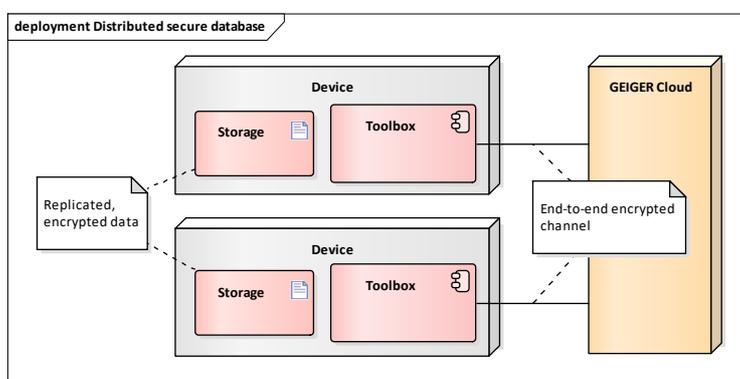


Figure 53: Distributed database consisting of local storages on the MSE’s devices and connected with a shared end-to-end encrypted channel.

Internal components and authorized external tools can read and write data without any code change in the toolbox backend, thus adding knowledge about the MSE and contributing to the GEIGER Indicator’s risk assessment of the MSE. The end-user may govern access to data by incorporating the traffic light protocol within the database. The traffic light reflects the consent obtained from the human end-user regarding sharing data with the GEIGER Cloud and Tools at the granularity of all data, respectively, for individual records.

To be useable in any of the EU Member States, the toolbox and database support internationalization.

3.2.2.1 Distributed Database Design

The GEIGER Toolbox requires generic storage to accommodate configuration, recommendations, and scoring data. This data needs to be protected against unauthorized sharing and must be kept locally or encrypted as end-to-end encrypted (E2EE) data blobs to guarantee data privacy.

The use of a strict set of predefined tables as in SQL tables was therefore not suitable. Instead, we decided to go for a tree-like structured database related to LDAP or Windows registry databases. Each node acts as a key/value store. We enriched this model with some node-specific attributes (ordinals) to accommodate required information, such as the level of sharing expressed in standardized Traffic Light Protocol¹⁴).

Figure 54 shows the hierarchical structure of the data.

```

:-----Node1
| ---Node2
| | ---Node21
| | | ---Node211
| | | ---Node212
| ---Node3

```

Figure 54: Hierarchical structure of the data.

Each node of the hierarchical structure consists of:

Ordinals (fixed set of attributes)

- Name.
- Owner (typically the plugin last writing to the node).
- Visibility (according to TLP).
- Key/value store
 - Key (string with a limited character set up to 1024 bytes).
 - Value (string; internationalized).
 - Type (optional; as a search filter).
 - Description (string; internationalized).

Common types of data are identities for devices and users, configuration settings, sensor values contributing to the MSE profile, indicator values and recommendations.

The toolbox replicates or shares data as follows. The end-user receives the opportunity to decide about the TLP levels as part of the toolbox settings and, for sensitive data, individually for each attempt of data transfer.

- TLP:BLACK – “secret”: data kept within the toolbox and not replicated or shared. The TLP:BLACK label will be applied to the replication keys and information marked as strictly confidential by the end-user.
- TLP:RED – “personal”: data replicated within the SME with end-to-end encryption but not shared.
- TLP:AMBER – “limited”, TLP:GREEN – “community”, TLP:WHITE – “unlimited”: data shared without encryption with the GEIGER Cloud. Here, the replication engine offered by the GEIGER Cloud and accessed through the GEIGER Cloud Adapter receives the user’s consent that the data sharing is OK.

The toolbox regulates the sharing of data with tools integrated as plug-ins into the toolbox as follows:

- TLP:WHITE – “unlimited”: data shared with tools.

¹⁴ Traffic Light Protocol (TLP): <https://www.first.org/tlp/>

- TLP:GREEN – “community”, TLP:AMBER – “limited”: data depends on constraints imposed on the tool, including that the tool can run standalone on the end-user’s device and does not need access to the tool vendor’s back-end cloud.

The toolbox user interface may define further rules that constrain the human end-user’s access to data. For example, to achieve simplicity and avoid unauthorized access to data within the MSE, the business logic may constrain access to data of the end-user’s devices only and scores of paired employees. Thus, potentially malicious employees will not see data that does not belong to them or has not been authorized to be seen.

3.2.2.2 Plugin Mechanism

The toolbox supports the addition and removal of tools as plug-ins at runtime and without knowledge of their inner workings and commonly used device types by the MSEs. A plugin may be a separate app to install on the end-user’s mobile platform.

A plugin can interoperate through the toolbox communication API that allows exchanging data with the toolbox. For the data exchange between the toolbox and the plugin, the following steps are required:

1. The plugin first registers itself to the toolbox. When doing so, the plugin declares whether it is sharing any information with third-party systems outside the GEIGER ecosystem or not.
2. If the user consents, the plugin gains the following capabilities through this API:
 - a. Share and request data with the Toolbox core and other plugins.
 - b. Pass the control of the visual stack (screen) to the core.
 - c. Receive the control of the visual stack from the core (e.g., when the user presses a recommendation or config item).
 - d. Add, remove, and disable menu entries within the core that would trigger actions within the plugin.
 - e. React to the pressing of the “Scan” button in the core.
 - f. Receive events to react to changed objects within the database.
 - g. Deregister the plugin.

Plugins sharing information outside the GEIGER ecosystem can access only “own” data, TLP:GREEN data, and TLP:WHITE data. This restriction contributes to the privacy and confidentiality of the stored data.

All plugin traffic is cryptographically secured to identify unauthorized access by a plugin. The securing is done in an accountless manner and only requires a plugin to request access and the toolbox to consent to the access. The exchange of cryptographic tokens is based on the Diffie-Hellmann Key Exchange and transparent to the plugin.

3.2.2.3 Scoring and Knowledge Overview

To accommodate the scoring of an unknown set of sensor values, UU developed a general scoring system. As scores are calculated based on the MSE profile, sensor values, and threats, the generic format for the data repository shown in Figure 54 is used to accommodate these values as node objects within the toolbox storage.

Event-based processing is used to keep the scores fresh. The storage allows a process, e.g., the GEIGER score calculation engine, to register for any new or updated values. Such an event is raised soon as a plugin updates the respective values. Upon such an event, the scoring engine may process the data. The TLP:BLACK label guarantees that a value is always calculated locally and not replicated. This control enables avoidance of race conditions due to replication conflicts resolved and delays in replication. When updated values are replicated on a different device, the score maintained on that device gets updated immediately. This approach ensures that the score is consistent on all devices of the same user.

The GEIGER score calculation engine performs the scoring based on knowledge about threats and recommendations stored locally in the toolbox under the “:Global” node. When installing the toolbox core, the toolbox retrieves threat and recommendation data from the GEIGER Cloud. The local storage of these data allows the scores to be calculated standalone without Internet access. Thus, the scoring and recommendations will work even if the user does not consent to use the GEIGER Cloud or share any data.

3.2.2.4 Cross-Platform Environment

According to the use case MSE requirements, support is needed for mobile devices with Android or iOS and PC-based devices with Windows, macOS, or Linux (decreasing order of priority). While on Windows or Linux, plugin mechanisms are common, Android and iOS prohibit applications to load code internally.

The GEIGER Toolbox uses TotalCross¹⁵ for supporting the mobile and pc-based platforms in use by the use case SMEs. TotalCross has been chosen due to the breadth of its offered platform support, low performance footprint, and minimum requirements for the end-user devices.

For interoperability with 3rd party tools integrated as plugins in the cross-platform environment, the toolbox supports a custom cross-platform-capable serialization mechanism. Still, two variants of the local API are required to work under TotalCross. One variant is offered for tools built with TotalCross and one for other tools. The transport mechanism for messaging and data exchange varies from platform to platform slightly depending on the framework used.

Differing is also the mechanism to ‘wake up’ a plugin on the platform. A plugin may update its state to ‘running’ or ‘not running’ anytime within the toolbox, and the platform tries to wake up ‘non-running’ plugins before contacting them. This sensing of whether a plugin is running allows fast response to user interactions. If a plugin does not update its state within a sub-second delay, the toolbox may conclude that the plugin is not yet running.

3.2.2.5 Internalization and Localization

To be used in an EU member state, each plugin and the core require the capability to express operations in the end-user’s local language. Such a requirement sounds as simple as replacing strings within an app but gets complicated when considering grammatical mechanisms as plural forms of phrases, different forms of numbers (e.g., decimal separators or separators for thousands), representations of dates, and similar phenomena. Internally, the GEIGER toolbox uses the standard internationalization and localization framework gettext¹⁶ to accommodate this problem.

For interoperating with plugins unknown to the toolbox, the toolbox labels values with internationalization information. The database allows storing any value in an internationalized form by indicating a language like “de” for German or a locale like “de_CH” for German in Switzerland. Any application may provide a preferred language or locale when reading data. Depending on the given value, the toolbox may present either the requested locale, a general translation, or a phrase in a default language (English). While this approach does not solve the issue of translating all contents into the EU member states’ languages, it allows to offer internationalization across all plugins and provides a controlled fallback in the case of missing translations.

¹⁵ <https://totalcross.com/>

¹⁶ <https://www.gnu.org/software/gettext/>

3.3 GEIGER Cloud

3.3.1 GEIGER Cloud Overview

The GEIGER Cloud is a key element in the GEIGER project. As a cybersecurity platform, GEIGER is designed to provide cybersecurity awareness for the end-user at any time. GEIGER Cloud provides support in terms of information and infrastructure available everywhere.

The Cloud offers an online infrastructure to support the GEIGER platform and data exchange.

3.3.2 GEIGER Cloud Function and Objectives

GEIGER Cloud is designed to provide several essential functions for the platform, including the following ones:

- Give support to the user on a 24x7 basis.
- Coordinate amongst internal Cloud components and the rest of the GEIGER environment (Toolbox, CERTs, external apps...).
- Host the online database storage as a point where information is gathered and shared amongst the components of the GEIGER platform.
- Deliver updated information when requested by other components (such as the GEIGER Indicator).
- Support GEIGER critical functions such as risk level calculation.
- Serve as a point of contact with the end-user by means of the Web Client.
- Provide endpoints of communication for external entities such as the CERTs and the GEIGER External tools.
- Ease the connection of coupling of new tools, that is, lessen updates in the GEIGER platform.

3.3.3 GEIGER Cloud Design and Architecture

Figure 55 shows a diagram of the complete infrastructure of the GEIGER Cloud:

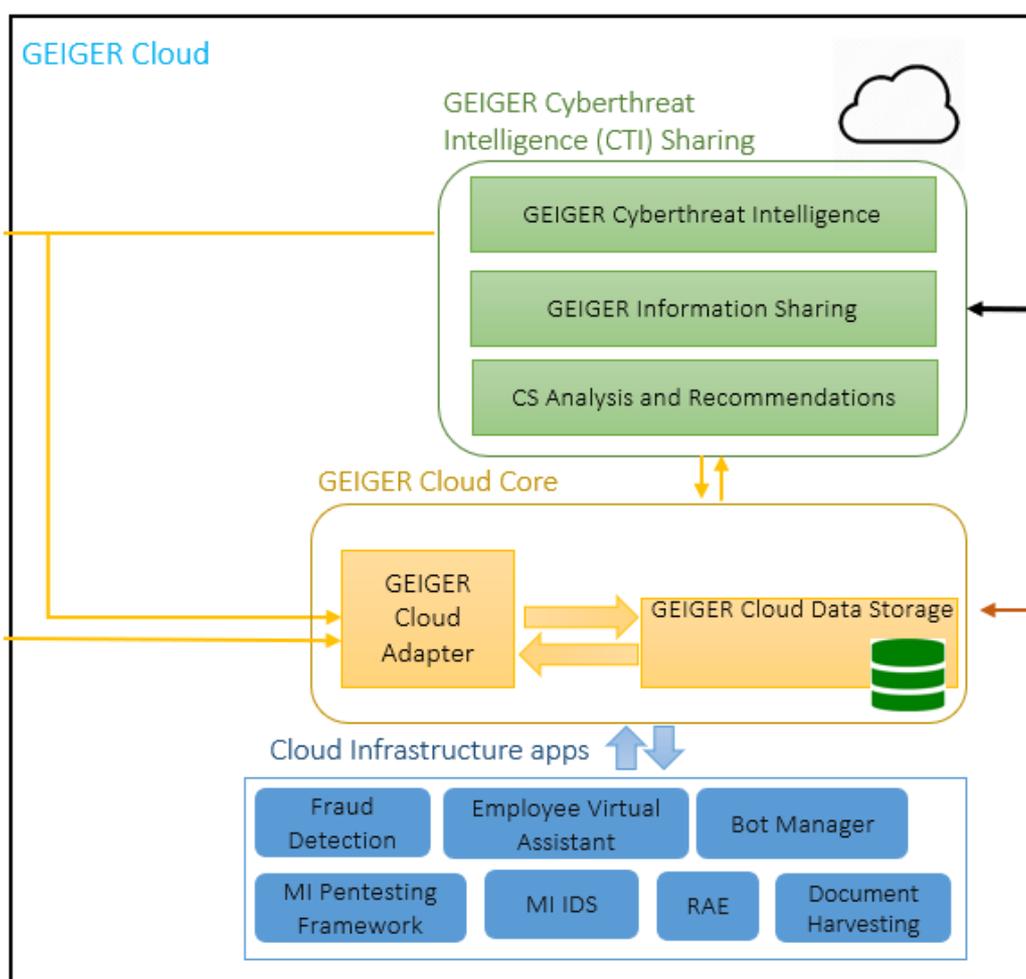


Figure 55 - Diagram of the GEIGER Cloud

GEIGER Cloud is made up of three important sub-components: the GEIGER Cloud Core, the Cyber Threat Intelligence (CTI) Sharing and the GEIGER infrastructure applications.

3.3.3.1 GEIGER Cloud Core

The Cloud Core is the central node of the GEIGER Cloud infrastructure. This is the place where Cloud Data Storage is placed. Therefore, the Core harbours the online database. Based on that condition, the Cloud Core plays a central role about collecting and exchanging information with the rest of the components, not only in the Cloud side but also with the Toolbox and external entities. All the online data flows are expected to pass through the GEIGER Cloud Core.

In addition, the online database storage must rely on an API for the data exchange. This API is the GEIGER Cloud Adapter and provides several methods for both sending and retrieving data from the GEIGER online storage.

3.3.3.2 GEIGER Cyber Threat Intelligence Sharing

The Cyber Threat Intelligence Sharing or simply CTI Sharing component performs the following tasks:

- Management of communication with external entities, that is, with the CERTs:
 - Gather updated cyber security information.
 - Push data to the CERTs to be analysed.
- Deliver security analysis (in relation to the information provided to CERTs) and broadcast security recommendations as required.

- Host the GEIGER Information Sharing component. This subcomponent is required to pass information to the CERTs according to the MISP platform.

3.3.3.3 GEIGER Infrastructure Applications

The GEIGER Infrastructure applications could be described as the internal applications of GEIGER. This refers to a group of valuable cybersecurity tools provided by the partners of the GEIGER project. The hallmark of these tools refers to their distribution within the platform: they are deployed in the GEIGER server, where:

- They provide various cybersecurity functionalities which enhance GEIGER capabilities.
- They deliver information to the GEIGER Cloud online database storage.
- They keep a fluent communication and data exchange with the GEIGER Cloud by means of the GEIGER Cloud Adapter.

This set of applications include various capabilities such as:

- Performing Risk assessment as the RAE tool performs.
- Aiding deal with fraud (Fraud Detection tool).
- Assisting the user with the help of AI techniques (Bot Manager).
- Gathering information from documents (Document harvesting).
- Perform network monitoring and analyse network traffic to detect attacks and anomalies (MI IDS).

3.3.4 GEIGER Cloud Data Exchange

Data exchange is a vital part for the GEIGER Cloud. Information is shared, collected, sent and received amongst the different components and third parties in the GEIGER platform. Relevant data to be exchanged include:

- Updates on cybersecurity information provided by CERTs.
- Information that the MSE grants to be stored online.
- Data for the risk score calculation.
- Information provided by both GEIGER Infrastructure and External apps.
- Recommendations generated by the platform.

All stored information is available upon request. The API provides a point of access to retrieve information easily. It is also important to notice that GEIGER Cloud storage keeps data needed for the risk score calculations, so it is necessary to keep information updated as much as possible to take advantage of it.

Although GEIGER Cloud is expected to store information of the MSE, this decision befalls only on the MSE owner. Therefore, just approved data are to be stored online according to TLP code described.

Interaction with tools feeds GEIGER platform. Regardless whether they are or not integrated in the GEIGER servers, these applications can provide valuable data to be considered in various scenarios, mainly focused on the risk score calculation.

3.3.5 KPMG GEIGER Conversation module

As part of the GEIGER bot's information set, it will store and fetch its data from and to the GEIGER Cloud Data Storage and locally through the GEIGER Toolbox Core.

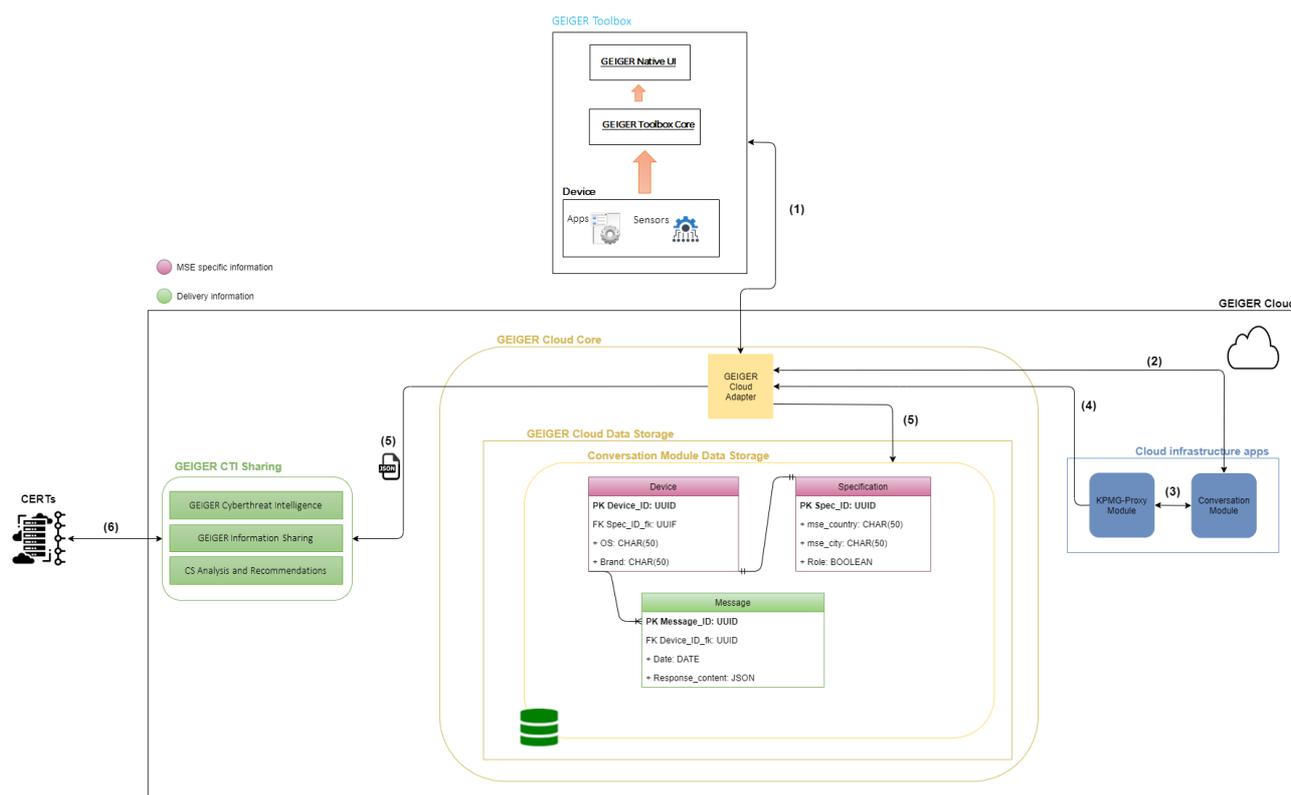


Figure 56 - MSE relations and data information

Triggers that invoke the chat bot might be caused by the client (MSE's end user), or by the KMS-SDK threats updates.

In a reference to the diagram above- Figure 56, the flow represented by numbers, demonstrate the process that the information goes through when the user begins a session with the chat bot.

1. The GEIGER Toolbox is connected to the GEIGER Cloud Adapter where it streams, delivers, and receives the relevant data and information that is necessary to form and build the conversation module.
2. As long the end-user communicates with the Conversation Module chatbot, on each turn of the conversation the relevant dataset is transmitted back and forth between the client and the Conversation Module that sits on the cloud.
3. When a session between the end user and the Conversation Module chatbot ends, the conversation content together with other relevant information is delivered to the KPMG Proxy Module.
 - a. The KPMG Proxy Module manipulate, extract, and process the data received by the Conversation Module.
4. The dataset received by the Proxy Module is delivered to:
 - a. To the Conversation Module Data Storage for further analytics and analysis.
 - b. To the GEIGER Cloud, from where the data will be used by the GEIGER CTI.

3.3.5.1 Local storage under the GEIGER Toolbox Core

Any personal information that is subject to privacy constraints and GDPR policy will be stored locally under the GEIGER Toolbox Core. The stored data at this phase is the relevant private information about the MSEs, their employees and their devices information. In addition, under the GEIGER Toolbox Core, the content of the conversation made by the Bot Manager in front of the client will be analysed and manipulated. This

manipulated data contains the relevant information that we allow to store in the cloud, and it contains the information processed by the GEIGER Logic and Analyses mechanics. The values regarding the MSEs, the clients and their devices will be saved as UUID without exposing their private information.

3.3.5.2 Conversation module GEIGER Cloud data storage

After receiving the processed information from the GEIGER Logic and Analyses mechanics, the data packet is sent to the GEIGER Cloud Adapter inside the Geiger Cloud Core and through it the packet is transferred to the Conversation Module Data Storage. The threat score is kept under the same storage.

3.3.6 KPMG ISP with Information Sharing

One of the requirements under this module is to transfer a JSON file containing the different types of threats from the relevant parties to the Information Sharing module. Accordingly, the relevant data is passed through the KPMG-Proxy Module to this same component for generating a MISP file that can be used for transferring the information to CERTs.

3.4 GEIGER Cyber Threat Intelligence (CTI) Sharing

3.4.1 GEIGER Information Sharing

The GEIGER Information Sharing component, also known as the Information Sharing and Analysis Center (ISAC) is a framework developed by ATOS with the purpose of increasing security on the intelligence data exchanged amongst stakeholders. ISAC can be described as a **secure communication API** for data exchange between the GEIGER Cloud components and external entities. While MISP standard helps obtaining cyber threat information, it does not allow for a proper management on how information is shared and what the addressees are. The ISAC covers this issue by means of:

- The use of symmetric **encryption** for all communications: any information is encrypted and subsequently included as an attribute inward the MISP data format.
- The **authentication of users** with a central third-party authentication. This includes the management of user roles, including different categories (Administrator, Publisher, Provider) and the need to be cleared to consume and publish events in the MISP instance connected to the ISAC.
- The establishment of **secure and fast communications** based on both encryption and the MISP Python API (or PyMISP).
- The possibility of achieving a **high level of configuration** regarding information shared: each user has his own profile so both accesses of users and roles can adapt to necessities of data exchange.

In addition, ISAC is easily scalable, since it is an open structure, which is recognised by several cybersecurity platforms and systems.

3.4.2 KPMG Conversation Module

Data is transmitted in a two-way manner, from and to the following modules

- Information Sharing: After the user completes a session in front of the chat bot, the processed information will be transferred to the GEIGER Cloud through the KPMG-Proxy Module for generation of MISP files that can be shared with CERTs.
- KMS-SDK sharing component: That component needs to update the Conversation Module with the threats relevant to it only. When the Conversation Module receives an update, the proactive session starts in front of the MSE's end user

3.4.3 KPMG Proxy Module

This module transfers data from a source party to its destination and serves as a central proxy channel through which the data passes between those channels. Through that proxy pipeline, the data is manipulated, extracted, and processed to ensure format consistency and efficiency improvement.

4. Conclusions and Future Work

GEIGER stands as an innovative platform to enhance resilience for the MSEs. The platform is a tool to increase cybersecurity awareness on a market segment where traditionally the companies not usually invest on security or, in the best scenario, consider it as something dispensable. The aim of GEIGER is, therefore, to help MSEs to achieve a higher cyber security level but trying to be as less invasive as possible for them. By means of details such as the risk level indicator or the cyber security training, end-users are more conscious of security issues and more compromised to take security of the business seriously.

GEIGER tries to enhance what is usually called the weakest link in the security chain, that is, the user. Through education, end-users will be increasing their knowledge, confidence and expertise on cybersecurity and that helps creating a more secure environment for the business.

One of the most important innovations of GEIGER is about the combination of several tools to provide more functionalities. Both external and internal tools help covering all possible needs from a security perspective. Besides, it is also possible to add, upon need, some other tools in the future. That means GEIGER could be updated with few efforts. In addition, offering the end-user an updated estimation of the risk level “at a glance” allows for a better, quicker and more effective management of security issues.

In addition, GEIGER innovates by means of providing cybersecurity training to the end-user. This functionality may go unnoticed but helps achieving, from a different perspective, the overall purpose of the platform: raising cybersecurity awareness. It is obvious how relevant is nowadays for the end-user to achieve a minimum cybersecurity knowledge and understanding, and GEIGER tries to assist in that issue.

The GEIGER architecture described in this document has been designed with the aim of being flexible, adaptable and, most important, to meet the objectives of the platform. The GEIGER Toolbox and the Cloud storage combine provide both online and offline functionality. Besides that, the support of applications and CERTs is vital in terms of data contribution.

The deliverable D1.2 is the M12 baseline of the work performed in the tasks T1.2 and T1.3 of WP1 “Requirements, Architecture, and Methodology”. It represents the developers’ view of the architecture building upon the requirements specified in D1.1 and taking into account relevant aspects of the training plan D3.1. The deliverable D1.2 also interacts with the work performed in T5.2 “Standardisation and Policy” as it builds upon the standards and external interfaces identified and documented in the impact plan D5.1. The deliverable D1.2 is building the base for to the deliverable D2.1, which was defined in parallel and D1.2 is the foundation for the entire development. The deliverable D1.2 moreover interacts with the approach defined for the protection of personal data (POPD) and reported in deliverable D7.2.

Finally, in terms of cybersecurity any solution must not be considered perfect. That is explained because the field encompass several challenges, issues and continuous change of the conditions. We plan to continue working in enhancing and refining the architecture especially on two aspects:

- How to integrate new tools, something that must be addressed considering user should not be affected.
- How effective GEIGER could be on a local-only mode, where the absence of updated information penalizes the experience of the end-user.

In any case, GEIGER brings a higher level of protection to MSEs and represents a huge step on the direction of making a more secure experience for these companies.

5. Annexes

Appendix A: Definitions

Table 16: Definitions employed in the GEIGER indicator development.

Term	Definition	Source
Security	The perceived or actual ability to prepare for, adapt to, withstand, and recover from dangers and crises caused by intentional or unintentional acts.	Adapted from Jore (2019)
Cyberspace	A collection of interconnected computerised networks, including services, computer systems, embedded processors, and controllers, as well as information in storage or transit.	Refsdal et al. (2015)
Cyber-system	A system that makes use of a cyberspace.	Refsdal et al. (2015)
Cybersecurity	The organisation and collection of resources, processes, and structures used to protect cyberspace and cyberspace-enabled systems from occurrences that misalign perceived from actual property rights.	Adapted from Craigen et al. (2014)
Cybersecurity metric	Any value resulting from the measurement of security-related properties of a cyber-system.	Borrowing from Böhme and Freiling (2008), Refsdal (2015), and Pendleton et al. (2016)
Threat	Any circumstance or event with the potential to adversely impact an asset through unauthorised access, destruction, disclosure, modification of data, and/or denial of service.	ENISA (2016a)
Countermeasure	An action, device, procedure, or technique that meets or opposes (i.e., counters) a threat, a vulnerability, or an attack by eliminating or preventing it, by minimizing the harm it can cause, or by discovering and reporting it so that corrective action can be taken.	Shirey (2007)
Risk	The potential that a given threat will exploit vulnerabilities of an asset or group of assets and thereby cause harm.	ENISA (2016a)
Vulnerability	The existence of a weakness, design, or implementation error that can lead to an unexpected, undesirable event compromising the security of the computer system, network, application, or protocol involved.	ENISA (2016a)

Appendix B: Threat mapping

In this section, we outline the guiding principles used to map the ENISA threat taxonomy (ENISA, 2016b) to the list of GEIGER threats presented in Table 3. Table 17 provides the detailed mapping, with an indication of the guiding principles.

Table 17: The guiding principles for mapping threats from the ENISA threat taxonomy (ENISA, 2016b) to the GEIGER setting.

Principle Label	Principle description
P1	All threats should be relevant to MSEs.

P2	Threat naming should be consistent and should capture the full scope of a type of threat.
P3	The overall threat list should stick as close as possible to an existing threat taxonomy standard.
P4	Threat concepts should be independent as much as possible.
P5	Countermeasures should apply to as few threats as possible.
P6	The threat list should facilitate straightforward mapping of MISP incident data as much as possible.
P7	The threat list should facilitate straightforward mapping of tool metrics to threats as much as possible.
P8	The threat list should facilitate straightforward mapping of metrics on knowledge gained in the education framework as much as possible.
P9	The threat list should be resistant to changes in the threat landscape.
P10	The threat list should contain as few threats as possible.
P11	The threat list should cover the complete set of available countermeasures for MSEs as much as possible.

Table 18: Mapping of threats from the ENISA Top 15 threats (ENISA, 2020) and the detailed ENISA threat taxonomy (ENISA, 2016b) to the GEIGER indicator threat concepts.

ENISA Threat	GEIGER Threat	Motivation	Guiding principles
1. Malware	Malware	-	P3
2. Web-based attacks	Web-based threats	Renamed to be consistent with the GEIGER focus on threats rather than attacks.	P2, P3
3. Phishing	Phishing	-	P3
4. Web application attacks	Web application threats	Renamed to be consistent with the GEIGER focus on threats rather than attacks.	P2, P3
5. Spam	Spam	-	P3
6. DDoS	Denial of service	Removing the abstraction of an abbreviation to promote MSE understanding. 'Distributed' DoS is still deemed a part of this threat, as in ENISA (2018).	P2, P3
7. Identity theft	Data breach	ENISA considers identity theft to be a type of data breach (ENISA, 2019). Countermeasures for both threats are practically identical.	P4, P5, P6, P7, P8, P10
8. Data breach	Data breach	-	P3
9. Insider threat	Insider threats	We choose to use the plural 'threats' to indicate to the user that this involves multiple threats. This category can originate from both intentional and unintentional actions.	P3
10. Botnets	Botnets	Botnets encompass more than just a relation to malware, but they are generally classed as malware, since they control a set of malware-infected devices. Nevertheless, we keep Botnets as a separate category, to better map to existing incident taxonomies and the education framework of GEIGER.	P3, P4, P6, P7, P8
11. Physical	Physical threats	-	P3
12. Information leakage	Insider threats	The countermeasures related to information leakage are essentially all applicable to the 'Insider threats' category.	P4, P5, P6, P7, P8, P10

13. Ransomware	Ransomware	Although ransomware is a type of Malware, we choose to not include it in the Malware category, but rather leave it as a separate category. This is mainly due to the high prevalence and relevance of ransomware in this time, which is expected to remain in coming years. Ransomware was classed in the top 2 threats for all MSE categories in a recent questionnaire.	P3, P9
14. Cyber espionage	-	Cyber espionage is generally considered to be more of a motive than a threat (ENISA, 2019). Additionally, it was found through surveying of experts to not be particularly relevant to MSEs.	P1, P8, P10
15. Cryptojacking	Malware	Threat resulting from the use of cryptominers, which are contained in the malware category (ENISA, 2020). Additionally, not nearly as relevant to MSEs as Ransomware, which motivates not leaving this as a separate category.	P1, P4, P5, P6, P7, P8, P10
[Extra] Erroneous use	Insider threats	Additional ENISA threat specified in the ENISA Threat Taxonomy (ENISA, 2016b). The main reason to include this threat in the mapping is to make clear that error and misuse form a part of the 'Insider threats' topic.	P7, P8, P9
[Extra] Third party	External environment threats	'Third party' is an additional ENISA threat specified in the ENISA Threat Taxonomy (ENISA, 2016b). Although this threat is not in the top 15 (ENISA, 2020), MSEs are especially dependent on third parties. Additionally, countermeasures related to this threat will not always be present in other threat categories. We choose to label this category 'External environment threats' as this encompasses the included elements and aligns with standard terminology used in the study of sociotechnical systems (Davis et al., 2014), which is what MSEs are.	P8, P9, P11
[Extra] Supply chain	External environment threats	'Supply chain' is an additional ENISA threat specified in the ENISA Threat Taxonomy (ENISA, 2016b). Similar arguments apply here as with the 'Third party' mapping.	P8, P9, P11
[Extra] Legal	External environment threats	'Legal' is an additional ENISA threat specified in the ENISA Threat Taxonomy (ENISA, 2016b). The legal and compliance sides of the threat landscape are important to MSEs. We saw this in the user requirement UR5 of Table 1. This is a threat that is certainly felt to be important by MSEs, so even though it is not generally found in the ENISA Top 15 (ENISA, 2020), we include it. This category also helps to align with the CERT-XLM category "conformity." Additionally, with data protection and cybersecurity getting increasing attention, it can be expected that this category will remain important for MSEs for years to come. Lastly, it is included in the education framework, so inclusion also helps in alignment.	P8, P9, P11, P12

Appendix C: Detailed recommendation information

Tables 19, 20, and 21 outline the procedure we followed to arrive at the global set of recommendations for the GEIGER indicator solution. From a broad set of cybersecurity sources (Table 20), we collected a set of recommendations which were subsequently merged into the global recommendation list (Table 21). We indicate whether these general recommendations can most aptly be applied to user or device scores. We demonstrate the completeness of this list by mapping the recommendations to the security control categories presented in Yigit Ozkan and Spruit (2021).

Table 19: Sources used to construct our global recommendation list.

Source Label	Source Reference	Source Description
S1	ENISA (2020)	ENISA reviews the cybersecurity threat landscape. Each threat has its own report, where proposed actions are given to counter the threats.
S2	Swiss NCSC (2021)	Directory of common cyber threats encountered by the Swiss NCSC, accompanied by measures to counter these threats.
S3	NCSC UK (2021)	Cybersecurity certification programme backed by the UK government. It aims to provide organisations simple steps to improve cybersecurity, with a focus on five basic security controls.
S4	NCSC UK (2018)	NCSC UK developed a series of infographics over the years. The infographics include basic cybersecurity tips and guides for both individuals and organisations.
S5	FTC (2018a)	The US Federal Trade Commission (FTC) covers various cybersecurity topics on their page aimed at small businesses. The FTC provides guidelines on how to protect your business for specific cybersecurity threats, such as ransomware and phishing.
S6	ACSC (2019)	The Small Business Cyber Security Guide of the ACSC covers cybersecurity threats and other security topics. It provides tips on how to prevent and recover from cybersecurity threats.

Table 20: The global recommendation list employed in the GEIGER indicator solution.

Recommendation Label	Recommendation	Sources	User/Device
R1	Employ mail filtering (sometimes referred to as spam filtering) and blocking to prevent malicious e-mails from reaching you.	S1, S2, S4	User
R2	Make use of security logging systems to be able to detect anomalies and incidents.	S1, S2, S4	Device
R3	Only provide privileged access to people who need it for their roles. Regularly review these and revoke privileges if no longer needed.	S1, S4	User
R4	Formulate an update policy and regularly review it to ensure it satisfies your needs.	S1, S2, S4, S5, S6	User
R5	Wherever they are available, enable default security measures on the devices within your organisation.	S3	Device
R6	Formulate an authentication (including passwords) policy and regularly review it to ensure it satisfies your needs.	S1, S3, S4, S5	User

R7	Formulate security practices for employees to follow when working remotely from home or on business travel.	S1, S3, S5	User
R8	Use whitelisting to prevent unknown applications and executables from being executed.	S1, S2, S3	Device
R9	Configure and enable firewalls within your enterprise.	S1, S2, S4	Device
R10	Prevent employees from downloading third party apps.	S4	Device
R11	When products reach the end of their supported life, they should be replaced.	S4	User
R12	Formulate guidelines for employees on how to use e-mail safely, and regularly review these guidelines.	S1, S2, S5	User
R13	Implement one of the standards for e-mail authentication.	S1, S4, S5	User
R14	Formulate a policy regarding information disclosure, both for individual employees and the organisation.	S1, S2, S4, S5	User
R15	Formulate a clear and strict policy regarding money transfers.	S1, S2, S4	User
R16	Website forms should be secured to prevent abuse by malicious actors.	S1	User
R17	Formulate, implement, and regularly review a backup policy.	S1, S5	User
R18	Track phishing practices and trends and update employees on practices and trends.	S4, S5	User
R19	When credentials are stolen or leaked, make sure to change them.	S4	User
R20	Ensure all important e-mail requests are verified using a second type of communication.	S4	User
R21	Formulate a communication plan and share this with your business partners and customers.	S4	User
R22	Have an effective security incident reporting process in place.	S4	User
R23	Use content-control software to prevent employees from accessing malicious websites.	S4	Device
R24	Formulate a security incident response plan, implement it, and ensure through practise that it works.	S1, S2, S4, S5	User
R25	Formulate and enact a security incident business continuity plan.	S1, S5	User
R26	Isolate and sandbox applications and systems that are vulnerable to attack.	S1, S2	Device
R27	Formulate and implement server and service hardening policies.	S1, S2	Device
R28	Limit the permissions of your browser, such as what it can execute.	S2	Device
R29	Use input validation and isolation techniques for injection type attacks.	S1	User
R30	Maintain an inventory of web application APIs and implement measures to ensure they are secure.	S1	User
R31	Deploy traffic and bandwidth management capabilities.	S1	Device

R32	Develop standard operating procedures and policies for handling (sensitive) data.	S1, S5	User
R33	Implement and regularly review IP filtering and blocking policies.	S1, S2	Device
R34	Publish services through content delivery networks to absorb volumetric attempts.	S1	User
R35	Have a clear communication plan and channel with your Internet Service- and Cloud Providers.	S1	User
R36	Gain understanding regarding your asset inventory and the actions that can be performed on these assets, together constituting your attack surface.	S1, S2	User
R37	Monitor the availability of your (customer) applications, including from your customers' viewpoint.	S2	User
R38	Use threat hunting and penetration testing techniques to strengthen your organisations' defences.	S1	User
R39	Install end-point protection on all devices and keep the protection updated.	S1, S2	Device
R40	Enforce the use of data loss prevention (DLP) solutions.	S1	Device
R41	Deploy border gateway protocol (BGP) feeds.	S1	Device
R42	Restrict cryptocurrency mining pools, protocols, and executables.	S1	Device
R43	Deploy challenge-based capabilities for website(s).	S1	User
R44	Formulate, implement, and regularly review a physical security policy at your organisation.	S1, S5	User
R45	Invest in a cybersecurity insurance policy that covers the damages caused by attacks.	S1	User
R46	Anonymise, pseudonymise, minimise and cipher data in accordance with the provisions of the EU GDPR.	S1	User
R47	Create a security operation centre (SOC) staffed by skilled security personnel within your organisation.	S1	User
R48	Conduct regular security audits to detect any security-related abnormalities within your organisation.	S1	User

Table 21: The global GEIGER Indicator recommendations are mapped to security control categories, where each recommendation maps to at least one security control category.

Security Control Category Label	Security control category	GEIGER Indicator Global Recommendations
SC1	Management Commitment and Policies	R4, R6, R7, R10, R12, R14, R15, R17, R18, R21, R22, R24, R25, R27, R32, R35, R44, R47, R48
SC2	Asset Management	R11, R36, R44
SC3	Patch Management	R4, R11
SC4	Access Control	R3, R6, R19, R44
SC5	Secure Computers, Servers, and Network Configuration	R5, R8, R10, R26, R27, R38, R39, R42
SC6	Log Management	R2, R37
SC7	E-mail and Web Security	R1, R12, R13, R20, R23, R28
SC8	Malware Protection	R5, R39

SC9	Network and Communications Security	R9, R31, R33, R38, R41, R42
SC10	Backup and Recovery Management	R17
SC11	Data Protection and Encryption	R14, R32, R40, R46
SC12	Awareness and Training	R7, R15, R18, R32
SC13	Secure Development	R16, R26, R29, R30, R34, R38, R43
SC14	Incident and Continuity Management	R19, R22, R24, R25, R45
SC15	Human Resource Security	
SC16	Improvement and Compliance	R15, R18, R47, R48
SC17	Supplier Relationships	R21, R35
SC18	Physical Security	R44

Appendix D: Algorithm variables

The GEIGER indicator algorithm presented in Section 3.1.4 includes many variables. Table 22 lists the employed variables and their definitions.

Table 22: GEIGER indicator algorithm variable definitions.

Variable	Definition
T	The set of GEIGER threats.
P	The set of MSE profiles.
S	The set of cyber-systems.
M	The set of metrics.
C	The set of countermeasures.
E	The set of employees.
D	The set of devices.
r_{pt}	Risk associated with a threat t for an MSE with profile p .
v_{ms}	Value of a metric m for cyber-system s , normalized to between 0 and 1.
i_{mt}	Impact of a metric m on a threat t . Can be low (0.1), medium (0.5), or high (1.0).
i_{ct}	Impact of a countermeasure c on a threat t . Can be low (0.1), medium (0.5), or high (1.0).
δ_{mt}	Boolean indicator variable, which equals 1 if metric m relates positively to the cybersecurity risk of threat t .
λ_{ms}	Boolean indicator variable, which equals 1 if metric m has been calculated for cyber-system s .
γ_{ms}	Boolean indicator variable, which equals 1 if countermeasure c has been calculated for cyber-system s .
λ_{cs}	GEIGER score for a threat t and cyber-system s , which is (a part of) an MSE with profile p .

γ_{cs}	The total GEIGER score for the cyber-system s , which is (a part of) an MSE with profile p .
G_{spt}	The number of metrics calculated for a cyber-system s .
G_{sp}	Aggregate GEIGER score for an employee e , in an MSE with profile p .
G_{ept}	The total number of metrics calculated and used in the aggregate GEIGER score of employee e .
G^{emp}_{pt}	The overall MSE GEIGER score, for an MSE with profile p .
β_s	Boolean indicator variable, which equals 1 if metric m has been calculated for cyber-system s .
G_{dpt}	Boolean indicator variable, which equals 1 if countermeasure c has been calculated for cyber-system s .
G^{dev}_{pt}	GEIGER score for a threat t and cyber-system s , which is (a part of) an MSE with profile p .
G^{MSE}_{pt}	The total GEIGER score for the cyber-system s , which is (a part of) an MSE with profile p .
G^{MSE}_p	The number of metrics calculated for a cyber-system s .

Appendix E: Data model entities

In this section we outline data model's entities and attributes.

Table 23: Data Model Entities and Attributes.

Entity	Attributes
Enterprise (All information regarding the MSE)	MSE ID: UUID Name: MSE Name RISK Profile: UUID of risk profile {Digitally dependent MSE, Digitally based MSE or Digital enabler} Sector: MSE Sector {Agriculture, Forestry and Fishing (A), Mining and Quarrying (B), Manufacturing (C), Etc.} Location: UUID of location
Users (Information concerning all users in a MSE)	User ID: UUID of user Name: User First and Last Name Implemented Recommendations: UUID of implemented device-related recommendations Owner (Main User): Boolean value to indicate if the user is the MSE owner
Device (Information concerning all devices linked to a MSE)	Device ID: UUID of device Name: Device Name Type: Device Type can be either: tablet, laptop, computer or phone OS: Type of OS OS Version: OS Version

	<p>Owner: UUID of user that owns the device</p> <p>Implemented Recommendations: UUID of implemented device-related recommendations</p>
<p>User Role (Contains all possible MSE employee roles)</p>	<p>Role ID: UUID of role</p> <p>Name: Role name (Main or Regular)</p>
<p>Risk Profile (Lists all the existing profiles, so an MSE can be associated to a specific Risk Profile)</p>	<p>Risk Profile ID: UUID of risk profile</p> <p>Name: Risk profile name, currently limited to: Digitally dependent MSE, Digitally based MSE or Digital enabler</p> <p>Thread UUID, Weight: Each threat is assigned a weight based on the risk profile. Threat is identified by its UUID, while the weight is a value from 0 to 1.</p>
<p>Plugin (Tools) (Information on the installed plugin)</p>	<p>Plugin ID: UUID of the installed tool</p> <p>Company: Company name of the tool.</p>
<p>Sensor Value (Metrics) (Metrics output of the plugins)</p>	<p>Plugin ID: UUID of the installed tool</p> <p>Value: Value of the metric.</p> <p>Min Value: Minimum achievable metric value.</p> <p>Max Value: Maximum achievable metric value.</p> <p>Value Type: Type of the provided value: Boolean, integer or double.</p> <p>Related Threats, Impact: A comma separated list of all related threats along with the recommendation impact on the threat. Impact can be either: low, medium, or high.</p> <p>Urgency: only tied to Boolean metrics, to send notifications to user based on how critical the metric value is. Possible values are 'low': no notification, 'medium': notification after 1 week, 'high': notification after 1 day and 'critical': immediate notification.</p> <p>Flag: Indicates the relation between the metric and effect on score, if its value is positively affecting the score it's set to 1, otherwise 0/</p> <p>Relation: Indicates if the metric value reflects a value related to Device or User.</p>
<p>Threats (List of threat)</p>	<p>Threat ID: Threat UUID</p> <p>Name: Threat name</p> <p>Description: Description of the threat.</p>
<p>Recommendations * (Global list of recommendations)</p>	<p>Recommendation ID: Recommendation UUID</p> <p>Short Description: Short description of the recommendation.</p> <p>Long Description: Short description of the recommendation.</p>
<p>Own Recommendations (Recommendation provided by tool owners)</p>	<p>Related Threats, Impact: A comma separated list of all related threats along with the recommendation impact on the threat. Impact can be either: 0.1, 0.5 or 1</p> <p>Recommendation Type: Type of recommendation is related to Device or User</p> <p>Steps: A comma separated list of a maximum of three steps to aid the user in implement the recommendations.</p> <p>Action/Config: Redirection URL.</p>
<p>Assets</p>	<p>Asset ID: UUID</p>

<p>(Contains all available assets that can be linked to an enterprise)</p>	<p>Asset Name: Name</p> <p>List of assets MSEs and tool owners can choose from, sourced from NIST (2011) and ISO/IEC (2004):</p> <ul style="list-style-type: none"> • Physical asset <ul style="list-style-type: none"> ○ Computer hardware ○ Communications facility ○ Building • IT asset <ul style="list-style-type: none"> ○ Document ○ Database ○ Circuit ○ Computing device ○ Data ○ Network ○ Service ○ Software ○ System ○ Website ○ Synthetic ID ○ Connection ○ IP address ○ Host • Ability to provide product or service. • Person • Intangible <ul style="list-style-type: none"> ○ Goodwill ○ Image • Organisation
<p>Cybersecurity Defenders</p> <p>(Contains a list of security defenders contact information)</p>	<p>Name: Role name (Main or Regular)</p> <p>Affiliated Company UUID: UUID of the associated affiliation</p> <p>Telephone number: (phone; intl. layout "+41 ..."; valid characters "+[0-9]{5,20}")</p> <p>Email: associated email</p>
<p>GEIGER Score</p> <p>(Output of the GEIGER indicator)</p>	<p>There are 3 instances of this entity for each cyber-system: Aggergate score, Device and User.</p> <p>GEIGER Score: The total GEIGER score for the corresponding cyber-system.</p> <p>Threat, Score: UUID of each threat along with the threat score.</p>
<p>GEIGER Recommendations</p> <p>(Output of the GEIGER indicator)</p>	<p>There are 2 instances of this entity for each cyber-system: Device and User. Each instance is further divided into T instances, where T is the number of threats.</p> <p>Threat ID: Threat UUID</p> <p>Recommendation, Rank: UUID of the recommendation along with its ranking. The rank is based on the impact the recommendation has on GEIGER score.</p>

6. References

6.1 GEIGER Indicator

1. Australian Cyber Security Centre (ACSC, 2019). Small Business Cyber Security Guide. URL: <https://www.cyber.gov.au/acsc/view-all-content/publications/small-business-cyber-security-guide>.
2. Australian Cyber Security Centre (ACSC, 2020). Cyber Security and Australian Small Businesses: Results from the Australian Cyber Security Centre Small Business Survey. URL: <https://www.cyber.gov.au/sites/default/files/2020-11/ACSC%20Small%20Business%20Survey%20Results.pdf>.
3. Barrett, M. P. (2018). Framework for improving critical infrastructure cybersecurity version 1.1. URL: <https://www.nist.gov/publications/framework-improving-critical-infrastructure-cybersecurity-version-11>.
4. Böhme, R., & Freiling, F. C. (2008). On metrics and measurements. In *Dependability metrics* (pp. 7-13). Springer, Berlin, Heidelberg.
5. Carré, H. (2008). Statistical classification of economic activities in the european community. *Publications Office of the European Union: Luxembourg*.
6. Centre for Internet Security (CIS, 2019). CIS Controls V7.1. URL: <https://www.cisecurity.org/controls/cis-controls-list/>.
7. Craigen, D., Diakun-Thibault, N., & Purse, R. (2014). Defining cybersecurity. *Technology Innovation Management Review*, 4(10).
8. Cyber Security Body of Knowledge (CyBOK, 2019). CyBOK Knowledgebase. URL: <https://www.cybok.org/knowledgebase/>.
9. Cybersecurity and Infrastructure Security Agency (CISA, 2009). Security Tip (ST014-015): Understanding Denial-of-Service Attacks. URL: <https://us-cert.cisa.gov/ncas/tips/ST04-015>.
10. Davis, M. C., Challenger, R., Jayewardene, D. N., & Clegg, C. W. (2014). Advancing socio-technical systems thinking: A call for bravery. *Applied ergonomics*, 45(2), 171-180.
11. Dutta, A., & Al-Shaer, E. (2019, June). "What", "Where", and "Why" Cybersecurity Controls to Enforce for Optimal Risk Mitigation. In *2019 IEEE Conference on Communications and Network Security (CNS)* (pp. 160-168). IEEE.
12. ENISA (2012). ENISA Threat Landscape 2012. URL: https://www.enisa.europa.eu/publications/ENISA_Threat_Landscape.
13. ENISA (2015). Cloud Security Guide for SMEs: Cloud computing security risks and opportunities for SMEs. URL: https://www.enisa.europa.eu/publications/cloud-security-guide-for-smes/at_download/fullReport.
14. ENISA (2016a). ENISA Threat and Risk Management Glossary. URL: <https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/current-risk/risk-management-inventory/glossary>.
15. ENISA (2016b). ENISA Threat Taxonomy. URL: <https://www.enisa.europa.eu/topics/threat-risk-management/threats-and-trends/enisa-threat-landscape/threat-taxonomy>.
16. ENISA (2016c). CSIRTs in Europe Glossary: Ransomware. URL: <https://www.enisa.europa.eu/topics/csirts-in-europe/glossary/ransomware>.
17. ENISA (2017). Guidelines for SMEs on the security of personal data processing. URL: <https://www.enisa.europa.eu/publications/guidelines-for-smes-on-the-security-of-personal-data-processing>.

18. ENISA (2018). ENISA Reference Incident Classification Taxonomy. URL: <https://www.enisa.europa.eu/publications/reference-incident-classification-taxonomy>.
19. ENISA (2019). ENISA Threat Landscape Report 2018. URL: <https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2018>.
20. ENISA (2020). ENISA Threat Landscape – The year in review. URL: <https://www.enisa.europa.eu/publications/year-in-review>.
21. European Digital SME Alliance (2020). The EU Cybersecurity Act and the role of standards for SMEs: Position Paper. URL: <https://www.digitalsme.eu/digital/uploads/The-EU-Cybersecurity-Act-and-the-Role-of-Standards-for-SMEs.pdf>.
22. Evesti, A., & Ovaska, E. (2013). Comparison of adaptive information security approaches. *International Scholarly Research Notices*, 2013.
23. Federal Trade Commission (FTC, 2018a). Cybersecurity for Small Business. URL: <https://www.ftc.gov/tips-advice/business-center/small-businesses/cybersecurity>.
24. Federal Trade Commission (FTC, 2018b). Cybersecurity for Small Business: Understanding the NIST Cybersecurity Framework. URL: <https://www.ftc.gov/tips-advice/business-center/small-businesses/cybersecurity/nist-framework>.
25. Goepel, K. D. (2013, June). Implementing the analytic hierarchy process as a standard method for multi-criteria decision making in corporate enterprises—a new AHP excel template with multiple inputs. In *Proceedings of the international symposium on the analytic hierarchy process* (Vol. 2, No. 10, pp. 1-10). Creative Decisions Foundation Kuala Lumpur.
26. Gollmann, D., Herley, C., Koenig, V., Pieters, W., & Sasse, M. A. (2015). Socio-Technical security metrics (Dagstuhl seminar 14491). *Dagstuhl reports*, 4(12), 28.
27. Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Issues in Information Warfare & Security Research*, 1(1), 80.
28. ISO/IEC (2004). ISO/IEC 13335-1:2004: Information technology - Security techniques - Management of information and communications technology security - Part 1: Concepts and models for information and communications technology security management. URL: <https://www.iso.org/standard/39066.html>.
29. Jore, S. H. (2019). The conceptual and scientific demarcation of security in contrast to safety. *European Journal for Security Research*, 4(1), 157-174.
30. Kim, W., Choi, B. J., Hong, E. K., Kim, S. K., & Lee, D. (2003). A taxonomy of dirty data. *Data mining and knowledge discovery*, 7(1), 81-99.
31. Menard, P., Bott, G. J., & Crossler, R. E. (2017). User motivations in protecting information security: Protection motivation theory versus self-determination theory. *Journal of Management Information Systems*, 34(4), 1203-1230.
32. Mijnhardt, F., Baars, T., & Spruit, M. (2016). Organizational characteristics influencing SME information security maturity. *Journal of Computer Information Systems*, 56(2), 106-115.
33. MISP (2021). MISP taxonomies and classification as machine tags: CERT-XLM. URL: <https://www.misp-project.org/taxonomies.html>.
34. National Institute of Standards and Technology (NIST, 2011). NISTIR 7693: Specification for Asset Identification 1.1. URL: <https://csrc.nist.gov/publications/detail/nistir/7693/final>.
35. National Institute of Standards and Technology (NIST, 2021). NIST Glossary. URL: <https://csrc.nist.gov/glossary>.

36. NCSC UK (2018), Infographics at the NCSC. URL: <https://www.ncsc.gov.uk/information/infographics-ncsc>.
37. NCSC UK (2021). Cyber Essentials. URL: <https://www.ncsc.gov.uk/cyberessentials/overview>.
38. Ortiz-de-Mandojana, N., & Bansal, P. (2016). The long-term benefits of organizational resilience through sustainable business practices. *Strategic Management Journal*, 37(8), 1615-1631.
39. OWASP (2016). OWASP Cyber Defense Matrix. URL: <https://owasp.org/www-project-cyber-defense-matrix/>.
40. OWASP (2020). OWASP Top 10. URL: <https://owasp.org/www-project-top-ten/#>.
41. Padayachee, K. (2012). Taxonomy of compliant information security behavior. *Computers & Security*, 31(5), 673-680.
42. Pendleton, M., Garcia-Lebron, R., Cho, J. H., & Xu, S. (2016). A survey on systems security metrics. *ACM Computing Surveys (CSUR)*, 49(4), 1-35.
43. Pfleeger, C.P., Pfleeger, S.L., Margulies, J. (2015). *Security in computing*. Prentice Hall, Upper Saddle River, NJ.
44. Refsdal, A., Solhaug, B., & Stølen, K. (2015). *Cyber-risk management*. In *Cyber-Risk Management* (pp. 33-47). Springer, Cham.
45. Sager, T. (2015). The Cyber OODA Loop: How Your Attacker Should Help You Design Your Defense. *2015 Cybersecurity Innovation Forum*. URL: <https://csrc.nist.gov/Presentations/2015/The-Cyber-OOA-Loop-How-Your-Attacker-Should-Help>.
46. Shirey, R. W. (2007). IETF (Internet Engineering Task Force) RFC 4949. Internet Security Glossary, Version 2.
47. Shojaifar, A., Fricker, S. A., & Gwerder, M. (2020, September). Automating the Communication of Cybersecurity Knowledge: Multi-case Study. In *IFIP World Conference on Information Security Education* (pp. 110-124). Springer, Cham.
48. Swiss National Cyber Security Centre (Swiss NCSC, 2021). Cyberthreats. URL: <https://www.ncsc.admin.ch/ncsc/en/home/cyberbedrohungen.html>.
49. Vocabulary for event recording and incident sharing (VERIS, 2017). VERIS Threat Actions. URL: <http://veriscommunity.net/actions.html>.
50. Vocabulary for event recording and incident sharing (VERIS, 2021). The VERIS Community Database. URL: <https://github.com/vz-risk/VCDB>.
51. Widmer, G., & Kubat, M. (1996). Learning in the presence of concept drift and hidden contexts. *Machine learning*, 23(1), 69-101.
52. Yigit Ozkan, B., Spruit, M. (2021). Cybersecurity Standardisation Essentials for European SMEs. URL: <https://webspacescience.uu.nl/~sprui107/download/Cybersecurity%20Standardisation%20Essentials%20for%20European%20SMEs.pdf>.