



# NextGEM

## **Next Generation Integrated Sensing and Analytical System for Monitoring and Assessing Radiofrequency Electromagnetic Field Exposure and Health**

### **D6.8: Trustworthy data management and compliance with ethics and legal aspects – Final report**

#### Document Summary Information

<b>Start Date</b>	01/07/2022	<b>Duration</b>	48 months
<b>Project URL</b>	<a href="https://www.nextgem.eu/">https://www.nextgem.eu/</a>		
<b>Deliverable</b>	D6.8: Trustworthy data management and compliance with ethics and legal aspects – Final report		
<b>Work Package</b>	WP6	<b>Task</b>	T6.3
<b>Contractual due date</b>	30/04/2025	<b>Actual submission date</b>	30/04/2025
<b>Type</b>	Report	<b>Dissemination Level</b>	PU - Public
<b>Lead Beneficiary</b>	SANL	<b>Deliverable Editor</b>	Stefanos Fafalios (SANL)



This project has received funding from the European Union's Horizon Europe research and innovation programme under the Grant Agreement No 101057527

## Contributors and Peer Reviewers

<b>Contributors</b>
Stefanos Fafalios (SANL), Panos Chatziadam, Nikolaos Petroulakis (FORTH), Nicolas Louca (EBOS), Dimitrios Laskaratos (ICOM), Maryse Ledent (SC), Anna Bodganova (UZH), Sam Aerts (THUAS)
<b>Peer Reviewers</b>
Schiavoni Andrea (FIBER), Dimitrios Laskaratos (ICOM)

## Revision history (including peer-reviewing and quality control)

Version	Issue Date	Changes	Contributor(s)
v0.1	22/01/2025	Table of Contents provided	Stefanos Fafalios (SANL)
v0.2	06/02/2025	Sections populated with the Task leaders	Stefanos Fafalios (SANL)
v0.3	13/03/2025	Section defined, assigned, and agreed	Stefanos Fafalios (SANL), Panos Chatziadam (FORTH), Nikolaos Petroulakis (FORTH)
v0.4	22/03/2025	First contributions	All partners
v0.5	04/04/2025	Integration and harmonization	Stefanos Fafalios (SANL)
v0.6	11/04/2025	Second contributions and updates	Dimitrios Laskaratos (ICOM), Panos Chatziadam (FORTH), Nicolas Louca (EBOS)
v0.7	16/04/2025	Complete version ready for peer review	Stefanos Fafalios (SANL)
v0.8	18/04/2025	Peer review	Schiavoni Andrea (FIBER), Dimitrios Laskaratos (ICOM)
v0.9	23/04/2025	Comments addressed from peer review, technical and quality assurance	Stefanos Fafalios (SANL), Mats-Olof Mattsson (SPi), Nicolas Louca (eBOS), Panos Chatziadam (FORTH)
v1.0	30/04/2025	Final review and submission	Nikolaos Petroulakis (FORTH)

## Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.”

While the information contained in the documents is believed to be accurate, the authors(s) or any other participant in the NextGEM consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the NextGEM Consortium nor any of its members, their officers, employees, or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the NextGEM Consortium nor any of its members, their officers, employees, or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

## Copyright message

© NextGEM Consortium. This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation, or both. Reproduction is authorised provided the source is acknowledged.

# Table of Contents

Executive Summary.....	8
1 Introduction.....	9
1.1 Mapping NextGEM Outputs .....	9
1.2 Deliverable overview and report structure .....	10
1.3 Updates from previous Deliverable D6.3 “Trustworthy data management and compliance with ethics and legal aspects – Initial report” .....	10
2 NIKH security requirements .....	11
2.1 GDPR requirements .....	11
2.2 CIA principles .....	11
2.2.1 Explicit security requirements and NIST SP 800-53 Controls.....	11
3 The NIKH ecosystem .....	13
3.1 The NIKH architecture.....	13
3.1.1 Application layer .....	14
3.1.2 Service layer.....	14
3.1.3 Third-party services .....	15
3.1.4 Infrastructure .....	16
3.2 Assets, procedures and relations.....	16
3.3 Ensuring GDPR compliance in data management.....	24
3.4 Potential threats .....	25
3.4.1 NIKH threat landscape overview .....	25
3.4.2 NIKH components’ potential threats and vulnerabilities.....	26
4 The Security and Privacy Assurance tool.....	31
4.1 Overview.....	31
4.1.1 Asset-based vulnerability assessments .....	34
4.1.2 Dynamic testing .....	35
4.1.3 Event reasoning toolkit.....	35
4.1.4 Event captors .....	36
4.2 Tool deployment and association with NIKH.....	37
4.2.1 Core deployment .....	37
4.2.2 Deployed captors on NIKH components .....	38
5 Vulnerability and security assurance.....	39
5.1 NVD vulnerability assessments .....	39
5.2 Continuous monitoring .....	41
5.2.1 Service availability monitoring.....	41
5.2.2 Access monitoring.....	42
5.3 Penetration testing.....	43
5.4 Assessing GDPR compliance .....	46
6 Risk minimization .....	48

6.1 Assessment: evaluating system vulnerabilities ..... 48

6.2 Action: implementing risk reduction measures ..... 48

6.3 Continuous improvement through feedback loops ..... 48

7 Conclusion ..... 49

## List of Figures

Figure 1. High-level view of the NIKH hub.....	13
Figure 2: The NextGEM community page on Zenodo.....	15
Figure 3: NIKH's Literature Search feature.....	16
Figure 4 : NIKH ecosystem component relations.....	17
Figure 5: High-level SPA tool diagram.....	31
Figure 6: Overview of the SPA tool architecture .....	33
Figure 7: Analyzer and synchronizer communication.....	35
Figure 8: Shipping NIKH logs to the SPA tool. ....	36
Figure 9: Interaction between SPA tool, Monitor and Event Captors.....	37
Figure 10: SPA GUI view: Defined asset list.....	39
Figure 11: SPA GUI view: NVD assessment results overview - 1.....	40
Figure 12: SPA GUI view: NVD assessment results overview - 2.....	40
Figure 13. SPA GUI view: NVD assessment results. ....	40
Figure 14. SPA GUI view: Details on identified vulnerability. ....	41
Figure 15: SPA GUI view: Availability monitoring results.....	42
Figure 16: SPA GUI view: IP confidentiality monitoring results. Both successful and unsuccessful attempts are shown.....	43
Figure 17: External/Internal Penetration Testing Workflow.....	44

## List of Tables

Table 1: Adherence to NextGEM's GA Tasks and Deliverables Descriptions .....	9
Table 2: NIKH assets - Controller.....	17
Table 3: NIKH assets - Risk Assessment Tool .....	19
Table 4: NIKH assets - GUI.....	20
Table 5: NIKH assets - Connectors & Data Spaces .....	20
Table 6: NIKH assets - Modelling tool.....	22
Table 7: NIKH assets - SPA tool.....	22
Table 8: Deployment Status of Event Captors on NIKH components .....	38
Table 9: 3rd party services .....	41
Table 10: Partner's Answers on Data Handling .....	46

## Glossary of terms and abbreviations used

Abbreviation / Term	Description
API	Application Programming Interface
CIA	Confidentiality, Integrity, Availability
CRUD	Create Read Update Delete
CVE	Common Vulnerabilities and Exposures
CWE	Common Weakness Enumeration
DoS	Denial of Service
DPO	Data Protection Officer
EEA	European Economic Area
ELK	Elastic Stack
EU	European Union
EVEREST	Event Reasoning Toolkit
GDPR	General Data Protection Regulation
GUI	Graphical User Interface
IP	Internet Protocol
NIKH	NextGEM Innovation and Knowledge Hub
NIST	National Institute of Standards and Technology
NVD	National Vulnerability Database
REST	Representational State Transfer
SPA	Security and Privacy Assurance
SSH	Secure Shell
STS	SPHYNX Technology Solutions

## Executive Summary

Deliverable D6.8 provides the final report for trustworthy data management and compliance with ethics and legal aspects, detailing the activities undertaken for Task 6.3 “Distributed and trustworthy data management and compliance with ethics and legal aspects”. The purpose of this deliverable is to assess the NIKH’s security posture. The deliverable provides an analysis of the NIKH’s components and enumerates possible security threats that these components could face. It then introduces the Security Assurance tool along with its capabilities, which will be used to assess the NIKH’s security posture from various perspectives. The report describes vulnerability assessments and continuous monitoring efforts conducted to date. It also outlines a penetration testing process, including a preliminary phase completed within the Task 6.3 timeline and a comprehensive test scheduled near the conclusion of Task 6.4, ensuring alignment with the completion of NIKH’s development cycle. Additionally, it details measures to mitigate the risk of identified vulnerabilities being exploited.



# 1 Introduction

The purpose of Task 6.3 is to evaluate and assess the security posture and GDPR compliance of the NextGEM Information and Knowledge Hub (NIKH), focusing on the core principles of the Confidentiality, Integrity, Availability (CIA) triad.

The evaluation considers how these principles align with platform security measures and GDPR compliance efforts. It aims to identify potential vulnerabilities and suggest improvements to strengthen the hub's overall security posture and GDPR adherence using a state-of-the-art Security and Privacy Assurance (SPA) tool tailored to the NIKH's requirements.

Following a comprehensive methodology, this task will provide reports on:

- **Identified vulnerabilities** or security weaknesses that could impact the hub's CIA.
- Possible **recommendations** for addressing these issues and enhancing security controls in each area.
- **Analysis** of adherence to GDPR requirements related to confidentiality, integrity, and availability of user data.

Deliverable D6.3 identified potential tools and methods for a comprehensive security evaluation of the NIKH. This deliverable also provides an initial assessment of the security posture of the hub using the chosen methods and tools and finally presented a first report summarizing evaluation methodology and high-level findings. Finally, it details a mitigation plan to address identified vulnerabilities and minimize security risks.

Following D6.3, the present deliverable, Deliverable D6.8, delves deeper by building upon the initial assessment. It includes a comprehensive security posture evaluation using the chosen tools and methods, identifying and documenting any discovered vulnerabilities. Additionally, it analyses the NIKH's current practices and compliance with GDPR, recommending possible enhancements to strengthen the NIKH's GDPR compliance posture. Finally, D6.8 delivers the final report summarizing the entire process, including the detailed methodology, a mitigation plan, and recommendations for improved GDPR compliance.

## 1.1 Mapping NextGEM Outputs

The purpose of this section is to map NextGEM's Grant Agreement (GA) commitments, both within the formal Task description and Deliverable, against the project's respective outputs and work performed.

Table 1: Adherence to NextGEM's GA Tasks and Deliverables Descriptions

TASKS	
Task Number & Title	Respective extract from formal Task Description
Task 6.3 - Distributed and trustworthy data management and compliance with ethics and legal aspects	This task deals with the security assurance for the NextGEM framework. For this purpose, the STS Security Assurance Platform will be tailored to NextGEM deployment to assess its security posture by means of vulnerability analysis, penetration testing and continuous monitoring. Moreover, this assessment and monitoring of the NextGEM framework will have a GDPR compliance flavour by: 1) inspect organisational and technical measures that are put in place to ensure compliance with GDPR requirements; 2) verify their effectiveness and; 3) record all information related to the handling of personal data offering their accountability.
DELIVERABLE	
Deliverable: D6.8: Trustworthy data management and compliance with ethics and legal aspects – Final report (M34)	
This deliverable will provide the final report to enable the capability of NextGEM for secure data exchange through the integration the STS security Assurance platform for trustworthy data management complying with GDPR and legal/ ethics requirements.	

## 1.2 Deliverable overview and report structure

Based on the objectives and work carried out under Task 6.3, the document starts with the Executive Summary followed by the introduction of the document in Section 1.

Section 2 deals with security and legal principles, such as CIA and GDPR.

Section 3 introduces the NIKH hub, its architecture and components, as well as what procedures have been put in place to ensure GDPR compliance of data handling within the NIKH.

Section 4 describes the SPA tool, which has been issued to provide and catalogue security assessments for the NIKH, while Section 5 details the deployment of the tool and includes a description of the assessments performed, along with assessment results.

In Section 6, an overview is provided on how vulnerabilities identified in Section 5 are mitigated to minimize the associated risks.

Finally, Section 7 concludes the deliverable.

## 1.3 Updates from previous Deliverable D6.3 “Trustworthy data management and compliance with ethics and legal aspects – Initial report”

This deliverable builds upon Deliverable D6.3, utilizing the infrastructure and software established during the initial cycle of security assessments detailed in that report. As a continuation, it integrates the foundational work from D6.3 and introduces the following updates and enhancements:

- Section 2.2.1 was updated with NIST security controls, mapping previous requirements (S1-S6) to NIST requirements, expanded with S7 and S8.
- Section 3.3 has been revised for improved readability and linked to deliverable D6.7.
- Section 4.1.2 now features an updated description of dynamic testing, refined to better align with the SPA tool implementation.
- Section 4.2.2 includes an updated event captor table containing additional components to be monitored, reflecting the latest deployment statuses.
- Section 5 has been expanded to map the security requirements from Section 2.2.1 into its subsections.
- Section 5.1 replaces outdated figures with new ones to accurately represent the current SPA tool GUI, supplemented by additional figures on result overviews and text describing these updates, along with a disclaimer on how results are shared with partners.
- Section 5.2.1 updates the monitored third-party databases table to list external databases currently utilized by NIKH, accompanied by a new figure illustrating the GUI view of availability monitoring results and descriptive text.
- Section 5.2.2 provides a rationale for shifting from the initial strategy, supported by a figure displaying IP confidentiality results and accompanying text.
- Section 5.3 has been significantly expanded, offering a detailed explanation of the penetration testing methodology, complemented by a figure depicting the workflow.
- Section 5.4 has been updated and linked with Section 3.3.
- Finally, Section 7 has been updated to reflect D6.8 as the final report, summarizing the completed Task 6.3 assessments (NVD analysis, initial monitoring, and GDPR compliance) and detailing the iterative improvement process from Section 6, with penetration testing and monitoring of two components scheduled for Task 6.4.

## 2 NIKH security requirements

In this section, the legal aspects governing the NIKH are introduced and explained, outlining the essential compliance obligations it must meet. Moreover, the security standards established to ensure alignment with these legal requirements are also introduced.

### 2.1 GDPR requirements

The General Data Protection Regulation (GDPR)<sup>1</sup> stands as one of the most significant pieces of data protection legislation in recent years. Enforced by the European Union (EU), it aims to safeguard the personal data of individuals within the EU and European Economic Area (EEA). GDPR was implemented on May 25, 2018, replacing the Data Protection Directive 95/46/EC. The principles that it governs are:

1. **Lawfulness, fairness and transparency:** Organizations must have a valid legal reason to collect and use personal data, and they need to be upfront with individuals about how their data is being used. This involves obtaining clear and informed consent.
2. **Purpose limitation:** Data can only be collected and used for the specific purposes that were communicated to the individual.
3. **Data minimization:** Organizations should only collect and use the minimum amount of personal data necessary for their specific purposes.
4. **Accuracy:** Personal data should be accurate and kept up-to-date whenever necessary.
5. **Storage limitation:** Data shouldn't be stored for longer than necessary for processing purposes. Organizations need to have clear guidelines for data retention and deletion.
6. **Integrity and confidentiality:** Appropriate technical and organizational measures must be implemented to safeguard personal data from unauthorized access, disclosure, alteration, or destruction.
7. **Accountability:** Organizations are ultimately responsible for ensuring compliance with the GDPR principles. This may involve appointing a Data Protection Officer (DPO) and conducting Data Protection Impact Assessments (DPIAs) for high-risk processing activities.

### 2.2 CIA principles

The CIA Triad<sup>2</sup> is a foundational framework in information security, representing three core principles aimed at safeguarding data assets:

- **Confidentiality:** Ensuring that information is only accessible to those who have the authority to access it. This involves measures such as encryption, access controls, and data classification to prevent unauthorized access to sensitive information.
- **Integrity:** Maintaining the accuracy and reliability of data throughout its lifecycle. Integrity measures focus on preventing unauthorized or unintended modification of data, ensuring that information remains accurate, complete, and trustworthy. Techniques such as checksums, digital signatures, and access controls contribute to maintaining data integrity.
- **Availability:** Ensuring that data and resources are available and accessible to authorized users when needed. Availability measures aim to prevent disruptions to services and ensure continuity of operations, often through redundancy, disaster recovery planning, and robust network infrastructure.

The CIA Triad provides a comprehensive framework for evaluating and implementing security controls to protect information assets from a wide range of threats, including unauthorized access, data breaches, and service interruptions. Organizations often use the CIA Triad as a guide to develop and maintain effective information security practices.

#### 2.2.1 Explicit security requirements and NIST SP 800-53 Controls

Due to the intricate nature of the NIKH, as outlined in Deliverable D2.5 Section 2.2 (“Stakeholders Requirements and scenarios”), explicit security requirements are essential to safeguard privacy and security at the user, infrastructure, and data-usage levels. These are complemented by controls from the National Institute of Standards

---

<sup>1</sup> <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

<sup>2</sup> <https://www.nccoe.nist.gov/publication/1800-25/VolA/index.html>

and Technology (NIST) SP 800-53<sup>3</sup>, a widely adopted framework that enhances NIKH's security posture and aligns with CIA principles and GDPR obligations. The following requirements are tailored to NIKH's role as a data-intensive research hub and matched to applicable NIST controls:

**S1: Authentication and Authorization (AC-2, AC-3):** NIKH must offer mechanisms for authentication and authorization of all involved actors/entities/stakeholders requiring access to restricted functionalities and/or data, supported by account management (AC-2) and access enforcement (AC-3) as implemented via tools like Keycloak (Section 3.3).

**S2: Role-Based Access (AC-5):** NIKH must support role-based access to ensure data security and privacy, aligned with the separation of duties (AC-5) to manage access based on defined roles through centralized authentication (Section 3.3).

**S3: Vulnerability Assessments (RA-3):** NIKH should perform vulnerability assessments of its constituent parts to ensure proper functionality and availability of services, corresponding to risk assessment (RA-3) integrated into the SPA tool's evaluations (Section 5.1).

**S4: Risk Assessment (RA-3):** NIKH should assess the impact and risk of the detected vulnerabilities, also supported by risk assessment (RA-3) to inform mitigation strategies within the risk minimization cycle (Section 6).

**S5: Trustworthy Data Exchange (SC-7, SC-8):** NIKH should adopt a sovereign and trustworthy process to perform data exchanges and transfers among interested parties, matching boundary protection (SC-7) and transmission confidentiality/integrity (SC-8) using encryption and firewalls (Section 3.1.2).

**S6: Policy Enforcement (MP-2):** NIKH should be able to encompass and enforce policy restrictions related to the usage of data, aligned with media access (MP-2) to secure data storage and access per policies (Section 3.3).

**S7: Audit and Accountability (AU-2, AU-6):** NIKH should maintain audit records of security-relevant events (e.g., login attempts, data access) and review logs to detect anomalies, supported by audit events (AU-2) and audit review (AU-6) via continuous monitoring (Section 5.2).

**S8: Attack Surface Reduction (CM-8, SI-2):** NIKH should reduce its attack surface by removing or disabling unused ports, services, and components, and ensure secure configurations of products used, supported by system component inventory (CM-8) to identify unnecessary elements and flaw remediation (SI-2) to address insecure configurations (Section 6.2).

These requirements ensure that NIKH protects its infrastructure, data, and stakeholder trust. Requirements S1-S7, address NIKH-specific needs, supplemented by S8 for attack surface reduction and secure configurations, while NIST controls (AC, RA, SC, MP, AU, CM, SI) provide corresponding standards, reinforcing CIA adherence and supporting GDPR compliance through structured security measures.

---

<sup>3</sup> <https://csrc.nist.gov/pubs/sp/800/53/r5/upd1/final>

### 3 The NIKH ecosystem

The NIKH aims to become a strategic key player in the area of electromagnetic fields (EMF) and health. Part of the broader CLUE-H (EMF Health Cluster)<sup>4</sup> initiative, the NIKH operates as a central hub for facilitating research, knowledge exchange, and the development of public policy in this critical area of study.

NIKH's mission is to serve as a collaborative space that brings together researchers, industry experts, healthcare professionals, and policymakers. The goal is to harness diverse expertise to understand better the potential health impacts of EMF exposure and to inform both the scientific community and the public. Through interdisciplinary research and dialogue, the NIKH aims to contribute to the development of evidence-based policies and practices that protect public health while supporting technological advancement.

The hub focuses on several key objectives, including the synthesis of existing knowledge on EMF and health, the facilitation of new research through the collection and organization of novel data, the translation of scientific insights into practical policy guidance, and the fostering of international cooperation. By addressing these objectives, the NIKH seeks to provide a comprehensive understanding of EMF's health implications and to promote informed decision-making across Europe and beyond.

In summary, the NIKH represents a concerted effort to bridge the gap between science, policy, and public understanding in the field of EMF and health. Its work is grounded in the principles of collaboration, transparency, and rigor, aiming to advance our knowledge and response to the challenges posed by EMF exposure.

#### 3.1 The NIKH architecture

In this section, the functional components comprising the NIKH hub, along with the architecture, are presented. Figure 1 presents a high-level view of the architecture of the NIKH.

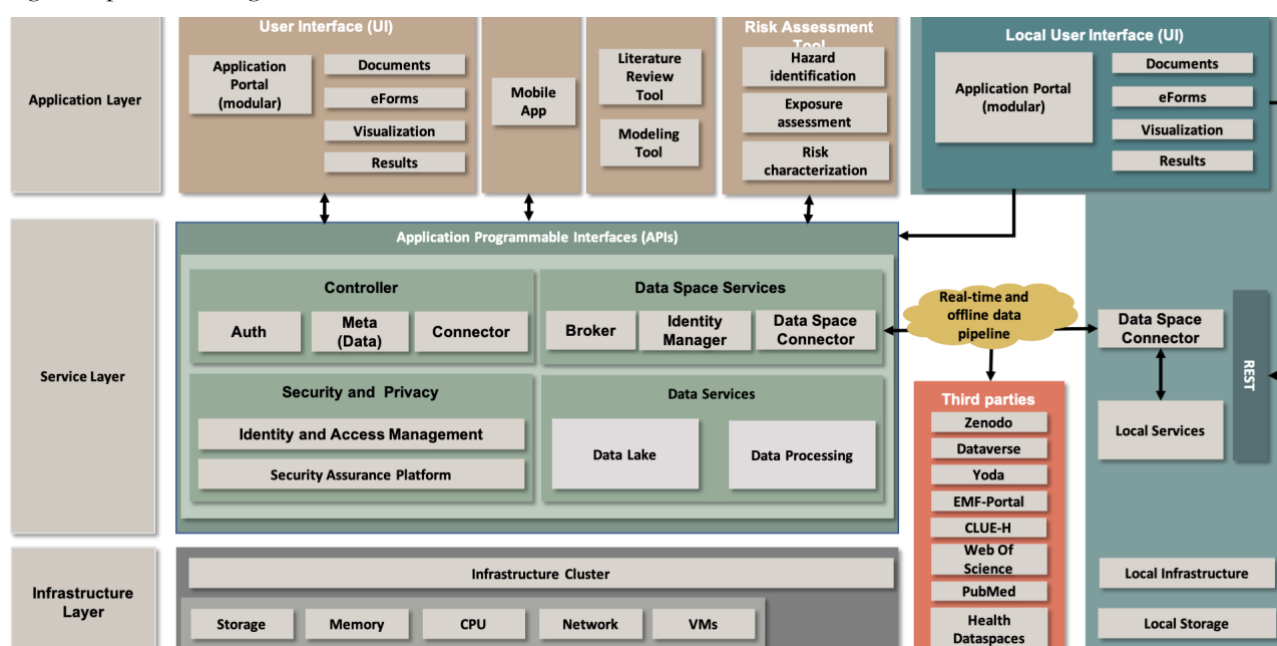


Figure 1. High-level view of the NIKH hub.

Each component serves a specific purpose and/or function, be it overall user experience, system management, or alignment with the project's requirements, to create a user-centric interface that facilitates user interaction, service management, and data access for optimized navigation through the ecosystem. An in-depth description of the NIKH's architecture is provided in D2.5 "NextGEM architectural framework – Final version", while the following subsections provide a summarised description of the main components of NIKH's architecture.

<sup>4</sup> <https://www.emf-health-cluster.eu/>

### 3.1.1 Application layer

- **User Interface (UI):** The UI includes the means for users to interact with the NIKH and perform all intended actions to be enabled by it.
  - **Application Portal:** The application portal provides a unified access point for both internal and external stakeholders of NextGEM. It offers a visual interface for interacting with the NIKH and its components, with functionality and access tailored according to the user's role and permissions.
  - **Documents:** Any content or data in the form of digital files that are available through the NIKH, which users can access or interact with.
  - **eForms:** eForms allow for data input to perform functions such as searching or filtering of content or initiating functions such as user sign-in.
  - **Visualisation:** the means of presenting data or information via graphs, gauges, maps, images, videos, or other graphical means.
- **Mobile Application:** The external stakeholders of NIKH are also offered a mobile-friendly experience that allows mobile-specific features that extend the core functionality while catering to the unique needs of mobile users.
- **Modelling Tool:** Provides a Graphical User interface (GUI) for the results taken from NextGEM such as from ERMES <sup>5</sup> modelling tool, allowing researchers to evaluate and compare the results from modelling.
- **Literature Review Tool:** Provides a tool for conducting literature searches across various internal and external resources, screen papers, data extraction, risk of bias as required for conducting literature, systematic, and umbrella reviews.
- **Risk Assessment Tool:** Provides a tool for conducting risk assessments based on EMF exposure scenarios.

### 3.1.2 Service layer

- **Application Programmable Interfaces (APIs):** The APIs have the critical role in facilitating communication and interaction between different components of NIKH, and enabling functions such as data retrieval, visualisation, security, authorisation and data manipulation.
  - **Controller:** responsible for orchestrating and managing received requests. The Controller contains all interfaces through which a user interacts with NIKH, offering the REST API endpoints. Every request is handled by the Controller, which performs the corresponding actions and generates the appropriate response. Regarding its architecture, the Controller implements various software components that are responsible for managing the various heterogeneous system resources that are available through other services on the hub.
  - **Security and Privacy:** This mechanism handles user authentication and authorization to ensure that only authorized users can access and interact with the NextGEM system from the dashboard. Specifically, this module is responsible for managing the use of the dashboard's functionality. This module handles user login, registration, and authentication and authorizes access based on roles and permissions.
  - **Data Space Services:** Data Spaces enable controlled, sovereign, and secure exchange and sharing of data between stakeholders. NIKH adopts International Data Spaces (IDS) to uphold trust, data sovereignty, and transparency. IDS, endorsed by the International Data Spaces Association (IDSA), facilitates secure data exchange by keeping data at its source until requested. Each participant maintains control over their data and can monitor transactions. Identity verification ensures security. IDS also provides metadata storage and search functionalities. NIKH integrates these principles, aiming for interoperability and diverse data source integration.
  - **Data Services:** Data services allow the ingestion, storage, and processing of data. The Data Lake serves as the central storage of NIKH and can accommodate structured and unstructured data of

<sup>5</sup> Otin, Ruben. "ERMES 20.0: Open-source finite element tool for computational electromagnetics in the frequency domain." Computer Physics Communications (2025): 109521.



any scale and format, while the Data Processing component handles any data-processing operations such as data profiling, parsing, filtering, transformation, and model training.

### 3.1.3 Third-party services

Through this component, external sources are integrated into NIKH, facilitating their provided APIs. The final data sources integrated into NIKH, in addition to information stored natively into NIKH, include several well established scientific stakeholders such as EMF-Portal <sup>6</sup>, Zenodo <sup>7</sup>, PubMed <sup>8</sup> and Web of Science (WOS) <sup>9</sup>.

Apart from that, the results from CLUE-H European projects <sup>10</sup> are included in the NIKH search. More specifically, the results from SEAWave are located in Zenodo <sup>11</sup>, GOLiAT <sup>12</sup> stores its results in the Dataverse <sup>13</sup> and ETAIN<sup>14</sup> stores its results in the YODA <sup>15</sup> where the models are in SketchFab <sup>16</sup>.

Especially in NextGEM stores all public results to Zenodo <sup>17</sup>. For that reason, a NextGEM community has been created in Zenodo, as shown in Figure 2, where the researchers participating in NextGEM upload research-related documents. Through its REST API, Zenodo enables third-party services to access publications uploaded onto its database. In the case of NIKH, its Metadata Manager will identify accessed NextGEM-related publications on Zenodo through metadata, and these publications will be available through NIKH.

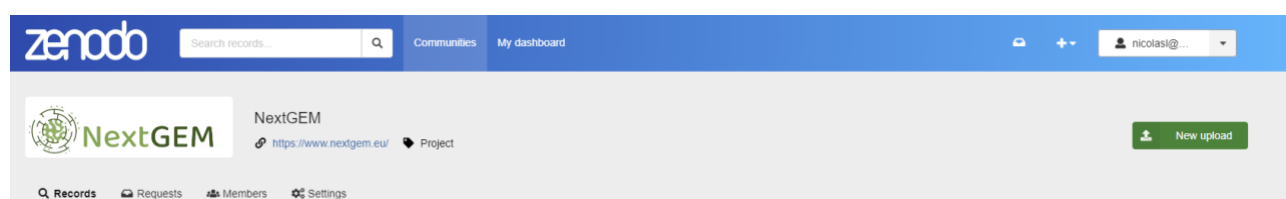


Figure 2: The NextGEM community page on Zenodo

A similar access methodology has been used for the remaining data services mentioned above, thus empowering NIKH to provide a unique and convenient “Literature Search” functionality, as depicted in Figure 3.

<sup>6</sup> <https://www.emf-portal.org/>

<sup>7</sup> <https://zenodo.org/>

<sup>8</sup> <https://pubmed.ncbi.nlm.nih.gov/>

<sup>9</sup> <https://www.webofscience.com/>

<sup>10</sup> <https://www.emf-health-cluster.eu>

<sup>11</sup> [https://zenodo.org/communities/seawave\\_data\\_managing](https://zenodo.org/communities/seawave_data_managing)

<sup>12</sup> <https://projectgoliat.eu/>

<sup>13</sup> <https://dataverse.csuc.cat/dataverse/goliat>

<sup>14</sup> <https://www.etaingroup.eu/>

<sup>15</sup> <https://www.uu.nl/en/research/yoda>

<sup>16</sup> <https://sketchfab.com/etaingroup/>

<sup>17</sup> [https://zenodo.org/communities/nextgem\\_project](https://zenodo.org/communities/nextgem_project)

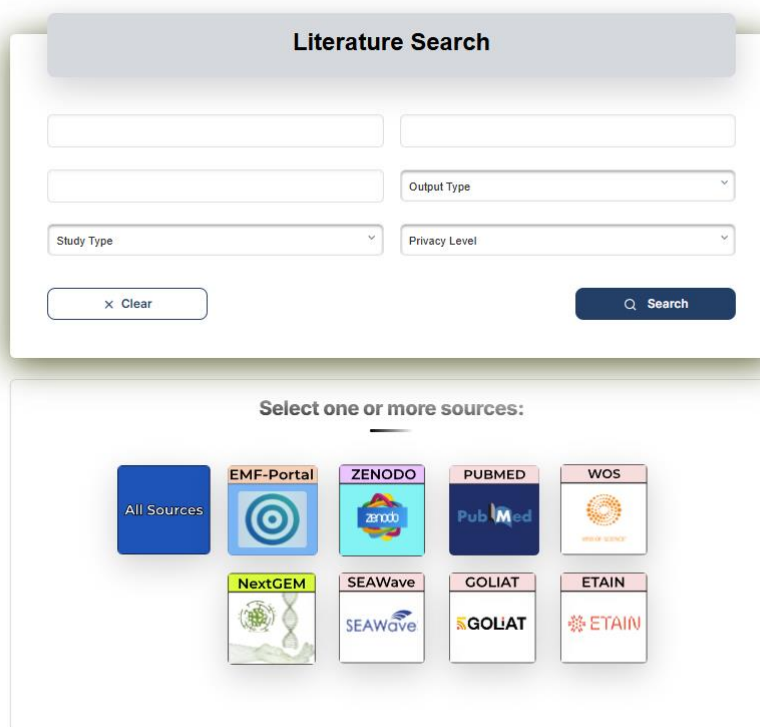


Figure 3: NIKH's Literature Search feature.

### 3.1.4 Infrastructure

- **Physical Infrastructure:** Comprises robust hardware selected for high-performance computing, storage, and security. Incorporates multiple layers of security measures such as firewalls, intrusion detection systems, and encryption technologies. To ensure reliable operation in a secure environment, the physical infrastructure is deployed at FORTH's main datacentre. This infrastructure is an essential part of FORTH and is actively maintained and expanded to support future needs beyond the scope and duration of any hosted projects.
- **Virtual Infrastructure:** Employs a hypervisor to manage multiple virtual machines (VMs), ensuring redundancy, incremental backups and built-in security features. In addition, VMs host various components, including the application portal, risk assessment tool and security assurance tool. NextGEM's virtual infrastructure is designed for long-term sustainability, with provisions for maintenance and potential expansion after the project's completion to support future research and development activities.

## 3.2 Assets, procedures and relations

Figure 4 depicts the components described in Section 3.1 and the interactions between them. In the context of trusted data management and security, these components can be viewed as assets whose operations and inter-communication are considered critical and must be closely monitored. Further, the building blocks of these components could also be viewed as assets, including but not limited to the software used to implement these components as well as the data that is handled from the operations of these components. Per-component lists of such identified assets of the NIKH architecture are available in Tables 2-7. The most important interactions between these components are the following:

- **Authentication Manager:** Interfaces with the Keycloak<sup>18</sup> software for user management and access control.
- **Metadata Manager:** Interfaces with the MongoDB database to store, retrieve, update and delete asset metadata.
- **Connector Manager:** Interfaces with Dataspace connectors for contract negotiation, policy management and asset retrieval.

<sup>18</sup> <https://www.keycloak.org/>



- Literature Review tool: Interfaces with Metadata Manager for data retrieval and Authentication Manager for access control.
- Risk Assessment tool: Interfaces with Metadata Manager for data retrieval and Authentication Manager for access control.
- ERMES tool: The ERMES tool comes with its own user interface that supports data importing from various sources, such as simulation outputs, experimental datasets, and third-party databases. In the case of NextGEM the source data will be retrieved from the NIKH Controller.
- GUI: Interfaces with all three components of the NIKH Controller for all user operations

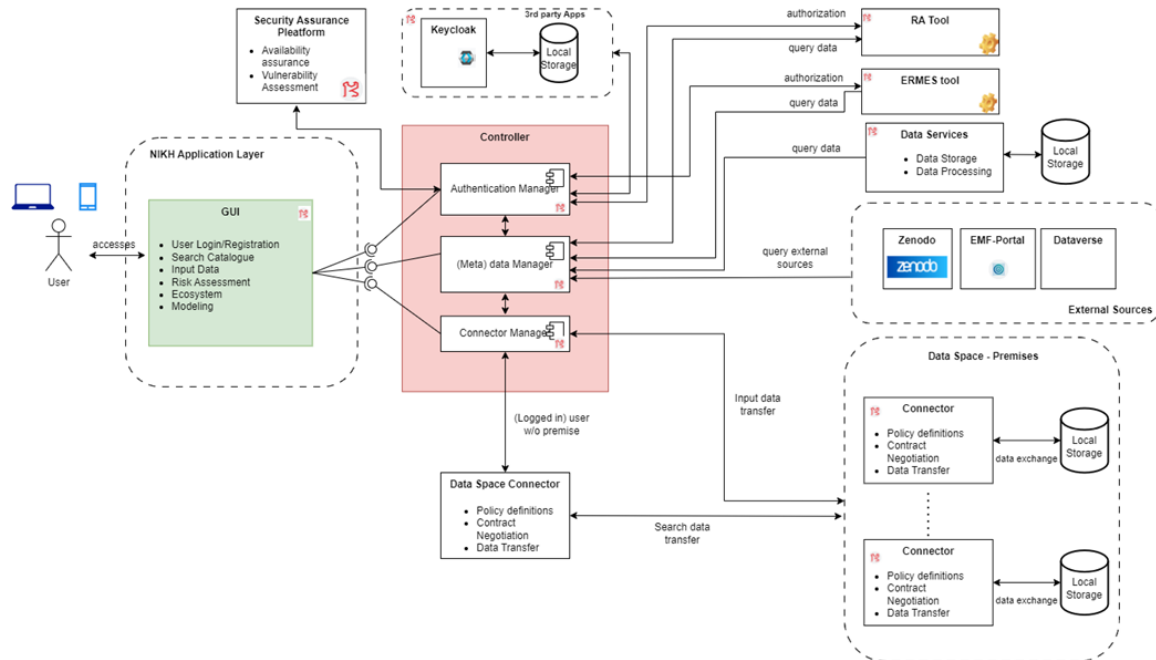


Figure 4: NIKH ecosystem component relations

Table 2: NIKH assets - Controller

Asset Name	Asset Type	Version	Dependencies	Description
<b>Authentication Manager</b>	Software	0.0.1	Keycloak	Uses Keycloak for user management and access control
<b>Metadata Manager</b>	Software	0.0.1	Mongo DB	Provides CRUD operations on data using the Mongo DB database
<b>Connector Manager</b>	Software	0.0.1	MongoDB, Eclipse Connectors	Handles data sharing via Dataspace Connectors
<b>Keycloak</b>	Software	23.0.2	PostgreSQL	Open-source Identity and Access Management (IAM) solution designed to modernize and secure applications
<b>MongoDB</b>	Software	7.0	-	Used for developing scalable applications with evolving data

				schemas of structured or unstructured data in JSON-like format
<b>Java (OpenJDK)</b>	Software	17.0.1	-	Popular programming language
<b>Kubernetes</b>	Software	1.25.4	-	Creates and manages containers on the cloud-based server systems
<b>Docker</b>	Software	25.0.2	-	Containerization platform and runtime for building and running containers
<b>PostgreSQL</b>	Software	16.24	-	Object-oriented relational database management system, used to define complex custom data types and communicate with the database servers using objects
<b>NGINX</b>	Software	1.25.4	-	Open-source web server software used for reverse proxy, load balancing, and caching
<b>Ubuntu</b>	Software	22.04	-	Operating system installed on all NIKH VMs
<b>Proxmox</b>	Software	8.4.1	-	Virtualization platform for NIKH Virtual Machines
<b>Connector information</b>	Data	N/A	PostgreSQL	Information stored in the NIKH controller database regarding connector location
<b>User information</b>	Data	N/A	Keycloak	Information stored in the keycloak database regarding NextGEM member details

Table 3: NIKH assets - Risk Assessment Tool

Asset Name	Asset Type	Version	Dependencies	Description
<b>Python</b>	Software	3.10.12	-	Popular programming language
<b>Ollama</b>	Software	0.1.28	-	Ollama offers a platform that enables running and customizing LLMs locally
<b>Playwright</b>	Software	1.40	Python	Playwright is an open-source automation Python library for browser testing and web scraping
<b>Next.js</b>	Software	14.1.0	Node.js	Framework that offers building blocks to create web applications and handles the tooling and configuration needed for React.js
<b>React.js</b>	Software	18	Next.js	JavaScript-based web UI development library
<b>Tailwind.css</b>	Software	3.4.1	-	Design system implementation in pure CSS
<b>Node.js</b>	Software	20.11	-	Used to create server-side web applications
<b>Ubuntu</b>	Software	22.04	-	Operating system installed on all NIKH VMs
<b>Proxmox</b>	Software	8.4.1	-	Virtualization platform for NIKH Virtual Machines

Table 4: NIKH assets - GUI

Asset Name	Asset Type	Version	Dependencies	Description
<b>MySQL</b>	Software	10.6.16	-	Relational database management system. It is free and open-source software under the terms of the GNU General Public License
<b>Apache</b>	Software	2.4.52	-	Free and open-source web server software developed by the Apache Software Foundation, known as the most widely used web server software on the Internet
<b>WordPress</b>	Software	6.8	-	WordPress is a content management system (CMS) that offers an online platform for building and hosting websites
<b>Ubuntu</b>	Software	22.04	-	Operating system installed on all NIKH VMs
<b>Proxmox</b>	Software	8.4.1	-	Virtualization platform for NIKH Virtual Machines

Table 5: NIKH assets - Connectors &amp; Data Spaces

Asset Name	Asset Type	Version	Dependencies	Description
<b>NIKH connector</b>	Software	0.01	Java JDK 11+, Gradle, Docker	Retrieves assets from partner's storages
<b>Ubuntu</b>	Software	22.04	-	Operating system installed on all NIKH VMs
<b>Proxmox</b>	Software	8.4.1	-	Virtualization platform for NIKH Virtual Machines

<b>Java (OpenJDK)</b>	Software	17.0.1	-	Popular programming language
<b>Kubernetes</b>	Software	1.25.4	-	Creates and manages containers on the cloud-based server systems
<b>Docker</b>	Software	25.0.2	-	Containerization platform and runtime for building and running containers
<b>PostgreSQL</b>	Software	16.24	-	Object-oriented relational database management system, used to define complex custom data types and communicate with the database servers using objects
<b>NGINX</b>	Software	1.25.4	-	Open-source web server software used for reverse proxy, load balancing, and caching
<b>NIKH metadata</b>	Data	N/A	-	Metadata defined in the NIKH platform to describe assets produced in the project
<b>Dataspace assets</b>	Data	N/A	NIKH connector	Files (datasets, images, videos) located in each organization's file storage and managed by dataspace connectors

Table 6: NIKH assets - Modelling tool

Asset Name	Asset Type	Version	Dependencies	Description
<b>Ubuntu</b>	Software	22.04	-	Operating system installed on all NIKH VMs
<b>Proxmox</b>	Software	8.4.1	-	Virtualization platform for NIKH Virtual Machines

Table 7: NIKH assets - SPA tool

Asset Name	Asset Type	Version	Dependencies	Description
<b>Ubuntu</b>	Software	22.04	-	Operating system installed on all NIKH VMs
<b>Java (OpenJDK)</b>	Software	17.0.1	-	Popular programming language
<b>Spring Boot</b>	Software	3.2.2	-	Java-based framework designed to simplify development of applications with minimal setup and configuration
<b>Jackson</b>	Software	2.11.4	-	Java library that provides core functionality for JSON parsing, generation, and manipulation
<b>Swagger</b>	Software	1.5.22	-	Framework that facilitates the generation, documentation, and consumption of RESTful APIs by providing annotations for describing API operations, models, and parameters in Java applications
<b>Micrometer</b>	Software	1.12.2	-	Java library that provides a simple facade over the instrumentation clients for various

				monitoring systems, allowing easy integration and consistent metrics collection within applications
<b>PostgreSQL</b>	Software	16.24	-	Object-oriented relational database management system, used to define complex custom data types and communicate with the database servers using objects
<b>PostgreSQL JDBC</b>	Software	42.6.0	PostgreSQL	Java library to connect to PostgreSQL
<b>Elasticsearch</b>	Software	8.13.0	-	Distributed, search and analytics engine, designed to provide real-time search, analysis of large volumes of data, and support for various use cases including text search, log analytics, and application monitoring
<b>jasypt</b>	Software	1.9.3	-	Java library which enables encryption in java apps with minimum effort
<b>Amqp</b>	Software	5.5.0	-	Allows Java applications to interface with RabbitMQ
<b>RabbitMQ</b>	Software	3.12.13	-	Open-source message-broker software that facilitates asynchronous messaging between distributed systems
<b>Keycloak</b>	Software	23.0.2	-	Open-source Identity and Access Management solution designed to

				modernize and secure applications
<b>KrakenD</b>	Software	2.6.0	-	Open-source API gateway and microservices aggregator designed to simplify the development, deployment, and management of complex distributed systems
<b>Apache Kafka</b>	Software	3.7.0	-	Apache Kafka is an open-source distributed event streaming platform used for building real-time data pipelines and streaming applications.
<b>Admin Credentials</b>	Data	N/A	Keycloak	Login credentials of the admin(s) which login to the SPA tool interface
<b>NIKH asset model</b>	Data	N/A	-	All the defined NIKH assets and their relations
<b>Assessment Results</b>	Data	N/A	-	Results generating from running assessments on the targeted assets
<b>Log information</b>	Data	N/A	-	Log and event information gathered by Log Shippers

Insights regarding the potential threats to these components can be found in Section 3.4.

### 3.3 Ensuring GDPR compliance in data management

NextGEM employs a policy of providing data as open as possible and as restricted as necessary, following rules and regulations to protect sensitive data while ensuring replicability of results. Open access to research data is ensured as per Horizon Europe guidelines, with embargo periods applied when necessary and restricted access justified based on confidentiality, security, privacy, or GDPR obligations.

NextGEM adheres to legal obligations, including the GDPR and national legislation, ensuring transparency in data management practices and compliance with ethical principles. Human participation in NextGEM is voluntary, with informed consent obtained from volunteers after clear communication of data sharing and preservation practices. Procedures involving human participation (such as sample collection from volunteers within the scope of Task 4.4) are submitted for approval to local ethics committees. Potential study participants receive an information sheet to ensure they are fully informed about the data collection process and can legally consent to participate, aligning with ethical standards. An ethics committee for NextGEM further ensures overall compliance with GDPR



principles. In studies involving humans, pseudonymization techniques preserve the anonymity of sensitive data, with special care taken to exclude such data from publicly shared research outputs. Only necessary data is collected, and raw datasets that cannot be shared due to constraints are synthesized into synthetic data to minimize risks.

The NIKH provides a standardized platform for the scientific community to store, evaluate, and access project outputs such as FAIR (Findable, Accessible, Interoperable, Reusable) data. As outlined in the NextGEM Data Management Plan, metadata is assigned to each research output and cataloged within the NIKH, which serves as a metadata catalogue to maximize the dissemination and impact of project results. The Data Space Connectors allow users to generate metadata records that are explicitly linked to usage policies, which are defined through contract specifications, as described in D6.7 (“Network provisioning and links with EU health data space – Final report”). These policies, mapped to NIKH’s framework by NextGEM members, include:

- **Public:** Actual data is fully accessible to participants without restrictions.
- **Public after embargo:** Data becomes publicly available after a specified embargo period, while metadata remains accessible.
- **Restricted access:** Data access is limited and may require negotiation, although metadata is publicly accessible.
- **Closed:** Both data and metadata are restricted to other participants.

Contract definitions can be negotiated with data space members requesting access, and upon mutual agreement, data is securely exchanged. Personal user data, however, is stored securely on local servers managed by participating organizations, with no metadata assigned. Encryption and access controls safeguard data integrity during storage and transfer, restricting access to authorized personnel only. Each organization in charge of sensitive data remains responsible for GDPR compliance, supported by a Data Protection Officer as stipulated in the ethics agreement, ensuring adherence to GDPR and ethical principles throughout the project lifecycle.

NextGEM has deployed Keycloak, an open-source identity and access management tool, to secure the NIKH platform. Keycloak provides Single-Sign-On, user role customization, and centralized authentication/authorization (D6.7). A dedicated “NextGEM” realm encapsulates all NIKH users, groups, and roles, with three levels defined: team member (search and view only), full member (create, search, view, edit), and supervisor (all permissions, including delete). Supervisors are manually assigned by the system administrator, while other roles are selected during registration and assigned via the Keycloak API upon approval. A registration request process ensures secure management of sensitive data by temporarily storing user input until an administrator approves or rejects it, creating a Keycloak account only upon acceptance. The user interface interacts with Keycloak for authentication and authorization.

Consequently, sensitive data from human studies is either pseudonymized or excluded from the NIKH entirely. Meanwhile, other sensitive information, specifically user credentials and restricted-access data, is protected within NIKH servers using encryption and access controls. This approach ensures GDPR-compliant management of sensitive data while supporting FAIR dissemination of pseudonymized and aggregated data through policy-driven access.

## 3.4 Potential threats

### 3.4.1 NIKH threat landscape overview

The NIKH is at the forefront of advancing research on EMF and health effects. Its role as a repository for potentially sensitive data can place it under the microscope of potential threats that could undermine its integrity, security, and compliance with regulations like the GDPR.

This analysis converges insights from comprehensive evaluations, providing a panoramic view of the threats that the NIKH might face, with a special focus on identifying and understanding these challenges.

#### Security and Vulnerability Threats:

1. **Data Breaches and Unauthorized Access:** NIKH's data, while mostly comprised of metadata, can potentially include proprietary research and intellectual property and, as such, is a prime target for cybercriminals. The threat extends from external hacking attempts to insider threats, where individuals with access might exploit their positions for unauthorized data access or theft. These breaches could devastate NIKH's reputation in the scientific community and legal standing under GDPR.
2. **Cyberattacks and System Disruptions:** The spectre of cyberattacks, including ransomware, phishing, and sophisticated exploits, poses a continuous threat to the NIKH. These attacks not only pose a risk to

data integrity but also threaten to disrupt NIKH's operational capabilities, potentially leading to significant downtime and hindering research progress.

3. **Insider Threats:** The human element within NIKH, whether through malice or negligence, represents a significant vulnerability. Privileged users could potentially abuse their access, leading to data leaks, operational sabotage, or fraudulent activities.

#### Operational and General Threats:

1. **Infrastructure Failures and Natural Disasters:** The physical and technical infrastructure supporting NIKH is susceptible to failures from both technological malfunctions and natural disasters. Such events can lead to service interruptions, data loss, and hinder the hub's overall reliability.
2. **Sustainability and Funding:** The long-term viability of the NIKH hinges on securing sustainable funding sources. A dependence on fluctuating funding streams can introduce operational vulnerabilities and impact the hub's ability to adapt and grow.
3. **Technological Evolution and Standards Compliance:** As digital technologies and data standards evolve, the NIKH must remain agile, ensuring its systems and practices are up-to-date. Failure to adapt could result in compatibility issues, impeding data exchange and collaboration.

#### GDPR Compliance and Privacy Threats:

1. **Data Protection and Privacy Violations:** In the realm of GDPR, the NIKH must navigate the tightrope of data protection, ensuring that stringent measures are in place to safeguard personal information. Any lapse in data protection practices could lead to erosion of stakeholder trust and even significant regulatory fines.
2. **Data Misuse and Integrity Challenges:** The misuse or unauthorized alteration of data within the NIKH poses a significant risk not only to privacy but also to the integrity of the research it supports. Ensuring data accuracy and preventing unauthorized access are paramount in maintaining the credibility of NIKH's outputs.
3. **Inadequate Security Measures:** The comprehensive nature of GDPR compliance underscores the need for the NIKH to maintain robust security measures. Weaknesses in these measures could expose the hub to data breaches, compromising the privacy and security of the information it holds.

### 3.4.2 NIKH components' potential threats and vulnerabilities

In this section, the vulnerabilities and threats, as applicable to each NIKH component as depicted in Figure 1, are analyzed and elaborated upon.

#### **3.4.2.1 Infrastructure layer**

The Infrastructure Layer consists of several components, as depicted in Figure 1. Each component will be individually analyzed, elaborating upon the threats and vulnerabilities associated with it.

##### **Physical infrastructure**

1. Servers:
  - **Hardware Failures:** Malfunction or failure of server hardware components, leading to service disruptions or data loss.
  - **Overheating:** Servers overheating due to inadequate cooling systems or environmental controls, causing hardware damage or system downtime.
2. Networking Equipment:
  - **Network Congestion:** High network traffic causing delays in data transmission or network outages.
  - **Network Security Flaws:** Exploitation of vulnerabilities in networking equipment firmware or configuration errors, leading to unauthorized access or Denial of Service (DoS) attacks.
3. Storage Devices:
  - **Data Corruption:** Malfunction or corruption of storage devices, leading to data loss or integrity issues.
  - **Data Theft:** Unauthorized access or data leakage due to weak access controls or improper data handling.
4. Power Infrastructure:
  - **Power Outages:** Disruptions causing unplanned downtime and potential data loss.

- Power Surges or Spikes: Voltage fluctuations damaging hardware components or causing data corruption.
- 5. Cooling Systems:
  - Cooling System Failure: Malfunction leading to overheating of server rooms or data centers.
  - Environmental Factors: Compromising the effectiveness of cooling systems and posing risks to infrastructure integrity.
- 6. Physical Security Measures:
  - Unauthorized Access: Weak physical access controls allowing unauthorized individuals to gain physical access.
  - Insider Threats: Trusted individuals with physical access misusing their privileges.

### Virtual infrastructure

1. Virtual Servers:
  - Hypervisor Vulnerabilities: Unauthorized access to virtual machines or compromise of VM isolation.
  - Resource Exhaustion: Performance degradation or service interruptions due to DoS attacks or resource-intensive workloads.
2. Virtual Networks:
  - Network Segmentation Issues: Unauthorized access or lateral movement between network segments due to inadequate segmentation or misconfiguration.
  - Network Traffic Interception: Man-in-the-middle attacks compromising the confidentiality and integrity of data.
3. Virtual Storage:
  - Data Leakage: Insecure configurations or weak access controls resulting in data breaches or compliance violations.
  - Data Corruption: Actions or bugs affecting virtual storage systems leading to data loss or integrity impact.
4. Virtualization Management Platform:
  - Management Interface Vulnerabilities: Unauthorized access, manipulation of virtual infrastructure settings, or disruption of virtualized environments.
  - Misconfigurations: Security gaps or weakened overall security posture due to improperly configured platforms or policies.

### Network infrastructure

1. Routers and Switches:
  - Firmware Vulnerabilities: Unauthorized access, DoS attacks, or interception of network traffic.
  - Misconfigurations: Insecure routing or forwarding rules enabling attackers to disrupt network operations.
2. Firewalls:
  - Inadequate Rule Management: Resulting in overly permissive or restrictive access policies.
  - Firewall Bypass Techniques: Evasion techniques or weaknesses exploited to bypass protections.
3. Intrusion Detection and Prevention Systems (IDPS):
  - Signature-based Detection Limitations: Evasion of known attack patterns or variants.
  - False Positives/Negatives: Inaccurate detection reducing the effectiveness of intrusion detection.
4. Virtual Private Networks (VPNs):
  - Encryption Weaknesses: Weak encryption algorithms exposing VPN communications to eavesdropping.
  - VPN Client Vulnerabilities: Exploited vulnerabilities leading to unauthorized access or malicious injections.
5. Domain Name System (DNS) Infrastructure:
  - DNS Spoofing and Cache Poisoning: Redirecting legitimate DNS queries to malicious domains.
  - DNS Tunneling: Bypassing network security controls, covertly exfiltrating data.

### 3.4.2.2 Service layer

#### Controller

NIKH's "Controller" component consists of sub-components "Authentication Service", "Metadata Service" and "Connector Service". The key potential vulnerabilities and threats associated with these sub-components are listed.

1. Authentication Service:
  - Weak Authentication Mechanisms: Unauthorized access due to weak or compromised authentication methods.
  - Credential Theft: Stealing user credentials through phishing attacks or malware.
2. Metadata Service:
  - Insecure Metadata Storage: Unauthorized access due to insecure storage practices.
  - Metadata Tampering: Unauthorized modification leading to integrity issues.
3. Connector Service:
  - Insecure Data Transmission: Data intercepted by man-in-the-middle attacks.
  - Data Leakage: Unintended data exposure due to improper configuration.

#### Common Vulnerabilities across Controller Sub-components:

- Lack of Encryption: Data transmitted may be intercepted or tampered with, leading to exposure.
- Insider Threats: Authorized users with malicious intent misusing privileges.

#### Data space services

1. Broker:
  - Insecure Authentication and Authorization Mechanisms: Unauthorized access to sensitive data or services.
  - Lack of Data Encryption: Data breaches or unauthorized access due to intercepted data.
  - Insufficient Input Validation: Injection attacks targeting interfaces.
2. Identity Manager:
  - Weak Password Policies: Brute-force attacks or password guessing.
  - Insecure Credential Storage: Unauthorized access due to inadequate encryption or storage practices.
  - Lack of Multi-factor Authentication: Unauthorized access via stolen or compromised credentials.
3. Data Space Connector:
  - Insecure Data Transmission: Man-in-the-middle attacks intercepting data.
  - Insufficient Access Controls: Unauthorized access to data or resources.
  - Lack of Data Validation: Injection attacks targeting interfaces.

#### Common Vulnerabilities across Data Space Services Sub-components:

- Lack of Patch Management: System compromise or unauthorized access due to unpatched software components.
- Insider Threats: Authorized users with malicious intent causing harm.
- Lack of Logging and Monitoring: Failure to detect and respond to security incidents in a timely manner.

#### Data services

1. Data Lake:
  - Insecure Data Storage: Unauthorized access due to inadequate access controls or encryption.
  - Lack of Data Validation: Injection attacks leading to data corruption.
  - Data Leakage: Unintentional exposure of sensitive data.
2. Data Processing:
  - Software Vulnerabilities: System compromise or unauthorized access.
  - Data Integrity Issues: Unauthorized modification leading to inaccurate results.
  - Resource Exhaustion: Service disruption or degradation due to DoS attacks.
3. Data Governance and Compliance:
  - Non-Compliance with Data Regulations: Legal and financial penalties.
  - Lack of Data Governance: Data misuse or unauthorized access due to inadequate practices.

### Common Vulnerabilities across Data Services Sub-components:

- Insider Threats: Authorized users with malicious intent causing harm during data processing operations.
- Lack of Monitoring and Logging: Failure to detect and respond to security incidents during data processing activities.
- Insecure Data Transmission: Data interception or tampering during exchanges between data processing components.

### **Security and privacy**

NIKH's "Security and Privacy" component consists of sub-components "Identity and Access Management" and "Security Assurance Tool". The key potential vulnerabilities and threats associated with these sub-components are listed.

1. Identity and Access Management:
  - Weak Authentication Mechanisms: Unauthorized access due to compromised methods.
  - Insider Threats: Misuse of access by authorized users.
  - Inadequate Access Controls: Unauthorized access to sensitive data.
2. Security Assurance Tool:
  - Insecure Configuration: Security gaps due to misconfiguration.
  - Data Leakage: Exposure of confidential information due to improper handling.
  - Lack of Continuous Monitoring: Delayed detection of cyber threats or vulnerabilities.

### **3.4.2.3 Application layer**

#### **User Interface (UI)**

- 1) Application Portal:
  - Insecure Authentication Mechanisms: Unauthorized access compromising user accounts and sensitive data.
  - Session Management Flaws: Attackers impersonating legitimate users or accessing their sessions.
- 2) Document Management:
  - Insecure File Uploads: System exposure to malicious file uploads.
  - Insufficient Access Controls: Unauthorized users accessing sensitive information.
- 3) Forms and Data Input:
  - Cross-Site Scripting (XSS): Injecting malicious scripts into web pages.
  - SQL Injection: Manipulating database queries to extract sensitive information.
- 4) Visualization Tools:
  - Client-Side Security Risks: Exploitation leading to arbitrary code execution or credential theft.
  - Data Privacy Concerns: Unintentional disclosure of confidential information.

### Common Vulnerabilities across UI Sub-components:

- Lack of Input Validation: Opening the system to various injection attacks.
- Insecure Configuration Settings: Introducing security weaknesses or exposing sensitive information.

### **Mobile application**

- 1) Insecure Data Storage:
  - Lack of Data Encryption: Exposing sensitive data to unauthorized access.
  - Insufficient Data Obfuscation: Allowing attackers to manipulate data.
- 2) Insecure Communication:
  - Insufficient Transport Layer Security: Exposing sensitive data to interception.
  - Man-in-the-Middle (MitM) Attacks: Intercepting and modifying communication.
- 3) Weak Authentication and Authorization:
  - Insecure Authentication Mechanisms: Compromising user accounts.
  - Insufficient Session Management: Hijacking active user sessions.
- 4) Code-Level Vulnerabilities:
  - Insecure Data Storage: Exposing sensitive data through reverse engineering.
  - Insecure Code Practices: Introducing exploitable vulnerabilities.

5) Lack of Secure Offline Functionality:

- Offline Data Storage Risks: Exposing data to unauthorized access.
- Insecure Offline Caching: Leading to data leakage or unauthorized access.

**Modelling tool**1) Software Vulnerabilities:

- Input Validation Issues: Allowing manipulation of input data and execution of arbitrary code.
- Memory Corruption Vulnerabilities: Potentially leading to denial of service or remote code execution attacks.

2) Data Security Risks:

- Insecure Data Storage: Exposing modeling data to unauthorized access or tampering.
- Data Leakage: Unintentional data leakage violating data privacy regulations.

3) Exploitation of Mathematical Model:

- Model Manipulation: Generating inaccurate or misleading modeling results.
- Model Poisoning: Generating biased or compromised modeling results.

4) Denial of Service (DoS) Attacks:

- Resource Exhaustion: Overwhelming the component leading to service quality degradation.
- Algorithmic Complexity Attacks: Triggering excessive resource consumption.

5) Insider Threats:

- Unauthorized Access: Misuse of privileges to tamper with modeling data or steal intellectual property.

**Risk assessment tool and literature review tool**1) Software Vulnerabilities:

- Lack of Patch Management: Leaving the tool vulnerable to security exploits.
- Insecure Configuration: Exposing the tool to unauthorized access or data leakage.

2) Data Security Risks:

- Inadequate Data Protection Measures: Resulting in data breaches, privacy violations, or regulatory non-compliance.
- Data Leakage: Exposing sensitive information to unauthorized access or disclosure.

3) Integrity Risks:

- Manipulation of Assessment Data: Leading to inaccurate or biased risk assessment outcomes.
- Data Falsification: Resulting in misleading risk characterization.

4) Denial of Service (DoS) Attacks:

- Service Disruption: Degrading the availability and performance of the tool.
- Resource Exhaustion: Degradation of service quality for legitimate users.

5) Insider Threats:

- Insider Misuse or Sabotage: Manipulating data, compromising integrity, or sabotaging functionality.



## 4 The Security and Privacy Assurance tool

The Security and Privacy Assurance (SPA) tool<sup>19</sup> is an integrated suite of services and components that provide comprehensive cyber security risk detection and management for enterprise systems. This tool provides the necessary state-of-the-art infrastructure to assess the security aspects of the NIKH hub effectively.

### 4.1 Overview

The tool boasts a diverse array of capabilities designed to cater to a broad spectrum of use cases. Among this extensive repertoire, the functionalities most pertinent and beneficial to the objectives of the NextGEM project include:

- Comprehensive client asset modelling and automated client asset discovery
- Automated threats and vulnerabilities detection
- Sophisticated event processing capabilities and continuous runtime monitoring (SIEM)
- Penetration testing and support for ingestion of penetration testing reports using third-party tools

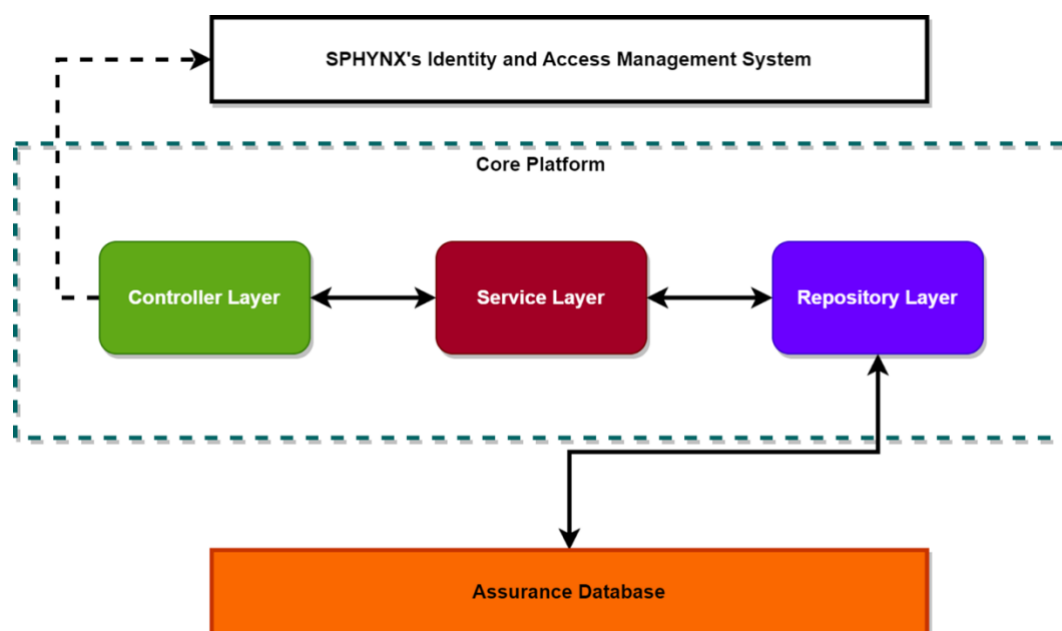


Figure 5: High-level SPA tool diagram

Figure 5 provides a high-level diagram of the core SPA tool, showing that it follows a controller, service and repository pattern. This pattern is a standard software architectural pattern used in building applications, particularly in web development. It helps separate concerns and improve code organization and maintainability. Here is a brief description of each component:

1. **Controller:** The controller is responsible for handling incoming requests, interpreting user input, and coordinating the overall flow of the application. It interacts with the external world (e.g., a web browser) and acts as an entry point for the application. Controllers receive requests, perform any necessary data validation or transformation, and delegate the processing to the appropriate service layer.
2. **Service:** The service layer contains the business logic and application-specific operations. It encapsulates the core functionality of the application and handles the processing and manipulation of data. Services are typically invoked by the controllers and can perform complex operations by leveraging one or more repositories. They provide a higher level of abstraction and reusability compared to directly interacting with the repositories.
3. **Repository:** The repository acts as an interface to the underlying data storage, such as a database or external APIs. It abstracts away the data access implementation details and provides a consistent interface

<sup>19</sup> <https://www.sphynx.ch/sphynx-products/>

for performing CRUD operations on the data. Repositories handle data persistence and retrieval, and they are typically responsible for querying, filtering, and transforming data based on specific requirements.

Using the controller, service, and repository pattern, the application separates concerns, with the controller handling the request/response flow, the service containing the business logic, and the repository managing the data access layer. This separation improves code maintainability and testability and allows for easier modification or extension of individual components without affecting the others.

The term used to refer to the targeted system by the SPA tool is the Asset Model. The Asset Model is an entity that is comprised of various assets and the relationships between them. In essence, it is a modeling of the target system. For the NextGEM case the Asset Model is the NIKH hub, while the assets that it consists of are the various hardware, services, processes, data and the software “building blocks” of the NIKH. Security assurance of the Asset Model is achieved via the collection of assessment criteria which evaluate and monitor the various assets defined within the Asset Model and can either be satisfied or violated.

Below is a list of possible assets that can be defined to model a targeted system:

- **Data:** In the context of a web system, a data asset refers to any type of data or information that is collected, processed, stored, and utilized by the system. It represents the valuable and often proprietary data that is integral to the functioning, operations and value proposition of the web system, such as customer information, financial data, intellectual property, and sensitive business data.
- **Hardware:** A hardware asset refers to a physical component or device that is utilized within an organization’s information technology (IT) infrastructure. It includes tangible items such as computers, servers, laptops, networking equipment, storage devices, printers, scanners and other peripherals. Hardware assets are the physical resources that enable the execution of software applications and the storage, processing and communication of data.
- **Network:** A network asset refers to the logical network of an organization and not the infrastructure or the hardware (e.g., router, switch). It encompasses the software resources used for communication, connectivity, and data transfer within a network environment. Network assets are critical for establishing and maintaining network connectivity, enabling data transmission, and facilitating communication between devices.
- **Process:** A process asset refers to a series of interrelated activities or steps that are designed to achieve a specific goal or outcome. Processes are systematic and structured ways of performing tasks, managing workflows, and delivering products or services within an organization. They provide a framework for coordinating and organizing work, ensuring consistency, efficiency and quality in the execution of tasks.
- **Software:** A software asset refers to any piece of software that that is owned, licensed, or utilized by an organization. It includes computer programs, applications, operating systems, libraries, scripts, plugins, and other software components that are used to support business operations, provide services, or enable specific functionalities.

The collection of services and components implemented and used in the tool, communicate under the hood in order to provide users (admins) with security insights regarding the targeted infrastructure (Figure 6).



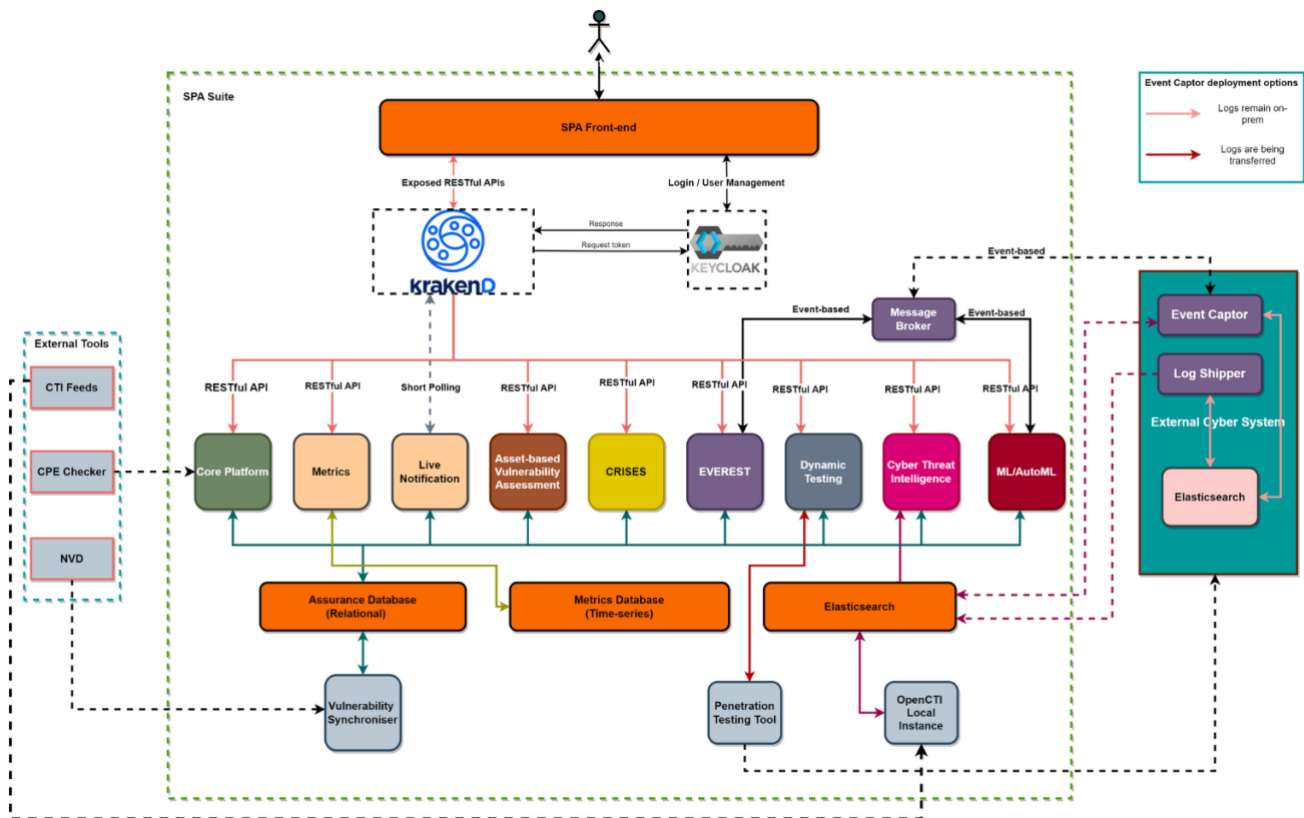


Figure 6: Overview of the SPA tool architecture

The core tool modules are:

- The **Cyber System Asset Loader**: the component responsible for receiving the Security Assurance Model for the target organization. This model includes the assets of the organization, security properties for these assets, threats that may violate these properties, and the security controls that protect the assets.
- The **Vulnerability Analyzer**: is responsible for identifying known vulnerabilities of assets defined within an organizations' asset model. This component automatically constructs the Common Platform Enumeration (CPE<sup>20</sup>, a structured naming scheme for information technology systems, software, and packages) per asset and then retrieves the relevant Common Vulnerabilities and Exposures (CVE<sup>21</sup>, a reference-method for defining unique, common identifiers for publicly known information-security vulnerabilities and exposures) entries, by searching in a local copy of the National Vulnerability Database (NVD<sup>22</sup>, a U.S. government repository of standards-based vulnerability management data, maintained by the National Institute of Standards and Technology, NIST<sup>23</sup>). This copy is continuously updated by utilizing an in-house component that fetches the latest known CVEs from NVD's JSON files (see Section 4.1.1).
- **Event Reasoning Toolkit (EVEREST)**: A runtime monitoring engine, built in Java, which offers an API for establishing the monitoring rules to be checked (see Section 4.1.3).
- **Event Captors**: The Event Captor is a tool, implemented in Java, which formulates a rule or a set of rules based on collected data and triggering events and then pushes them towards the monitoring module for evaluation (see Section 4.1.4).
- **Dynamic Tester**: the component responsible for initiating the testing assessment. This module consists of two components: (a) the dynamic tester or manager and (b) the dynamic testing tool (see Section 4.1.2).
- **Security Component**: is comprised of two main components:

<sup>20</sup> <https://nvd.nist.gov/products/cpe>

<sup>21</sup> <https://cve.mitre.org/>

<sup>22</sup> <https://nvd.nist.gov/>

<sup>23</sup> <https://www.nist.gov/>

- **Keycloak:** which is used for user and client authentication. Keycloak<sup>24</sup> is an open-source identity and access management platform that provides a single point of access for modern applications, APIs, and microservices. It offers features such as single sign-on, identity brokering, and social login, as well as robust user management and authentication capabilities.
- **KrakenD:** an API Gateway that handles authorization in conjunction with Keycloak. KrakenD<sup>25</sup> is a lightweight, high-performance API Gateway that helps expose internal and external microservices to the world while keeping the complexity and routing logic out of core services. It allows to easily build and deploy API Gateway configurations using a simple, declarative configuration file and provides features such as rate limiting, circuit breaking, and caching to help manage the performance and reliability of your APIs.

Moreover, the operation of the tool relies on the following databases:

- **EVEREST Database:** Holds the monitoring rules and the overall process done by EVEREST if the templates and other values are important for the monitoring procedure to conclude.
- **Security Assurance Database:** Holds the cyber system asset model, its components, and the results of the assessments.
- **Vulnerabilities Database:** Holds the known vulnerabilities, as retrieved from the NVD database.

Finally, the Assurance tool interfaces with the RabbitMQ<sup>26</sup> message broker, which is a messaging bus that allows communication between external components of the core tool using the AMPQ<sup>27</sup> protocol.

#### 4.1.1 Asset-based vulnerability assessments

The Asset-Based Vulnerability Assessment is responsible for identifying known vulnerabilities of the assets defined within the asset model. The main result of such a vulnerability assessment is the risk probability of an identified vulnerability to be exploited.

Two main components are responsible for orchestrating an Asset-Based Vulnerability Assessment:

- **Vulnerability synchronizer:** used to retrieve and store all new and updated Common Vulnerabilities and Exposures (CVEs).
- **Vulnerability analyzer.** This tool is responsible for running the assessment, which uses the CVEs retrieved from the synchronizer.

The **Vulnerability synchronizer** is the tool that fetches the new/modified CVEs. It runs on a bi-hourly interval but can also be run on request, making it possible to sync whenever a user wants to. It uses polling to fetch the new/modified CVEs, by using the NIST API<sup>28</sup>.

The **Vulnerability analyzer** is the tool running the assessments. It maps assets with possible vulnerabilities and weaknesses. It follows a standardized assessment flow and expects the assets as input.

Figure 7 presents a high-level diagram of both the synchronizer and the analyzer tools in communication with the assurance database tool.

- The **vulnerability synchronizer** communicates with the external NVD API and saves the fetched CVEs in the Assurance Database.
- The **Vulnerability analyzer**, which is initiated from the SPA GUI, retrieves the CVEs affecting the targeted assets and saves the results in the Assurance DB.

Internally, both tools follow the Controller, Service, and Repository layer pattern. Each request is handled in the controller layer, processed in the service layer and saved/retrieved using the repository layer.

<sup>24</sup> <https://www.keycloak.org/>

<sup>25</sup> <https://www.krakend.io/>

<sup>26</sup> [www.rabbitmq.com](http://www.rabbitmq.com)

<sup>27</sup> <https://www.rabbitmq.com/tutorials/amqp-concepts>

<sup>28</sup> <https://nvd.nist.gov/developers/vulnerabilities>

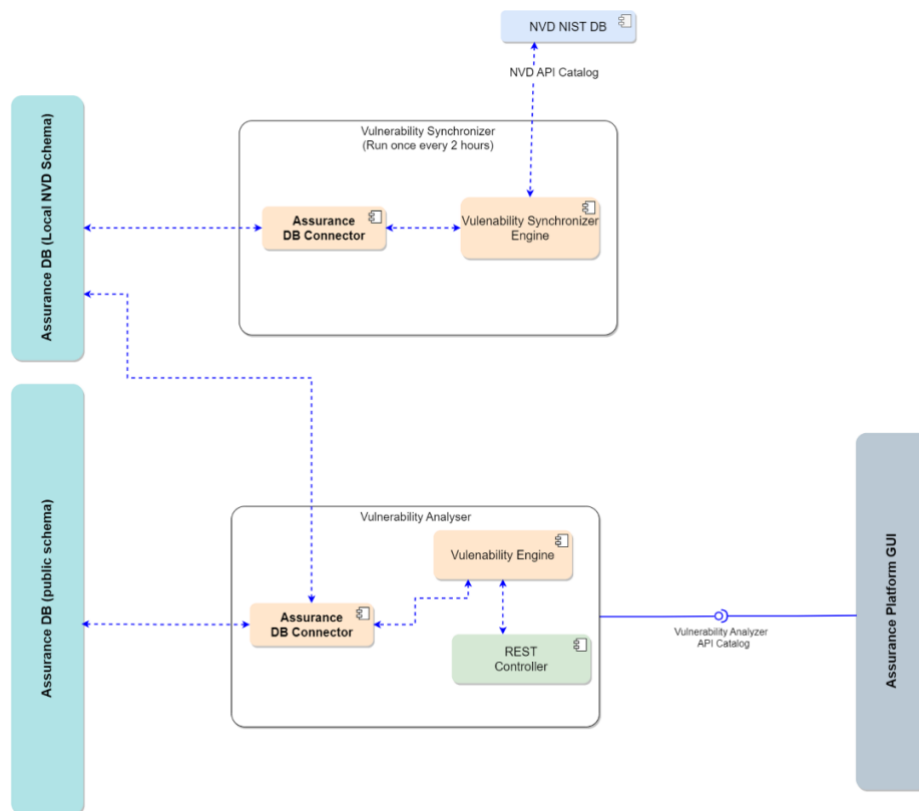


Figure 7: Analyzer and synchronizer communication.

### 4.1.2 Dynamic testing

The Dynamic Testing tool, integrated within the SPA tool, facilitates the import of scan results from security tools, currently supporting Greenbone (OpenVAS<sup>29</sup>). This functionality enhances the management of penetration testing outcomes during engagements by providing security analysts with structured and actionable insights.

Key features of the Dynamic Testing tool include:

- Extraction of security properties (Confidentiality, Integrity, Availability) from each finding
- Identification of severity levels for each finding
- Automatic creation of new assets based on detected findings
- Options for users to adjust the severity of a finding with a documented justification
- Options for users to designate a finding as a False Positive with a documented justification
- Generation of detailed metrics for individual imported reports
- Compilation of aggregated metrics from multiple imported reports
- Enrichment of findings with additional context, such as retrieving detailed information (e.g., Common Weakness Enumeration (CWE) name, related CVEs, and associated security properties) when a CWE is identified in a report

This tool empowers security analysts with efficient result management and enriched data, supporting more informed decision-making during security assessments.

### 4.1.3 Event reasoning toolkit

The Event Reasoning Toolkit (EVEREST) is the monitoring tool of the SPA tool, designed to address the evolving challenges of cybersecurity and data protection in contemporary cyber systems. EVEREST provides the necessary guarantees and measures to ensure data sustainability and protect against cyber-attacks by alerting for violations of cybersecurity space, in a continuous runtime manner.

<sup>29</sup> <https://www.openvas.org/>

The EVEREST Assessment Tool is a Dockerized <sup>30</sup> solution that executes assessments based on rules written in Event Calculus. It serves a multifaceted purpose, encompassing the creation of assessment criteria, running assessments on defined assets, and the generation of comprehensive assessment results which denote possible satisfactions or violations on the assessment criteria.

Organizations collecting and handling sensitive information in today's data-driven world must adhere to stringent data protection regulations and security standards.

Key features of the EVEREST Tool:

1. **Continuous Monitoring Assessments:** EVEREST conducts thorough continuous runtime monitoring assessments to check for violations of security and dependability properties in a system. It is used to ensure that all security solutions in place are functioning correctly and efficiently.
2. **Event Stream Analysis:** The tool excels in complex event recognition and reasoning, tracking and analysing streams of events to detect patterns of special significance. Event streams from various sources, including mainly Elasticsearch stack log aggregators, sensors, computer networks, and different log files are processed with exceptional accuracy.
3. **Logical Reasoning System:** EVEREST is based on Drools, a logical reasoning system that utilizes the Event Calculus formalism. It employs Business Rules Management Language for logical operations related to security policy assessment, ensuring accurate and effective event recognition.
4. **Interconnection with Other Tools:** EVEREST seamlessly integrates with a wide range of other cybersecurity tools within the SPA tool. It enables evaluation of assessment results with additional layers of analysis and facilitates the initiation of mitigation processes. With EVEREST's flexible interconnection capabilities, organizations can enhance their cybersecurity ecosystem and optimize their defence strategies.

#### 4.1.4 Event captors

An Event Captor is a tool that aggregates log and event information from the targeted infrastructure based on specifications set by EVEREST and encapsulates that information in a specific format that can be consumed by the EVEREST model. The collection of data and events mainly occurs through the Elastic Stack<sup>31</sup> (ELK). The ELK is a group of products that can reliably and securely take data from any source, in any format, and then search, analyze, and visualize it in real-time. From this stack, **Elasticsearch** and **Beats** are used. **Elasticsearch** is a distributed, RESTful search engine designed for working with large volumes of data. **Beats**, on the other hand, are lightweight data shippers that are used to collect different types of operational data. They can be installed on servers, containers, or even edge devices to gather information about various aspects of system and application performance. Beats monitor, collect, and send data to Elasticsearch for further analysis.

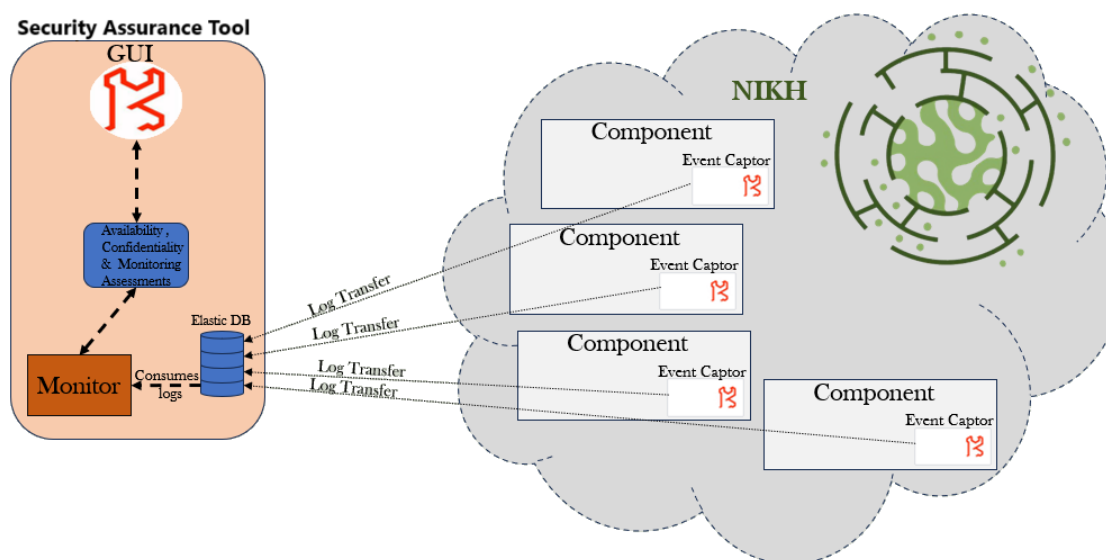


Figure 8: Shipping NIKH logs to the SPA tool.

<sup>30</sup> <https://www.docker.com/>

<sup>31</sup> <https://www.elastic.co/>

The Event captor tool is activated from the monitor module. This tool is crucial for EVEREST functionality since, through Elastic Beats, it collects essential asset information as events and supplies them to be stored in a dedicated database.

As shown in Figure 8, the Event captors, which are activated inside NIKH components, “ship” log and event information to be stored to the Elastic Database. This information is subsequently consumed by the Monitor module so that monitoring assessments can be evaluated.

Figure 9 presents a small diagram, where the sequence and interactions between Security Assurance Core tool, the Monitor tool, and the Event Captors are depicted.

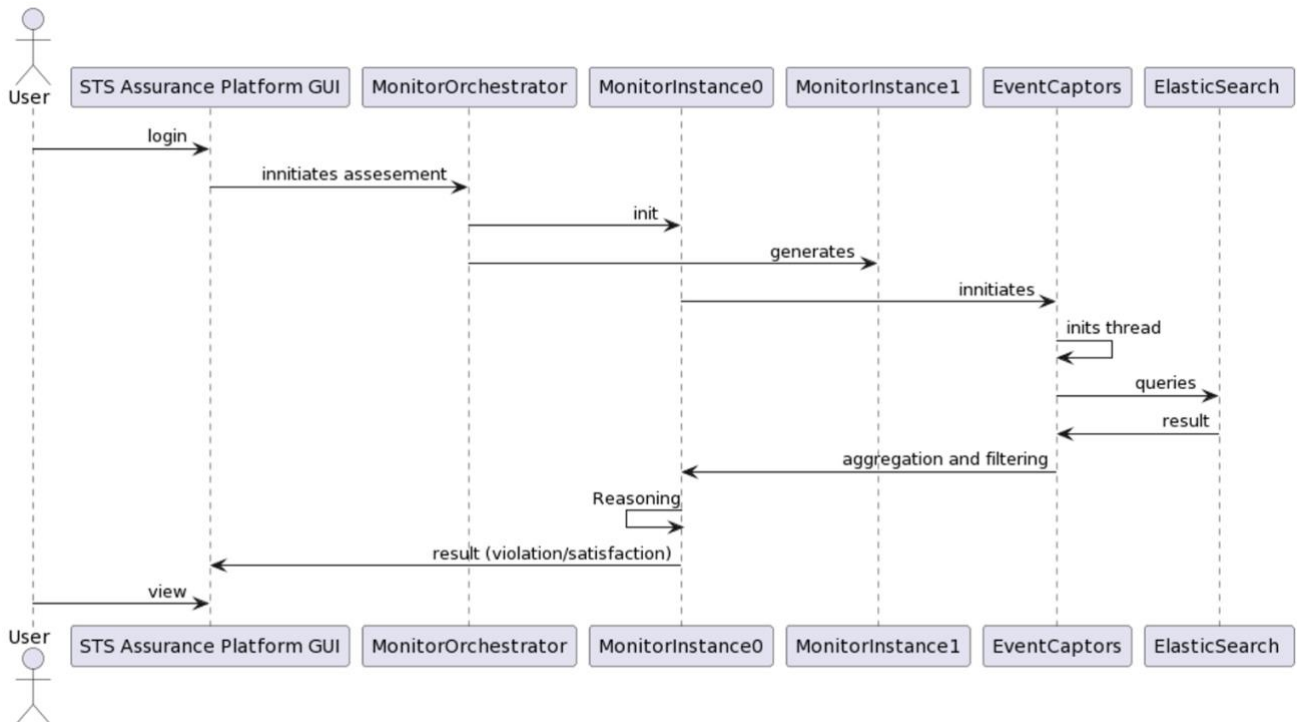


Figure 9: Interaction between SPA tool, Monitor and Event Captors.

In this diagram, after logging in, the user with access to the SPA tool, initiates an assessment. This action creates two Monitor instances. The first (Monitorinstance0) that is created during initialization of monitor which interacts with the Event Captors and the second (Monitorinstance1) that is instantiated and waits for the next assessment. Each assessment corresponds to an individual threat, and the Event Captors that work for the specific assessment assigned. The captors gather the necessary information required by the assessment via the ELK stack and provide this information back to the monitoring instance to be assessed for possible violations or satisfactions. Finally, the results are shown to the user via the SPA tool GUI.

## 4.2 Tool deployment and association with NIKH

### 4.2.1 Core deployment

To guarantee uninterrupted access to the SPA tool, even in scenarios where specific NIKH components may become unresponsive, the SPA tool has been strategically implemented within a dedicated Virtual Machine (VM). Within this VM, essential components such as the core tool, the administrative Graphical User Interface (GUI), and the databases responsible for storing assets, assessment criteria, and results are seamlessly integrated. Additionally, the Elastic Database plays a pivotal role in housing targeted component information, while Event Captors, embedded within each targeted component, facilitate ongoing and continuous monitoring. Further details on these aspects will be expounded upon in the subsequent sections.

#### 4.2.1.1 The ELK stack

To enable the SPA tool to access log and event information generated within targeted NIKH components, an elastic database on the VM that houses the SPA tool. This database serves as a repository for storing such information. The monitoring SPA component accesses this database to consume and assess the stored entries. Additionally, event shippers were instantiated on targeted NIKH components. These shippers collect and transmit

event and log information to the elastic database for future utilization. Figure 8 illustrates the interaction between these shippers and the core SPA tool.

#### 4.2.2 Deployed captors on NIKH components

Within the scope of NextGEM there are two kinds of event captors deployed to monitor NIKH components, the Availability Event Captor and the Internet Protocol (IP) Confidentiality Event Captor.

The **Availability Event Captor** is the only Event Captor that belongs to the native captors (i.e. it does not need information from Elasticsearch and Beats to gather data). It is based on ping/curling the respective asset (by IP or URL) to observe if the service is up or down.

The **IP Confidentiality Captor** is based on the Filebeat<sup>32</sup>. It gathers relevant information regarding Secure Shell (SSH) logins that occur on the targeted component. By whitelisting a set of allowed IP addresses, notifications can be provided when an unlisted IP address attempts to connect to the targeted asset through SSH.

Table 8 shows the NIKH components on which the event captors have been enabled. The table shows which components already have deployed captors, as well as which components are scheduled to have captors deployed on them in the future.

Table 8: Deployment Status of Event Captors on NIKH components

NIKH Component	Deployed Captors	
	IP Confidentiality	Availability
<b>SPA tool</b>	Yes	N.A.
<b>Controller</b>	Yes	Yes
<b>GUI</b>	Yes	Yes
<b>RA tool</b>	Yes	Yes (*)
<b>Literature Review Tool</b>	Yes	Yes (*)
<b>Modelling Tool</b>	Yes	Yes (*)
<b>3<sup>rd</sup> Party DBs</b>	N.A.	Yes (**)
<b>Local Storage</b>	Yes (**)	Yes (***)
<b>Data accessed via Connectors</b>	N.A.	N.A.

N.A. = Not Applicable

(\*) Achieved through monitoring the VM that houses these tools.

(\*\*) 3<sup>rd</sup> party DBs do not have a dedicated captor to assess their availability. Their availability is assessed directly since they are accessible via the public web.

(\*\*\*) Achieved through monitoring of the VM that houses local storage services.

<sup>32</sup> <https://www.elastic.co/beats/filebeat>



## 5 Vulnerability and security assurance

This section presents a comprehensive overview of all Security Assessments conducted or scheduled to run through the Security and Privacy Assurance (SPA) tool, aimed at cataloguing findings and detailing future assessments to be executed. These assessments are designed to ensure that the NextGEM Innovation and Knowledge Hub (NIKH) adheres to the explicit security requirements (S1-S8) outlined in Section 2.2.1 and aligned with CIA principles and GDPR obligations. The SPA tool facilitates a multi-faceted evaluation of NIKH's security posture, addressing vulnerabilities, monitoring system behaviour, and verifying compliance with legal and ethical standards. Each subsection below corresponds to specific security requirements as follows:

- Section 5.1 (NVD Vulnerability Assessments) aligns with S3: Vulnerability Assessments and S4: Risk Assessment, identifying known vulnerabilities in NIKH assets and assessing their potential impact to inform mitigation strategies, as well as S7: Audit and Accountability, by maintaining records of security-relevant results.
- Subsection 5.2 (Continuous Monitoring), supports S7 by maintaining records of security-relevant events and contributes to S1: Authentication and Authorization and S3 by monitoring access and availability to ensure operational integrity.
- Subsection 5.3 (Penetration Testing) addresses S3 and S4 by simulating real-world attacks to uncover exploitable weaknesses and evaluate risks, while also testing S1, S2: Role-Based Access, S5: Trustworthy Data Exchange, S6: Policy Enforcement, and S8: Attack Surface Reduction through proactive measures like system hardening and access control validation.
- Subsection 5.4 (Assessing GDPR Compliance) ensures adherence to S1, S2, S5, and S6 by evaluating data handling practices, access controls, and secure data exchange mechanisms to meet GDPR requirements.

By integrating these assessments, this section demonstrates how NIKH systematically addresses the security requirements (S1-S8) to safeguard its infrastructure, data, and stakeholder trust, with findings and actions detailed in the subsections that follow.

### 5.1 NVD vulnerability assessments

The NVD assessment will be run multiple times throughout the course of Task 6.3. Below a view of findings regarding known vulnerabilities of NIKH assets (described in Section 3.2) are presented.

#### Assets Dashboard

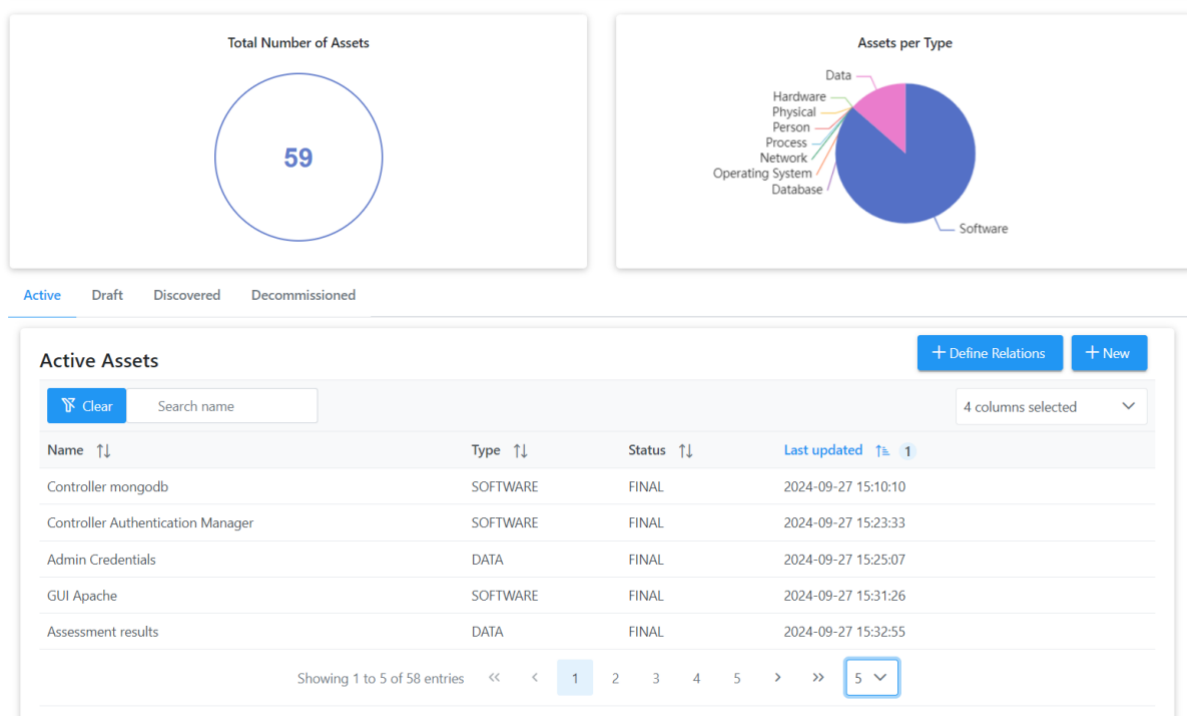


Figure 10: SPA GUI view: Defined asset list

Figure 10 shows the GUI view of the SPA tool where a list of the defined assets is provided. Relationships for multiple assets have been identified to undergo evaluation for NVD assessments.

### Assessment Results

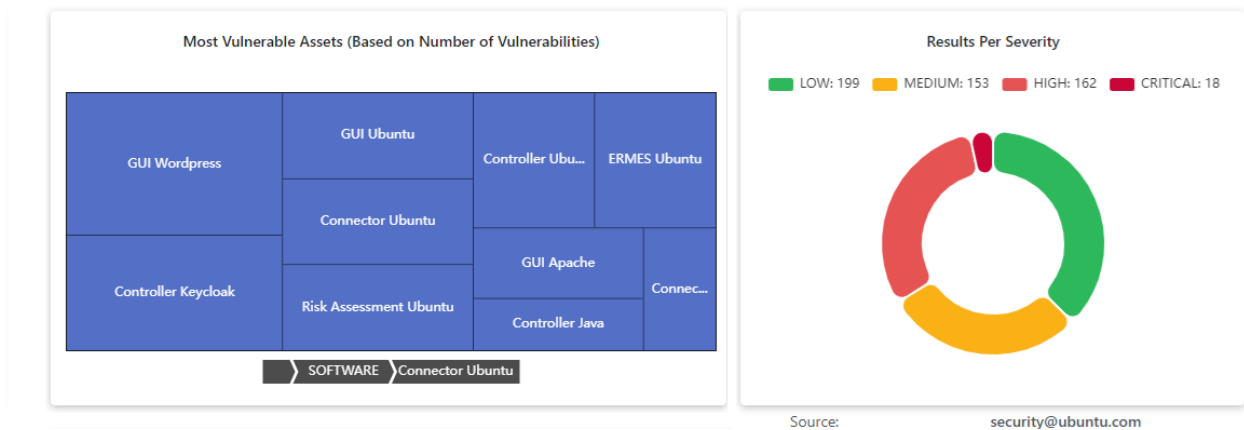


Figure 11: SPA GUI view: NVD assessment results overview - 1.

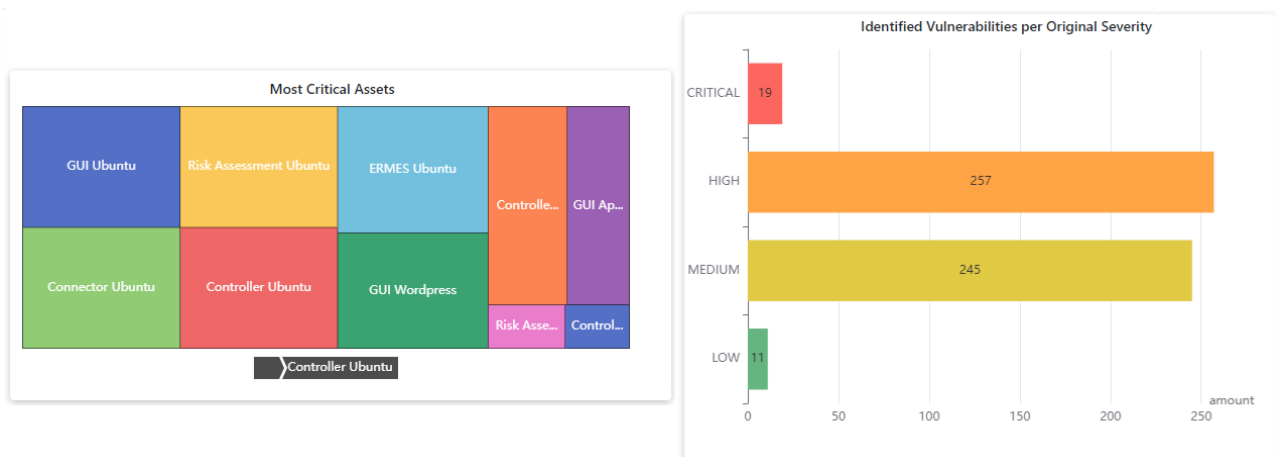


Figure 12: SPA GUI view: NVD assessment results overview - 2.

Result ID	Computed Type	Asset ID	Property	Original Severity	Assessed Severity	Last checked
> 231	CVSS_v3.1	13	INTEGRITY, AVAILABILITY, CONFIDENTIALITY	100/100	CRITICAL	2025-03-14 07:58:14
> 264	CVSS_v3.1	63	INTEGRITY, AVAILABILITY, CONFIDENTIALITY	98/100	CRITICAL	2025-03-14 07:58:14
> 170	CVSS_v3.1	19	INTEGRITY, AVAILABILITY, CONFIDENTIALITY	98/100	CRITICAL	2025-03-14 07:58:14
> 107	CVSS_v3.1	19	INTEGRITY, AVAILABILITY, CONFIDENTIALITY	96/100	CRITICAL	2025-03-14 07:58:13
> 194	CVSS_v3.1	19	INTEGRITY, CONFIDENTIALITY	91/100	CRITICAL	2025-03-14 07:58:14
> 223	CVSS_v3.1	19	INTEGRITY, AVAILABILITY, CONFIDENTIALITY	88/100	HIGH	2025-03-14 07:58:14
> 120	CVSS_v3.1	19	INTEGRITY, AVAILABILITY, CONFIDENTIALITY	88/100	HIGH	2025-03-14 07:58:13
> 152	CVSS_v3.1	19	INTEGRITY, AVAILABILITY, CONFIDENTIALITY	88/100	HIGH	2025-03-14 07:58:14
> 186	CVSS_v3.1	19	INTEGRITY, AVAILABILITY, CONFIDENTIALITY	88/100	HIGH	2025-03-14 07:58:14
> 145	CVSS_v3.0	19	INTEGRITY, AVAILABILITY, CONFIDENTIALITY	81/100	HIGH	2025-03-14 07:58:13

Figure 13. SPA GUI view: NVD assessment results.



Figure 11 and Figure 12 provide an overview of the vulnerability assessment findings. These figures depict the most vulnerable software assets as well as the per-severity type number of vulnerabilities identified. Further, in Figure 13, a list of identified vulnerabilities on the targeted NIKH assets is presented. The vulnerabilities are sorted to prioritize the most severe issues, which are displayed at the top of the returned list.

Name	CVE-2023-1523	Cisa Vulnerability Name	
Source Identifier	security@ubuntu.com	Description	Using the TIOCLINUX ioctl request, a malicious snap could inject contents into the input of the controlling terminal which could allow it to cause arbitrary commands to be executed outside of the snap sandbox after the snap exits. Graphical terminal emulators like xterm, gnome-terminal and others are not affected - this can only be exploited when snaps are run on a virtual console.
Created	2025-03-12 16:00:30		
Published	2023-09-01 19:15:42	Cve References	Url: <a href="https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2023-1523">https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2023-1523</a>
Last Modified	2024-11-21 07:39:21		Source: security@ubuntu.com
Vuln Status	Modified		Tags: Third Party Advisory
Evaluator Comment			Url: <a href="https://github.com/snapcore/snapd/pull/12849">https://github.com/snapcore/snapd/pull/12849</a>
Evaluator Solution			Source: security@ubuntu.com
Evaluator Impact			Tags: Issue Tracking
Cisa Exploit Add			Patch
			Url: <a href="https://marc.info/?l=oss-security&amp;m=167879021709955&amp;w=2">https://marc.info/?l=oss-security&amp;m=167879021709955&amp;w=2</a>
			Source: security@ubuntu.com

Figure 14. SPA GUI view: Details on identified vulnerability.

Figure 14 depicts the details of an identified vulnerability that a SPA tool user can see. In this figure, the most severe vulnerability identified has been selected. As described by the NVD, this vulnerability has to do with Kubernetes, and the security issue that is reported is that users who can create pods and volumes on Windows nodes may be able to escalate to admin privileges on these nodes. Further, as stated in the description of the vulnerability, Kubernetes clusters are only affected if they are using an in-tree storage plugin for Windows nodes. As the NIKH currently does not utilize any Windows nodes, the identified vulnerability is not exploitable within its infrastructure. Similarly, other vulnerabilities identified cannot be exploited as they rely on specific use cases or plugins that are not employed for NIKH's purposes at present.

Given the sensitive nature of the identified vulnerability results, they have been shared directly with the partners responsible for developing and maintaining the NIKH components. This approach ensures they are informed and can implement appropriate mitigation measures as needed while maintaining confidentiality by excluding the details from this report.

## 5.2 Continuous monitoring

Section 4.2.2 describes the continuous monitoring schemes that are employed within NIKH components, mostly to provide assurances regarding availability and confidentiality. The monitoring of NIKH component events is facilitated by the Event Captors detailed in Section 4.1.4. Subsequently, the SPA tool's monitoring component receives and consumes these events, enabling a thorough assessment. Finally, Table 8 presents all the components that are monitored through availability monitoring and IP confidentiality.

### 5.2.1 Service availability monitoring

In relation to the Service Availability assessment, possible periods of downtime observed in the monitored services and components will be reported within the SPA tool user interface.

Table 9 shows the 3<sup>rd</sup> party services to be monitored (described in Section 3.1.3).

Table 9: 3rd party services

Name	Web Address
<b>Zenodo</b>	<ul style="list-style-type: none"> <li><a href="https://zenodo.org/">https://zenodo.org/</a></li> <li><a href="https://zenodo.org/communities/nextgem_project/">https://zenodo.org/communities/nextgem_project/</a></li> <li><a href="https://zenodo.org/communities/seawave_data_managing/">https://zenodo.org/communities/seawave_data_managing/</a></li> </ul>

<b>EMF-Portal</b>	<a href="https://www.emf-portal.org/">https://www.emf-portal.org/</a>
<b>Goliat Dataverse</b>	<a href="https://dataverse.csuc.cat/dataverse/goliat">https://dataverse.csuc.cat/dataverse/goliat</a>
<b>PubMed</b>	<a href="https://pubmed.ncbi.nlm.nih.gov/">https://pubmed.ncbi.nlm.nih.gov/</a>
<b>Web of science</b>	<a href="https://access.clarivate.com/login?app=wos">https://access.clarivate.com/login?app=wos</a>
<b>SketchFab</b>	<a href="https://sketchfab.com/etainproject/">https://sketchfab.com/etainproject/</a>
<b>Yoda</b>	<a href="https://www.uu.nl/en/research/yoda">https://www.uu.nl/en/research/yoda</a>

Result ID	Computed Type	Property	Original Severity	Assessed Severity	Initial detection
68	EVEREST	AVAILABILITY	100	CRITICAL	2025-03-11 05:45:31
Assessment Criterion ID: 7 Criterion Description: Website: <a href="http://emf-portal.org">http://emf-portal.org</a> is down Result: Satisfaction Events: <div>             ID: 27              Description: availability              Severity: CRITICAL              Time: 2025-03-11 05:44:57              Alert Name: Website: <a href="http://emf-portal.org">http://emf-portal.org</a> is down              Elastic Record ID: LXS8g5UB8vDroVDfkLXx           </div> <a href="#">Show more</a>					
> 67	EVEREST	AVAILABILITY	100	CRITICAL	2025-03-11 05:35:36
> 66	EVEREST	AVAILABILITY	100	CRITICAL	2025-03-11 05:15:35
> 65	EVEREST	AVAILABILITY	100	CRITICAL	2025-03-11 01:45:36
> 64	EVEREST	AVAILABILITY	100	CRITICAL	2025-03-11 01:35:40
> 60	EVEREST	AVAILABILITY	100	CRITICAL	2025-03-10 05:05:37
> 59	EVEREST	AVAILABILITY	100	CRITICAL	2025-03-10 04:45:39
> 58	EVEREST	AVAILABILITY	100	CRITICAL	2025-03-10 01:45:36

Figure 15: SPA GUI view: Availability monitoring results.

At the current stage of NIKH monitoring, only minor instances of downtime were detected in the monitored components. The majority of catalogued downtime events occurred in the EMF-Portal external database. This is noteworthy due to its systematic pattern: most incidents took place between 1:00 and 6:00 AM UTC (Figure 15), suggesting that maintenance or updates are scheduled during these hours. Beyond this, no other noteworthy downtime events (either of significant duration or occurring with high frequency) were observed in the monitored components thus far.

### 5.2.2 Access monitoring

In the context of Access Monitoring assessments, SSH logins to monitored components are screened to identify attempts originating from IP addresses not included in a predefined whitelist. This whitelist consists of IP addresses belonging to partners involved in developing each NIKH component. The primary goal of this screening is to strengthen security by generating alerts for login attempts from unlisted IP addresses. Once the whitelist is fully

standardized, this section will present results from the Access Monitoring assessments, detailing any login attempts from IP addresses absent from the whitelist.

This was the original strategy, as outlined in Section 4.2.2. However, to enhance clarity and accountability, we revised our strategy. We now monitor all connection attempts, both successful and failed, rather than relying solely on a whitelist. This allows for comprehensive logging of login activities, facilitating robust oversight and enhanced accountability. Figure 16 illustrates the SPA GUI displaying three SSH login attempts: two unsuccessful attempts, assessed as 'CRITICAL' severity, and one successful attempt, assessed as 'LOW' severity. Note that the assessed severity is indicative only and may not directly provide actionable insights. For instance, if a malicious actor obtains a legitimate user's credentials (e.g., via social engineering or phishing) and logs in successfully, the system will still assign a 'LOW' severity. However, an administrator, upon noticing unusual activity, can review the logs to identify the user and component involved, taking appropriate action as needed.

Clear

Search keyword

Property	Original Severity	Assessed Severity	Initial detection
CONFIDENTIALITY	100	CRITICAL	2025-03-05 13:40:36
Assessment Criterion ID: 8			
Criterion Description: Unsuccessfull login from user [dpallass] with IP [2.87.150.196] in host [subra].			
Result: Violation			
Events: <div><div>ID: 1</div><div>Description: Unsuccessful SSH login</div><div>Severity: CRITICAL</div><div>Time: 2025-03-05 13:40:22</div><div>Alert Name: Unsuccessfull login from user [dpallass] with IP [2.87.150.196] in host [subra].</div><div>Elastic Record ID: h06jZpUB8vDroVDfVCWy</div><div>Show more</div></div>			
CONFIDENTIALITY	100	CRITICAL	2025-03-05 13:40:36
CONFIDENTIALITY	0	LOW	2025-03-05 13:42:16
Assessment Criterion ID: 8			
Criterion Description: Successfull login from user [nkardoulakis] with IP [85.75.37.144] in host [belel].			
Result: Satisfaction			
Events: <div><div>ID: 3</div><div>Description: Successful SSH login</div><div>Severity: LOW</div><div>Time: 2025-03-05 13:41:59</div><div>Alert Name: Successfull login from user [nkardoulakis] with IP [85.75.37.144] in host [belel].</div><div>Elastic Record ID: nk6KZpUB8vDroVDfzCXv</div><div>Show more</div></div>			

Figure 16: SPA GUI view: IP confidentiality monitoring results. Both successful and unsuccessful attempts are shown.

### 5.3 Penetration testing

Penetration testing is a security evaluation in which assessors simulate real-world cyberattacks to identify methods for bypassing the security features of an application, system, or network. In the context of the NIKH system, a specialist conducts a deliberate cyberattack to assess its resilience, potentially resulting in outcomes such as component downtime, unauthorized access leading to data breaches, or unexpected data modification. This

comprehensive assessment aims to proactively identify and mitigate vulnerabilities, strengthening the NIKH's security posture against a broad spectrum of cyber threats.



Figure 17: External/Internal Penetration Testing Workflow.

The Penetration Testing Methodology employed is based on the NIST SP800-115<sup>33</sup> and PTES<sup>34</sup> frameworks, with its workflow illustrated in Figure 17. Each phase of this methodology is outlined below, followed by the specific application to NIKH.

### Penetration Testing Methodology

1. **Planning:** The planning phase establishes the foundation for a successful penetration test by defining rules, securing management approval, and setting testing objectives. No actual testing occurs at this stage. A key deliverable is the Rules of Engagement (ROE), which authorizes the test team to perform specified activities without requiring additional permissions. The ROE outlines the scope, timing, and methods of testing, ensuring clarity and protection for both the testers and the client.

<sup>33</sup> <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-115.pdf>

<sup>34</sup> [http://www.pentest-standard.org/index.php/Main\\_Page](http://www.pentest-standard.org/index.php/Main_Page)

2. **Scoping Document/Questionnaire:** Initial discussions with the client address critical questions to define the engagement's scope, including goals, motivations, and preferences for specific tests.
3. **Information Gathering (Reconnaissance):** Assessors collect extensive data about the target system—such as network topology, operating systems, applications, and user accounts to inform an effective attack strategy. This phase includes passive reconnaissance (using publicly available resources) and active reconnaissance (interacting directly with the system), both of which are typically combined to build a comprehensive vulnerability profile.
4. **Port Scanning:** This phase involves sending packets to specific ports on the target system to determine which are open and potentially vulnerable. Techniques such as ping sweeps, traceroutes, and automated discovery tools are used to identify exploitable entry points.
5. **Service Enumeration:** Assessors identify the versions of services running on open ports. This information is critical for researching known security vulnerabilities associated with specific software versions.
6. **Vulnerability Identification:** Vulnerability analysis compares the target's services, applications, and operating systems against vulnerability databases (e.g., the National Vulnerability Database) and the assessors' expertise. Automated scanners provide rapid results, while manual analysis can uncover novel or obscure vulnerabilities, albeit at a slower pace.
7. **Vulnerability Exploitation:** Using specialized tools or custom code, assessors attempt to exploit identified vulnerabilities to gain access to the system. Although precautions are taken to use reliable exploits, there remains a risk of temporary system or connection disruptions.
8. **Privilege Escalation:** Successful exploits may allow assessors to elevate their privileges, granting access to restricted resources or enabling further unauthorized actions. This step requires additional analysis to assess the full extent of risk, such as identifying accessible or modifiable data.
9. **Reporting:** Upon completion, assessors compile a detailed report summarizing the identified vulnerabilities, their operational impact, and mitigation recommendations. The report includes an executive summary for senior management and a technical section for IT personnel.

## Application to NIKH

For the NIKH, the penetration test will simulate attacks that could lead to the following potential outcomes:

- **Component Downtime:** Attacks targeting specific components may cause temporary or prolonged disruptions in their operation, impacting the availability and functionality of essential services within the NIKH.
- **Data Breach and Leakage:** Unauthorized access to sensitive data could be achieved, leading to data breaches where confidential information is accessed, copied or transmitted to unauthorized entities. This could compromise the privacy and security of users' personal or proprietary data.
- **Data Corruption / Data Modification:** Unauthorized manipulation or corruption of data stored within the system, leads to inaccuracies, inconsistencies, or loss of integrity in critical information. This can have cascading effects on decision-making processes or system functionality.
- **Privilege Escalation:** Successful exploitation of vulnerabilities may grant elevated privileges within the system, enabling unauthorized access to restricted resources, possible execution of unauthorized commands, or compromise of additional components or accounts.
- **Denial of Service (DoS):** A possible flood of the NIKH with a high volume of traffic or malicious requests, could lead to service degradation or complete unavailability. This would disrupt legitimate user access and overwhelm system resources.
- **Execution of Malicious Code:** Implanting and executing malicious software or scripts within the NIKH environment, would lead to further exploitation, reconnaissance, or persistence within the system.

As NIKH is in its final development cycles, the full penetration test has been deferred to avoid disrupting ongoing work. The current focus is on finalizing core functionalities, as early testing might misrepresent the mature system's security posture. Instead, preliminary security measures are being applied to address immediate vulnerabilities, ensuring a stable platform for a more effective future test.

These measures include fortifying the NIKH environment by optimizing and securing its virtual machines. A detailed inventory of applications and configurations is being compiled to identify critical components, remove unnecessary software, and disable non-critical services, thereby reducing vulnerabilities and enhancing security.

In parallel, the network configuration is under review, with open ports limited to essential ones and the latest security patches applied. Key focus areas include secure configuration (e.g., system hardening, disabling unnecessary services), patch management, strict access control, comprehensive logging and monitoring, and

network security (e.g., firewall configuration, port management). This proactive approach minimizes the attack surface and establishes a baseline, ensuring that a full penetration test, scheduled towards the conclusion of Task 6.4, will provide actionable insights to enhance NIKH's resilience.

## 5.4 Assessing GDPR compliance

NextGEM ensures GDPR compliance for data involving human participation, such as sample collection from volunteers in Task 4.4. Experiments are approved by local ethics committees, participants provide informed consent via information sheets, and sensitive data is pseudonymized, with oversight from the NextGEM ethics committee (see Section 3.3 for details). The NIKH supports FAIR data management, cataloguing research outputs with metadata while securing sensitive information, such as user credentials, on servers of the partner that manages NIKH using encryption and access controls. Keycloak enhances NIKH security with centralized authentication and user role management, as outlined in Section 3.3. To evaluate data handling practices, a questionnaire was developed to assess how partners manage data stored locally and accessed via NIKH connectors without direct storage within NIKH. Focusing on data integrity and availability, it targets organizations with limited local premises. Responses, summarized in Table 10, reflect several of the partners' practices.

Table 10: Partner's Answers on Data Handling

Data Management Practice	Partner Answers			
	CSIC	IMBEI	SPi	CNR
<b>Access Control Mechanisms</b>	Access through a private intranet  Firewall blocks unauthorized personnel	Standard Microsoft Active Directory account access control. Defined user groups. Accounts with a limited lifespan unless explicitly extended.  Firewall antivirus.	Access to data only via admins with username/password-protected accounts	Only Researchers of the BioEM laboratory have access to the data
<b>Encryption</b>	No	Yes	No	No
<b>Backup Schedule</b>	Daily or weekly	Daily	Weekly	Monthly
<b>Backup Location</b>	ICMAB's servers	NextGEM Repo, central NAS	External Storage	instruments-dedicated PC, researchers' PC, external hard disk and CNR-IREA Cloud
<b>Data Quality Checks</b>	Yes	Yes	No	Yes
<b>Data Storage Location</b>	Windows 2019 server	Central NAS, accessed via personal computer with Windows 10	Personal Computer with Windows 11 and Dropbox	instruments-dedicated PC, researchers' PC, external hard disk and CNR-IREA Cloud running any of the Mac OS, Windows 10/11
<b>System Update Schedule</b>	Weekly	Microsoft updates their systems when necessary	As soon as updates are available	System updates on PC are performed when available. Instruments-connected PC are not subjected to systems

				updates because they are not connected to the web
--	--	--	--	---

Responses show a robust approach to data security. Most partners implement regular backups (daily to monthly) across varied storage solutions, reducing data loss risks. System updates are prioritized to address vulnerabilities, with schedules from immediate to weekly. Access controls, including intranets, firewalls, and restricted accounts, prevent unauthorized access. However, encryption is inconsistently applied, with only one partner using it, suggesting an area for improvement. These findings, alongside the measures in Section 3.3, demonstrate NextGEM’s commitment to GDPR compliance in data handling practices.



## 6 Risk minimization

Risk minimization within NIKH is a structured and ongoing process aimed at safeguarding the NIKH against vulnerabilities that could compromise its integrity, functionality, and the confidentiality of the data it manages. This process is critical given NIKH's pivotal role in EMF and health research, where the security and reliability of data are paramount. The cycle of risk minimization at NIKH encompasses two fundamental phases: i) assessment and ii) action, each phase designed to enhance NIKH's resilience iteratively.

### 6.1 Assessment: evaluating system vulnerabilities

The first phase involves comprehensive assessments to evaluate the NIKH's systems for potential vulnerabilities. This step is not a one-time event but a continuous effort to proactively identify and understand the range of risks that the NIKH might face. Assessments are conducted through a variety of methods, including but not limited to, security audits, penetration testing, and vulnerability scanning. These evaluations also consider the evolving nature of cyber threats, changes in technology, and shifts in the regulatory landscape that could impact the NIKH's operations.

This phase extends beyond technical assessments to include reviews of operational procedures, employee access controls, and third-party interactions. By adopting a holistic view, the NIKH aims to identify any weak links in its chain of operations that could be exploited by malicious actors or lead to unintentional data exposures. Stakeholder feedback is also an integral part of the assessment phase, offering invaluable insights into potential risks from the perspective of those who interact with the system regularly.

### 6.2 Action: implementing risk reduction measures

Following the assessment phase, the next step involves taking decisive actions to mitigate identified risks and reduce the NIKH's overall vulnerability to exposure. These actions are informed by the assessment results and tailored to address specific vulnerabilities, ensuring that resources are allocated efficiently to bolster the NIKH's defences.

The action phase can include a wide array of measures, such as:

- Strengthening cybersecurity defences with advanced encryption, multi-factor authentication, and intrusion detection systems.
- Updating and patching software and systems to close off vulnerabilities.
- Revising access controls and data management policies to ensure that only authorized personnel have access to sensitive information, and that data is handled in accordance with best practices for privacy and security.
- Implementing employee training programs focused on cybersecurity awareness and data protection protocols to minimize the risk of insider threats.
- Developing and refining incident response plans to ensure swift and effective action in the event of a breach or other security incidents.

### 6.3 Continuous improvement through feedback loops

The risk minimization cycle is characterized by its iterative nature, where feedback from the action phase is used to inform subsequent assessments. This feedback loop allows the NIKH to adapt to new threats, technologies, and regulatory requirements, ensuring that its risk minimization strategies remain dynamic and effective over time.

By engaging in this cycle of assessment and action, the NIKH demonstrates its commitment to maintaining a secure and reliable hub for EMF and health research. This proactive approach to risk minimization is essential for protecting the valuable data entrusted to the NIKH and for ensuring that it can continue to facilitate important scientific and policy-oriented work without interruption.

Navigating the intricate landscape of potential threats to the NIKH underscores the need for a holistic and proactive approach to risk management. From cyber threats to operational vulnerabilities and regulatory challenges, each facet requires meticulous attention and strategic planning.

The future of the NIKH, in safeguarding its data and operations against these multifaceted threats, hinges on continuous vigilance, adaptation, and a commitment to excellence in security and privacy practices. By understanding and acknowledging these potential threats, the NIKH can navigate the complexities of the digital age, ensuring it remains at the forefront of scientific advancement while protecting the interests and trust of its global stakeholder community.



## 7 Conclusion

This document, Deliverable D6.8, represents the final report on trustworthy data management and compliance with ethics and legal aspects for Task 6.3, establishing a robust framework for security requirements within the NIKH hub aligned with CIA principles and GDPR standards. It provides a comprehensive overview of the NIKH ecosystem, detailing its components, assets, and potential vulnerabilities, as explored in Section 3. The Security and Privacy Assurance (SPA) tool has been pivotal in identifying security issues through static analysis against the National Vulnerability Database (NVD) and dynamic analysis via continuous monitoring, implemented within Task 6.3, with penetration testing scheduled for Task 6.4. Additionally, GDPR compliance has been evaluated through data management practices and partner assessments, as detailed in Sections 3.3 and 5.4.

The assessments conducted reveal a set of potential vulnerabilities, which have been shared with responsible partners for awareness and proactive management. Building on the initial findings from D6.3, this final version reflects the completion of NVD assessments, initial monitoring efforts, and GDPR compliance checks, ensuring a secure and compliant platform as NIKH reaches its final development cycles.

Central to this effort is the iterative risk minimization process outlined in Section 6, which drives continuous improvement of NIKH's security posture. Through regular assessments (Section 6.1), vulnerabilities have been identified using the SPA tool, followed by targeted actions (Section 6.2) such as system hardening, patch management, and enhanced access controls, implemented within Task 6.3. Feedback loops (Section 6.3) ensure that insights from these actions refine subsequent efforts, enabling NIKH to adapt to emerging threats and technological changes. This cycle continues into Task 6.4, with a full penetration test planned to validate the hub's resilience (Section 5.3). These iterations strengthen NIKH's security, GDPR adherence, and operational reliability, ensuring it meets the explicit security requirements (S1-S8) from Section 2.2.1 while remaining a trustworthy hub for EMF and health research.