

**ADVANCE'2025** 23-25 June 2025, Sophia Antipolis, France



# Proceedings of the 12th International Workshop on ADVANCEs in ICT Infrastructures and Services

23-25 June 2025

Polytech Nice Sophia

Sophia Antipolis, France



# Table of contents

|  |          |
|--|----------|
| <b>FOREWORD</b>  | <b>1</b> |
| <b>FULL PAPERS</b>   | <b>3</b> |
| Automated Flash Audit for Web Cybersecurity: Design of a Proactive Vulnerability Scanner, Martin Amandine [et al.]   | 3        |
| Comparison of Machine Learning and the ELK Stack for Automated Cyber Attack Detection: A Comparative Approach, Rodrigues Da Silva Gabrielly [et al.]           | 14       |
| Leveraging Large Language Models for the Generation of Stochastic Models in System Availability and Reliability Analysis, Martins Maciel Paulo Romero [et al.] | 21       |
| Shade-aware Routing For Sunburn Prevention, Tsintsadze Magda [et al.]  | 30       |
| Towards Sustainability in 6G Network Slicing with Energy-Saving and Optimization Methods, Moreira Rodrigo [et al.]   | 38       |
| Meeting Healthcare Security Standards: From Legal Requirements to Technical Implementation, Journot Ambre [et al.]   | 46       |
| Data Fusion of Observability Signals to Detect Anomalies and Failures in Microservices, Chaves Adalberto [et al.]  | 52       |
| Rapid Review of Digital Sovereignty and Open-Source Software, Martins José [et al.]  | 60       |

|  |            |
|--|------------|
| From male dominated-Telecommunication, Network and Cyber jobs to more win-win jobs and companies, Houssain Corinne [et al.]    | 67         |
| <b>SHORT PAPERS</b>  | <b>75</b>  |
| Data Beyond Control, Pantano Loic [et al.]   | 75         |
| Modeling the Complexity of the Georgian Verbal System through Linked Data and Machine Learning, Archil Elizbarashvili [et al.] | 84         |
| Role of AI driven Internet of Things (IoT) and Wireless Sensor Networks in developing Smart Cities, Garg Tanya [et al.]        | 88         |
| User-Centric Challenges in Digital Identity Wallets: Insights from Industry Experimentation, Raffin Louis [et al.]             | 93         |
| Next Step Towards Intent-based Scheduling for Performance and Security in Edge-to-Cloud Networks, Boutih Karim [et al.]        | 98         |
| Modeling the Complexity of the Georgian Verbal System through Linked Data and Machine Learning, Tsintsadze Magda [et al.]      | 103        |
| <b>LIST OF AUTHORS</b>   | <b>107</b> |

## Foreword

It is a great pleasure to welcome you to ADVANCE'2025: the 12th International Workshop on ADVANCES in ICT Infrastructures and Services, held this year in Sophia Antipolis, France. ADVANCE continues to serve as a dynamic forum for researchers, practitioners, and engineers from academia and industry to share and discuss the latest developments and future directions in ICT infrastructures, networking, and distributed services.

After the successful organization of the 1st ADVANCE workshop in 2012 in Canoa Quebrada (Brazil) with the support of IFCE Aracati, the 2<sup>nd</sup> edition was held in the city of Morro de Sao Paulo (Brazil). In 2013, with the support of IFCE, the 3<sup>rd</sup> edition was held in Miami (USA). In 2014, with the support of IFU, the 4<sup>th</sup> edition was held in Recife (Brazil). In 2015, with the support of UFPE, the 5<sup>th</sup> edition was held in the city of Evry Val d'Essonne (France). In 2017, with the support of UEVE/Paris-Saclay, the 6<sup>th</sup> edition of the workshop was held in the beautiful city of Santiago de Chile (Chile), with the support of the Universidad De Chile (UC). The 7<sup>th</sup> edition was held in Cape Verde Islands with the support of the Universidad de Cabo Verde. In 2020, the 8<sup>th</sup> edition was held in the city of Cancun (Mexico) with the support of the Universidad del Caribe and the Universidad Autonoma de Yucatan. The 9<sup>th</sup> edition was held online, with the support of the University of Zaragoza (Spain) and University College Cork (Ireland). After the pandemy, the 10<sup>th</sup> edition was held in-person, with the support of the Federal University of Ceará (UFC) and the Federal Institute of Education, Science and Technology of Ceará (IFCE-Fortaleza) (Brazil). The 11<sup>th</sup> edition of the workshop was held in Vietnam in the city of Hanoi in 2024, with the support of the Information Technology Institute (ITI) - (VNU - Vietnam National University). Each edition has been made possible thanks to the generous support of local academic partners and the active participation of the research community.

The 2025 edition of the workshop includes five technical sessions, featuring 10 full papers and 6 short papers, as well as a keynote talk by Prof. Abdelhakim Hafid (University of Montreal, Canada) on Scalability, healthcare, decentralized ID, and postquantum security, and an invited talk by Mrs. Mônica Simões Bandeira Cunha (SETD – Pernambuco State Department of Administration, Brazil) on the Digital Agents Program: innovation, academic immersion, and leadership for public sector transformation.

We thank all the authors who submitted their work, the participants joining us at Sophia-Antipolis, and our distinguished speakers. We are also deeply grateful to the 33 members of the Technical Program Committee for their careful review and selection of the contributions. Special thanks to our colleagues at University of Nice Côte d'Azur and Université d'Évry Paris-Saclay for their commitment and support in organizing this edition.

We hope ADVANCE'2025 offers an enriching and collaborative experience for all.

Prof. Paulo Nazareno Maia Sampaio and Prof Nazim Agoulmine  
TPC Co-chairs



---

# Automated Flash Audit for Web Cybersecurity: Design of a Proactive Vulnerability Scanner

Amandine Martin  
Polytech Nice Sophia  
Sophia Antipolis, FR  
amandine-martin@sfr.fr

Karima Boudaoud  
Université Côte d'Azur - Laboratoire I3S-CNRS  
Sophia Antipolis, France  
karima.boudaoud@univ-cotedazur.fr

Maxime Billy  
Polytech Nice Sophia  
Sophia Antipolis, FR  
ozeliurs@gmail.com

Christian Delettire  
Deltekzen  
Valbonne, France  
christian.delettire@deltekzen.com

## ABSTRACT

Web application security is now a major strategic challenge for companies, whether large corporations or very small/small and medium-sized enterprises (VSEs/SMEs). Threats and attacks are diversifying and multiplying, including data breaches, ransomware, defacements, and intellectual property theft. These cyberattacks have severe repercussions for VSEs and SMEs, which are often not sufficiently protected. Even if reliable solutions such as penetration testing and vulnerability scanning tools exist they are generally expensive, not easy to use and not suitable for non-specialist users. Moreover, European legislation (such as GDPR, NIS2 Directive, Cyber Resilience Act and the Digital Operational Resilience Act) imposes increasing cybersecurity obligations and requirements to companies that must adopt security solutions that are compliant with the legislation. To offer a response to VSEs and SMEs companies (i.e. a solution that is affordable and can be used easily), we propose a platform, called WeakSpotter. This latter was developed to enable these companies to understand cybersecurity challenges and thus correct the vulnerabilities they encounter on their websites.

## CCS CONCEPTS

• **Security and privacy** → **Vulnerability scanners; Web application security; Authentication.**

## KEYWORDS

Web Cybersecurity, Vulnerability Scanner, Automated Audit, VSE/SME, Application Security, Open Source, WeakSpotter, GDPR, NIS2

## 1 INTRODUCTION

Web application security is now a major strategic challenge for all businesses, regardless of their size. Large structures, as well as very small and medium-sized enterprises (VSEs/SMEs), face constantly evolving threats: data breaches, ransomware, website defacements, and intellectual property theft.

According to ANSSI [4], 40% of ransomware attacks, in 2022, targeted VSEs and SMEs. This trend intensified in 2023, with cybermalveillance.gouv.fr [9] reporting a significant increase in assistance requests for various types of attacks: account hacking (+26%), phishing (+21%), and ransomware (+17%).

Cyberattacks have serious consequences for their victims, especially for VSEs and SMEs, which are generally less well protected. According to francenum.gouv.fr[14], an attack costs a company an average of 58,600 euros and causes an interruption of IT services for 26 to 29 days.

Professional penetration testing and vulnerability scanning tools exist. However, they remain inaccessible for many small businesses, even though they are comprehensive and reliable. In fact, they are expensive and are complex (i.e. not easy to use), making them unsuitable for non-specialist users. As for free or open-source solutions, they require advanced technical skills, often beyond the internal capabilities available within VSEs/SMEs.

Furthermore, the European legislative framework imposes increasing cybersecurity requirements. The GDPR (2018)[12] imposes strict measures for personal data protection, while the NIS2 Directive (2022) extends cybersecurity obligations to certain SMEs and VSEs operating in critical sectors. Moreover, the Cyber Resilience Act[11] and the Digital Operational Resilience Act[13] establish security standards for software and digital infrastructures.

Faced with these regulatory constraints and the intensification of threats, a significant gap persists between the current offerings and the real needs of VSEs/SMEs. These organizations require a solution allowing them to perform security audits, without complex technical interventions, and at an affordable cost.

In this context, we propose a platform called WeakSpotter, specifically developed to help VSEs/SMEs understand cybersecurity challenges and effectively improve security of their websites. WeakSpotter, that has been initially developed in the context of a final student project[5], focus on detection of web vulnerabilities.

The paper is structured as follows. We first give an overview about the European legislation and the existing frameworks and vulnerability scanners. Then, we describe more deeply our platform WeakSpotter. Afterwards, we briefly underline some potential improvements. Finally, we conclude and outline some future works.

## 2 STATE OF THE ART

To design a solution adapted to the cybersecurity needs of VSEs and SMEs, it is essential to conduct a thorough analysis of the existing landscape. This section focus on current regulations, recognized vulnerability frameworks, available vulnerability detection tools, and existing vulnerability analysis models. The objective is

to identify the challenges these companies face, pinpoint the shortcomings of current solutions, and define the foundations necessary for developing WeakSpotter.

## 2.1 Regulations

Regulations play a key role in protecting systems and data against growing cyber threats. They require companies to follow best practices aimed at strengthening their resilience and protecting personal data.

These regulations apply to all companies, regardless of their size, and aim to ensure data protection, security of digital infrastructures, and trust between users and businesses. They are essential for ensuring digital security and preventing cyberattacks.

### 2.1.1 . GDPR: General Data Protection Regulation

The GDPR[12], which came into force in May 2018, imposes strict rules concerning the protection of personal data within the European Union. This regulation applies to all companies processing personal data, regardless of their size.

Companies must ensure the confidentiality and integrity of data by implementing adequate technical and organizational measures to prevent any unauthorized access or processing. In case of a data breach, they are required to notify the competent authorities within a maximum of 72 hours. Organizations must also maintain a record of processing activities and assess the risks associated with these activities.

Failure to comply with these obligations can result in severe financial penalties, potentially reaching 4% of the global annual turnover or 20 million euros, whichever is higher.

### 2.1.2 . NIS2 Directive: Security of Network and Information Systems

Adopted in 2022, the NIS2 Directive[25] (Network and Information Security) strengthens the framework established by the 2016 NIS Directive. It raises cybersecurity requirements for critical sectors and expands its scope to a larger number of companies, including certain VSEs and SMEs. Affected companies must identify and manage risks related to their network and information systems by implementing appropriate technical and organizational measures.

The directive also mandates the notification of any major incident likely to compromise service continuity or data security. An initial notification must be made within 24 hours, followed by a detailed report within 72 hours. Company executives are directly responsible for compliance with these obligations, which includes cybersecurity training and conducting regular audits.

The entities concerned by this directive are classified into two categories: essential entities and important entities.

#### *Essential Entities.*

Essential entities include public and private organizations operating in critical sectors such as energy, transport, finance, health, and water supply. These sectors play a crucial role in the functioning of society and the economy, making enhanced protection against cyber threats imperative. In case of non-compliance, these entities risk penalties of up to 10 million euros or 2% of their global annual turnover.

#### *Important Entities.*

Important entities cover sectors such as food production, digital services, postal service providers, and waste management. Although these sectors are considered less critical than those of essential entities, their proper functioning remains vital for society. Penalties for non-compliance are also substantial, potentially reaching up to 7 million euros or 1.4% of the global annual turnover.

### 2.1.3 . Cyber Resilience Act : Resilience of Digital Products

Proposed by the European Commission in 2022, the Cyber Resilience Act[11] (CRA) establishes mandatory cybersecurity standards for all digital products marketed within the European Union, whether software, applications, or connected devices.

Manufacturers must ensure the security of their products from the design phase, integrating preventive measures to counter vulnerabilities. They are also required to provide regular updates to fix flaws identified after market release. Clear documentation regarding product risks and the measures taken to mitigate them must be made available to users.

The CRA requirements apply to all companies, including VSEs and SMEs, which must ensure the compliance of their products to avoid financial penalties or the withdrawal of their offerings from the European market. In case of non-compliance, a fine of up to 15 million euros or 2.5% of the global annual turnover may be imposed, depending on the infringement. Companies failing to meet security standards or neglecting to report incidents within the specified deadlines also risk additional penalties.

### 2.1.4 . Digital Operational Resilience Act

Adopted in 2022, the Digital Operational Resilience Act (DORA[13]) strengthens the digital resilience of companies operating in the financial sector. This regulation imposes strict standards for technology risk management, requiring the identification of vulnerabilities related to information and communication technologies (ICT).

Companies must conduct regular tests of their systems to assess their robustness against cyber threats and proactively manage risks associated with their digital service providers. Furthermore, significant incidents must be notified to the competent authorities within specific deadlines, thereby enhancing transparency and responsiveness of companies to cyberattacks.

Failure to comply with DORA[13] obligations can lead to significant financial penalties, potentially up to 1% of the global daily turnover for each day of non-compliance. This underscores the importance for financial sector players, including their service providers, to adhere to these requirements.

### 2.1.5 . Support Solutions

In addition to European regulatory frameworks, many EU member states have implemented national initiatives to strengthen the cybersecurity of businesses, particularly VSEs and SMEs.

These initiatives include subsidies for acquiring cybersecurity solutions, awareness campaigns, and guides to best practices adapted to local specificities. Some countries also offer free assessment tools or simplified audits to help companies identify their weaknesses and comply with European regulations.

In France, a support program has been established to assist companies in their cybersecurity efforts. Cyber PME[6], a joint initiative between the Directorate General for Enterprise (DGE) and Bpifrance[6], offers a tool to identify and prioritize security actions. For a cost of €8,800 (50% subsidized), this tool facilitates diagnosis, the development of a security plan, and the purchase of suitable solutions.

These measures aim to alleviate the financial and technical difficulties of small structures by providing them with concrete support to improve their resilience against cyber threats.

2.1.6 . Summary

All the discussed legislation converges on one essential conclusion: it is indispensable to invest in cybersecurity, both in France and at the European level. Neglecting this aspect exposes companies not only to heavy financial penalties but also to service interruptions that can incur costs far exceeding those of preventive compliance.

2.2 Frameworks

Web application security relies on protecting three fundamental pillars: **confidentiality**, **integrity** and **availability**, often summarized by the acronym CIA. This principle ensures that sensitive data is accessible only to authorized individuals (confidentiality), that information and systems are protected against unauthorized modification (integrity), and that services are available without interruption for users (availability).

2.2.1 . OWASP Top 10

The OWASP[29] (Open Web Application Security Project) framework is an essential international reference for web application cybersecurity. Its "Top 10" ranking identifies the most critical security flaws (OWASP, 2021) (see Table 1).

Table 1: OWASP Top 10 - 2021 (Source: OWASP).

| ID  | Vulnerability Name                         |
|-----|--|
| A01 | Broken Access Control                      |
| A02 | Cryptographic Failures                     |
| A03 | Injection                                  |
| A04 | Insecure Design                            |
| A05 | Security Misconfiguration                  |
| A06 | Vulnerable and Outdated Configurations     |
| A07 | Identification and Authentication Failures |
| A08 | Software and Data Integrity Failures       |
| A09 | Security Logging and Monitoring Failures   |
| A10 | Server-Side Request Forgery                |

2.2.2 . CWE Top 25

The **Common Weakness Enumeration**(CWE) [44] framework lists the 25 most critical security flaws in software. This ranking is based on the prevalence and severity of vulnerabilities and helps developers and security professionals focus on the most frequent and impactful issues for application security.

2.3 Vulnerability Detection

Detecting vulnerabilities in web applications relies on several tools and techniques. These can be classified into three broad categories: commercial scanners, open-source scanners, and open-source tools. Each category has advantages and limitations, and it is crucial to consider the specific constraints of VSEs and SMEs when choosing a solution.

2.3.1 . Commercial Vulnerability Scanners

Commercial cybersecurity solutions , often offered as turnkey services, combine effectiveness, support, and regular updates. However, their adoption remains difficult for VSEs and SMEs, often limited by an average annual budget of €2,000 (according to a survey by Usine Digitale in 2024). In 80% of cases, the manager handles IT themselves.

Analyzing the advantages and limitations of these tools is essential for companies with limited resources. We examined six major solutions available on the market (see Table 2). Paid offerings generally provide comprehensive technical support, sometimes accompanied by training tailored for non-specialists. These tools are often easier to use but still require qualified personnel, which 72% of small businesses lack, according to a study by Cybermalveillance.gouv.fr[9].

Table 2: Summary Table of Commercial Scanners (Simplified).

| TOOL                       | PRICE     | SIMPLICITY | COVERAGE  | SUPPORT   |
|----------------------------|-----------|------------|-----------|-----------|
| <b>Acunetix</b> [3]        | High      | High       | Very High | Good      |
| <b>Burp Suite Pro</b> [30] | Medium    | Medium     | Very High | Very Good |
| <b>Nessus VRx</b> [41]     | High      | Medium     | High      | Good      |
| <b>(Vicar-ius)</b> [47]    | Very High | Medium     | Very High | Very Good |
| <b>Qualys TruRisk</b> [32] | Variable  | Medium     | High      | Good      |
| <b>Intruder</b> [20]       | Medium    | Very High  | Medium    | Good      |

These solutions cover a wide range of known vulnerabilities. For example, Nessus detects 41.8% of identified vulnerabilities, compared to 37.4% for OpenVAS. However, this difference remains modest for companies seeking basic protection. Their cost, ranging from €500 to €10,000 per year, is a major obstacle, especially when 68% of VSEs and SMEs have an annual budget under €2,000.

Although these tools offer undeniable advantages, their cost and the lack of internal skills make their adoption difficult for many small businesses.

2.3.2 . Free and Open-Source Vulnerability Scanners

Faced with the high costs of commercial solutions, many free and open-source alternatives exist. Although these tools are free to use, they are often less comprehensive and more complex to handle. This is particularly problematic given that 72% of VSEs/SMEs lack qualified cybersecurity personnel, according to a study for Cybermalveillance.gouv.fr[9].

**Table 3: Summary Table of Free/Open-Source Scanners (Simplified)**

| TOOL                  | SIMPLICITY | COVERAGE  | SUPPORT |
|-----------------------|------------|-----------|---------|
| <b>OWASP ZAP</b> [27] | Complex    | Very High | Medium  |
| <b>Nikto</b> [38]     | Very Low   | Low       | Low     |
| <b>WPScan</b>         | High       | Low       | Good    |
| <b>OpenVAS</b> [19]   | Low        | Very High | Good    |

In this comparative analysis, we selected four open-source scanners, two of which specialize in website analysis (cf. Table 3). These community-created tools are free and accessible to everyone but do not guarantee dedicated support or systematic reliability.

Unlike commercial solutions, the complexity of use begins with installation, which often requires manual configuration of specific packages. Some tools, like Nikto, operate only via command line, demanding terminal proficiency and the ability to interpret technical results. This complexity, combined with the lack of structured support, makes their adoption difficult for small structures. Furthermore, community documentation can be insufficient or too technical, further hindering their use.

Finally, vulnerability coverage is generally less comprehensive than commercial solutions. However, for VSEs and SMEs primarily seeking basic security, this gap is often negligible.

### 2.3.3 . Other Free and Open-Source Vulnerability Tools

In addition to general vulnerability scanners, specialized tools exist for detecting flaws in specific areas (see Table 4), such as server configuration or vulnerabilities related to certain software. Among these, some focus on CMS (Content Management Systems), popular platforms for creating websites without technical skills. These solutions are particularly suited for VSEs and SMEs, which often lack resources to develop custom sites. According to AFNIC, 83% of VSEs/SMEs and 63% of micro-enterprises have a website, the vast majority using CMS like WordPress (62.7% market share).

However, this popularity makes CMS prime targets for cyberattacks. A Sucuri study reveals that 94% of hacked sites used WordPress. Configuration errors and the use of vulnerable plugins significantly increase the risk of attack. It is therefore crucial for these platforms to have dedicated security tools. Among these, WPScan, Joomscan[28], and Droopescan help detect vulnerabilities specific to WordPress, Joomla, and Drupal sites, respectively. Unfortunately, similar tools do not exist for CMS like Shopify, Wix, or Squarespace, limiting analysis possibilities.

Besides these CMS scanners, other specialized tools exist, but their use often requires executing specific command lines and a good understanding of the results. The lack of automation and the diversity of configurations make their use difficult for a VSE/SME manager, who must also be able to interpret the obtained results.

### 2.3.4 . Comparative Summary

The table 5 highlights the gap between commercial solutions, which are comprehensive but costly and complex, and free or open-source tools, which are affordable but require technical skills and manual integration. SMEs seeking a simplified approach (e.g., URL

**Table 4: Summary Table of Tools by Category (Simplified)**

| Category                                    | Tools  | Description   |
|---|--|---|
| <b>Infrastructure Mapping and Discovery</b> | SUBLIST3R, CSmap[46], Whois[2], Dig[1], Oralyzer | Subdomain enumeration, technology identification, retrieval of administrative information (Whois), DNS queries, and analysis of robots.txt files. |
| <b>Configuration and Service Analysis</b>   | Nmap[15], SSH Audit[7], Wapiti[39]               | Detection of active services, SSH security audit, analysis of configurations and web protections (CSP, HSTS...).                                  |
| <b>Exploration and Access Testing</b>       | GoBuster[34], FUZZ, Cloudflare Detect            | Searching for hidden directories, injection and access tests, identification of Cloudflare protections.   |
| <b>Vulnerability and Exploit Research</b>   | Searchsploit[26], Metasploit[33]                 | Searching for exploits in public databases and automated exploitation of vulnerabilities.   |
| <b>Information Gathering (OSINT)</b>        | Email Harvester[21]                              | Automatic extraction of email addresses from public sources.  |

field, one click, clear report) struggle to find a suitable solution among the current market offerings.

## 2.4 Cybersecurity for All

### 2.4.1 . Analysis of the Search Engine Optimization Model

Search Engine Optimization (SEO analysis tools, although focused on search engine optimization, share similarities with vulnerability scanners, particularly in terms of operation and user experience. Traditionally, SEO analysis was complex and required specialized skills. However, publicly accessible tools, free or paid, have simplified this analysis.

Here, we focus on the interface of PageSpeed Insights[18] (Google Lighthouse), a free Google tool with a clean design. Simply entering a site's URL launches the analysis, requiring no installation or configuration, making it accessible to everyone. After the analysis, an overall score is provided, accompanied by a color code and a site preview. The interface stands out for its simplicity, offering only the essentials while providing technical details on possible

Table 5: Summary Table of the Three Categories

| Category                         | Cost                | Complexity | Ease of Use                                   | SME Suitability |
|----------------------------------|---------------------|------------|---|-----------------|
| <b>Commercial Scanners</b>       | €500/yr - \$6000/yr | Medium     | Professional interface, complex configuration | Low             |
| <b>Free/Open-Source Scanners</b> | €0                  | Medium     | Manual configurations                         | Medium          |
| <b>Free/Open-Source Tools</b>    | €0                  | High       | Command line and technical configuration      | Low             |

improvements. The user can consult guides and resources to resolve identified issues.

This clear and concise user experience serves as a model for simplification that could make cybersecurity more accessible, especially for less skilled users, like those in VSEs and SMEs.

### 3 THE PROPOSED ARCHITECTURE: WEAKSPOTTER

#### 3.1 Specifications

Weakspotter has been built upon a previous work done by students of Networks and Telecommunications department (R&T: Réseaux et Télécommunications) of IUT (Institut Universitaire Technologique) Nice Côte d'Azur, supervised by one of the authors. In this work, they have built an interface and proposed the required features for a vulnerability scanning platform that borrows its simplicity from the SEO domain.

Based on this initial work, we built a first mock-up and drafted the specifications structuring the project requirements. The platform to be designed had to adopt a Software as a Service (SaaS) architecture, implying cloud hosting and permanent access from anywhere with an internet connection. This approach required ensuring storage reliability, data isolation, and security, in compliance with current data protection standards.

Another important requirement concerned the legal aspect of the solution: the platform had to be developed and distributed under an open-source license. We chose the GNU General Public License Version 3 (GPLv3) due to its strong presence among the analysis tools we use and its simplicity, which ensured compatibility with other free software in the project, even though the MIT license offers more flexibility.

Finally, special attention had to be paid to the user experience. The application needed to be intuitive, use accessible vocabulary, and offer a clear interface facilitating the understanding of results. Two types of analysis had to be integrated: a "simple scan" mode for a quick assessment, and a "complex scan" mode offering an in-depth and detailed analysis of the collected data.

#### 3.2 Mockup and Proof of Concept (PoC)

To better define the functional and ergonomic needs of users, we began with a mockup. This step allowed us to compare our initial hypotheses with the expectations of a real client. Discussions during this phase led to the validation of a first set of essential functionalities. Although the final design evolved later, this preparatory phase was crucial for establishing the system's functional architecture.

Subsequently, a Proof of Concept (PoC) was developed to demonstrate the technical feasibility of the proposed solution, WeakSpotter. This intermediate version integrated the initial components of the scan engine and a draft of the report generation module. These two essential system components were presented to the client, thus confirming the relevance of the development undertaken.

In parallel, an analysis of existing solutions was conducted to identify relevant tools to integrate into the platform (see section State of the Art). This study highlighted the need to orchestrate the different modules in a coherent execution order. An initial attempt at manually modeling this logic was undertaken. However, the increasing complexity of this process quickly revealed the benefit of automating this orchestration.

At the end of this design phase, the main functionalities to be implemented in WeakSpotter were precisely defined. They are structured as follows:

- **Analysis Interface:** A dedicated web interface allowing the user to enter a target URL, with a choice between several levels of analysis depth (simple or advanced scan).
- **Real-time Tracking:** A dynamic feedback mechanism to visualize the progress of the ongoing scan. This feature aims to improve the user experience by providing transparency about the process steps.
- **Results Presentation:** The tool generates two levels of output:
  - a synthetic overall score, facilitating a quick understanding of the vulnerability level;
  - a detailed report, listing all detected vulnerabilities, each accompanied by a severity score and remediation recommendations.
- **User Management:** A complete system including:
  - a secure authentication module;
  - a history of previously performed analyses, viewable by each user.

#### 3.3 Software Design

WeakSpotter's architecture was designed following the classic client-server model to facilitate future development and maintenance. This section details the technical choices made, the organization of the different system parts, their interaction, and their deployment.

**3.3.1 . Architecture** The architecture relies on several interconnected components, presented as follows:

- **Frontend:** User interface to initiate a scan by entering a URL.
- **Backend:** Manages frontend requests and orchestrates scan execution.
- **OCI Containers:** Executes an open-source tool to scan vulnerabilities.

- **AI Backend:** Transmits scan results to an artificial intelligence model.
- **LLM (Large Language Model):** Analyzes results and provides recommendations.

The architectural diagram shown in Figure 1 describes the operational order of the different components:

- (1) **Enter a URL to scan:** The user enters the target URL via the web interface.
- (2) **Scan request:** The frontend sends the request to the backend.
- (3) **Scan with an open-source tool:** The backend runs a Docker[10] container that launches the vulnerability scan on the URL.
- (4) **Return tool scan result:** The container returns the raw results to the backend.
- (5) **Results processing:** The AI backend transmits the results to the LLM model for analysis.
- (6) **Return processed results (CVEs, recommendations):** The LLM identifies vulnerabilities (CVEs) and generates tailored recommendations.
- (7) **Display results:** The backend sends the analyzed results back to the frontend, where they are presented to the user.
- (8) **Read scan report:** The user can consult a detailed report of detected vulnerabilities and recommended actions.

### 3.3.2 . Frontend

The frontend part of our platform is responsible for client-side display (see Figure 2). We chose to develop it in TypeScript[24] with the React framework[23]. This combination allowed us to speed up development while ensuring a solid and maintainable foundation.

To simplify the design and focus on cybersecurity, we integrated the DaisyUI[31] component library and the Tailwind CSS[40] framework. These tools enabled us to create a responsive and aesthetic site quickly, while adhering to web development best practices. Additionally, all these tools are distributed under an open-source license compatible with the GPL.

Our platform is designed as a Single Page Application (SPA). React Router[35] is used to organize navigation between pages smoothly. We structured the frontend into reusable React components, promoting a modular architecture and facilitating long-term project maintenance.

To establish effective communication between the frontend and backend, we use the Axios[22] library. Axios[22] allows us to define an API client with specific routes, HTTP methods (GET, POST, etc.), parameters, and well-defined return types. Each page component calls these routes, and the data is then passed as state to child components, ensuring fluid hydration and consistent data management within the application.

Regarding authentication, we use a React context to manage the global user state across all pages. This centralized context is crucial to ensure the user state is accessible all the time, especially during interactions with the backend that require authentication. This approach guarantees efficient and secure user management, regardless of the page the user is on.

### 3.3.3 . Backend

#### Persistence and Models.

For managing our database, we chose the sqlmodel[37] ORM, which integrates perfectly with the FastAPI[36] framework. The underlying database is SQLite. However, it is possible to easily replace it with MariaDB or PostgreSQL, simply by modifying an environment variable.

We defined three main entities to structure our database:

- **User:** This entity stores user information (name, password, etc.).
- **Scan:** This entity is linked to one or more users and represents the scans performed.
- **Result:** This entity is linked to a Scan and contains the scan results.

#### REST API.

Our REST API routes were developed based on the entities and specific needs of our platform. Here are the main routes for managing scans and results (see Table 6, Table 7 and Table 8).

**Table 6: API Routes for Scan Management**

| Method | Route                | Description                     |
|--------|----------------------|---------------------------------|
| GET    | /api/scans           | List scans for a user           |
| GET    | /api/scans/{scan_id} | Read a scan by its identifier   |
| POST   | /api/scans           | Create and launch a scan        |
| DELETE | /api/scans/{scan_id} | Delete a scan from the database |

**Table 7: API Routes for Scan Results**

| Method | Route                        | Description              |
|--------|------------------------------|--------------------------|
| GET    | /api/scans/{scan_id}/results | Read results from a scan |
| GET    | /api/scans/{scan_id}/report  | Generate a PDF report    |

**Table 8: API Routes for Authentication**

| Method | Route           | Description         |
|--------|-----------------|---------------------|
| POST   | /api/users      | Register a new user |
| POST   | /api/auth/login | Log in a user       |

*Scan Orchestration.* After setting up the backend API, the next step was to design the orchestration of scan tasks. We have chosen to isolate each scanning tool within distinct containers. This approach aims to facilitate dependency management while ensuring the functional independence of the tools, allowing them to be executed and tested individually.

The backend is responsible for creating, configuring, and executing these tools through a task orchestrator. The orchestrator relies on a configuration file describing both simple and complex scans,

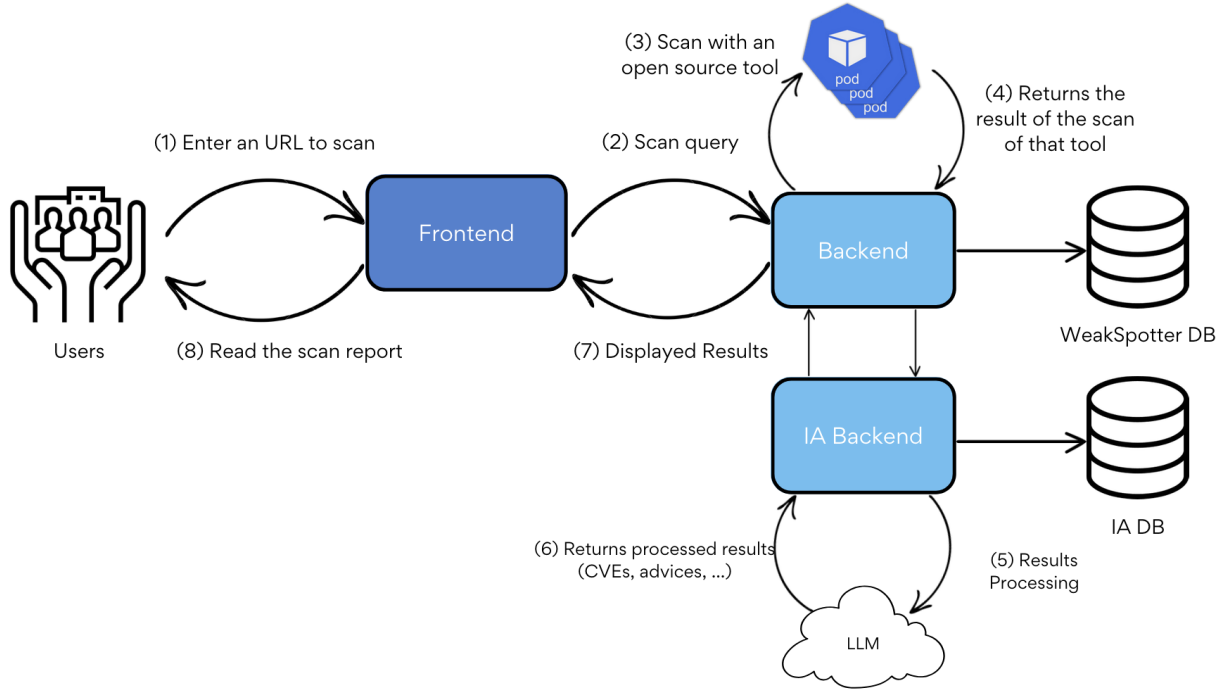


Figure 1: Diagram of the Architecture of the Weakspotter Platform

from which it dynamically generates processing pipelines. The adopted architecture is based on a modular model, where different modules are imported dynamically at runtime. To ensure system extensibility, a standard Job format was defined. All scan modules must inherit from a common abstract class, Job, which specifies the basic methods to implement. This mechanism allows adding new jobs without requiring structural changes to the application.

The choice of deploying on Kubernetes[8] and the widespread use of containers introduced certain constraints, particularly concerning local development. Indeed, working with a Kubernetes cluster in a development environment presents practical difficulties. To overcome this, we opted for Docker as the development environment, which required implementing an abstraction layer allowing the orchestrator to interface interchangeably with Docker (via the local socket) or the Kubernetes API, depending on the target environment (development, staging, or production).

Once this architecture was in place, it was necessary to adapt some open-source tools to our infrastructure. Several projects available on GitHub did not offer containerized images. A specific effort was therefore dedicated to containerizing these applications. To date, 13 pull requests have been submitted to open-source projects to automate their containerization process. These contributions include integrating GitHub Actions[16] and using the GitHub Container Registry[17] (GHCR) to automatically publish Docker[10] images upon updates to the main branch. This work is part of a

broader community contribution effort, making available improvements that we would have had to implement for our own use anyway.

#### Results Interpretation.

The final link in the scan execution chain involves interpreting and simplifying the results provided by the various tools. Our system integrates about twenty security tools, whose output formats vary considerably. While some tools offer structured results (in JSON, YAML, or XML format), others simply provide raw text output via standard output.

To standardize the results, specific parsing logic for each tool was implemented. This logic is integrated directly into the Job module corresponding to the tool and transforms heterogeneous outputs into a common JSON format, thus facilitating their subsequent processing and exposure via the API.

Once the results are extracted and standardized, a simplification phase is applied to make technical information accessible to non-expert users. We defined a common structure for each result, composed of four key elements:

- **Title:** concise summary of the detected problem or configuration.
- **Score:** positive or negative value reflecting the result's impact on the audited system's security.

- **Description:** detailed explanation of the problem or context.
- **Recommendation:** suggested action(s) to correct or improve the situation.

The approach chosen for this step is hybrid. Some tools, like Wapiti[39], directly provide usable descriptions and recommendations, which are simply extracted and integrated into the results. However, other tools like Nikto[38] or Nmap[15] generate purely technical results without contextualization. For these cases, a dedicated microservice based on artificial intelligence was developed. This service is responsible for automatically generating descriptions and recommendations from the tools' raw output. This component is designed modularly so it can be easily replaced or improved, notably by integrating more powerful or specialized AI models.

#### Score Calculation.

Within our scanning platform, we implemented a global scoring mechanism to provide users with a synthetic assessment of their website's security status. Each scan starts with an initial score of 100 points, representing the maximum attributable security level. As results are produced by the different tools, score modifications are applied based on the nature and severity of the detected elements.

Each Job module assigns an individual score to its results, based on deterministic rules defined by developers according to a pre-established evaluation grid. For example, if the Nmap tool reports that the SSH port (port 22) is open, this is considered a poor configuration practice, and a penalty of -3 points is applied.

Individual scores are guided by the following directives:

- **Critical:** -40 to -20 points
- **Errors:** -20 to -10 points
- **Warnings:** -10 to -3 points
- **Information:** -3 to +10 points
- **Debugging:** 0 points

This deterministic system, not based on artificial intelligence, offers the advantage of being transparent, reproducible, and explainable. It also allows integrating a scan history mechanism, offering the possibility to track the score evolution of a given site over time. This feature aims to raise user awareness by showing the concrete impact of their corrective actions and encouraging them to adopt more robust and sustainable cybersecurity practices.

#### Report Export.

To facilitate consultation, archiving and sharing of scan results, a report export feature was integrated into the platform. In addition to accessing results via the web interface, users can generate a report in PDF format, suitable for printing or distribution through external channels such as email.

#### 3.3.4 . AI Microservice

One of the most critical components of our platform concerns the simplification of vulnerabilities, a central objective outlined in the specifications. To ensure maximum scalability and modularity, this functionality was designed as an independent microservice, capable of evolving autonomously from the main backend.

One of the main challenges identified early in development is the extremely high volume of known vulnerabilities. Indeed, databases

such as the National Vulnerability Database (NVD) and the MITRE CVE repository currently list over 200,000 entries. A manual simplification approach being unrealistic at this scale, we opted for an automated solution, accepting a partial compromise on description quality to ensure comprehensive coverage.

To meet this need, we chose to use large language models (LLMs). Although these models are not necessarily specialized in cybersecurity, their ability to rephrase and generate natural language makes them particularly suitable for translating technical language into accessible language.

Our simplification microservice combines these LLMs with data extracted from the MITRE and NVD databases. The latter provide for each CVE (Common Vulnerabilities and Exposures) a standardized technical description as well as information on remediation methods. The system thus relies on a fusion between structured data from official vulnerability databases and the linguistic capabilities of LLMs to automatically produce descriptions understandable by a non-expert audience, while retaining the essence of the technical content.

This approach allows maintaining a balance between scalability, accuracy, and information accessibility, while ensuring a flexible, scalable, and easily maintainable architecture.

#### Prompt Engineering.

Several techniques today allow enriching artificial intelligence model inferences with external data, including approaches like Retrieval-Augmented Generation (RAG) or tool calling. In our project, we opted for a more direct strategy: explicit data injection into the initial prompt. Although this method is considered naive, it proved sufficiently effective and suitable for our use case, where data volumes remain moderate and well-structured.

The principle relies on constructing an enriched prompt, into which relevant information about a given vulnerability is integrated, extracted notably from the MITRE CVE and NVD databases. This data, encoded in JSON format, includes:

- the CVE identifier
- its technical description
- its severity score

An example of a prompt used in our microservice is presented below: *"Simplify this CVE vulnerability in simple terms, understandable for someone with no cybersecurity knowledge. Use clear language and don't be lengthy in your response if unnecessary. Only return the simplified explanation as plain text, without JSON or complex technical information."*

By directly integrating structured data into this prompt, we leverage the language model's synthesis and reformulation capabilities while maintaining precise control over the generation context. This method offers the advantage of simple implementation, without requiring additional document retrieval infrastructure, while allowing fine customization of the tone and length of the generated responses.

#### Interfacing with Groq.

In parallel with prompt engineering work, a comparative analysis of language models (LLMs) available on the market was conducted



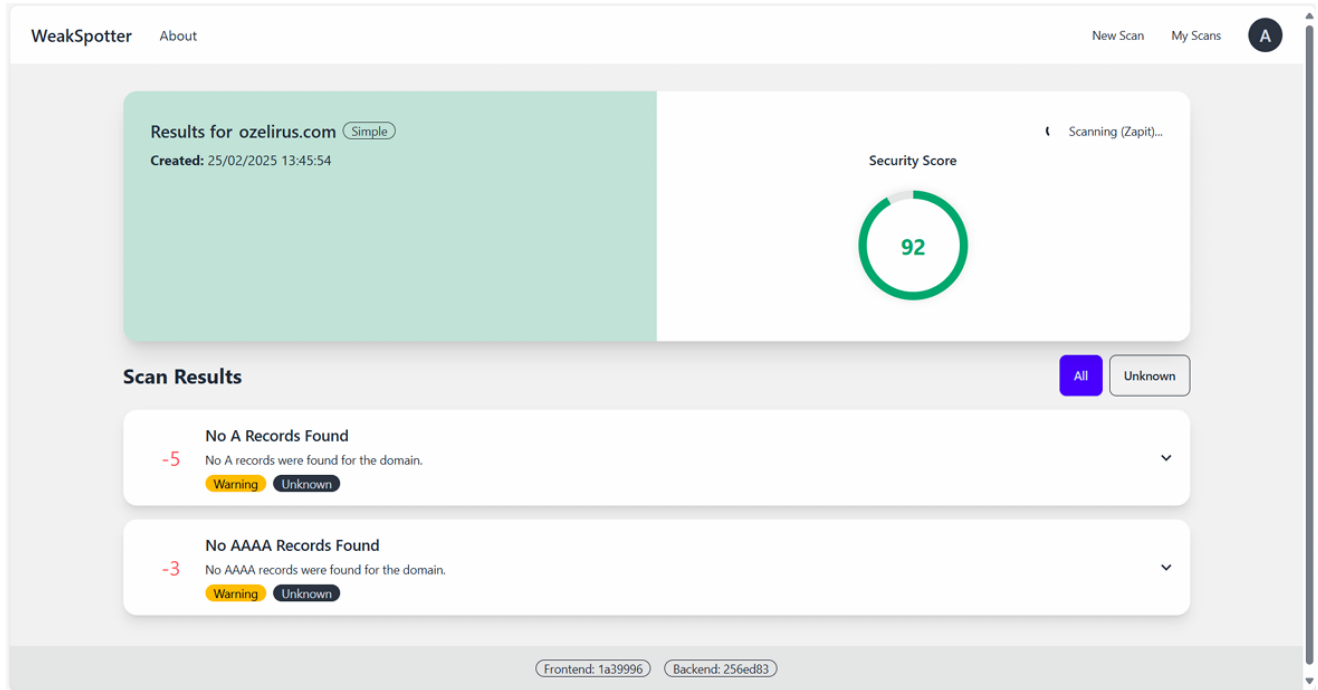


Figure 2: scan result on frontend

to identify the most suitable solution for our simplification microservice. This analysis aimed to evaluate the feasibility of local inference as well as the performance of inference services accessible via API.

The comparison criteria used included:

- **response speed,**
- **quality of generated responses,**
- **cost of use.**

Among the solutions studied were offerings from OpenAI, Anthropic, Mistral AI, Groq (not to be confused with Grok), and Ollama, a solution allowing local model execution.

Models deployed locally via Ollama offered a cost advantage (completely free) but showed insufficient performance in terms of speed and quality, especially compared to large models available via API.

The Groq service, on the other hand, offered a particularly competitive solution, with very low latency and generous free usage up to 3 million tokens per month. Although the generated results were not always the best in terms of relevance or linguistic richness, the execution speed and zero cost for moderate use made it a balanced solution for our application.

The model ultimately chosen for our microservice is “llama-3.3-70b-versatile”, an open-source model developed by Meta and hosted via Groq’s infrastructure. Our microservice includes a dedicated Python client for interfacing with the Groq API and ensuring smooth integration of responses into the main backend.

*Database and REST API.*

To optimize our microservice’s performance and reduce excessive token consumption, we implemented a caching system to store the simplification results of already processed vulnerabilities. This system aims to prevent redundant calls to language models for already processed vulnerabilities, thereby reducing inference costs and improving system responsiveness.

Simplification results are thus saved in a dedicated database, where each record contains the necessary information to quickly retrieve results for previously analyzed vulnerabilities. When the microservice receives a request for a vulnerability already present in the cache, it can simply return the stored result, thus avoiding costly reprocessing.

This caching mechanism also ensures response consistency while optimizing resources, particularly in the context of the frequent use of the same vulnerabilities across successive scans.

### 3.3.5 . Deployment

For our infrastructure, we chose Kubernetes[8] for deploying our platform due to its advantages in container management and scalability. To facilitate operational management and promote smooth collaboration, we adopted a GitOps approach, using ArgoCD[42] to ensure automatic synchronization between our source code, stored on GitHub, and our Kubernetes[8] cluster.

The platform deployment can be described in three distinct parts:

The Frontend includes a deployment, a service, and an ingress. The Backend includes these elements plus a persistent volume and a service account with the necessary rights to create pods in its namespace. Secrets, on the other hand, are stored on the server, but

we could have used a solution like “sealed secrets” to store them on GitHub and version them with the rest of the deployment.

To ensure secure communication between clients and the server, we configured an HTTPS server. This includes managing DNS records for the domain: “\*.weakspotter.ozeliurs.com”. For obtaining and automatically renewing the SSL certificates necessary for securing exchanges, we installed a Kubernetes[8] operator named cert-manager[43].

Finally, we used the Traefik reverse proxy[45] to manage redirections and exchanges between the frontend and backend, while maintaining a simple organization of the single domain. The Traefik configuration is as follows:

- weakspotter.ozeliurs.com/api → http://backend/api
- weakspotter.ozeliurs.com → http://frontend

It should be noted that the order of these redirection rules is important, with the first rule being processed first.

## 4 TO GO FURTHER

Several components of the platform can be improved. For example, on the frontend, we currently use the Axios library to define an API client manually. A significant improvement would be to automate the generation of this client using the OpenAPI specification provided by the backend, ensuring consistency and reducing the risk of discrepancies between the two parts of the application. Furthermore, a centralized state management approach could replace the multiple local states currently used for certain components, such as the scan status. Moreover, on the backend, authentication is handled via JSON Web Tokens (JWT), a widely adopted standard. However, outsourcing this functionality to a dedicated service like Keycloak could reduce security risks and lighten the maintenance burden.

In future work, we plan to evaluate the platform’s usability through real-world user testing to ensure it meets the needs of small business owners with diverse technical backgrounds. We also aim to assess and report on the accuracy of the system’s security analyses. Furthermore, we intend to implement a feedback mechanism that allows users to flag unclear explanations, enabling continuous refinement of the AI-generated outputs.

## 5 CONCLUSION

WeakSpotter delivers a lightweight, open-source SaaS platform that enables VSEs/SMEs to audit website security from a simple URL input. It automates vulnerability detection using containerized tools, standardizes the results with a unified JSON schema, and leverages LLMs for non-technical summaries and remediation tips. The architecture supports modular scans, real-time progress, PDF export, and historical tracking. Deployment is handled via Kubernetes with GitOps. This project bridges the gap between over-engineered enterprise tools and unusable open-source scripts by focusing on accessibility, automation, and clarity.

As future work, we plan to conduct a user testing, to collect essential insights to guide further development. For example, regarding the AI microservice, although we employ a large language model (LLM) followed by manual review, some explanations may still be too complex or unclear for non-expert users. Introducing a reporting system would allow users to flag content that is too

technical or imprecise, enabling iterative refinement and improving the overall accessibility and adaptability of the platform.

Additionally, the absence of testing in the current implementation means that the backend functionalities have not been thoroughly validated. Future work should also include both unit and end-to-end tests to ensure robustness.

## ACKNOWLEDGMENTS

We would like to thank Evan TRANVOUEZ and Mohamed BOUCHEN-GUOUR for their participation in this project and all the students of the Professional Licence (2021-2022 class) of the Networks and Telecommunications department (R&T: Réseaux et Télécommunications) of IUT (Institut Universitaire Technologique) Nice Côte d’Azur, who initially started this work.

## REFERENCES

- [1] Accessed 2025. *dig utility*. Mentioned for DNS queries[cite: 147]. Standard network utility..
- [2] Accessed 2025. *whois utility*. Mentioned for retrieving administrative information[cite: 147]. Standard network utility..
- [3] Acunetix. Accessed 2025. *Acunetix*. <https://www.acunetix.com/> Mentioned as a commercial vulnerability scanner[cite: 120].
- [4] Agence nationale de la sécurité des systèmes d’information (ANSSI). 2023. *Cyber Threat Overview 2022*. <https://cyber.gouv.fr/en/publications/cyber-threat-overview-2022> Accessed: April 16, 2025.
- [5] Amandine MARTIN, Maxime BILLY, Evan TRANVOUEZ, Mahomed BOUCHEN-GUOUR. 2025. *Flash audit automatisé pour la Cybersécurité Web : Conception d’un scanner de vulnérabilités proactif*. Technical Report. Polytech Nice Sophia.
- [6] Bpifrance. [n. d.]. *Cyber PME*. Accessed: April 16, 2025.
- [7] Arturo ‘Buanzo’ Busleiman. Accessed 2025. *SSH Audit*. <https://github.com/jtesta/ssh-audit> Mentioned for SSH security audit[cite: 148].
- [8] Cloud Native Computing Foundation. [n. d.]. *Kubernetes*. Accessed: April 16, 2025.
- [9] Cybermalveillance.gouv.fr. 2024. *Dispositif national d’assistance aux victimes de cybermalveillance - Rapports d’activité*. Accessed: April 16, 2025. (Source for general cyberattack trends and assistance requests in France).
- [10] Docker, Inc. [n. d.]. *Docker*. Accessed: April 16, 2025.
- [11] European Commission. 2022. *Proposal for a Regulation of the European Parliament and of the Council on horizontal cybersecurity requirements for products with digital elements and amending Regulation (EU) 2019/1020 (Cyber Resilience Act)*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0454> Accessed: April 16, 2025. (Note: Citing the proposal; final text adopted later).
- [12] European Union. 2016. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. <https://eur-lex.europa.eu/eli/reg/2016/679/oj> Accessed: April 16, 2025.
- [13] European Union. 2022. *Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector and amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014, (EU) No 909/2014 and (EU) 2016/1011 (DORA)*. <https://eur-lex.europa.eu/eli/reg/2022/2554/oj> Accessed: April 16, 2025.
- [14] France Num. 2024. *France 2030 - Cyber PME*. Accessed: April 16, 2025. (Cites Cabinet Asteres, 2022 for cost figure of approx. €59,000).
- [15] Gordon Lyon (Fyodor). Accessed 2025. *Nmap (Network Mapper)*. <https://nmap.org/> Mentioned for service detection and configuration analysis (NSE)[cite: 148, 263, 271].
- [16] GitHub, Inc. [n. d.]. *GitHub Actions*. Accessed: April 16, 2025.
- [17] GitHub, Inc. [n. d.]. *GitHub Container Registry (GHCR)*. Accessed: April 16, 2025.
- [18] Google. [n. d.]. *PageSpeed Insights*. Accessed: April 16, 2025.
- [19] Greenbone Networks. Accessed 2025. *OpenVAS (Open Vulnerability Assessment Scanner)*. <https://www.openvas.org/> Mentioned as a free/open-source vulnerability scanner[cite: 122, 130].
- [20] Intruder. Accessed 2025. *Intruder*. <https://intruder.io/> Mentioned as a commercial vulnerability scanner[cite: 121].
- [21] Christian Martorella. Accessed 2025. *theHarvester (Email Harvester)*. <https://github.com/laramies/theHarvester> Mentioned for automatic extraction of email addresses (OSINT)[cite: 151]. ‘theHarvester’ is a common tool for this..
- [22] Matt Zabriskie and contributors. [n. d.]. *Axios*. Accessed: April 16, 2025.
- [23] Meta Platforms, Inc. [n. d.]. *React*. Accessed: April 16, 2025.
- [24] Microsoft Corporation. [n. d.]. *TypeScript*. Accessed: April 16, 2025.

- [25] National Institute of Standards and Technology (NIST). [n. d.]. *National Vulnerability Database (NVD)*. Accessed: April 16, 2025. (Database providing details on CVEs).
- [26] Offensive Security. Accessed 2025. *Searchsploit*. <https://www.exploit-db.com/searchsploit> Mentioned for searching exploits in public databases[cite: 150]. Part of Exploit Database..
- [27] Open Web Application Security Project (OWASP). Accessed 2025. *OWASP Zed Attack Proxy (ZAP)*. <https://www.zaproxy.org/> Mentioned as a free/open-source vulnerability scanner[cite: 129]..
- [28] OWASP. Accessed 2025. *JoomScan*. <https://github.com/OWASP/joomscan> Mentioned as a Joomla vulnerability scanner[cite: 145]..
- [29] OWASP Foundation. 2021. *OWASP Top 10:2021*. Accessed: April 16, 2025.
- [30] PortSwigger. Accessed 2025. *Burp Suite Professional*. <https://portswigger.net/burp/pro> Mentioned as a commercial vulnerability scanner[cite: 120]..
- [31] Pouya Saadeghi. [n. d.]. *DaisyUI*. Accessed: April 16, 2025.
- [32] Qualys. Accessed 2025. *Qualys TruRisk*. <https://www.qualys.com/> Mentioned as a commercial vulnerability scanner[cite: 121]..
- [33] Rapid7 Metasploit Community. Accessed 2025. *Metasploit Framework*. <https://www.metasploit.com/> Mentioned for automated exploitation of vulnerabilities[cite: 150]..
- [34] OJ Reeves. Accessed 2025. *GoBuster*. <https://github.com/OJ/gobuster> Mentioned for searching hidden directories[cite: 149]..
- [35] Remix Software, Inc. [n. d.]. *React Router*. Accessed: April 16, 2025.
- [36] Sebastián Ramírez. [n. d.]. *FastAPI*. Accessed: April 16, 2025.
- [37] Sebastián Ramírez. [n. d.]. *SQLModel*. Accessed: April 16, 2025.
- [38] Chris Sullo and David Lodge. Accessed 2025. *Nikto Web Scanner*. <https://cirt.net/Nikto2> Mentioned as a free/open-source vulnerability scanner[cite: 129, 133, 263]..
- [39] Nicolas Surribas. Accessed 2025. *Wapiti*. <https://wapiti.sourceforge.net/> Mentioned for analysis of configurations and web protections[cite: 148, 262]..
- [40] Tailwind Labs Inc. [n. d.]. *Tailwind CSS*. Accessed: April 16, 2025.
- [41] Tenable. Accessed 2025. *Nessus Professional*. <https://www.tenable.com/products/nessus> Mentioned as a commercial vulnerability scanner[cite: 120, 122]..
- [42] The Argo Project. [n. d.]. *Argo CD - Declarative GitOps CD for Kubernetes*. Accessed: April 16, 2025.
- [43] The cert-manager contributors. [n. d.]. *cert-manager*. Accessed: April 16, 2025.
- [44] The MITRE Corporation. 2024. *2024 CWE Top 25 Most Dangerous Software Weaknesses*. Accessed: April 16, 2025. (Using 2024 list as latest available).
- [45] Traefik Labs. [n. d.]. *Traefik Proxy*. Accessed: April 16, 2025.
- [46] Unknown. [n. d.]. .
- [47] Vicarius. Accessed 2025. *VRx*. <https://vicarius.io/> Mentioned as a commercial vulnerability scanner[cite: 121]..

# Comparison of Machine Learning and the ELK Stack for Automated Cyber Attack Detection: A Comparative Approach

Gabrielly R. da Silva  
gabyrlsilva@gmail.com  
Brazilian Federal Institute of  
Education, Science and Technology  
Fortaleza, Ceará, Brazil

Bruno M. C. de Oliveira  
bruno.oliveira@outlook.de  
Technische Universitat Dresden  
Dresden, Germany

Reinaldo B. Braga  
reinaldobraga@gmail.com  
Brazilian Federal Institute of  
Education, Science and Technology  
Fortaleza, Ceará, Brazil

## Abstract

The increasing complexity of cyberattacks has heightened the demand for effective detection mechanisms beyond traditional SIEM systems, which still depend on manual analysis. This study evaluates common classification algorithms—Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Naive Bayes (NB)—and addresses a key limitation in existing datasets: the lack of detail on attack techniques.

To overcome this, we simulated nine attack techniques and applied both manual and automated detection approaches using the ELK Stack and Artificial Intelligence (AI) models. While the manual method enhanced system visualization, it yielded accuracy below 39%.

In contrast, automated classification achieved near-perfect results: DT and NB reached 100% accuracy, RF and SVM 99% in multi-class tasks; for binary classification, DT, RF, and NB reached 100%, with SVM at 99.98%. NB proved the most efficient in balancing accuracy, scalability, and resource use.

## Keywords

SIEM, Machine Learning, Intrusion Detection System, ELK Stack.

### ACM Reference Format:

Gabrielly R. da Silva, Bruno M. C. de Oliveira, and Reinaldo B. Braga. 2025. Comparison of Machine Learning and the ELK Stack for Automated Cyber Attack Detection: A Comparative Approach. In . ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

## 1 Introduction

The evolution from analog to digital media has profoundly transformed the dissemination of information and redefined the landscape of communication [9]. Initially, security threats were confined to physical interception or eavesdropping. With the shift to digital environments [8], attack techniques have become increasingly sophisticated.

While technological progress has enhanced information exchange, it has also introduced new vulnerabilities. A 2023 Kaspersky report [6] indicates a 3% increase in cyber threats compared to the previous

year, with an average of 411,000 malicious files detected daily—most targeting the *Windows* operating system. These statistics underscore the critical need for effective detection and mitigation systems managed by cybersecurity professionals.

As noted by Ayala et al. [1], cybersecurity decision-making is constrained by three key factors: time, due to the demand for rapid responses; dependence on manual processes, since many systems require human intervention; and subjectivity, as mitigation often relies on the expertise of individual analysts.

To address these challenges, professionals rely on Security Information and Event Management (SIEM) systems for real-time monitoring and historical data analysis. According to Kimathi [7], SIEM platforms aggregate information from sources such as access control logs, vulnerability scanners, compliance tools, and application logs.

Analysts define custom alert rules to filter and interpret collected data. However, as networks scale, the volume of logs and alerts grows proportionally, increasing the risk of alert fatigue and inefficient threat prioritization. This paper proposes a comparative evaluation between manual methods—using the Elasticsearch, Logstash, and Kibana (ELK) Stack—and automated approaches based on Artificial Intelligence (AI), aiming to assess their strengths and limitations.

The remainder of this paper is structured as follows: Section 2 reviews related work; Section 3 details the methodology; Section 4 describes the proposed system; Section 5 presents experimental results; and Section 6 concludes the study and outlines future work.

## 2 Related Research

Intrusion Detection Systems (IDS) are designed to identify and alert users to potential security breaches based on network traffic or system logs [11, 16]. Manual analysis has often been supported by Security Information and Event Management (SIEM) platforms, which aggregate network data for real-time analysis and event storage [4]. When used for anomaly detection, SIEM solutions enable the creation of custom rules and alerts. Myers et al. [3] compared three open-source SIEM platforms—Prelude, OSSIM, and the ELK Stack—highlighting ELK’s superior performance and extensive documentation.

Automated approaches increasingly rely on machine learning techniques for intrusion detection, trained on publicly available datasets. Silva et al. [15] evaluated algorithms such as Multilayer Perceptron (MLP), SVM, DT, and RF on the CIC-DDoS2019 dataset, comparing binary and multiclass classification scenarios with and without synthetic data. Braga et al. [14] conducted a survey identifying DT, RF, and SVM as consistently effective across multiple

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
*Conference’17, Washington, DC, USA*

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM  
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

datasets, including CIC IDS 2018, UNSW-NB15, NSL-KDD, and CIC IDS 2017.

Berhili et al. [2] emphasized that while NSL-KDD improves upon KDD 99, it still lacks coverage of modern threats such as zero-day and advanced persistent attacks. They noted the prevalence of classical and deep learning methods, including SVM, RF, DNN, and CNN. Similarly, Giraddi et al. [5] used the INSDN dataset to compare DT, Gaussian Naive Bayes (GNB), SVC, and ANN, with SVC demonstrating the best performance.

Turcato et al. [16] focused on feature extraction from PROFINET traffic, applying techniques such as SVM with Recursive Feature Elimination (RFE) and Autoencoders (AE). Their findings indicated that classical classifiers can achieve performance comparable to deep learning models, often with lower complexity and fewer false positives.

Myneni et al. [10] proposed a semi-synthetic dataset using an Employee Behavior Generation (EBG) model to simulate daily activities and Advanced Persistent Threat (APT) scenarios aligned with the MITRE ATT&CK framework. The dataset was evaluated using models like LSTM\_SAE, Isolation Forest (IF), and SAE, revealing the difficulty in detecting APTs. Park et al. [12] addressed the same challenge with a fast detection approach employing GRR, Audiobeat, and MITRE ATT&CK.

Despite advancements, most datasets represent completed attacks without detailing the techniques used during intrusion. To address this, our study evaluates the performance of DT, RF, SVM, and GNB using a custom dataset built from simulated attack techniques. This approach aims to improve detection granularity by targeting the intrusion process itself.

### 3 Methodology

The proposed methodology comprises four main stages: Data Collection, Pre-processing, Method Implementation, and Metrics and Evaluation, as illustrated in Figure 1.

#### 3.1 Data Collection

**3.1.1 Attack Techniques.** Cyberattacks consist of multiple tactics and techniques defined in the MITRE ATT&CK framework<sup>1</sup>, which catalogs 14 tactical categories. To simulate these techniques, we used Atomic Red Team<sup>2</sup>, a library aligned with MITRE ATT&CK, providing reproducible attack modules with unique IDs, descriptions, and execution parameters.

Based on a Red Team report [13], we selected nine prevalent attack techniques from 2023, listed in Table 1.

Each technique follows five execution stages: (i) identify and select the technique from the module, (ii) verify and install required prerequisites, (iii) configure execution parameters, (iv) execute the technique, and (v) clean up post-execution. Resulting logs are forwarded to the ELK Stack for further analysis.

**3.1.2 Event Logs.** Logs were collected from Windows system components, including application, system, security, operational, and PowerShell logs. These records contain over 230 attributes such as timestamps, file paths, executed commands, integrity levels, and descriptive metadata.

<sup>1</sup><https://attack.mitre.org/>

<sup>2</sup><https://github.com/redcanaryco/atomicredteam/wiki>

**Table 1: Simulated Attack Techniques and IDs**

| Technique                          | ID        |
|------------------------------------|-----------|
| Obfuscated Files or Information    | T1027     |
| Rundll32                           | T1218.011 |
| Rename System Utilities            | T1036.003 |
| PowerShell                         | T1059.001 |
| Windows Command Shell              | T1059.003 |
| Windows Management Instrumentation | T1047     |
| Local Accounts                     | T1078.003 |
| Ingress Tool Transfer              | T1105     |
| OS Credential Dumping              | T1003     |

**Table 2: Log Event Attributes**

| Attributes               | Attributes                                   |
|--------------------------|--|
| event.action             | process.parent.args                          |
| event.category           | process.parent.name                          |
| event.code               | registry.data.type                           |
| event.module             | related.user                                 |
| event.outcome            | rule.name                                    |
| event.provider           | user.name                                    |
| event.type               | winlog.channel                               |
| file.extension           | winlog.event_data.CountOfCredentialsReturned |
| file.name                | winlog.event_data.Description                |
| log.level                | winlog.event_data.IntegrityLevel             |
| powershell.provider.name | winlog.event_data.ParentUser                 |
| winlog.activity_id       | winlog.event_data.ProcessCreationTime        |
| process.command_line     | winlog.event_data.TerminalSessionId          |
| process.executable       | winlog.event_data.PrivilegeList              |
| process.name             | winlog.task                                  |

**3.1.3 Beats Agents.** Beats are lightweight agents that collect and ship data to Elasticsearch. Winlogbeat is used on Windows systems, while Filebeat is used on Linux, both maintaining consistent functionality.

**3.1.4 Elasticsearch.** Elasticsearch acts as the central data repository in the ELK Stack, enabling text-based search, indexing, and real-time analytics using the Apache Lucene engine.

#### 3.2 Preprocessing

**3.2.1 Data Filtering.** An initial sample of 500 log entries—comprising benign and malicious activity—was analyzed. From 213 original attributes, 39 were found to be constant, and 144 were either duplicates, invalid categories, or null-filled. The remaining 30 relevant attributes used for model training are shown in Table 2.

**3.2.2 Natural Language Processing (NLP).** Event log messages contain unstructured text. To transform them into numerical vectors, we used Term Frequency-Inverse Document Frequency (TF-IDF), which assigns weights to terms based on their frequency and rarity:

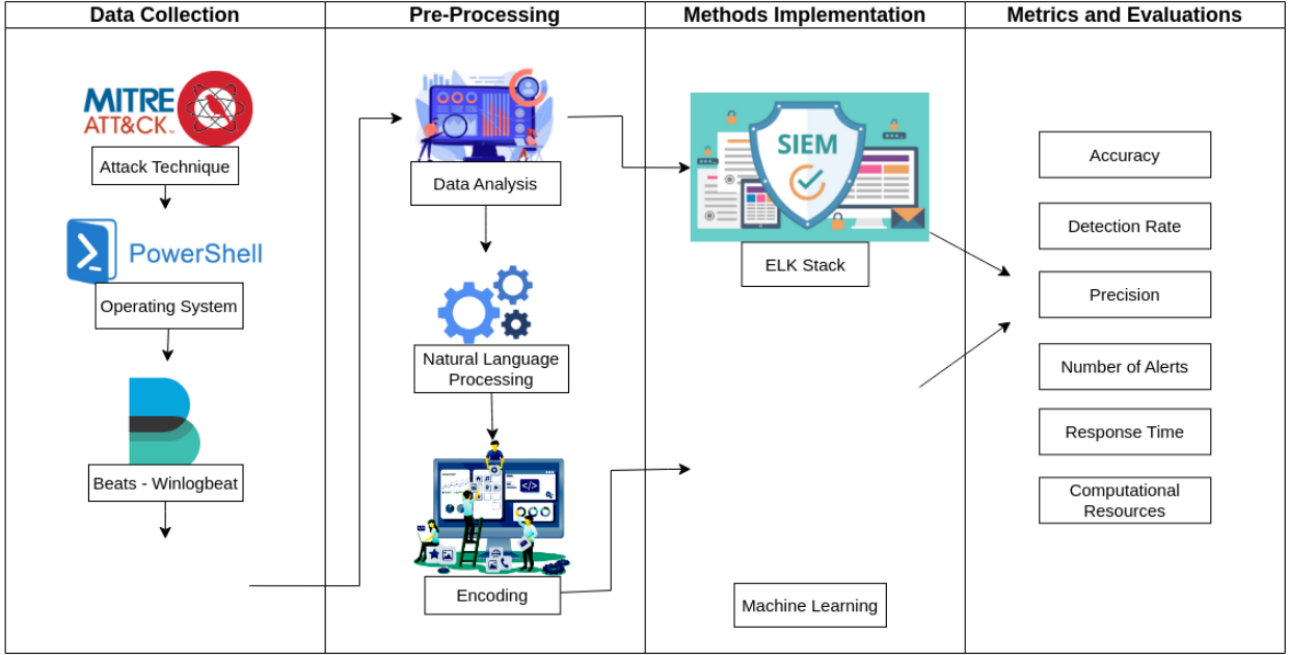


Figure 1: Overview of the proposed methodology.

$$TFIDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (1)$$

where  $t$  is the term,  $d$  the document containing  $t$ , and  $D$  the full corpus.

3.2.3 *Encoding*. All preprocessing and ML algorithms were implemented in Python using the scikit-learn library.

### 3.3 Method Implementation

3.3.1 *ELK Stack*. The ELK Stack (Elasticsearch, Logstash, and Kibana) enables real-time collection, transformation, and visualization of logs. Figure 2 illustrates the architecture, where logs are processed by Beats or Logstash and indexed by Elasticsearch. Analysts access visual dashboards and alerts via Kibana.

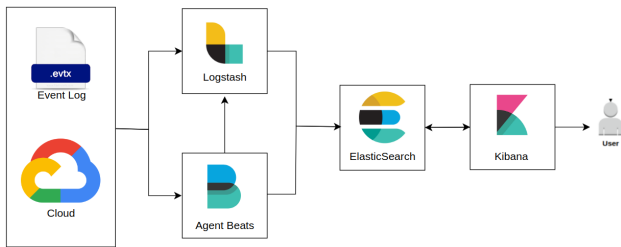


Figure 2: ELK Stack architecture.

3.3.2 *Machine Learning Algorithms*.

*Decision Tree (DT)*. is a hierarchical model where nodes represent features, and leaves correspond to output classes. It recursively partitions the feature space using entropy, calculated as:

$$H(S) = - \sum_{i=1}^n p_i \log_2 p_i \quad (2)$$

*Random Forest (RF)*. is an ensemble of decision trees built from bootstrap samples. Final predictions are obtained via majority voting:

$$\hat{y} = \frac{1}{N} \sum_{i=1}^N \hat{y}_i \quad (3)$$

*Support Vector Machine (SVM)*. seeks an optimal hyperplane separating data with maximum separation margin, typically used in nonlinear problems:

$$w \cdot x + b = 0 \quad (\text{hyperplane}) \quad (4)$$

$$M = \frac{2}{\|w\|} \quad (\text{margin}) \quad (5)$$

*Naive Bayes (NB)*. is a probabilistic classifier based on Bayes' Theorem:

$$P(C_k|X) = \frac{P(X|C_k) \cdot P(C_k)}{P(X)} \quad (6)$$

NB assumes feature independence and is efficient for pattern recognition and text classification, particularly when combined

with NLP techniques which seeks to transform text data into numeric data.

### 3.4 Evaluation Metrics

3.4.1 *Accuracy*. Proportion of correctly classified instances:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (7)$$

3.4.2 *Detection Rate*. Percentage of attacks correctly identified:

$$DetectionRate = \frac{TP}{TP + FN} \times 100 \quad (8)$$

3.4.3 *Precision*. Percentage of true positives among all positive predictions:

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (9)$$

3.4.4 *Alert Volume*. Measures the number of alerts generated and the potential workload for analysts.

3.4.5 *Response Time*. Measures the average time between attack execution and detection.

3.4.6 *Computational Resources*. Analyzes system resource consumption such as CPU and memory usage for manual and automated methods.

## 4 Experiment

Figure 3 illustrates the communication flow of the proposed system. The experimental setup consists of two virtual machines (VMs): one hosts the client environment running Windows, while the other runs a Linux-based ELK Stack server. Both VMs operate on a host machine that also executes the attack simulations and automated analysis routines.

The nine selected attack techniques are executed on the client VM, which has the Winlogbeat agent installed to capture and forward event logs to Elasticsearch. The host system accesses Kibana via port 5601 to visualize and analyze data, while Elasticsearch is available through port 9200. Manual and automated detection methods are initiated after log ingestion.

### 4.1 Manual Method

The manual approach utilizes Kibana's native capabilities for event analysis, rule creation, and alert management. Among the available services—Search, Observability, Security, and Analytics—this study leverages the Security and Analytics modules.

The Analytics module provides real-time visualization of system and network events, enabling granular inspection. The Security module supports rule-based detection and alerting. Rules define filtering criteria using the Kibana Query Language (KQL), while alerts are generated when conditions are met.

In this study, three detection rules were defined:

- **Rule 1:** Filters tasks such as "Provider Lifecycle", "Engine Lifecycle", and "Run Remote Command" via the `winlog.task` field.
- **Rule 2:** Monitors the `process.name` field for suspicious executables (e.g., `net.exe`, `reg.exe`, `rundll32.exe`, etc.).

- **Rule 3:** Correlates the usage of `cmd.exe` and `powershell.exe` by applying a conjunction filter on `process.name` and `process.parent.name`.

Each rule is defined with a severity level, risk score, and execution interval. Once activated, Kibana continuously monitors incoming logs and generates alerts when rule conditions are satisfied. Analysts then review alerts, classifying them as true positives, false positives, or candidates for further use-case analysis.

For this experiment, we assumed that all triggered alerts correspond to malicious events. Filtering was applied using the `signal.rule.name` attribute to correlate alerts with specific rules.

### 4.2 Automated Method

The automated pipeline begins with data extraction from Elasticsearch. During preprocessing, null values are replaced with zero, and categorical data is transformed into numerical vectors using Natural Language Processing (NLP) techniques.

TF-IDF is applied to textual attributes, converting each message into a numerical vector based on term frequency (TF) and inverse document frequency (IDF). A new attribute is then created by summing TF-IDF scores across all terms in a given event sample.

The dataset is split into 70% for training and 30% for testing. A 5-fold cross-validation strategy is employed during training to improve generalization and minimize overfitting, eliminating the need for synthetic balancing. The final model is built from the averaged training results and validated using the remaining test set.

## 5 Results

The Windows client machine was subjected to nine attack techniques from the Atomic Red Team. A total of 48,266 event samples were collected, as summarized in Table 3, which includes both benign and malicious data, labeled accordingly.

**Table 3: Number of Samples per Attack**

| Technique ID | Samples       | Label    |
|--------------|---------------|----------|
| T1027        | 4,886         | 1        |
| T1218.001    | 1,021         | 2        |
| T1036.003    | 2,300         | 3        |
| T1059.001    | 4,680         | 4        |
| T1059.003    | 2,409         | 5        |
| T1047        | 3,365         | 6        |
| T1078.003    | 2,007         | 7        |
| T1105        | 2,187         | 8        |
| T1003        | 2,530         | 9        |
| Benign       | 22,881        | 0        |
| Malicious    | 25,385        | 1        |
| <b>Total</b> | <b>48,266</b> | <b>–</b> |

After data collection, the detection rules from the manual method were activated. Table 4 shows the number of alerts generated by each rule. The *Malware Detection* rule is an ELK built-in rule for

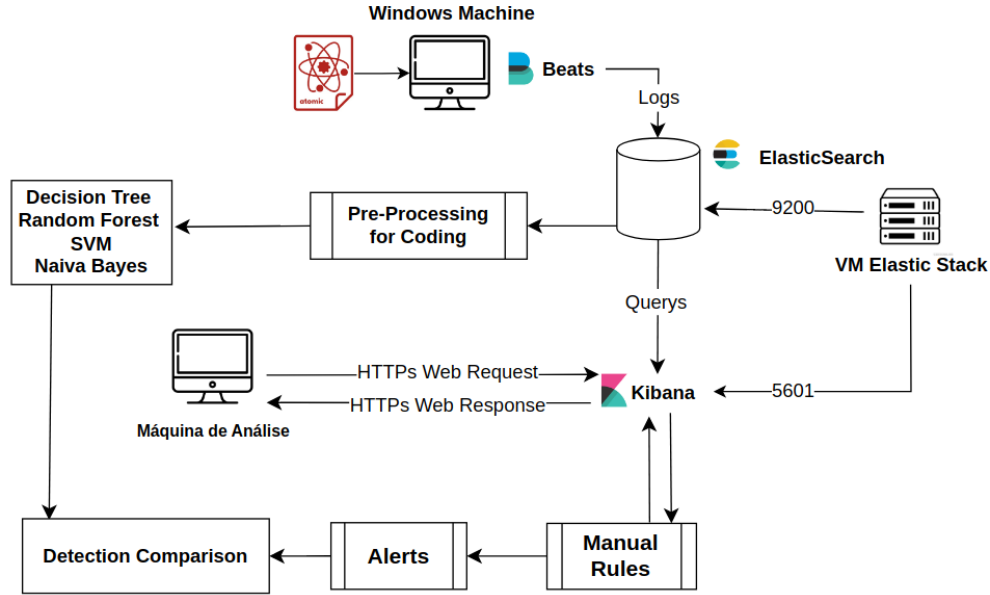


Figure 3: Experiment Flow

identifying traces of malicious scripts. In total, 9,953 alerts were generated. For simplicity, all alerts were considered malicious, resulting in a binary classification setting.

Table 4: Number of Alerts per Rule

| Rule              | Alerts |
|-------------------|--------|
| Malware Detection | 41     |
| Rule 1            | 3,286  |
| Rule 2            | 1,992  |
| Rule 3            | 4,634  |

Figure 4 shows the distribution of alerts over time. Alert frequency varied across the monitoring period, reflecting the dynamic nature of attack execution.

Based on Table 3, only 39.2% of the 25,385 malicious samples were detected manually, indicating that approximately 15,432 events remained undetected (false negatives). While the manual approach supports near real-time monitoring, it requires a virtual environment with at least 4 GB RAM and 2 CPU cores to run the ELK Stack.

### 5.1 Multiclass Classification

The automated approach was evaluated in multiclass and binary settings. Table 5 presents the results for multiclass classification. All models performed well, with DT and NB achieving 100% accuracy. However, DT is more prone to overfitting, which may compromise generalizability. SVM achieved 99.6% accuracy, but generated more false positives, potentially increasing analyst workload.

Table 5: Multiclass Results

| Algorithm | Training |               |        |        |
|-----------|----------|---------------|--------|--------|
|           | Acc (%)  | Precision (%) | Hits   | Errors |
| DT        | 100      | 100           | 33,786 | 0      |
| RF        | 99.97    | 99.98         | 33,778 | 8      |
| NB        | 99.98    | 99.99         | 33,782 | 4      |
| SVM       | 99.46    | 99.18         | 33,604 | 182    |
| Testing   |          |               |        |        |
| DT        | 100      | 100           | 14,480 | 0      |
| RF        | 99.98    | 99.99         | 14,478 | 2      |
| NB        | 100      | 100           | 14,480 | 0      |
| SVM       | 99.60    | 99.38         | 14,423 | 57     |

### 5.2 Binary Classification

Binary results showed significant improvement over the manual method. As shown in Table 6, average accuracy across models exceeded 99.99%. Only NB and SVM presented false positives, but the quantity was negligible.

### 5.3 Resource Consumption

Table 7 shows the resource usage for training and testing each model. RF and SVM exhibited longer execution times. DT consumed more CPU in the multiclass scenario (54.3%), while SVM reached 29.8% during binary testing. Memory usage remained stable across models.



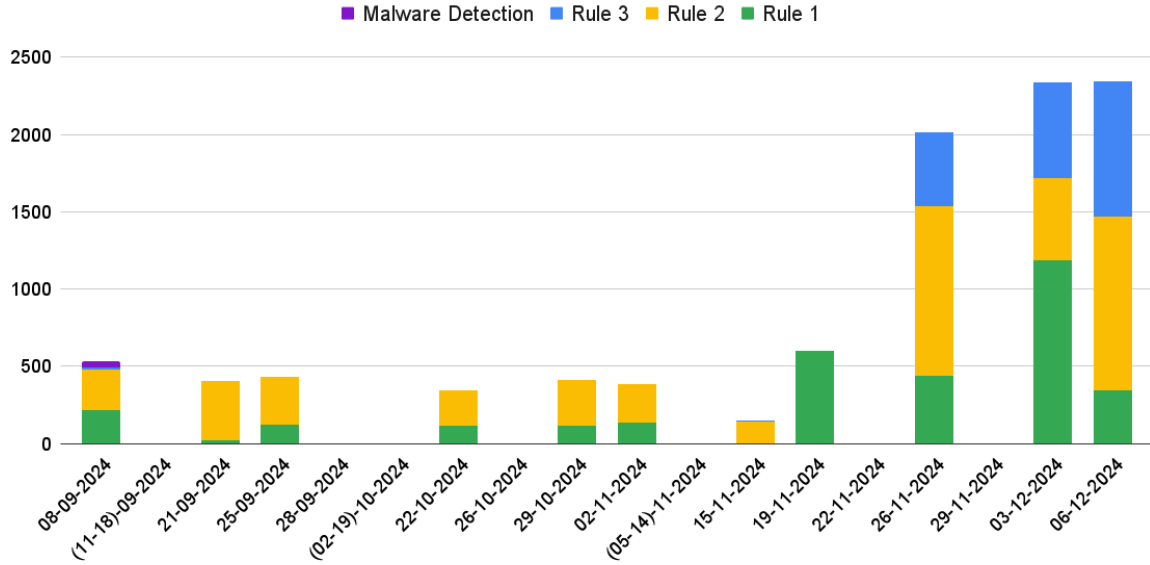


Figure 4: Distribution of alerts over the monitoring period.

Table 6: Binary Results

| Algorithm | Training |               |        |        |
|-----------|----------|---------------|--------|--------|
|           | Acc (%)  | Precision (%) | Hits   | Errors |
| DT        | 100      | 100           | 33,786 | 0      |
| RF        | 100      | 100           | 33,786 | 0      |
| NB        | 99.99    | 99.99         | 33,783 | 3      |
| SVM       | 99.98    | 99.98         | 33,781 | 5      |
| Testing   |          |               |        |        |
| DT        | 100      | 100           | 14,480 | 0      |
| RF        | 100      | 100           | 14,480 | 0      |
| NB        | 100      | 100           | 14,480 | 0      |
| SVM       | 99.98    | 100           | 14,478 | 2      |

Table 7: Computational Resources

| Multiclass Training |       |       |       |       |
|---------------------|-------|-------|-------|-------|
|                     | DT    | RF    | NB    | SVM   |
| Time (s)            | 11.23 | 21.21 | 10.94 | 54.34 |
| RAM (%)             | 16.6  | 16.9  | 16.7  | 15.8  |
| CPU (%)             | 54.3  | 3.0   | 2.5   | 3.0   |
| Binary Testing      |       |       |       |       |
| Time (s)            | 10.62 | 17.74 | 10.47 | 14.75 |
| RAM (%)             | 16.2  | 16.3  | 16.1  | 16.2  |
| CPU (%)             | 2.5   | 2.5   | 2.5   | 29.8  |

## 6 Conclusion

This study conducted a comparative analysis between manual and automated methods for cyberattack detection. The manual approach was implemented using the ELK Stack as a SIEM solution, while the automated method employed supervised machine learning algorithms—Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Naive Bayes (NB).

Unlike most public datasets, which focus on final-stage attack artifacts, we generated a synthetic dataset based on nine MITRE ATT&CK techniques using Atomic Red Team. This allowed for the analysis of intermediate attack behaviors, resulting in over 48,000 labeled samples.

In the manual method, Kibana facilitated real-time log visualization and rule-based filtering. A total of 9,953 alerts were detected, corresponding to approximately 39% of malicious samples. While the approach benefits from immediacy and flexibility, it requires expert oversight and manual validation, and its scalability is limited in more complex scenarios. However, Kibana’s capabilities in filtering, visualization, and dashboard generation proved valuable for operational use.

The automated method was evaluated in both multiclass and binary classification scenarios. DT and NB achieved 100% accuracy, while RF and SVM also performed well (99%). Resource usage analysis revealed that DT, despite its high accuracy, consumed over 50% CPU in multiclass tasks, suggesting potential scalability issues. NB demonstrated a strong balance between performance and efficiency, making it the most resource-effective option.

In the binary classification setting, all algorithms achieved near-perfect accuracy. However, SVM exhibited higher CPU consumption during testing. While DT is interpretable and effective, it is

**Table 8: Advantages and Disadvantages**

| Methods   | Advantages  | Disadvantages   |
|-----------|---|---|
| Manual    | No pre-processing required, only identification of relevant attributes; Nearly real-time detection. | Large volume of alerts;<br>Detection rate less than 39%;<br>Rules created manually; |
| Automatic | Detection rate higher than 90%;<br>Fewer false positives.   | Requires pre-processing;<br>Manual data classification.                             |

prone to overfitting. RF offers robustness and generalization by aggregating multiple decision trees. SVM excels in high-dimensional spaces but is computationally expensive. NB remains simple, scalable, and effective for large datasets, though it may underperform against more complex models.

As summarized in Table 8, the manual method yielded lower accuracy but supports immediate detection and alert generation upon rule definition. In contrast, automated methods achieved higher accuracy but require periodic retraining and offline analysis.

In conclusion, both approaches offer complementary strengths. Manual methods are suitable for real-time incident response, while automated techniques provide high-accuracy detection with reduced human effort. Hybrid approaches that integrate both can enhance detection capabilities across different operational contexts.

**Future Work.** The manual approach depends heavily on rule configuration, which may not generalize well to large-scale or evolving environments. Future studies could explore dynamic rule generation and integration with threat intelligence feeds. In the automated pipeline, more complex scenarios demand advanced models, including deep learning and neural network architectures. This study did not explore these models, but their application presents a promising direction for future research.

## References

- [1] Carlos Ayala, Kevin Jimenez, Edison Loza-Aguirre, and Roberto O. Andrade. 2021. A Hybrid Recommender System for Cybersecurity Based on a Rating Approach. In *Advances in Security, Networks, and Internet of Things*, Kevin Daimi, Hamid R. Arabnia, Leonidas Deligiannidis, Min-Shiang Hwang, and Fernando G. Tinetti (Eds.). Springer International Publishing, Cham, 397–409.
- [2] Mohammed BERHILLI, Omar CHAIEB, and Mohammed BENABDELLAH. 2024. Intrusion Detection Systems in IoT Based on Machine Learning: A state of the art. *Procedia Computer Science* 251 (2024), 99–107. doi:10.1016/j.procs.2024.11.089
- [3] Fidel E. Christopher and Kurt J. Myers. 2018. *Siem-Enabled Cyber Event Correlation (What and How)*. Technical Report. Naval Postgraduate School Monterey CA, Monterey, Califórnia.
- [4] Mohammad Rasool Fatemi and Ali A. Ghorbani. 2020. *Threat Hunting in Windows Using Big Security Log Data*. Security, Privacy, and Forensics Issues in Big Data. 168–188 pages. doi:10.4018/978-1-5225-9742-1.ch007
- [5] Vishal Giraddi, Shantala Giraddi, Narayan D G, Anupama Bidaragaddi, and Suvarna G Kanakareddi. 2024. Machine Learning Approach to Intrusion Detection: Performance Evaluation. *Procedia Computer Science* 235 (2024), 1851–1859. doi:10.1016/j.procs.2024.04.176
- [6] Kaspersky. 2024. *Kaspersky: mais de 400 mil malware foram descobertos por dia em 2023*. [https://www.kaspersky.com.br/about/press-releases/2024\\_kaspersky-mais-de-400-mil-malware-foram-descobertos-por-dia-em-2023](https://www.kaspersky.com.br/about/press-releases/2024_kaspersky-mais-de-400-mil-malware-foram-descobertos-por-dia-em-2023)
- [7] Collins Chandi Kimathi. 2017. *A Platform for monitoring of security and audit events: a test case with windows systems*. Master's thesis. Strathmore University.
- [8] Daniele Verza Marcon. 2023. *Seguro contra riscos cibernéticos: desafios para delimitar a garantia e promover a cibersegurança na era digital*. (2023).
- [9] Sérgio Augusto Soares Mattos. 2013. *A revolução digital e os desafios da comunicação*. Editora UFRB.
- [10] Sowmya Myneni, Kritshekhar Jha, Abdulhakim Sabur, Garima Agrawal, Yuli Deng, Ankur Chowdhary, and Dijiang Huang. 2023. Unraveled — A semi-synthetic dataset for Advanced Persistent Threats. *Computer Networks* 227 (2023), 109688. doi:10.1016/j.comnet.2023.109688
- [11] Seab-Philip Oriyano. 2017. Working with Defensive and Detection Systems. In *Penetration Testing Essentials*. John Wiley Sons, Ltd, 171–191. doi:10.1002/9781119419358.ch11
- [12] Na-Eun Park, Yu-Rim Lee, Soyoung Joo, So-Yeon Kim, So-Hui Kim, Ju-Young Park, Seo-Yi Kim, and Il-Gu Lee. 2023. Performance evaluation of a fast and efficient intrusion detection framework for advanced persistent threat-based cyberattacks. *Computers and Electrical Engineering* 105 (2023), 108548. doi:10.1016/j.compeleceng.2022.108548
- [13] Security Report. 2023. *Ataques cibernéticos no trabalho remoto mais que triplicaram durante a pandemia*. <https://www.securityreport.com.br/ataques-ciberneticos-no-trabalho-remoto-mais-que-triplicam-durante-a-pandemia/>
- [14] Gabrielly Silva, Carina Oliveira, and Reinaldo Braga. 2023. Um Mapeamento Sistemático sobre Detecção de Ataques em Redes de Computadores. In *XI Escola Regional de Computação do Ceará, Maranhão e Piauí* (Aracati/CE). SBC, Porto Alegre, RS, Brasil, 11–20. doi:10.5753/ercemapi.2023.236238
- [15] Rodrigo Silva, Felipe Henriques, Igor Moraes, and Dalbert Mascarenhas. 2024. Detecção de Ataques de Negação de Serviço Distribuídos com Algoritmos de Aprendizado de Máquina. In *Anais do XXIV Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais* (São José dos Campos/SP). SBC, Porto Alegre, RS, Brasil, 226–241. doi:10.5753/sbseg.2024.241684
- [16] Afonso Celso Turcato. 2020. *Desenvolvimento de método para detecção de intrusão em redes PROFINET baseado em técnicas de Aprendizado de Máquina*. Doutorado em Sistemas Dinâmicos. Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos. doi:10.11606/T.18.2020.tde-16072021-172236

Received 15 April 2025

---

# Leveraging Large Language Models for the Generation of Stochastic Models in System Availability and Reliability Analysis

Paulo R. M. Maciel  
prmm@cin.ufpe.br  
Centro de Informática - Cin  
Universidade Federal de Pernambuco  
Recife-PE, Brazil

Erick B. Nascimento  
ebn@cin.ufpe.br  
Centro de Informática - Cin  
Universidade Federal de Pernambuco  
Recife-PE, Brazil

Jamilson R. Dantas\*  
jrd@cin.ufpe.br  
Centro de Informática - Cin  
Universidade Federal de Pernambuco  
Recife-PE, Brazil

## Abstract

This paper investigates the ability of large language models (LLMs) such as ChatGPT, Gemini, Claude, and Mistral to generate and expand Stochastic Petri Net (SPN) models specifically for system availability and reliability analysis. The study evaluates how LLMs can assist in automating the expansion of formal models in these domains, following a structured methodology: (a) modeling a basic availability and reliability system in Mercury-script; (b) providing detailed prompts to guide the extension of the model; (c) obtaining the expanded models from LLMs; (d) validating the generated models in Mercury for syntactic correctness, semantic accuracy, and dependability metrics; and (e) conducting a comparative analysis of the results across the different LLMs. The findings reveal the strengths and limitations of LLM-generated models regarding availability and reliability, providing insights into the potential of AI-assisted formal modeling for improving the efficiency and accuracy of system dependability evaluations.

## CCS Concepts

• **Do Not Use This Code → Generate the Correct Terms for Your Paper**; *Generate the Correct Terms for Your Paper*; Generate the Correct Terms for Your Paper; Generate the Correct Terms for Your Paper.

## Keywords

Large language models; Stochastic Petri Net; Mercury-script language; Availability and Reliability analysis

## 1 Introduction

The increasing complexity of modern systems, particularly those embedded in the Internet of Things (IoT) and characterized by their ubiquity, presents significant challenges for dependability analysis [9, 10]. As systems grow more interconnected and pervasive, ensuring their reliability, availability, and safety becomes increasingly critical. These systems are often expected to function continuously, autonomously, and in unpredictable environments, making the need for robust modeling and evaluation techniques more urgent than ever. As such, system dependability—encompassing reliability and availability—remains a fundamental concern, especially as systems scale in both complexity and scope [2, 15].

To address these challenges, stochastic modeling has long been a cornerstone of dependability analysis [15]. Among the various tools available, Stochastic Petri Nets (SPNs) have proven to be

particularly effective [9, 10]. SPNs extend traditional Petri Nets by incorporating probabilistic timing in transitions, which allows for modeling systems with complex failure behaviors, repair processes, and redundancy mechanisms. SPNs are thus an ideal formalism for evaluating system reliability and availability, offering insights into critical systems' functional behavior and performance under uncertain conditions [2, 5, 8].

However, as the complexity of systems grows, traditional methods for generating and analyzing SPN models can become increasingly labor-intensive and error-prone. Here, the emergence of Large Language Models (LLMs) such as ChatGPT [14], Gemini [4], Claude [1], and Mistral [6] offers a promising solution. These models are designed to understand and generate human-like text, but their potential extends beyond natural language processing. Recent advances suggest that LLMs could be leveraged to automate the generation of formal models for complex systems, including Stochastic Petri Nets, thus alleviating some of the burdens associated with manual model creation [3, 7].

While LLMs have demonstrated their capability in generating software code and formal specifications, their application to dependability analysis, particularly in the context of SPNs for availability and reliability modeling, remains an underexplored area. This paper aims to fill this gap by evaluating the ability of LLMs to generate and expand SPN models specifically for analyzing system availability and reliability. We assess how well these models can automate the expansion of formal models and whether they meet the necessary criteria for syntactic correctness, semantic accuracy, and dependability metrics [7].

The primary objective of this study is to evaluate the effectiveness of LLMs in generating and expanding SPN models specifically tailored for system availability and reliability analysis. By using the Mercury tool [11, 13], a well-established platform for formal validation [9, 10], we validate the models generated by the LLMs, ensuring they adhere to correct stochastic principles and accurately model the intended system behaviors. Through a systematic comparative analysis of LLM-generated models, we aim to highlight their strengths, identify potential limitations, and gain insights into the feasibility of AI-driven formal modeling in system dependability assessments.

This paper fills a gap in research by examining how Large Language Models (LLMs) can generate and expand Stochastic Petri Net (SPN) models for system availability and reliability analysis. It highlights how LLMs can enhance automation and efficiency in SPN model generation.

---

\*Both authors contributed equally to this research.

The paper is organized as follows: In Section 2, we provide background information on Stochastic Petri Nets, their application in dependability analysis, and on AI-Assisted Model Generation. Section 3 presents the Mercury tool. Section 4 outlines the methodology used for evaluating the LLMs’ ability to generate SPN models for availability and reliability analysis. In Section 6, we present and analyze the models generated by the selected LLMs, followed by a comparative analysis of their performance. Finally, Section 8 discusses the study’s conclusions and suggests directions for future research.

## 2 Background

The growing complexity of modern systems, especially those in the context of the Internet of Things (IoT) and ubiquity, makes system dependability analysis crucial for ensuring that systems meet their reliability and availability requirements [9, 10]. This section presents foundational concepts related to Stochastic Petri Nets (SPNs), a widely used tool in dependability analysis, and the role of Artificial Intelligence (AI), specifically Large Language Models (LLMs), in automating formal modeling tasks.

### 2.1 Stochastic Petri Nets in Dependability Analysis

Stochastic Petri Nets (SPNs) are an extension of traditional Petri Nets that incorporate stochastic timing in transitions. This allows SPNs to model systems with probabilistic behavior, such as failure rates, repair processes, and redundancy mechanisms. SPNs are particularly effective for system dependability analysis, as they provide a detailed representation of a system’s behavior under uncertainty. In SPNs, the system is modeled as a set of places, transitions, and arcs, with tokens representing the system’s state. The transition rates, such as failure and repair rates, are modeled using exponential distributions, allowing for the incorporation of probabilistic timing.

SPNs have been successfully applied in various domains, including the availability and reliability evaluation of systems [9, 15]. In such applications, SPNs help assess the system’s steady-state behavior, transient states, and the impact of different configurations or failure mechanisms on the system’s overall dependability. The primary advantage of SPNs is their ability to capture both the structural and dynamic aspects of a system’s behavior. This makes them highly suitable for modeling complex systems’ performance, availability, and reliability, including those found in IoT and cloud computing environments.

### 2.2 AI-Assisted Model Generation

The rise of Large Language Models (LLMs), such as ChatGPT, Gemini, Claude, and Mistral, has opened new possibilities for automating various aspects of formal modeling. LLMs are designed to understand and generate human-like text, and their potential extends to structured tasks such as code generation, specification writing, and model generation. These models can analyze complex prompts and generate corresponding outputs based on the provided context, making them a valuable tool for augmenting traditional modeling workflows [3, 7].

The application of LLMs in formal modeling, including the generation of SPN models, is an emerging area of research. LLMs can assist in automating the model creation process, from defining the system’s structure to expanding existing models with additional components, failure states, and redundancy mechanisms. By leveraging LLMs for this task, researchers and engineers could potentially reduce the time and effort spent on manual model development while improving the consistency and accuracy of the generated models.

Despite the promising potential of LLMs in formal modeling, their application to dependability analysis, particularly for generating SPN models in the context of system availability and reliability, remains underexplored. While a growing body of work on using LLMs for software code generation and natural language processing, their effectiveness in generating accurate, syntactically correct, and semantically sound SPN models remains largely untested. This paper aims to evaluate how well LLMs can be employed to generate SPN models specifically tailored to the analysis of availability and reliability, providing insights into the feasibility of AI-driven formal modeling for dependability evaluations.

Formal modeling techniques such as SPNs have long been used to analyze system reliability, availability, and performance [9, 12, 15]. SPNs extend traditional Petri Nets by incorporating stochastic timing, which allows for the modeling of probabilistic system behavior, including failure rates and repair processes. Despite their effectiveness, the manual construction and extension of SPN models can be labor-intensive and error-prone, especially as system complexity grows. Recent research has focused on improving the efficiency and accuracy of these models, but little work has been done on automating the model generation process using AI-driven tools like LLMs.

This paper contributes to bridging this gap by systematically evaluating the ability of LLMs to generate and expand SPN models in the context of system dependability. By leveraging LLMs, we aim to automate the generation of complex formal models for reliability and availability analysis, providing insights into the potential of AI-assisted modeling in this critical area.

## 3 Mercury Tool

Mercury is a powerful and specialized tool designed to analyze performance, reliability, and dependability in complex systems [11, 13]. It supports a wide range of modeling formalisms, including Stochastic Petri Nets (SPNs), Markov Chains, Reliability Block Diagrams (RBDs), Fault Trees (FTs) and Event Tree (ET). This versatility allows users to perform multi-modeling evaluations, where multiple formal models can be integrated and analyzed simultaneously, offering a comprehensive approach to system dependability analysis.

Mercury provides both graphical and script-based modeling interfaces, making it suitable for novice users and experienced researchers. The graphical interface enables intuitive model building with drag-and-drop elements, while the script-based interface (using Mercury script) offers greater flexibility for advanced users to define complex models programmatically. These two modes are fully integrated, allowing seamless transitions between visual modeling and script-driven simulations.

One of the key strengths of Mercury is its automated validation and evaluation capabilities. Once a model is defined, Mercury uses its engine to validate the system's behavior, ensuring that it adheres to formal dependability principles. This includes checking the model's syntax, verifying that it is logically consistent and correctly formatted, and performing semantic analysis to ensure that the system behavior represented by the model matches the intended real-world system behavior. Additionally, Mercury computes a variety of dependability metrics, such as availability, reliability, mean time to failure (MTTF), and mean time to repair (MTTR), among others, to evaluate the system's overall performance.

Mercury has proven to be a powerful tool in automating formal modeling tasks for system dependability analysis. In this study, we use Mercury to validate the LLM-generated SPN models, ensuring they meet the necessary formal criteria for correctness and dependability. Specifically, we employ Mercury to assess the syntax and semantics of the models, calculate key dependability metrics, and evaluate the feasibility of using LLMs for the automated generation of formal models.

## 4 Methodology

This study follows a structured workflow for generating, expanding, and evaluating SPN models, focusing on leveraging Large Language Models (LLMs) to automate the model expansion process. The workflow ensures that the generated models adhere to formal dependability principles while systematically evaluating their effectiveness in capturing system availability and reliability characteristics.

The methodology consists of the following key steps:

- (1) **Basic Model Generation:** A simple, fundamental system component is initially modeled in Mercury-script to define an initial SPN representation. This basic model typically includes essential system states, such as operational and failed, to create a baseline for further expansion. The aim is to create a minimal yet functional representation of the system's behavior.
- (2) **Prompt Engineering:** A detailed and carefully crafted prompt specifies the intended extensions for the base model. These extensions include adding failure states, introducing redundancy mechanisms, or incorporating performance enhancements. The prompt provides clear guidance for the LLMs to generate meaningful modifications to the initial SPN model, ensuring the expansions align with the intended system behavior.
- (3) **LLM Model Expansion:** The prompt and baseline model are provided to the selected LLMs (ChatGPT, Gemini, Claude, and Mistral) for automated expansion. These LLMs are chosen based on their ability to generate high-quality formal models, and their outputs are carefully reviewed. The goal is to obtain expanded models that accurately reflect the added complexity and new behaviors defined by the prompt.
- (4) **Model Validation:** The extended SPN models generated by the LLMs are validated using Mercury, a performance and dependability analysis tool. This step ensures the models are syntactically correct, semantically accurate, and follow formal stochastic principles (such as valid probabilistic

timing). Validation also involves checking that the model behavior matches the intended system behaviors, ensuring the consistency and correctness of the LLM-generated models.

- (5) **Model Evaluation:** Once validated, the extended models are evaluated based on several key criteria: complexity growth (how the model scales with the added components), dependability metrics (such as availability, reliability, and MTTF), and their overall suitability for modeling system availability and reliability. A comparative analysis is performed to assess the strengths, limitations, and trends across the different LLMs, providing valuable insights into the effectiveness of each model generation approach.

By following this structured methodology, the study ensures that the LLM-generated SPN models are rigorously evaluated and validated, providing valuable insights into the feasibility and effectiveness of AI-assisted formal modeling for system dependability analysis.

## 5 Basic Model and Prompt

This section presents the baseline SPN model and the prompts used to guide the LLMs in generating extended models. The baseline model represents a minimal system configuration, while the prompts specify how the models should be expanded.

### 5.1 Baseline SPN Model

The baseline SPN model serves as the foundation for this study, representing a single-component system where the component can either be operational or fail. This simplified model allows for analyzing system availability and reliability by focusing on the two primary states of the system: operational and failed. By modeling a basic system with simple transitions between these states, we can calculate the probability that the system remains operational over time. This model is a starting point for expanding more complex models in subsequent study steps.

The single-component system availability model represents the behavior of a system with a single operational component that can fail and be repaired. The system has two primary states: operational and failed. The system starts in the operational state, and the goal is to compute the probability that the system remains operational over time.

Let MTTF denote the Mean Time To Failure, which is the expected time before the system fails. Let MTTR denote the Mean Time To Repair, which is the expected time to repair the system once it has failed. In this model, we use two places in the Petri net: UP, representing the operational state of the system, and DOWN, representing the failed state.

The system starts in the UP state with one token, indicating the system is fully operational. The transition from UP to DOWN is governed by the failure transition, which has an associated delay of MTTF. This means that, on average, after MTTF time units, the system will fail and move to the DOWN state.

Once in the DOWN state, the repair transition becomes active. The system will transition back to the UP state after MTTR time units,

where it remains operational until it fails again. The repair time is modeled by the delay of the transition from DOWN to UP, and the transition follows an exponential distribution with a mean repair time of MTTR.

The system's availability is defined as the steady-state probability that the system is in the UP state, which is the operational state. The expression gives this:

$$A = P(\#UP > 0)$$

To calculate this availability, we use the fact that the system alternates between the UP and DOWN states. The average time spent in the operational state (UP) is MTTF, and the average time spent in the failed state (DOWN) is MTTR. Therefore, the availability can be computed as the ratio of the time the system is operational to the total time of a cycle (the sum of MTTF and MTTR):

$$A = \frac{MTTF}{MTTF + MTTR}$$

This formula provides a measure of the system's availability, representing the fraction of time the system is operational compared to the total time spent in both the operational and failed states. The availability value is between 0 and 1, where one means the system is always operational, and zero means it is never operational.

In the model, the availability metric  $A$  is calculated using stationary analysis, which evaluates the system's long-term behavior. The solution of the model provides the steady-state probability of the system being operational, allowing for an assessment of its reliability and performance over time.

## 5.2 Prompt Provided to LLMs

To guide the Large Language Models (LLMs) in generating the desired model extensions, a detailed and carefully crafted prompt is provided. This prompt specifies the task, which involves expanding a basic SPN model to incorporate additional features, such as redundancy mechanisms. The LLMs are tasked with transforming the base model—representing a single-component system availability model—into a more complex system, such as a hot standby system. The prompt ensures that the generated models meet the required specifications for system dependability analysis. Below is the exact prompt provided to the LLMs:

*You are an expert in system dependability analysis using Stochastic Petri Nets (SPNs) defined in the book Performance, Reliability, and Availability Evaluation of Computational Systems by Maciel. Given the following base model in Mercury-script that represents a single-component system availability model. Please provide a hot-standby system availability model in Mercury-script notation. The single-component system availability model:*

```
mttf = 1000;
mttr = 10;
SPN Model{
  place DOWN;
  place UP( tokens= 1 );

  timedTransition Failure(
    inputs = [UP],
```

```
    outputs = [DOWN],
    delay = mttf
  );

  timedTransition Repair(
    inputs = [DOWN],
    outputs = [UP],
    delay = mttr
  );

  metric A = stationaryAnalysis(
    method = "direct",
    storage = "memory",
    error = 1e-12,
    expression = "P{#UP>0}" );
}

main {
  A = solve( Model,A );
  println(A);
}
```

Provide the modified SPN model in Mercury-script format.

## 6 Generated SPN Models from LLMs

This section presents the SPN models generated by each LLM in response to the provided prompt. Although core components are generally consistent, notable differences emerge in structure, syntax, and the specification of transition rates. These variations reflect the distinct modeling tendencies of each LLM. A detailed analysis of these aspects is provided in the following sections.

### 6.1 ChatGPT-Generated SPN Model

The following Stochastic Petri Net (SPN) model was generated using ChatGPT in response to a prompt requesting a representation of a simple fault-tolerant system. The purpose of including this model is to illustrate the type of output that can be produced by a large language model when tasked with generating availability-related structures.

**Listing 1: ChatGPT-Generated SPN Model**

```
mttf = 1000;
mttr = 10;

SPN Model{
  place DOWN;
  place UP( tokens= 1 );
  place STANDBY( tokens= 1 );

  // Standby system ready to take
  // over immediately
  // Transition for failure of the primary
  // system (UP)
  timedTransition Failure(
    inputs = [UP],
    outputs = [DOWN],
    delay = mttf
  );

  // Transition for repair of the primary
  // system (UP)
```

```

    timedTransition Repair(
        inputs = [DOWN],
        outputs = [UP],
        delay = mtrr
    );
// Transition for the backup system
(STANDBY) to take over immediately
after failure

    timedTransition SwitchToStandby(
        inputs = [DOWN],
        outputs = [STANDBY],
        delay = 0
    );
// No delay for immediate switch-over

// Transition for the backup system (STANDBY)
to be repaired and return to standby

    timedTransition RepairStandby(
        inputs = [STANDBY],
        outputs = [STANDBY],
        delay = mtrr
    );
// Availability metric: system is available
if either UP or STANDBY is operational

metric A = stationaryAnalysis(
    method = "direct",
    storage = "memory",
    error = 1e-12,
    expression = "P{#UP > 0 || #STANDBY > 0}"
);
}
main {
    A = solve( Model, A );
    println(A);
}

```

## 6.2 Gemini-Generated SPN Model

The following SPN model was generated using Gemini, another large language model, in response to the same modeling prompt. The model represents a fault-tolerant structure using replicated components and includes failure and repair transitions for each. As with the previous example, no manual correction or interpretation has been applied at this stage—the model is shown exactly as generated to support later analysis and comparison.

**Listing 2: Gemini-Generated SPN Model**

```

mttf = 1000;
mtrr = 10;

SPN Model {
    place UP1(
        tokens = 1
    );

    place UP2(
        tokens = 1
    );

```

```

    place DOWN1;
    place DOWN2;

    timedTransition Failure1(
        inputs = [UP1],
        outputs = [DOWN1],
        delay = mttf
    );

    timedTransition Failure2(
        inputs = [UP2],
        outputs = [DOWN2],
        delay = mttf
    );

    timedTransition Repair1(
        inputs = [DOWN1],
        outputs = [UP1],
        delay = mtrr
    );

    timedTransition Repair2(
        inputs = [DOWN2],
        outputs = [UP2],
        delay = mtrr
    );

metric Availability =
    stationaryAnalysis(
        method = "direct",
        storage = "memory",
        error = 1e-12,
        expression = "P{#UP1>0 || #UP2>0}"
    );
}
main {
    Availability = solve(
        Model,
        Availability
    );

    println(Availability);
}

```

## 6.3 Claude-Generated SPN Model

The SPN model below was generated using Claude, a large language model, following the same prompt provided to the other models. The response includes failure and repair transitions for both a primary and a backup component, as well as availability metrics for the overall system and individual components. As before, the model is presented without manual editing to support an unbiased comparative analysis.

**Listing 3: Claude-Generated SPN Model**

```

// Mean Time To Failure for primary component
mttf_primary = 1000;

// Mean Time To Failure for backup component

```

```

(slightly more reliable)

mttf_backup = 1200;

// Mean Time To Repair (same for both
components)
mttr = 10;

SPN HotStandbyModel {
// Places representing component states

place PRIMARY_UP(
    // Primary component starts in UP state
    tokens = 1
);

// Primary component in DOWN state
place PRIMARY_DOWN;

// Backup component starts in UP state
place BACKUP_UP(
    tokens = 1
);

// Backup component in DOWN state
place BACKUP_DOWN;

// Transitions for the primary component
timedTransition PrimaryFailure(
    inputs = [PRIMARY_UP],
    outputs = [PRIMARY_DOWN],
    delay = mttf_primary
);
timedTransition PrimaryRepair(
    inputs = [PRIMARY_DOWN],
    outputs = [PRIMARY_UP],
    delay = mttr
);

// Transitions for the backup component
timedTransition BackupFailure(
    inputs = [BACKUP_UP],
    outputs = [BACKUP_DOWN],
    delay = mttf_backup
);

timedTransition BackupRepair(
    inputs = [BACKUP_DOWN],
    outputs = [BACKUP_UP],
    delay = mttr
);

// System availability metric - system is
available if at least one component is UP

metric SystemAvailability = stationaryAnalysis(
    method = "direct",
    storage = "memory",
    error = 1e-12,
    expression = "P{#PRIMARY_UP>0 || #BACKUP_UP>0}"
);

```

```

// Individual component availability
metrics (for analysis)

metric PrimaryAvailability = stationaryAnalysis(
    method = "direct",
    storage = "memory",
    error = 1e-12,
    expression = "P{#PRIMARY_UP>0}"
);

metric BackupAvailability = stationaryAnalysis(
    method = "direct",
    storage = "memory",
    error = 1e-12,
    expression = "P{#BACKUP_UP>0}"
);

main {
// Solve for system and component availabilities

SysA=solve(HotStandbyModel,SystemAvailability);
PrimA=solve(HotStandbyModel,PrimaryAvailability);
BackA=solve(HotStandbyModel,BackupAvailability);

// Print results
println("System Availability:" + SysA);
println("Primary Component Availability:"+PrimA);
println("Backup Component Availability:"+BackA);

}

```

## 6.4 Mistral-Generated SPN Model

The SPN model below was generated using Claude, a large language model, following the same prompt provided to the other models. The response includes failure and repair transitions for both a primary and a backup component, as well as availability metrics for the overall system and individual components. As before, the model is presented without manual editing to support an unbiased comparative analysis.

### Listing 4: Mistral-Generated SPN Model

```

// Mean time to failure for the active
component
mttf = 1000;

// Mean time to repair for the failed
component
mttr = 10;

// Mean time to failure for the standby
component
mttf_standby = 1000;

//Time to switch to the standby
component
switchover_time = 1;

```



```

SPN Model {
    place DOWN;
    place UP(tokens = 1);

    // Standby component is initially available
    place STANDBY(tokens = 1);

    place ACTIVE_FAILED;

    // Token to represent the switchover process
    place SWITCHOVER(tokens = 1);

    timedTransition Failure(
        inputs = [UP],
        outputs = [ACTIVE_FAILED],
        delay = mttf
    );

    timedTransition Repair(
        inputs = [DOWN],
        outputs = [STANDBY],
        delay = mtrr
    );

    timedTransition Failure_Standby(
        inputs = [STANDBY],
        outputs = [DOWN],
        delay = mttf_standby
    );

    immediateTransition SwitchToStandby(
        inputs = [ACTIVE_FAILED,
                  STANDBY, SWITCHOVER],
        outputs = [UP],
        priority = 1
    );

    timedTransition SwitchoverDelay(
        inputs = [SWITCHOVER],
        outputs = [SWITCHOVER],
        delay = switchover_time
    );

    metric A = stationaryAnalysis(
        method = "direct",
        storage = "memory",
        error = 1e-12,
        expression = "P{#UP>0}"
    );
}

main {
    A = solve(Model, A);
    println(A);
}

```

## 7 Comparative Analysis of LLM-Generated SPN Models

This section presents a comparative analysis of the SPN models generated by different LLMs, evaluating their outputs with respect to syntax, semantics, complexity, and accuracy. The assessment considers how effectively each model adhered to the prompt, the structural correctness of the generated SPN code, and the interpretability of system behavior. Additionally, transition logic and correctness of availability metrics were evaluated.

Key evaluation criteria include:

- **Syntax and Semantics:** Compliance with SPN formalism and correctness in model logic.
- **Complexity:** Appropriateness of model structure and abstraction level.
- **Accuracy:** Coherence in state transitions and parameter use.
- **Trends and Limitations:** Recurrent issues, modeling limitations, and room for improvement in AI-assisted modeling.

The results provide valuable insights into the strengths and limitations of large language models in generating SPN models, particularly in modeling fault-tolerant systems like the Hot-Standby configuration. The findings reveal nuanced differences in how each model interprets prompts and translates them into code.

**ChatGPT 6.1** presented inconsistencies, most notably in the syntax of the availability metric using “||” instead of the expected logical construct OR, and in the omission of necessary parentheses in expression = “P{#UP>0 || #STANDBY>0}”. The expression is correctly declared as expression = “P{(#UP>0) OR (#STANDBY>0)}”. Semantically, it introduced the RepairStandby loop and a faulty logic that allows indefinite token accumulation in STANDBY after triggering SwitchToStandby. Despite these issues, the model could be manually corrected and interpreted.

**Gemini 6.2** also produced a syntax error in the availability expression. However, once corrected, the model ran successfully on Mercury, indicating sound comprehension of the Hot-Standby system’s semantics. Among the evaluated LLMs, Gemini generated the most accurate representation with minimal need for intervention. Once again, the availability expression is correctly declared as expression = “P{(#UP>0) OR (#STANDBY>0)}”.

**Claude 6.3** presented both syntactic and semantic issues. It introduced the variable MTTF\_BACKUP with a value that diverged from the baseline specification and used an invalid Println method. When compiled, the line println(“System Availability:” + SysA) triggered an exception. Despite these shortcomings in code generation, the model produced structurally sound logic. When transcribed into Mercury, the underlying model yielded coherent results, indicating that the essential structure was correctly captured even though the implementation was flawed.

**Mistral 6.4** displayed the most critical shortcomings. Despite producing syntactically valid code, it failed to capture the correct behavior of a Hot-Standby system. It introduced a deadlock due to an improperly modeled Repair transition and an infinite failure-recovery cycle. The misdirection of repair tokens to the STANDBY place rather than UP compromised the model’s validity.

Table 1: Comparison of LLMs in Generating SPN Models

| Feature                | ChatGPT   | Gemini  | Claude  | Mistral  |
|------------------------|---|---|---|--|
| Syntax and Semantics   | Syntax errors in the metric and semantic issues compromise model logic.       | Syntax error in the metric, but semantics of the Hot-Standby model are well captured. | Syntax errors in the metric and semantic issues in code generation. | No syntax errors, but severe semantic flaws lead to incorrect Hot-Standby modeling and deadlock. |
| Complexity             | Capable of generating a model structure, though with logical faults.          | Produces a correct Hot-Standby model after minor syntactic corrections.               | Generates a Hot-Standby model with semantic issues in the code.     | Fails to generate a valid Hot-Standby model, introducing a deadlock.                             |
| Accuracy               | Low accuracy due to semantic and logical inconsistencies.                     | High accuracy after resolving minor syntax issues.                                    | Moderate accuracy, with semantic issues in implementation.          | Low accuracy; semantic flaws invalidate the model.   |
| Trends and Limitations | Tendency to produce semantically flawed models requiring manual intervention. | Generates mostly correct models with minor syntactic adjustments.                     | Constructs model structure but fails in implementation semantics.   | Struggles to generate complex models without introducing critical errors.                        |

In summary, all LLMs demonstrated at least a partial ability to generate SPN model structures based on a common modeling prompt. While semantic inconsistencies were observed across most outputs—including those from ChatGPT, Claude, and Mistral—each model captured relevant aspects of the system description. Gemini produced the model requiring the least intervention, limited to syntactic adjustments. Mistral, although free of syntax errors, introduced structural inconsistencies that affected the model’s correctness, highlighting the challenges LLMs face in capturing more intricate behavioral semantics. Table 1 provides a structured comparison of the key aspects evaluated in this study.

This study represents a first effort in evaluating the capability of large language models to support the dependability modeling process through the generation and expansion of formal models such as SPNs. By prompting LLMs to construct models based on a common fault-tolerant system description, we assessed their ability to interpret system-level requirements and express them in a formal, tool-compatible specification. While current limitations—especially semantic inconsistencies—still necessitate human oversight, the results highlight the potential of LLMs to assist in model prototyping and expansion during the early phases of system design. Future work may explore the integration of LLMs into modeling environments, where they can serve as co-modeling agents in more complex and iterative dependability analysis workflows.

## 8 Conclusion and Future Work

This paper presented an initial investigation into the applicability of Large Language Models (LLMs) to the generation of formal dependability models, specifically Stochastic Petri Nets (SPNs) for fault-tolerant systems. By providing a consistent modeling prompt to four state-of-the-art LLMs—ChatGPT, Gemini, Claude, and Mistral—, we analyzed the syntactic validity, semantic consistency, and correctness of the generated SPN models within the Mercury tool environment.

The results show that all LLMs were capable of generating structurally coherent SPN code but with varying degrees of syntactic

and semantic correctness. Gemini stood out as the most accurate model after minimal syntactic correction, while Mistral generated the only model free of syntax errors yet semantically incorrect due to deadlock and logical inconsistencies. ChatGPT and Claude produced models that reflected a partial understanding of the Hot-Standby concept but required post-processing to be functional and valid.

This evaluation represents a first step toward understanding how LLMs can support the dependability modeling process through code generation and model expansion. While promising for rapid prototyping and educational support, the use of LLMs in critical modeling workflows still requires cautious interpretation and manual oversight.

As future work, we plan to (i) investigate prompt engineering techniques tailored for formal modeling tasks; (ii) explore the fine-tuning of LLMs on domain-specific modeling corpora; (iii) embed validation and correction mechanisms within modeling environments such as Mercury; and (iv) expand the scope of evaluation to include more complex architectures and additional modeling formalisms.

## References

- [1] Anthropic. 2023. Claude: An AI assistant by Anthropic. <https://www.anthropic.com/index/introducing-claude>. Accessed: 2025-05-03.
- [2] Gianfranco Balbo. 1995. On the success of stochastic Petri nets. In *Proceedings 6th International Workshop on Petri Nets and Performance Models*. IEEE, 2–9.
- [3] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.
- [4] Google DeepMind. 2023. Gemini: A family of multimodal AI models. <https://deepmind.google/technologies/gemini/>. Accessed: 2025-05-03.
- [5] Peter W Glynn and Peter J Haas. 2012. On simulation of non-Markovian stochastic Petri nets with heavy-tailed firing times. In *Proceedings of the 2012 Winter Simulation Conference (WSC)*. IEEE, 1–12.
- [6] Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, Léo Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timothée Lacroix, and William El Sayed. 2023. Mistral 7B. *arXiv preprint arXiv:2310.06825* (2023). <https://arxiv.org/abs/2310.06825> Accessed: 2025-05-03.

- [7] Lukasz Kaiser, Aidan N Gomez, Noam Shazeer, Ashish Vaswani, Niki Parmar, Llion Jones, and Jakob Uszkoreit. 2017. One model to learn them all. *arXiv preprint arXiv:1706.05137* (2017).
- [8] L. Kleinrock. 1976. Queueing Systems, Volume 1: Theory. Wiley (1976).
- [9] Paulo Maciel. 2023. *Performance, Reliability, and Availability Evaluation of Computational Systems*. Taylor & Francis.
- [10] Paulo Maciel. 2023. *Performance, Reliability, and Availability Evaluation of Computational Systems, Volume II*. Taylor & Francis.
- [11] Paulo Maciel, Rubens Matos, Bruno Silva, Jair Figueiredo, Danilo Oliveira, Iure Fé, Ronierison Maciel, and Jamilson Dantas. 2017. Mercury: Performance and Dependability Evaluation of Systems with Exponential, Exponential, and General Distributions. In *2017 IEEE 22nd Pacific Rim International Symposium on Dependable Computing (PRDC)*. 50–57. doi:10.1109/PRDC.2017.16
- [12] M. Ajmone Marsan, G. Balbo, G. Conte, S. Donatelli, and G. Franceschinis. 1998. Modelling with Generalized Stochastic Petri Nets. *SIGMETRICS Perform. Eval. Rev.* 26, 2 (Aug. 1998), 2. doi:10.1145/288197.581193
- [13] Danilo Oliveira, Rubens Matos, Jamilson Dantas, João Ferreira, Bruno Silva, Gustavo Callou, Paulo Maciel, and André Brinkmann. 2017. Advanced stochastic petri net modeling with the mercury scripting language. In *Proceedings of the 11th EAI International Conference on Performance Evaluation Methodologies and Tools*. 192–197.
- [14] OpenAI. 2023. ChatGPT: GPT-4 model. <https://openai.com/chatgpt>. Accessed: 2025-05-03.
- [15] Kishor S. Trivedi. 2006. *Probability and Statistics with Reliability, Queueing, and Computer Science Applications*. Wiley.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009

---

# Shade-aware Routing For Sunburn Prevention

Levan Gogorikidze

Tbilisi State University

Tbilisi, Georgia

levan.gogorikidze874@ens.tsu.edu.ge

Giorgi Svanadze

Tbilisi State University

Tbilisi, Georgia

giorgi.svanadze070@ens.tsu.edu.ge

Magda Tsintsadze

Tbilisi State University

Tbilisi, Georgia

magda.tsintsadze@tsu.ge

## ABSTRACT

Walking is a widely favored mode of travel, but pedestrians face significant exposure to ultraviolet radiation, which can cause skin discoloration and irritation. UV exposure can exceed healthy levels all year round, leading to short-term discomfort and long-term damage to the skin. The public is often advised to seek shade to avoid prolonged exposure to UV radiation. The practicality of this recommendation for pedestrians is not obvious as the shade coverage is highly variable in urban environments depending on the date and time. To test the feasibility of shade-aware routing, we model the approximate shade coverage with OpenStreetMap data and simulate pedestrian paths that minimize sun exposure. We quantify the opportunities to use shaded navigation in urban environments as it relates to avoidance of UV radiation. We show that it is useful to seek shade in urban environments even in situations where the shade coverage is the lowest. As an extreme case, we simulate shade on the day the UV index was the highest throughout 2024 in Paris, we show that it was 43% safer for pedestrians of skin type I to maximize shade as opposed to distance. Safe walking distance increased by 15%. Even on the day of minimal shade throughout the year, shade-aware routing was 34% safer compared to shortest-path routing, increasing safe walking distance by 4%.

## KEYWORDS

Shade-aware Routing, Sunburn Prevention, UV radiation

## 1 INTRODUCTION

Most trips in urban areas are taken by cars and public transport. Walking is mainly reserved for short distance travel. Approximately 15% of the population walks or cycles for 30 minutes or more each day in Europe [6]. In some cities, walking is the most common mode of daily transport [7]. There are many factors that contribute to the comfort of the pedestrian, such as the presence of stores or drinking fountains [12], and more importantly, the thermal control, which is mostly influenced by the sun exposure and wind [20]. Pedestrians avoid excessive UV radiation as it is harmful to the skin. Failure to do so can result in sunburns and long-lasting discomfort. In addition, there is a health risk associated with excessive sun exposure, as a significant portion of melanoma cases can be attributed to UV radiation [2].

Ultraviolet radiation (UV) is solar radiation with wavelengths of 10-400nm. This range is further divided, UVB radiation of wavelength range 280-315nm being of most interest to us when discussing erythema. Erythema is inflammation, redness sometimes caused by excessive exposure to sun's radiation. The susceptibility of a subject to skin damage is dependent on many variables, but the skin tone is the property used when determining the minimal dose of radiation required to cause discoloration. This value is called the

Minimal Erythema Dose. It is measured in  $mJ/cm^2$ . Fitzpatrick skin phototyping is a method of classifying people with white skin [8]. Types I-VI experience sunburns with different times of exposure to UVA and UVB radiation. In various studies the participants are classified with this scale and then subjected to solar simulators, the results showing high correlation between Fitzpatrick skin type and Minimal Erythema Dose for the subject [19].

UV radiation reaching the Earth's surface is measured in  $mW/m^2$ , or in UV index where 1 UVI unit =  $25mW/m^2$ . Radiation data is collected and archived by Royal Netherlands Meteorological Institute (KNMI) [9]. UV radiation at earth's surface is measured, then the UV spectrum is weighted, with more weight given to UVB radiation and less to UVA radiation. UVI value is approximated at local solar noon, which is the point in time when the sun's altitude is the highest. Compiled yearly archives for the entire earth can be easily downloaded.

One way to block UV rays is to increase shade cover. Shade cover can be created by trees, building awnings, and tall structures. according to [17], tree and other canopy shade might not be entirely sufficient to protect skin because of the scattered and reflected UV radiation. Through their testing they determined that the tree cover is insufficient in blocking UV rays, the same results were observed for building awnings. Shade cast from Dense materials such as concrete protect against UV radiation better, providing heavy shade, but subjects are still exposed to sky radiation and reflected UV radiation.

In high UV-index situations, public is advised to seek shade. To test the usefulness of this recommendation in urban environments, we performed a series of experiments. The simulations are run assuming the absence of sunscreen, which offers the best protection against UV radiation. We design efficient shade-seeking agents and compare their ability to avoid skin damage with a control group of efficient distance-minimizing agents. To illustrate the differences, we introduce simple metrics. Popular navigation assistants (Google, Apple) do not account for shade, instead opting to optimize distance. Efficient distance-minimizing navigation is readily available to the public, while shade-seeking routing is not.

## 2 LITERATURE REVIEW

Numerous projects have explored the shade and thermal comfort of walkways in urban environments. [10] introduces a metric called Cool Walkshed Index (CWI) to evaluate walkways in the city, suitable for urban planning applications to improve pedestrian thermal comfort. Additionally, the work describes existing walkway networks, showcasing the social, economic and environmental benefits of sustainable transportation.

Wolf et al. uses building footprint datasets provided by certain cities [21]. city is assumed to be flat, and shadows are cast with a simple and performant geometric algorithm. Such maps are then

used to run navigation simulations, the results of which are averaged and a metric called CoolWalkability is calculated, which quantifies the opportunities to use shade in routing throughout the day. It differs from the overall shade cover metric and is not dependent on the personal preference of sun avoidance. The geometries of the cities and road networks are studied in context of sun avoidance.

Many existing software solutions approximate shade. Most of these implementations use OpenStreetMap (OSM) data [16] for road network information. Some further process the network to synthesize pedestrian routes or to reduce the number of nodes. Road networks are turned into graphs where many standard routing algorithms can be used for planning. The edges of these graphs are weighted sums of sunny and shaded lengths of the road segments.

Alternative solutions are also available. "Parasol Navigation" [13] uses a pre-downloaded Lidar data (available for some states of USA) and sun.r package to approximate a shade map, then using Dijkstra's algorithm, they navigate through the graph with a weighted cost function. Lidar height information is very expensive to collect, so the data does not have universal coverage. One interesting way to extract shade information is to use satellite imagery and machine learning vision classification models [4]. This method is computationally expensive and inflexible. Satellite images need to be high-resolution and time-stamped. One such dataset is the NAIP imagery dataset, but it only collects data in some parts of the United States and in a very limited time range [14].

After acquiring the data, there are many ways to calculate shade, WebGL and its wrappers like Three.js [3] allow us to cast shadows on 3d environments, which some projects have demonstrated [18]. [15] calculates shade by assuming all buildings are tall enough to cast shadows and uses mathematical approximations for shade coverage evaluation. This assumption is made because building height information is not universal in open datasets. [5] surveys the mobile apps that help pedestrians plan cool routes.

The projects we surveyed do not consider long-distance travel. Typical pedestrian trips have an average length of 3 km [7], so it might be beneficial to create a procedure to account for the change of shade distribution on the routes.

## 3 METHODOLOGY

### 3.1 Data processing

We show the complete pipeline for shade-aware routing in Figure 1. The process starts with data extraction. OSM Road and Buildings data we access is made of ways, polygons and nodes, which we process to create a model of the city and cast shade on the roads. The walkway network and building data is annotated with tags. The tags are hand picked by the authors according to the relevancy to our problem. A graph is constructed, weighted according to the shade cast on street segments. Using the standard routing algorithms, such as A\* and Dijkstra, we optimize routes between pairs of nodes. The results are then aggregated and the metrics are calculated

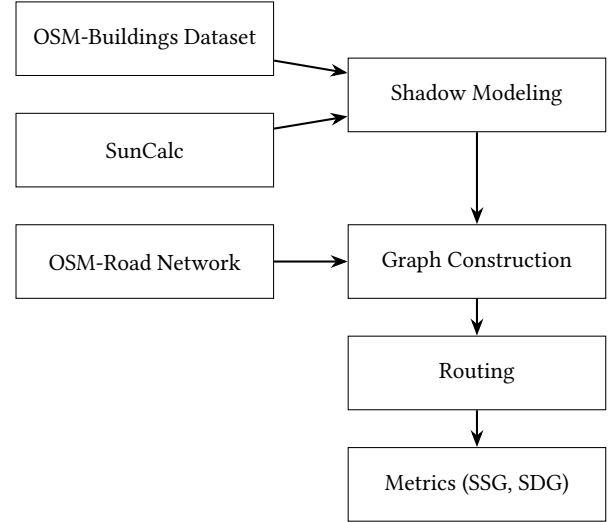


Figure 1: Pipeline overview of shade-aware routing process.

### 3.2 Sun Position and Shadow Geometry

Sun position is found using JS SunCalc library [1], which incorporates astronomical calculations for stellar bodies. The sun vector is created according to azimuth and altitude of the Sun at the solar noon of input day.

For the shade calculation we take the footprint of the building and raise it to the building height, then we descend the shape of the building along the sun vector until  $z = \text{ground\_plane}$  is reached (the horizontal ground surface is raised for calculations), then the footprint and the descended shape are connected with triangles. This projection gives accurate shadows only if buildings have no vertical variation. We observed that objects in OSM data are highly rectangular in general, the assumption holding for most buildings.

### 3.3 Road-Shade Intersections

The shade array, calculated from the buildings array is a collection of triangles, which are added into a grid with 100x100 meter cells, so that the road segments only have to be checked for intersections in their vicinity. The triangles are chosen as the primitive to represent shade due to the ease of performing intersection checks. Using shadow triangles array and roads graph adjacency matrix we calculate how shaded each walkway is. The method used to find the shaded road segments uses several geometric algorithms. After the comparison of axis-aligned bounding boxes for each triangle, we discard ones that do not intersect with the bounding box of the road. Cross-product signs are used to determine if the point of the road lies inside the shadow triangle. Road and triangle edge intersections are found using a determinant based formula. After finding the intersection points of the road and each triangle, they are collected along the parameterized road. These values are sorted, duplicate points are removed using a small tolerance for floating-point precision. Overlapping intervals are merged to form a continuous shaded segment. The total length of the shaded segment is

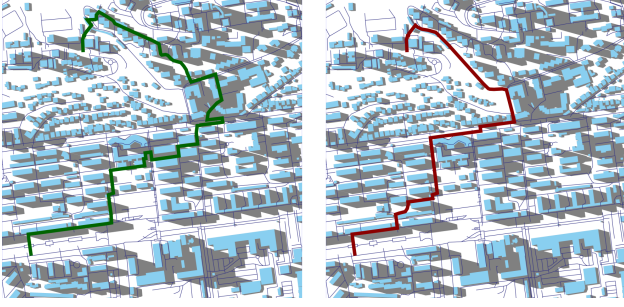


Figure 2: Example of shade-aware and shortest-path routes

calculated by summing the merged intervals. The fraction of road covered by shadow is computed as the ratio  $S$ ,  $S_{i,j} \in [0, 1]$  of the shadow length to the total road length. Finally, we build a new matrix containing the shading ratios for each road. Road network shade coverage is the sum of shaded lengths of all edges divided by the sum of all edge lengths in the constructed graph.

### 3.4 Routing

As described above, using OSM building data and a simple geometric shading algorithm we find all shade in the city for a fixed timestamp  $t$ , which we omit from notation for brevity. We then construct a graph  $G(V, E)$ ,  $E = (w)_{i,j}$ .

$$w_{i,j} = \alpha * S_{i,j} * d_{i,j} + (1 - S_{i,j}) * d_{i,j}$$

$d_{i,j}$  is real distance of the graph edge and  $\alpha$  is sun avoidance, which is a personal preference for a pedestrian. This means the pedestrian is content to lengthen their route by factor  $\alpha$  for the benefits of sun avoidance.

For long routes, where the shade changes over the traversal duration, many graphs can be constructed  $[S^{t_1}, S^{t_2}, S^{t_3} \dots]$ . In our calculations, walkable area is limited to less than  $3km^2$ , where the change in the direction of the sun is insignificant enough to be completely ignored. We also disregard cases where the agent wonders as the agents are penalized for traversing unnecessary distances, and the city networks are designed for connectivity.

We then use standard optimization algorithms such as Dijkstra and A\* to minimize the distance in the reweighted graph.

### 3.5 UV Exposure Modeling

Let  $T$  be the maximum time a person can stay exposed in the sun without visible damage to the skin.  $T$  depends on the skin type of the person and the UV radiation at the location at that particular time. It should be noted that reporting exposure doses in time units is generally not recommended as sensitivity to sun radiation is variable and aggregate values might mislead the public. We calculate these times  $T$  from Minimal Erythral Dose values from the original Fitzpatrick paper, concretely, we set 30, 35, 50, 60  $mJ/cm^2$  max doses for skin types I-IV. We use the maximal values of the ranges to underestimate the damage caused to the skin in some fixed time. The source of these ranges is a study in which the values are aggregated from a survey with self-reported values.

### 3.6 Shade Safety Gain (SSG)

SSG is a metric that measures how safe shade-aware routing is in a city compared to shortest path routing.

We sample source/destination node pairs  $(i, j)$  from a city randomly. For each path we fix the skin type of the simulated pedestrian, we minimize the length of the path (or the total exposure to the sun in case of shade-aware routing), and we measure if the amount of radiation the pedestrian received exceeded the minimal erythral dose.

$$k_{shortest-path}^{(i,j)} = \begin{cases} 1 & M_{shortest-path}^{(i,j)} / \text{Speed} \leq T \\ 0 & \text{else} \end{cases}$$

$M_{shortest-path}^{(i,j)}$  is summed length of all sunny segments on the shortest path from node  $i$  to  $j$ .  $k_{shade-aware}^{(i,j)} \in \{0, 1\}$  is calculated in similar a way. We then average the route safety measure.

$$K_{shortest-path} = \sum_{i \in src, j \in dest} k_{shortest-path}^{(i,j)}$$

With  $K_{shortest-path}$  and  $K_{shade-aware}$  the Shade Safety Gain is calculated.

$$SSG = \frac{K_{shade-aware} - K_{shortest-path}}{K_{shade-aware}}$$

### 3.7 Shade Distance Gain (SDG)

SDG is a metric that measures how much more Euclidean distance can be safely traversed in the city if the pedestrian diverges from the shortest routes.

We calculate the average distance between start/end points in the city where the path is safely traversable using shortest-path and shade-aware routing.

$$L_{shortest-path} = \sum_{i \in src, j \in dest} \text{dist}(i, j) * k_{shortest-path}^{(i,j)}$$

$L_{shade-aware}$  is also averaged and Shade Distance Gain is calculated.

$$SDG = \frac{L_{shade-aware} - L_{shortest-path}}{L_{shade-aware}}$$

## 4 LIMITATIONS

Tree shade is an important portion of shade in the green cities, we do not take vegetation into account when calculating shade because of non-uniform coverage in public datasets. We assume that the cities are flat. This helps us calculate shadows in a more efficient manner. A rough approximation can be made with algorithms discussed in the literature review section, and by overlaying our city shade with elevation shade, we can achieve a reasonable trade-off between efficient and exact computation.

We are also limited by the accuracy of data in OSM. other map providers seem to give better descriptions of the building shapes, whereas buildings in OSM are highly rectangular with minimal variation. When calculating the metrics, we choose the source / destination points randomly in the city, This might skew the results in favor of cities with highly shaded minor walkways. We also assume that the agents don't discriminate between paths for any reason other than shade coverage. The way pedestrians choose

routes is a complex process, modelling of which exceeds the scope of this project. It should be noted that the measurement of MED and skin photosensitivity is not a consistent process, there are many variables, which are hard to account for. Measurements are also subject to errors because of the skin developing resistance and calibration of measuring tools and machinery.

## 5 RESULTS

### 5.1 Experiment setup

We assume that all buildings are made of thick concrete and cast a heavy shadow, blocking most of UV radiation. There is still a considerable amount of UV radiation that can be incurred in the shade so we set 0.7 as the blocking ratio. Metrics are calculated at solar noon when the altitude is the highest throughout the day, and thus the agents have the least amount of shade possible to work with. For each location, we choose the dates that had the highest UVI throughout the year 2024. In addition, we also run simulations for days when the sun altitude was the highest throughout the year. The dates are given in Table 4. Two chosen days often overlap with small differences. For cities between latitudes 23.4 and  $-23.4$  there is a day every year when the sun altitude reaches 90 degrees. This day is known as Zero Shadow Day or Lahaina Noon. In Table 1, the latitudes for such cities are in bold. The shade coverage for such cities on the day with maximum sun altitude is essentially zero. For other cities, the shade cast by buildings is minimal on solstices, the shade coverage on this day is non-zero and thus the shade-seeking agents can have an advantage.

For shade-seeking agents, shade-avoidance factor of 2 is used. For route simulations, we assume that pedestrians move with a comfortable constant speed of 1.4 meters per second [11]. With a naive intersection check, in extreme cases, only a small part of human body might be in shade. We observed that more often than not the walkways are adjacent to the buildings in the datasets, so an assumption can be made that pedestrians can safely move into more shade, but we instead chose to compensate for average human height of 1.8m when calculating shade coverage of roads. Sample size of 3000 (source/destination) was used for every city. Routes were calculated within 1.5km radius from center of the city.

### 5.2 Results Analysis

We calculated SSG and SGD for skin types I through IV for various cities. The results are given in Table 1.

We observed that shade coverage is tied to the geometry of the city, this was shown in prior works too. We illustrate this relation in 3, where we display the shade coverage for cities modeled with empirical height information and constant 3m heights on days of maximum UV and maximum altitude. The shade seeker's advantage is almost nonexistent in cities where there is no building height information or where the buildings are not tall enough to cast adequate shade.

The results suggest that shade-seeking agents have an advantage compared to length-minimizing agents even in the most extreme and disadvantageous conditions. In urban environments outside the tropics, seeking shade is always a good strategy to minimize UV exposure. For cities between latitudes 23.4 and  $-23.4$  it is

not entirely unfruitful to seek shade for most days, except for the day when the sun's altitude is 90 degrees.

## 6 CONCLUSION AND FUTURE WORKS

UV radiation is a source of danger, especially for common sustainable transportation methods. In order to determine how useful urban shading is for prevention of skin damage, we ran simulations in different cities and introduced metrics to compare shade-aware and distance-minimizing navigation. The findings outline the circumstances in which shade-aware routing has a safety advantage. Considering current limitations of the simulation, such as not taking vegetation and elevation into account, incomplete building data for some regions, and simplicity of OSM building shapes, future work may focus on finding better data sources and improving the simulation with more complex systems making assumptions about the data. The framework presented can be used with minor modifications to plan long-distance routes. An important advance would be modeling the comfort of the pedestrian, which requires simulating a set of complex systems that cannot be efficiently approximated, unlike the exposure to UV radiation. A more data-centric approach can be taken to measure pedestrian satisfaction along the routes planned with different algorithms, taking into account variables such as thermal comfort, exposure to UV radiation, and more. These tools can then be used for real-time navigation in urban environments if the system is modeled and validated with personalized data, planning more pleasant and safer routes. Such models may be used for sustainable urban planning, creating more shade protection in cities for pedestrians and cyclists.

Table 1: SSG values for skin types I-IV

| City          | LAT    | Max UVI |      |      |      |      |      |      |      | Max altitude |      |      |      |      |      |      |      |
|---------------|--------|---------|------|------|------|------|------|------|------|--------------|------|------|------|------|------|------|------|
|               |        | SSG     |      |      |      | SDG  |      |      |      | SSG          |      |      |      | SDG  |      |      |      |
|               |        | I       | II   | III  | IV   | I    | II   | III  | IV   | I            | II   | III  | IV   | I    | II   | III  | IV   |
| Helsinki      | 60.17  | 0.2     | 0.15 | 0.17 | 0.14 | 0.12 | 0.11 | 0.07 | 0.05 | 0.19         | 0.13 | 0.17 | 0.13 | 0.07 | 0.05 | 0.01 | 0.01 |
| Oslo          | 59.91  | 0.31    | 0.21 | 0.28 | 0.21 | 0.1  | 0.1  | 0.02 | 0.03 | 0.26         | 0.19 | 0.21 | 0.16 | 0.04 | 0.04 | 0.0  | 0.01 |
| Tallinn       | 59.44  | 0.12    | 0.08 | 0.1  | 0.08 | 0.07 | 0.05 | 0.04 | 0.04 | 0.11         | 0.07 | 0.11 | 0.08 | 0.07 | 0.06 | 0.02 | 0.02 |
| Stockholm     | 59.33  | 0.33    | 0.24 | 0.28 | 0.21 | 0.1  | 0.09 | 0.03 | 0.03 | 0.3          | 0.21 | 0.21 | 0.16 | 0.05 | 0.05 | 0.01 | 0.01 |
| Riga          | 56.95  | 0.08    | 0.06 | 0.06 | 0.05 | 0.03 | 0.03 | 0.02 | 0.02 | 0.06         | 0.05 | 0.04 | 0.04 | 0.02 | 0.02 | 0.01 | 0.01 |
| Moscow        | 55.76  | 0.15    | 0.1  | 0.15 | 0.1  | 0.1  | 0.07 | 0.05 | 0.04 | 0.16         | 0.1  | 0.13 | 0.08 | 0.08 | 0.06 | 0.03 | 0.03 |
| Minsk         | 53.9   | 0.1     | 0.06 | 0.09 | 0.06 | 0.06 | 0.05 | 0.03 | 0.03 | 0.06         | 0.03 | 0.06 | 0.04 | 0.04 | 0.03 | 0.01 | 0.01 |
| Berlin        | 52.52  | 0.28    | 0.18 | 0.26 | 0.17 | 0.14 | 0.12 | 0.05 | 0.05 | 0.26         | 0.18 | 0.25 | 0.17 | 0.14 | 0.11 | 0.05 | 0.05 |
| Warsaw        | 52.23  | 0.29    | 0.2  | 0.29 | 0.21 | 0.17 | 0.14 | 0.08 | 0.07 | 0.31         | 0.21 | 0.29 | 0.2  | 0.17 | 0.13 | 0.08 | 0.07 |
| London        | 51.51  | 0.17    | 0.12 | 0.15 | 0.1  | 0.08 | 0.07 | 0.04 | 0.04 | 0.12         | 0.09 | 0.1  | 0.07 | 0.03 | 0.03 | 0.0  | 0.0  |
| Astana        | 51.18  | 0.0     | 0.0  | 0.01 | 0.01 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Kiev          | 50.45  | 0.09    | 0.06 | 0.07 | 0.05 | 0.05 | 0.04 | 0.03 | 0.02 | 0.09         | 0.05 | 0.08 | 0.05 | 0.04 | 0.03 | 0.02 | 0.01 |
| Paris         | 48.85  | 0.43    | 0.3  | 0.36 | 0.27 | 0.15 | 0.14 | 0.05 | 0.05 | 0.34         | 0.26 | 0.24 | 0.2  | 0.04 | 0.04 | 0.0  | 0.0  |
| Vienna        | 48.21  | 0.22    | 0.14 | 0.19 | 0.12 | 0.16 | 0.12 | 0.09 | 0.07 | 0.21         | 0.14 | 0.19 | 0.13 | 0.16 | 0.12 | 0.09 | 0.07 |
| Bratislava    | 48.15  | 0.17    | 0.11 | 0.15 | 0.1  | 0.11 | 0.08 | 0.07 | 0.06 | 0.18         | 0.11 | 0.16 | 0.11 | 0.11 | 0.07 | 0.08 | 0.06 |
| Budapest      | 47.5   | 0.0     | 0.0  | 0.01 | 0.0  | 0.01 | 0.01 | 0.01 | 0.0  | 0.01         | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.0  |
| Zagreb        | 45.81  | 0.12    | 0.08 | 0.13 | 0.08 | 0.11 | 0.08 | 0.09 | 0.07 | 0.14         | 0.08 | 0.13 | 0.08 | 0.12 | 0.09 | 0.09 | 0.07 |
| Ottawa        | 45.42  | 0.19    | 0.14 | 0.19 | 0.13 | 0.16 | 0.12 | 0.1  | 0.08 | 0.2          | 0.14 | 0.18 | 0.13 | 0.12 | 0.1  | 0.08 | 0.07 |
| Bishkek       | 42.87  | 0.04    | 0.02 | 0.03 | 0.02 | 0.03 | 0.02 | 0.02 | 0.01 | 0.03         | 0.01 | 0.05 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 |
| Sofia         | 42.7   | 0.14    | 0.1  | 0.13 | 0.09 | 0.1  | 0.08 | 0.06 | 0.05 | 0.14         | 0.09 | 0.13 | 0.09 | 0.11 | 0.08 | 0.07 | 0.05 |
| Tirana        | 41.33  | 0.01    | 0.01 | 0.02 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01         | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| Madrid        | 40.42  | 0.05    | 0.05 | 0.05 | 0.04 | 0.07 | 0.05 | 0.04 | 0.04 | 0.03         | 0.02 | 0.04 | 0.03 | 0.04 | 0.03 | 0.03 | 0.02 |
| Yerevan       | 40.18  | 0.1     | 0.06 | 0.09 | 0.05 | 0.07 | 0.05 | 0.04 | 0.03 | 0.11         | 0.06 | 0.08 | 0.05 | 0.06 | 0.04 | 0.03 | 0.03 |
| P'yongyang    | 39.03  | 0.0     | 0.0  | 0.01 | 0.01 | 0.01 | 0.0  | 0.01 | 0.01 | 0.01         | 0.01 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Washington DC | 38.9   | 0.1     | 0.07 | 0.1  | 0.07 | 0.07 | 0.05 | 0.04 | 0.04 | 0.1          | 0.06 | 0.1  | 0.07 | 0.05 | 0.04 | 0.02 | 0.02 |
| Lisbon        | 38.72  | 0.02    | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.02         | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| Valletta      | 35.9   | 0.0     | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Rabat         | 34.01  | 0.03    | 0.02 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02         | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.0  | 0.0  |
| Dhaka         | 23.71  | 0.0     | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Hong Kong     | 22.28  | 0.11    | 0.08 | 0.09 | 0.08 | 0.06 | 0.06 | 0.05 | 0.05 | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Bangkok       | 13.72  | 0.1     | 0.06 | 0.1  | 0.07 | 0.08 | 0.06 | 0.05 | 0.04 | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Willemstad    | 12.11  | 0.0     | 0.0  | 0.01 | -0.0 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Caracas       | 10.49  | 0.0     | 0.0  | 0.01 | 0.01 | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| San José      | 9.93   | 0.0     | 0.0  | 0.0  | 0.0  | 0.01 | 0.01 | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Addis Ababa   | 9.02   | 0.0     | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Panama City   | 9.0    | 0.0     | 0.0  | 0.0  | 0.0  | 0.0  | -0.0 | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Kuala Lumpur  | 3.14   | 0.0     | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Singapore     | 1.29   | 0.05    | 0.03 | 0.03 | 0.02 | 0.03 | 0.02 | 0.03 | 0.02 | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Quito         | -0.23  | 0.03    | 0.03 | 0.03 | 0.02 | 0.01 | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Port Moresby  | -9.44  | 0.0     | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Lima          | -12.04 | 0.0     | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Mamoudzou     | -12.78 | 0.0     | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Lusaka        | -15.41 | 0.0     | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0          | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  | 0.0  |
| Cape Town     | -33.93 | 0.16    | 0.12 | 0.15 | 0.11 | 0.1  | 0.08 | 0.08 | 0.07 | 0.12         | 0.1  | 0.11 | 0.08 | 0.07 | 0.06 | 0.05 | 0.05 |
| Canberra      | -35.28 | 0.05    | 0.04 | 0.03 | 0.02 | 0.04 | 0.03 | 0.02 | 0.02 | 0.02         | 0.01 | 0.02 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 |
| Wellington    | -41.29 | 0.01    | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01         | 0.01 | 0.01 | 0.01 | 0.0  | 0.0  | 0.01 | 0.0  |



**Table 2: Objects with height information attached in Open-StreetMaps**

| City          | Median height<br>* In available data | Objects | Availability |
|---------------|--------------------------------------|---------|--------------|
| Tirana        | 9m                                   | 5177    | 30.6%        |
| Yerevan       | 9m                                   | 5955    | 39.26%       |
| Canberra      | 6m                                   | 1525    | 32.52%       |
| Vienna        | 18m                                  | 6749    | 86.07%       |
| Dhaka         | 9m                                   | 27972   | 41.13%       |
| Minsk         | 12m                                  | 2430    | 75.19%       |
| Sofia         | 12m                                  | 8030    | 39.44%       |
| Ottawa        | 12m                                  | 7237    | 29.72%       |
| Hong Kong     | 36m                                  | 2908    | 59.73%       |
| San José      | 6m                                   | 4985    | 31.53%       |
| Zagreb        | 12m                                  | 6400    | 23.44%       |
| Willemstad    | 6m                                   | 1529    | 31.26%       |
| P'yongyang    | 15m                                  | 1714    | 46.91%       |
| Quito         | 6m                                   | 1313    | 75.25%       |
| Tallinn       | 13m                                  | 5876    | 80.94%       |
| Addis Ababa   | 3m                                   | 1494    | 34%          |
| Helsinki      | 15m                                  | 3303    | 57.28%       |
| Paris         | 18m                                  | 13240   | 64.82%       |
| Berlin        | 15m                                  | 5511    | 78.5%        |
| Budapest      | 12m                                  | 4279    | 56.32%       |
| Astana        | 3m                                   | 2352    | 44.81%       |
| Bishkek       | 6m                                   | 2890    | 37.02%       |
| Riga          | 9m                                   | 2761    | 56.28%       |
| Kuala Lumpur  | 12m                                  | 2439    | 29.36%       |
| Valletta      | 9m                                   | 2662    | 61.91%       |
| Mamoudzou     | 3m                                   | 9157    | 20.66%       |
| Rabat         | 15m                                  | 3387    | 29.44%       |
| Wellington    | 6m                                   | 6725    | 45.61%       |
| Oslo          | 15m                                  | 4979    | 71.8%        |
| Panama City   | 6m                                   | 4666    | 79.83%       |
| Port Moresby  | 6m                                   | 1364    | 91.2%        |
| Lima          | 6m                                   | 6429    | 29.8%        |
| Warsaw        | 12m                                  | 5592    | 95.05%       |
| Lisbon        | 12m                                  | 12752   | 25.74%       |
| Moscow        | 11.5m                                | 16774   | 91.42%       |
| Singapore     | 12m                                  | 4146    | 87.92%       |
| Bratislava    | 9m                                   | 7077    | 45.06%       |
| Cape Town     | 8.5m                                 | 5157    | 32.56%       |
| Madrid        | 18m                                  | 12219   | 33.31%       |
| Stockholm     | 15m                                  | 3726    | 86.82%       |
| Bangkok       | 12m                                  | 7682    | 43.58%       |
| Kiev          | 12m                                  | 3959    | 78.88%       |
| London        | 12m                                  | 10517   | 60.53%       |
| Washington DC | 24m                                  | 1908    | 47.01%       |
| Caracas       | 24m                                  | 971     | 25.23%       |
| Lusaka        | 3m                                   | 1797    | 45.85%       |

**Table 3: Road network shade coverage. Real heights on max UV day, Real heights on max altitude day and constant height of 3m on max UV day**

| City          | Real | Real(max Alt) | Const 3m |
|---------------|------|---------------|----------|
| Tirana        | 0.02 | 0.02          | 0        |
| Yerevan       | 0.04 | 0.04          | 0        |
| Canberra      | 0.02 | 0.01          | 0        |
| Vienna        | 0.11 | 0.11          | 0        |
| Dhaka         | 0    | 0             | 0        |
| Minsk         | 0.04 | 0.04          | 0        |
| Sofia         | 0.04 | 0.04          | 0        |
| Ottawa        | 0.04 | 0.04          | 0        |
| Hong Kong     | 0.07 | 0             | 0        |
| San José      | 0.01 | 0             | 0        |
| Zagreb        | 0.03 | 0.03          | 0        |
| Willemstad    | 0.01 | 0             | 0        |
| P'yongyang    | 0.02 | 0.01          | 0        |
| Quito         | 0    | 0             | 0        |
| Tallinn       | 0.1  | 0.1           | 0        |
| Addis Ababa   | 0.01 | 0             | 0        |
| Helsinki      | 0.09 | 0.09          | 0.01     |
| Paris         | 0.21 | 0.21          | 0.09     |
| Berlin        | 0.09 | 0.09          | 0        |
| Budapest      | 0.02 | 0.02          | 0        |
| Astana        | 0.01 | 0.01          | 0        |
| Bishkek       | 0.02 | 0.02          | 0        |
| Riga          | 0.05 | 0.05          | 0        |
| Kuala Lumpur  | 0.02 | 0             | 0        |
| Valletta      | 0.01 | 0             | 0        |
| Mamoudzou     | 0    | 0             | 0        |
| Rabat         | 0.01 | 0             | 0        |
| Wellington    | 0.04 | 0.03          | 0        |
| Oslo          | 0.13 | 0.13          | 0        |
| Panama City   | 0.01 | 0             | 0        |
| Port Moresby  | 0    | 0             | 0        |
| Lima          | 0    | 0             | 0        |
| Warsaw        | 0.1  | 0.1           | 0        |
| Lisbon        | 0.02 | 0.02          | 0.01     |
| Moscow        | 0.09 | 0.09          | 0        |
| Singapore     | 0.05 | 0.01          | 0.01     |
| Bratislava    | 0.07 | 0.07          | 0        |
| Cape Town     | 0.04 | 0.03          | 0.01     |
| Madrid        | 0.02 | 0.02          | 0        |
| Stockholm     | 0.13 | 0.13          | 0        |
| Bangkok       | 0.09 | 0             | 0        |
| Kiev          | 0.06 | 0.06          | 0        |
| London        | 0.11 | 0.11          | 0.02     |
| Washington DC | 0.04 | 0.04          | 0        |
| Caracas       | 0    | 0             | 0        |
| Lusaka        | 0    | 0             | 0        |

**Table 4: Max UV and Max altitude days used in calculations with their respective UVI**

| City          | max UV                  | max Altitude            |
|---------------|-------------------------|-------------------------|
| Tirana        | Jun 21 2024 (UVI:9.87)  | Jun 21 2024 (UVI:9.87)  |
| Yerevan       | Jun 20 2024 (UVI:10.99) | Jun 21 2024 (UVI:10.11) |
| Canberra      | Jan 29 2024 (UVI:15.1)  | Dec 22 2024 (UVI:12.05) |
| Vienna        | Jun 21 2024 (UVI:8.95)  | Jun 21 2024 (UVI:8.95)  |
| Dhaka         | May 27 2024 (UVI:13.49) | Jun 21 2024 (UVI:12.48) |
| Minsk         | Jul 14 2024 (UVI:7.14)  | Jun 21 2024 (UVI:6.25)  |
| Sofia         | Jun 21 2024 (UVI:9.86)  | Jun 21 2024 (UVI:9.86)  |
| Ottawa        | Jun 19 2024 (UVI:9.42)  | Jun 21 2024 (UVI:8.22)  |
| Hong Kong     | Aug 16 2024 (UVI:13.45) | Jun 03 2024 (UVI:12.15) |
| San José      | Mar 06 2024 (UVI:16.4)  | Apr 15 2024 (UVI:13.4)  |
| Zagreb        | Jun 20 2024 (UVI:9.63)  | Jun 21 2024 (UVI:9.45)  |
| Willemstad    | Mar 07 2024 (UVI:14.96) | Apr 21 2024 (UVI:14.02) |
| P'yongyang    | Jul 02 2024 (UVI:10.77) | Jun 21 2024 (UVI:8.77)  |
| Quito         | Mar 10 2024 (UVI:20.21) | Sep 23 2024 (UVI:14.15) |
| Tallinn       | Jun 27 2024 (UVI:5.59)  | Jun 21 2024 (UVI:4.93)  |
| Addis Ababa   | Mar 04 2024 (UVI:18.25) | Aug 30 2024 (UVI:13.78) |
| Helsinki      | Jul 02 2024 (UVI:5.41)  | Jun 21 2024 (UVI:4.71)  |
| Paris         | Jun 24 2024 (UVI:7.78)  | Jun 21 2024 (UVI:6.29)  |
| Berlin        | Jun 21 2024 (UVI:7.17)  | Jun 21 2024 (UVI:7.17)  |
| Budapest      | Jun 21 2024 (UVI:8.95)  | Jun 21 2024 (UVI:8.95)  |
| Astana        | Jun 28 2024 (UVI:7.66)  | Jun 21 2024 (UVI:6.76)  |
| Bishkek       | Jul 30 2024 (UVI:10.06) | Jun 21 2024 (UVI:8.59)  |
| Riga          | Jul 13 2024 (UVI:6.33)  | Jun 21 2024 (UVI:5.34)  |
| Kuala Lumpur  | Mar 06 2024 (UVI:17.35) | Mar 28 2024 (UVI:15.13) |
| Valletta      | Jun 12 2024 (UVI:10.89) | Jun 21 2024 (UVI:10.4)  |
| Mamoudzou     | Feb 26 2024 (UVI:17.3)  | Oct 27 2024 (UVI:12.62) |
| Rabat         | Jul 12 2024 (UVI:11.12) | Jun 21 2024 (UVI:10.56) |
| Wellington    | Jan 10 2024 (UVI:12.77) | Dec 22 2024 (UVI:10.72) |
| Oslo          | Jun 26 2024 (UVI:5.94)  | Jun 21 2024 (UVI:5.23)  |
| Panama City   | Mar 23 2024 (UVI:16.21) | Apr 12 2024 (UVI:14.25) |
| Port Moresby  | Feb 29 2024 (UVI:18.83) | Feb 25 2024 (UVI:16.99) |
| Lima          | Feb 27 2024 (UVI:18.5)  | Feb 17 2024 (UVI:16.7)  |
| Warsaw        | Jun 21 2024 (UVI:7.5)   | Jun 21 2024 (UVI:7.5)   |
| Lisbon        | Jul 11 2024 (UVI:9.93)  | Jun 21 2024 (UVI:8.71)  |
| Moscow        | Jun 29 2024 (UVI:6.57)  | Jun 21 2024 (UVI:6.17)  |
| Singapore     | Mar 06 2024 (UVI:17.65) | Mar 24 2024 (UVI:14.59) |
| Bratislava    | Jun 21 2024 (UVI:8.91)  | Jun 21 2024 (UVI:8.91)  |
| Cape Town     | Jan 14 2024 (UVI:14.28) | Dec 21 2024 (UVI:12.36) |
| Madrid        | Jul 18 2024 (UVI:9.78)  | Jun 21 2024 (UVI:8.89)  |
| Stockholm     | Jun 26 2024 (UVI:5.84)  | Jun 21 2024 (UVI:5.11)  |
| Bangkok       | Mar 11 2024 (UVI:14.82) | Aug 16 2024 (UVI:13.75) |
| Kiev          | Jun 23 2024 (UVI:7.91)  | Jun 21 2024 (UVI:7.04)  |
| London        | Jun 23 2024 (UVI:7.67)  | Jun 21 2024 (UVI:6.04)  |
| Washington DC | Jul 08 2024 (UVI:9.97)  | Jun 21 2024 (UVI:9.25)  |
| Caracas       | Mar 09 2024 (UVI:16.42) | Aug 25 2024 (UVI:12.82) |
| Lusaka        | Jan 25 2024 (UVI:17.47) | Nov 04 2024 (UVI:13.09) |

## REFERENCES

- [1] Vladimir Agafonkin. 2011. SunCalc: A tiny JavaScript library for calculating sun/moon positions and phases. <https://github.com/mourner/suncalc>
- [2] Melina Arnold, Esther de Vries, David C. Whiteman, Ahmedin Jemal, Freddie Bray, Donald Maxwell Parkin, and Isabelle Soerjomataram. 2018. Global burden of cutaneous melanoma attributable to ultraviolet radiation in 2012. *International Journal of Cancer* 143, 6 (2018), 1305–1314. <https://doi.org/10.1002/ijc.31527> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/ijc.31527>
- [3] Three.js Authors. 2010. Three.js – JavaScript 3D library. <http://threejs.org>.
- [4] Longchao Da, Rohan Chhibba, Rushabh Jaiswal, Ariane Middel, and Hua Wei. 2024. Shaded Route Planning Using Active Segmentation and Identification of Satellite Images. arXiv:2407.13689 [cs.CY] <https://arxiv.org/abs/2407.13689>
- [5] Kaveh Deilami, Julie Rudner, Andrew Butt, Tania MacLeod, Geoff Williams, Harmen Romeijn, and Marco Amati. 2020. Allowing Users to Benefit from Tree Shading: Using a Smartphone App to Allow Adaptive Route Planning during Extreme Heat. *Forests* 11, 9 (2020). <https://doi.org/10.3390/f11090998>
- [6] Eurostat. 2022. Walking and cycling at least 30 minutes per day by sex, age and educational attainment level. [https://doi.org/10.2908/HLTH\\_EHIS\\_PE6E](https://doi.org/10.2908/HLTH_EHIS_PE6E)
- [7] Davide Fiorello, Angelo Martino, Loredana Zani, Panayotis Christidis, and Elena Navajas-Cawood. 2016. Mobility Data across the EU 28 Member States: Results from an Extensive CAWI Survey. *Transportation Research Procedia* 14 (2016), 1104–1113. <https://doi.org/10.1016/j.trpro.2016.05.181> Transport Research Arena TRA2016.
- [8] Thomas B. Fitzpatrick. 1988. The Validity and Practicality of Sun-Reactive Skin Types I Through VI. *Archives of Dermatology* 124, 6 (06 1988), 869–871. <https://doi.org/10.1001/archderm.1988.01670060015008>
- [9] Royal Netherlands Meteorological Institute (KNMI). 2017?. Tropospheric Emission Monitoring Internet Service. [www.temis.nl](http://www.temis.nl)
- [10] Todd Litman. 2023. Cool walkability planning: Providing pedestrian thermal comfort in hot climate cities. *Journal of Civil Engineering and Environmental Sciences* 9, 2 (2023), 79–86. <https://doi.org/10.17352/2455-488X.000073>
- [11] Olfat Mohamed and Heather Appling. 2020. Clinical Assessment of Gait. In *Orthotics and Prosthetics in Rehabilitation (Fourth Edition)* (fourth edition ed.), Kevin K. Chui, Milagros "Millee" Jorge, Sheng-Che Yen, and Michelle M. Lusardi (Eds.). Elsevier, St. Louis (MO), 102–143. <https://doi.org/10.1016/B978-0-323-60913-5.00005-2>
- [12] Filipe Moura, Paulo Cambra, and Alexandre Gonçalves. 2014. IAAPE - Pedestrian accessibility and attractiveness assessment tool when planning for walkability. (10 2014).
- [13] All NaNs! 2018. Parasol Navigation: Optimizing Walking Routes to Keep You in the Sun or Shade. <https://www.allnans.com/jekyll/update/2018/08/07/introducing-parasol.html>. <https://github.com/keithfma/parasol>
- [14] U.S. Department of Agriculture. [n. d.]. NAIP: National Agriculture Imagery Program. <https://naip-usdaonline.hub.arcgis.com/>.
- [15] Cristina Olaverri Monreal, Matthias Pichler, Gerd Krizek, and Sebastian Naumann. 2016. Shadow as Route Quality Parameter in a Pedestrian-Tailored Mobile Application. *IEEE Intelligent Transportation Systems Magazine* 8, 4 (2016), 15–27. <https://doi.org/10.1109/ITS.2016.2605358>
- [16] OpenStreetMap contributors. 2017. Planet dump retrieved from <https://planet.osm.org>. <https://www.openstreetmap.org>.
- [17] Pete Parsons, Ray Neale, Penny Wolski, and Adèle C. Green. 1998. The shady side of solar protection. *Medical Journal of Australia* 168 (1998). <https://api.semanticscholar.org/CorpusID:36819592>
- [18] StrandedKitty. 2020. streets.gl: OpenStreetMap 3D renderer powered by WebGL2. <https://github.com/StrandedKitty/streets-gl/>.
- [19] Martha Valbuena, John Nova, and G. Sánchez. 2020. Minimal Erythema Dose: Correlation with Fitzpatrick Skin Type and Concordance Between Methods of Erythema Assessment in a Patient Sample in Colombia. *Actas Dermosifiliográficas (English Edition)* 111 (06 2020). <https://doi.org/10.1016/j.adengl.2019.12.010>
- [20] Carolina Vasilikou and Marialena Nikolopoulou. 2020. Outdoor thermal comfort for pedestrians in movement: thermal walks in complex urban morphology. *International journal of biometeorology* 64 (2020), 277–291.
- [21] Henrik Wolf, Ane Rahbek Viero, and Michael Szell. 2024. CoolWalks: Assessing the potential of shaded routing for active mobility in urban street networks. arXiv:2405.01225 [physics.soc-ph] <https://arxiv.org/abs/2405.01225>

# Towards Sustainability in 6G Network Slicing with Energy-Saving and Optimization Methods

Rodrigo Moreira  
rodrigo@ufv.br  
Universidade Federal de Viçosa (UFV)  
Viçosa, Minas Gerais, Brazil

Tereza C. M. Carvalho  
terezacarvalho@usp.br  
Universidade de São Paulo (USP)  
São Paulo, Brazil

Flávio de Oliveira Silva  
flavio@di.uminho.pt  
Universidade do Minho  
Braga, Portugal

Nazim Agoulmine  
nagoulmine@gmail.com  
Université Paris-Saclay - Évry  
Évry, France

Joberto S. B. Martins\*  
Universidade Salvador (UNIFACS)  
Salvador, Brazil  
joberto.martins@gmail.com

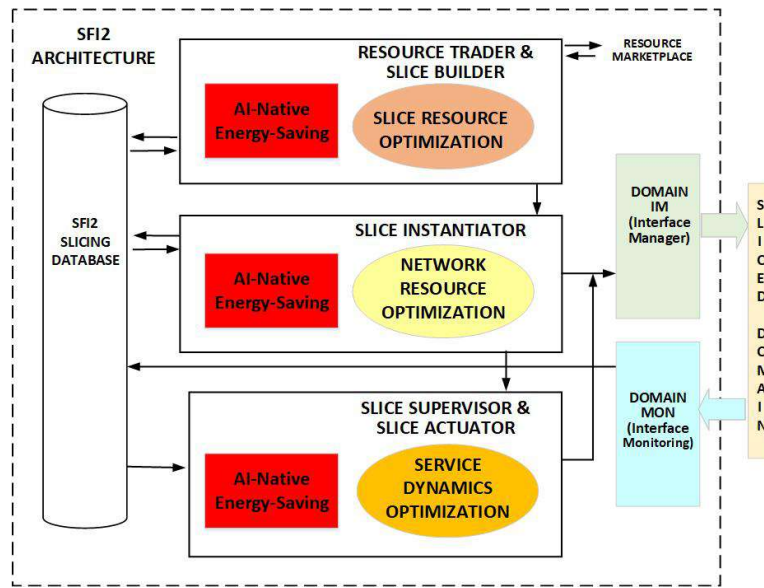


Figure 1: TEASER: 6G Energy-Saving Methods for Network Slicing Architectures

## ABSTRACT

The 6G mobile network is the next evolutionary step after 5G, with a prediction of an explosive surge in mobile traffic. It provides ultra-low latency, higher data rates, high device density, and ubiquitous coverage, positively impacting services in various areas. Energy saving is a major concern for new systems in the telecommunications sector because all players are expected to reduce their carbon footprints to contribute to mitigating climate change. Network slicing is a fundamental enabler for 6G/5G mobile networks and various other new systems, such as the Internet of Things (IoT), Internet of Vehicles (IoV), and Industrial IoT (IIoT). However, energy-saving methods embedded in network slicing architectures are still a research gap. This paper discusses how to embed energy-saving methods in network-slicing architectures that are a fundamental enabler for nearly all new innovative systems being deployed worldwide. This paper's main contribution is a proposal to save energy

in network slicing. That is achieved by deploying ML-native agents in NS architectures to dynamically orchestrate and optimize resources based on user demands. The SFI2 network slicing reference architecture is the concrete use case scenario in which contrastive learning improves energy saving for resource allocation.

## CCS CONCEPTS

• **Computing methodologies** → Machine learning; Distributed artificial intelligence; Multi-agent systems; Cooperation and coordination; • **Networks** → Programmable networks; Network management; Network design principles; Network dynamics.

## KEYWORDS

6G, Energy-Efficiency, Sustainability, Network Slicing, SFI2, Contrastive Learning

\* All authors contributed equally to this research.

# 1 INTRODUCTION

The 6G mobile network is the next step toward a new generation of high-speed mobile services, bringing new capabilities such as ultra-fast connection at terabit/second, high reliability, and ultra-low latency, support to new types of communication services such as immersive holographic and virtual reality communications, tactile communications, ubiquitous connectivity encompassing air, ground, and sea, self-sustainable network, etc. [1]. The 6G will be globally disseminated and improve support for many productive sectors and verticals. Its global reach is to positively impact many future businesses as an enabler for highly innovative services, and its adoption is a matter of time [2].

Regarding the 6G telecommunications industry and climate change mitigation, most actions seek more sustainable solutions through carbon footprint reduction. In effect, carbon footprint reduction is one of the various approaches for climate change mitigation that has been extensively considered by various areas, including the telecommunications sector [3].

In the 6G telecommunications sector, energy consumption optimization and reduction are two of the main contributors to reducing greenhouse gas emissions and, as such, contributing to climate change mitigation (Figure 2) [4]. As such, energy-efficient solutions are the main target and approach to promoting sustainable telecoms solutions. In this regard, the telecommunications sector is aligned with climate change mitigation by defining and promoting a net zero<sup>1</sup> target in which the telecommunications sector's carbon footprint is drastically reduced [5]. Complementary to the net zero targets, new key performance indicators (KPIs) are necessary to monitor the outcomes of carbon footprint solutions based on energy consumption reduction [6].

<sup>1</sup>Zero carbon-dioxide emissions

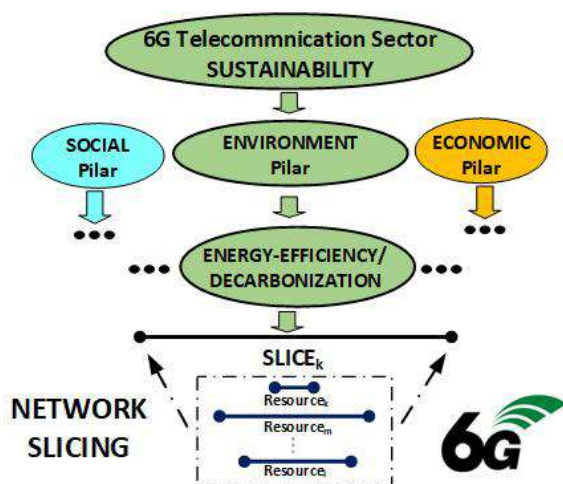


Figure 2: 6G Telecommunication Sector - Sustainability.

However, in the current 6G research and development scenario, there is a gap in exploring opportunities for defining more ambitious new energy-saving solutions and energy-efficiency (EE) key performance indicators (KPIs).

This paper explores and addresses the gap in new energy-saving methods for 6G mobile services by focusing on the network slicing architectures and capabilities inherently necessary to support the 6G infrastructure. Network slicing is a fundamental component of 6G infrastructure that enables the multiplexing of virtualized and independent logical networks on the same physical network infrastructure. This work is, therefore, fundamental since promoting new energy-saving methods for slicing addresses the sustainability of 6G services.

Network slicing (NS) is, therefore, a cornerstone for 6G mobile network deployment because it supports virtual networks' planning, commissioning, configuration, operation, and management of all phases [7]. By virtualizing physical and virtual resources like machines, communication links, switches, and radio access networks (RAN), among others, Network slicing allows service customization and infrastructure optimization that is crucial for 6G networks sustainable energy-saving mechanisms [8] [9] [10].

Network Slicing is a flexible approach to provide the optimization capabilities that are necessary to address the highly dynamic and variable requirements imposed by future 6G mobile users [11], future Vehicular Networks [12], experimental networks [13], industrial IIoT [14] and so on. Thanks to its capabilities to optimize and customize the network delivered to the end users, energy-saving should be addressed accordingly to reduce the CO2 footprint in these various areas.

This paper's objective is, therefore, to present the energy-efficient solutions that are envisioned in the context of the Slicing Future Internet Infrastructures (SFI2) reference architecture<sup>2</sup>. It benefits from its innovative characteristics as described in Martins et al. in [15].

This paper is organized as follows: The introduction section 1 presents the vision of future telecommunications networks and highlights the need for energy-saving mechanisms for 6G mobile networks. Sections 2 and 3 present how the sustainability goal is achieved using energy-saving mechanisms for 6G deployments that are highly dependent on network slicing as an enabler. Section 3 complementary proposes an ML-embedded energy-saving solution in the network slicing architecture frame. Section 4 presents a use case to save energy in network slicing architectures, having SFI2 architecture as the architectural reference. Finally, Section 5 closes the discussion with the final considerations.

## 2 6G SLICING WITH ENERGY-SAVING FOR SUSTAINABILITY

6G networks must provide enhanced speeds, global coverage, improved service capacity, high reliability, and ultra-low latency, to name a few stringent capabilities and requirements, while minimizing energy consumption. Sustainability through energy-efficient methods is a foundational element of 6G research and design, which implements new technologies and architectures with a sustainability mindset [16].

<sup>2</sup><https://sites.google.com/view/sfi2/home>

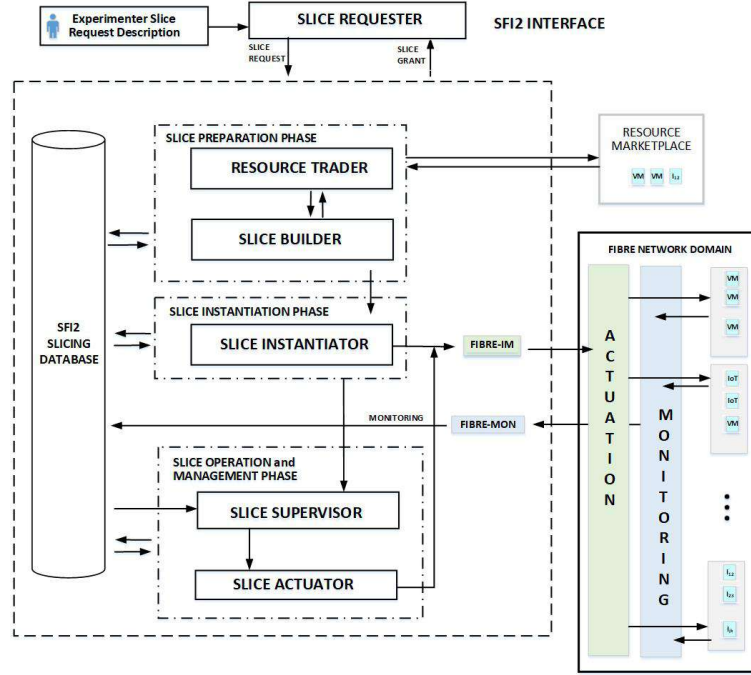


Figure 3: The Slicing Future Internet Infrastructures (SFI2) Network Slicing Reference Architecture [13].

Standardization institutions like 3GPP (3rd Generation Partnership Project), ITU-T (International Telecommunications Union - Telecommunications), and ETSI (European Telecommunications Standards Institute) have proposed NS architectures with various features and distinct target domains [15]. Although this plurality of options, network slicing architectures have a common group of functional phases. The 3GPP network slicing standard and initiative defines the following phases [7]:

- The *Preparation Phase* in which the user's slice requests are received, and the necessary resources are identified and localized in the target domain.
- The *Commissioning Phase* in which the NS service provider makes choices and orchestrates among the available resources aiming to configure the requested slice.
- The *Operation Phase* in which the deployed slice is operational and dynamic orchestration of resources may occur due to the user's time-varying demands and traffic fluctuation.
- The *Decommissioning Phase* in which the allocated slice resources from single or multi-domains are liberated.

The SFI2 project [13] defines the SFI2 network slicing reference architecture to create and manage multi-domain and multi-technology end-to-end slicing services. This architecture incorporates advanced network slicing concepts like ML-native optimizations, energy-efficient slicing, and slicing-tailored security functionalities [17] [18] as highlighted in Figure 3.

In the SFI2 architecture, the slice builder component builds the requested slice, considering the available resources in the domain

marketplace, and optimizes their utilization in each built slice. In sequence, the slice instantiation component deploys the configured slice in the target and optimizes the resources made available to the provider's network. At this phase, the optimization concerns the ensemble of resources available by the slicing service provider. Once the slice is deployed, the SFI2 slice supervisor component manages the slice's operation, allowing slice reconfiguration to cope with changes in the traffic, changes in user dynamics, or SLA (Service Level Agreement) tuning, among other possibilities [13].

As previously highlighted, 6G infrastructure and services that telecommunication service providers will deploy will extensively use network slicing to virtualize and share resources. In this context, the main question is how NS architectures will contribute to and promote sustainability.

In this research paper, we approach this question from the perspective of energy-saving methods. In other words, to promote sustainability in 6G network slicing architectures, we aim to propose a novel solution to render the setup and operation of network slicing energy-saving aware.

The idea is that energy consumption and RAN resource optimization are minimized during the preparation, deployment, and supervision phases of slice deployment, taking into account the users' demands.

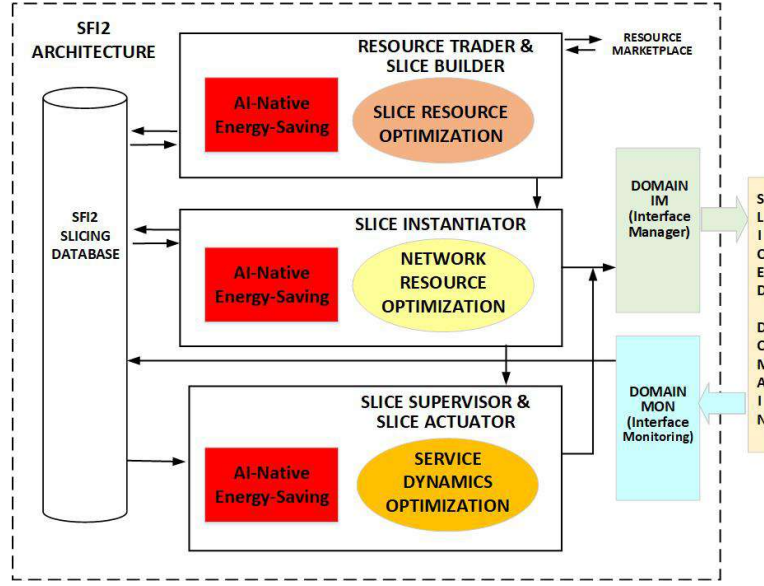


Figure 4: 6G Resource Orchestration and Optimization towards Energy Saving in SFI2 Architecture.

### 3 6G RESOURCE ORCHESTRATION AND ENERGY-SAVING IN SFI2 NETWORK SLICING ARCHITECTURE

The computer industry has multiple pathways to energy efficiency and sustainable solutions since the sustainable strategy adopted will differ among sectors and embrace various social, economic, and environmental aspects. This leads to potentially multiple deployed solutions across the industry.

In the context of the SFI2 network architecture, the 6G resource orchestration optimizations consider the following three optimization possibilities as illustrated in (Figure 4):

- Slice Resource Optimization (SRO): occurs while resources are selected and deployed for a specific slice to provide the required service in the slicing-building process. The approach consists of optimizing resources considering a slice individually.
- Network Resource Optimization (NRO): takes place during the instantiation phase of the network slicing process and concerns the set of slices provided by the service provider. In this case, the optimization process considers the set of slices active and the pool of resources available for the slice provider.
- Service Dynamic Optimization (SDO): This is executed during the slice operation and will deal with optimizing resources concerning the real-time services provided by the slice. This approach will consider a set of slices that may comprise either all active slices or any set of them, depending on the actual services being deployed by the provider.

### 4 PROPOSED SOLUTION: NETWORK SLICING ENERGY SAVING USING CONTRASTIVE LEARNING

In this paper, we propose using Contrastive Learning (CL) [19] to optimize network slicing and save energy by improving resource allocation, load balancing, and decision-making processes.

The idea is to use contrastive learning models to compare traffic patterns and demand forecasts for different slices and identify similarities and differences in traffic behaviors. Based on that, system resources like computing power, bandwidth, and others can be allocated efficiently.

The proposed solution aims to introduce ML agents in the SFI2 architecture to monitor the energy consumption of the processes allocated to instantiate and supervise slices.

The system model considers the SFI 2 architecture's Service Dynamic Optimization (SRO) and Network Resource Optimization (NRO), along with embedded ML agents (using CL) to monitor the allocated process's energy consumption and provide forecasting information to SRO and NRO to achieve their objectives, taking sustainability into account.

#### 4.1 Contrastive Learning for Detecting Anomalies in Unlabeled Energy Consumption Time Series

Contrastive learning is a robust method for learning data representations without labels. It aims to bring similar data points closer together in the representation space while pushing dissimilar points apart. This approach is effective for detecting anomalies in time series data, which often lack labeled examples. Given a set of time series data  $\{x_i\}_{i=1}^N$ , where  $x_i$  represents an instance, the objective



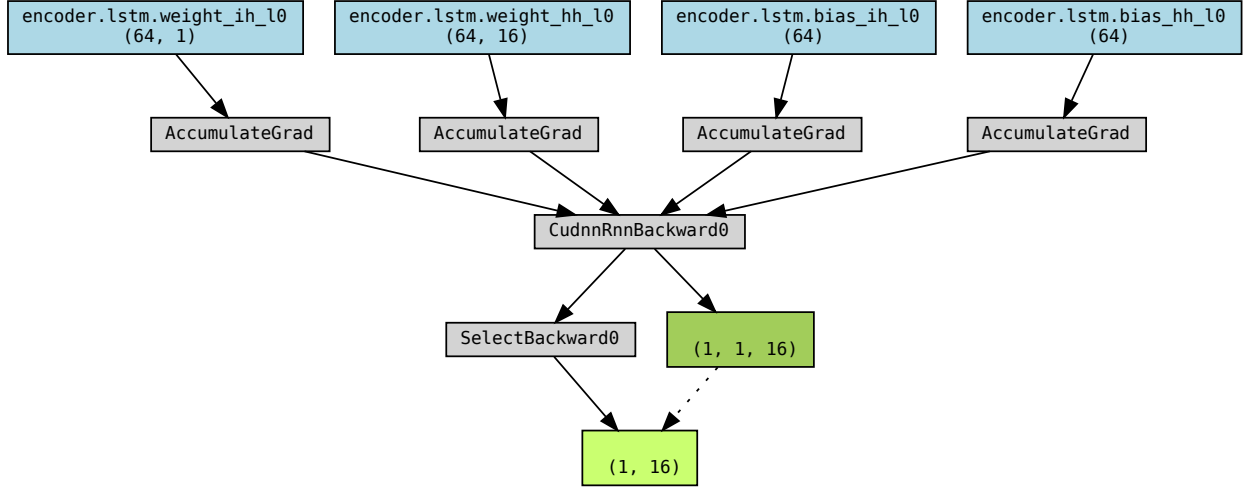


Figure 5: A LSTM model for Contrastive Learning.

is to learn a representation  $f(x_i)$  that clusters normal patterns and separates anomalies. The contrastive loss function is according to Equation 1.

$$\mathcal{L}_{\text{contrastive}} = \sum_{i=1}^N \sum_{j=1}^N \ell(f(x_i), f(x_j), y_{ij}) \quad (1)$$

where  $\ell(f(x_i), f(x_j), y_{ij})$  is the loss between instances  $x_i$  and  $x_j$ , and  $y_{ij}$  indicates similarity (1) or dissimilarity (0). A common form of the loss is according to Equation 2.

$$\ell(f(x_i), f(x_j), y_{ij}) = y_{ij} \|f(x_i) - f(x_j)\|^2 + (1 - y_{ij}) \max(0, m - \|f(x_i) - f(x_j)\|)^2 \quad (2)$$

Here,  $m$  is a margin parameter,  $\|\cdot\|$  is the Euclidean norm, and  $f(x)$  is the learned representation. The function  $f$  is typically implemented using deep neural networks like CNNs or RNNs, which capture temporal dependencies. To detect anomalies, measure the distance between representations. Instances outside the normal cluster are considered anomalies. For a new instance  $x'$ , the anomaly score  $s(x')$  is according to Equation 3.

$$s(x') = \min_{i=1}^N \|f(x') - f(x_i)\| \quad (3)$$

If  $s(x')$  exceeds a threshold,  $x'$  is an anomaly. We defined empirically a threshold of 0.5.

For our use case, we employed a datacenter energy-consumption dataset containing 15 features [20]. Data were collected over a period of 120 days, with a sampling frequency of one measurement per hour. Among these features, we selected the five best ones by the correlation method to be used as an object of anomaly analysis, namely, Voltage, Current, Power, Frequency, and Energy, in

the order of relevance measured by the method. We conducted a preliminary analysis to validate how anomaly detection methods with unlabeled data are capable of identifying anomalies with the least amount of noise possible.

Our Contrastive Learning method utilizes a Time Series Encoder with an LSTM layer to capture the temporal dependencies in time-series data using 16 hidden units. The encoded output was fed into a contrastive model to enhance feature discrimination by contrasting positive and negative pairs. An Adam optimizer was employed to adjust the parameters efficiently. A visualization of the model architecture is shown in Fig. 5, which illustrates the sequential connections and operations, highlighting gradient flow and backpropagation.

Detecting anomalies using contrastive learning involves comparing pairs of data points to determine their similarity. The process can be summarized in the Algorithm 1.

---

**Algorithm 1** Detect Anomalies

---

- 1: **Input:** Column *column*, Model *model*, DataFrame *df*, Threshold *threshold*
  - 2: **Output:** Vector of Anomalies
  - 3: *model.eval()*
  - 4:  $(X_i, X_j) \leftarrow \text{create\_pairs}(\text{column}, df)$
  - 5:  $X_i, X_j \leftarrow X_i.\text{to}(\text{device}), X_j.\text{to}(\text{device})$
  - 6:  $(h_i, h_j) \leftarrow \text{model}(X_i.\text{unsqueeze}(-1), X_j.\text{unsqueeze}(-1))$
  - 7:  $\text{distances} \leftarrow \text{torch.norm}(h_i - h_j, \text{dim} = 1)$
  - 8:  $\text{anomalies} \leftarrow \text{distances} > \text{threshold}$
  - 9: **return** *anomalies*
- 

Algorithm 1 begins by receiving the input parameters: a column of data (*column*), a machine learning model (*model*), a DataFrame



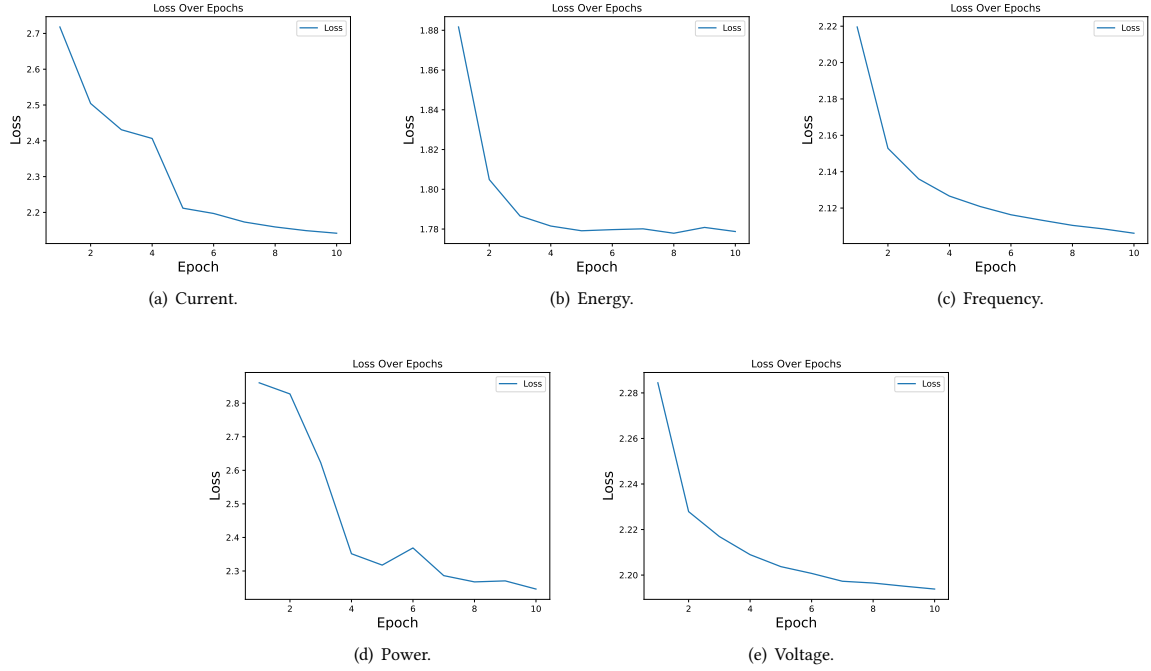


Figure 6: Losses over training time.

(df), and an anomaly detection threshold (threshold). The model was set to the evaluation mode (`model.eval()`) to ensure proper functioning during inference. Data pairs ( $X\_i, X\_j$ ) were created from the specified column in the DataFrame. These pairs were then transferred to an appropriate computational device. The model processes pairs to generate their respective embeddings ( $h\_i, h\_j$ ). The distance between these embeddings was calculated using the Euclidean norm. Anomalies were identified by comparing these distances to a predefined threshold, with distances exceeding the threshold indicating potential anomalies. The algorithm then returns a vector of the identified anomalies.

In the context of unlabeled time series data, k-means clustering and skewness can be effectively used to identify anomalies. K-means clustering involves partitioning the time series data into  $k$  clusters by minimizing the variance within each cluster. The formula for updating the centroid of a cluster is given by Equation 4

$$\mu_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i \quad (4)$$

where  $\mu_j$  is the centroid of cluster  $j$ ,  $C_j$  is the set of points assigned to cluster  $j$ , and  $x_i$  represents the data points. Anomalous points can be identified as those that are far from their assigned cluster centroids. Skewness, a measure of the asymmetry of the data distribution, can further enhance anomaly detection. The skewness  $\gamma$  is calculated according to Equation 5

$$\gamma = \frac{n}{(n-1)(n-2)} \sum \left( \frac{x_i - \bar{x}}{s} \right)^3 \quad (5)$$

where  $n$  is the number of data points,  $\bar{x}$  is the mean,  $s$  is the standard deviation, and  $x_i$  represents each data point. High skewness values can indicate the presence of outliers or anomalies in the data. By combining the clustering results from k-means with skewness analysis, one can robustly detect and interpret anomalies in unlabeled time series data.

We report these results in Table 1, where it can be observed that Contrastive Learning captured anomalies with the lowest incidence of noise.

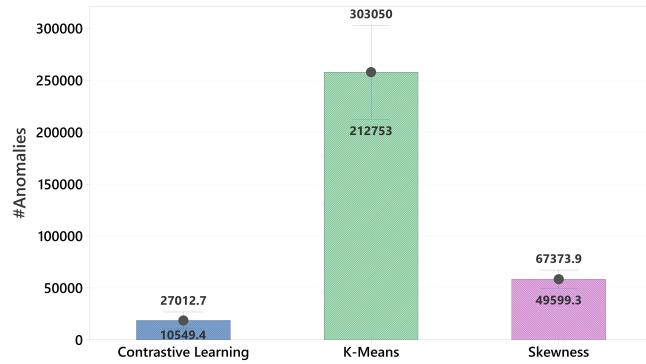
| Feature   | Contr. Learning with LSTM | K-Means             | Skewness      |
|-----------|---------------------------|---------------------|---------------|
| Voltage   | $2093.2 \pm 874$          | $112972 \pm 26295$  | $69821 \pm 0$ |
| Current   | $8115 \pm 4271$           | $87731 \pm 34729$   | $80380 \pm 0$ |
| Power     | $65073 \pm 35727$         | $262117 \pm 105794$ | $50560 \pm 0$ |
| Frequency | $18597 \pm 8021$          | $332186 \pm 12262$  | $89431 \pm 0$ |
| Energy    | $26 \pm 11$               | $494502 \pm 17353$  | $2241 \pm 0$  |

Table 1: Detected Anomalies by Method and Feature

With respect to contrastive learning with LSTM, we confirm in Fig. 6, there was learning in the model without overfitting in the task of detecting anomalies in the five features considered. We empirically used a learning rate of  $10^{-3}$  and Adam optimizer.

Finally, we measured the quantity and anomalies detected by each method, namely Contrastive Learning, K-means, and Skewness. We report in Fig. 7 and verify that contrastive learning focuses

on distinguishing meaningful differences by learning a representation that groups similar patterns while separating dissimilar ones. This method reduces noise sensitivity and captures more genuine structural variations in the time series data, leading to visually coherent and plausible anomaly detections. Although the lower number of detected anomalies might seem conservative, the model's improved alignment with visual expectations suggests it may more accurately represent true anomalies. This outcome underscores contrastive learning's potential in balancing anomaly detection sensitivity with robustness to noise, presenting a promising direction for handling unlabeled time series anomaly detection tasks.



**Figure 7: Amount of noised anomalies detected using this method.**

K-means, for instance, is sensitive to outliers since its cost function minimizes the distance between data points and cluster centers, which can lead to clustering anomalies alongside noisy data and obscure precise anomaly patterns. Skewness, which measures data distribution asymmetry, can reveal significant deviations in the shape of time series but is also influenced by extreme variations, often amplifying noise presence. Together, these effects result in anomaly detection that frequently flags noise as relevant deviations, raising false positive rates. This behavior underscores the potential need for more robust methods to distinguish noise patterns from genuine anomalies, suggesting an exploration of preprocessing techniques or algorithms that are less sensitive to non-representative data fluctuations.

## 5 FINAL CONSIDERATIONS

Energy-efficient and energy-saving strategies are crucial in the contemporary era of network architectures. In the state of the art, there are directions towards a network for more than connectivity; while guaranteeing stringent metrics, it is imperative to consider energy efficiency in the early stages of network architecture design. This paper elucidates how modern network slicing architectures can benefit from machine learning-based energy-aware slicing control plane approaches. This study incorporates results from embedding contrastive learning in the SFI2 Slicing Architecture to identify energy consumption anomalies in data centers where slicing and applications are deployed. The findings indicate that contrastive learning is less susceptible to noise, enabling AI algorithms to accurately capture energy demands throughout the slicing lifecycle.

Future work will aim to integrate this approach with real-time energy probes to accurately estimate and inform control plane slicing for modern network architectures.

## ACKNOWLEDGMENTS

The authors thank the FAPESP MCTIC/CGI cooperation agreement under the thematic research project 2018/23097-3 - Slicing Future Internet Infrastructures (SFI2), Brazilian National Council for Scientific and Technological Development (CNPq), grant # 421944/2021-8, FAPEMIG (Grant APQ00923-24) and the ANIMA Institute for scholarship support 2024/2025.

## REFERENCES

- [1] Maurizio Murrone, Matteo Anedda, Mauro Fadda, Pietro Ruiu, Vlad Popescu, Corneliu Zaharia, and Daniele Giusto. 6G—Enabling the New Smart City: A Survey. *Sensors*, 23(17):7528, January 2023.
- [2] Emmanuel Ekene Okere and Vipin Balyan. Sixth Generation Enabling Technologies and Machine Learning Intersection: A Performance Optimization Perspective. *Future Internet*, 17(2):50, February 2025.
- [3] ITU and World Bank. *Measuring the Emissions and Energy Footprint of the ICT Sector*. World Bank and International Telecommunication Union, 2024.
- [4] Ashu Taneja, Nitin Saluja, Neeti Taneja, Ali Alqahtani, M. A. Elmagzoub, Asadullah Shaikh, and Deepika Koundal. Power Optimization Model for Energy Sustainability in 6G Wireless Networks. *Sustainability*, 14(12):7310, January 2022.
- [5] Hoi-Lam Lou and Shang-Hsien Hsieh. Towards Zero: A Review on Strategies in Achieving Net-Zero-Energy and Net-Zero-Carbon Buildings. *Sustainability*, 16(11):4735, January 2024.
- [6] ATIS Alliance for Telecommunications Industry Solutions. 6G Sustainability KPI Assessment Introduction and Gap Analysis. Technical Report, Alliance for Telecommunications Industry Solutions, 2023.
- [7] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Management and Orchestration; Concepts, Use Cases and Requirements. Technical Specification 3GPP TS 28.530 V15.0, 3GPP, 2019.
- [8] Rodrigo Moreira, Flavio de Oliveira Silva, Tereza Cristina Melo de Brito Carvalho, and Joberto S. B. Martins. Intelligent Data-Driven Architectural Features Orchestration for Network Slicing. In *Proceedings of the International Workshop on ADVANCES in ICT Infrastructures and Services (ADVANCE)*, pages 1–12, Hanoi, Vietnam, 2024.
- [9] Prashant Subedi, Abeer Alsadoon, P. W. C. Prasad, Sabih Rehman, Nabil Giweli, Muhammad Imran, and Samrah Arif. Network Slicing: A Next Generation 5G Perspective. *EURASIP Journal on Wireless Communications and Networking*, 2021(1):102, April 2021.
- [10] Alcardo Alex Barakabitze, Arslan Ahmad, Rashid Mijumbi, and Andrew Hines. 5G Network Slicing Using SDN and NFV: A Survey of Taxonomy, Architectures and Future Challenges. *Computer Networks*, 167:106984, February 2020.
- [11] Een-Kee Hong, Inkyu Lee, Byonghyo Shim, Young-Chai Ko, Sang-Hyo Kim, Sangheon Park, Kyunghan Lee, Sunwoo Kim, Jae-Hyun Kim, Yoan Shin, Young-han Kim, and Haejoon Jung. 6G R&D Vision: Requirements and Candidate Technologies. *Journal of Communications and Networks*, 24(2):232–245, April 2022.
- [12] Abdul Waheed, Munam Ali Shah, Syed Muhammad Mohsin, Abid Khan, Carsten Maple, Sheraz Aslam, and Shahab Shamshirband. A Comprehensive Review of Computing Paradigms, Enabling Computation Offloading and Task Execution in Vehicular Networks. *IEEE Access*, 10:3580–3600, 2022. Conference Name: IEEE Access.
- [13] Joberto S. B. Martins, Tereza C. Carvalho, Rodrigo Moreira, Cristiano Bonato Both, Adnei Donatti, João H. Corrêa, José A. Suraagy, Sand L. Corrêa, Antonio J. G. Abelem, Moisés R. N. Ribeiro, José-marcos S. Nogueira, Luiz C. S. Magalhães, Juliano Wickboldt, Tiago C. Ferreto, Ricardo Mello, Rafael Pasquini, Marcos Schwarz, Leobino N. Sampaio, Daniel F. Macedo, José F. De Rezende, Kleber V. Cardoso, and Flávio De Oliveira Silva. Enhancing Network Slicing Architectures With Machine Learning, Security, Sustainability and Experimental Networks Integration. *IEEE Access*, 11:69144–69163, 2023.
- [14] Yulei Wu, Hong-Ning Dai, Haozhe Wang, Zehui Xiong, and Song Guo. A Survey of Intelligent Network Slicing Management for Industrial IoT: Integrated Approaches for Smart Transportation, Smart Energy, and Smart Factory. *IEEE Communications Surveys & Tutorials*, 24(2):1175–1211, 2022. Conference Name: IEEE Communications Surveys & Tutorials.
- [15] Adnei Donatti, Sand L. Correa, Joberto S. B. Martins, Antonio Abelem, Cristiano B. Both, Flavio Silva, José A. Suraagy, Rafael Pasquini, Rodrigo Moreira, Kleber V. Cardoso, and Tereza C. Carvalho. Survey on Machine Learning-Enabled Network Slicing: Covering the Entire Life Cycle. *IEEE Transactions on Network and Service*

- Management*, 21(3):1–18, 2023. Conference Name: IEEE Transactions on Network and Service Management.
- [16] Marja Matinmikko-Blue, Seppo Yrjölä, and Petri Ahokangas. Multi-Perspective Approach for Developing Sustainable 6G Mobile Communications. *Telecommunications Policy*, 48(2):102640, March 2024.
  - [17] Rodrigo Moreira, Joberto S. B. Martins, Tereza C. M. B. Carvalho, and Flávio de Oliveira Silva. On Enhancing Network Slicing Life-Cycle Through an AI-Native Orchestration Architecture. In Leonard Barolli, editor, *Advanced Information Networking and Applications*, Lecture Notes in Networks and Systems, pages 124–136, Cham, 2023. Springer International Publishing.
  - [18] Rodrigo Moreira, Rodolfo S. Villaça, Moisés R. N. Ribeiro, Joberto S. B. Martins, João Henrique Corrêa, Tereza C. Carvalho, and Flávio de Oliveira Silva. An Intelligent Native Network Slicing Security Architecture Empowered by Federated Learning. *Future Generation Computer Systems (FGCS)*, 163:1–19, 2025.
  - [19] Ashish Jaiswal, Ashwin Ramesh Babu, Mohammad Zaki Zadeh, Debapriya Banerjee, and Fillia Makedon. A Survey on Contrastive Self-supervised Learning, February 2021. arXiv:2011.00362 [cs].
  - [20] Rebeca Estrada, Víctor Asanza, Danny Torres, Adrian Bazaruto, and Irving Valeriano. Learning-based energy consumption prediction. *Procedia Computer Science*, 203:272–279, 2022.

Received 12 March 2025; revised 30 April 2025; accepted 15 05 25

---

# Meeting Healthcare Security Standards: From Legal Requirements to Technical Implementation

Ambre Journot  
Université Côte d'Azur - Laboratoire  
I3S-CNRS  
Sophia Antipolis, France  
ambre.journot@etu.univ-cotedazur.fr

Karima Boudaoud  
Université Côte d'Azur - Laboratoire  
I3S-CNRS  
Sophia Antipolis, France  
karima.boudaoud@univ-cotedazur.fr

Christian Delette  
Deltekzen  
Valbonne, France  
christian.delette@deltekzen.com

## Abstract

Ensuring the security of healthcare data continues to pose a significant challenge, particularly for professionals operating beyond the confines of institutional frameworks. Existing tools frequently prove inadequate in meeting the regulatory requirements and usability expectations of these practitioners. This paper undertakes a comprehensive analysis of the legal and normative framework governing the protection of sensitive medical information. The analysis culminates in the derivation of a set of essential criteria that any compliant and practical solution must fulfill. We then proceeded to evaluate several secure messaging and health-focused platforms, highlighting their limitations with respect to these criteria. To address these criteria, we introduce U-HealthSec, an architecture is designed to ensure strong data protection while remaining accessible to non-expert users. U-HealthSec integrates established cryptographic methodologies and security protocols with regulatory compliance, thereby addressing the operational and legal requirements of healthcare professionals in a comprehensive and scalable manner.

## CCS Concepts

• **Social and Professional topics** → **Professional topics**; • **Security and privacy** → **Security services**; *Usability in security and privacy*; *Access control*; **Multi-factor authentication**; • **Applied computing** → *Health informatics*.

## Keywords

Medical data security; End-to-end encryption (E2EE); GDPR compliance; Attribute-Based Keyword Search (ABKS); Secure file sharing; Health information systems; Privacy-by-design; Authentication (OAuth2, MFA); Healthcare compliance; Independent practitioners; HDS hosting; Data minimization; Auditability and non-repudiation

## 1 Introduction

Ensuring the security of healthcare data represents a significant challenge within the medical field. Moreover, confidentiality, integrity, and traceability of information exchanged between healthcare professionals and their patients is paramount to cultivate a foundation of trust. However, a significant number of practitioners, particularly those engaged in private practice or alternative medicine, lack adequate solutions for securing their exchanges of sensitive data.

The utilization of unsecured communication channels, such as conventional e-mail or instant messaging, renders this information vulnerable to risks of leakage, compromise, or malicious use. This

challenge is further increased by the emergence of stringent regulations, such as the General Data Protection Regulation (GDPR), which impose strict security standards that healthcare professionals must adhere to. However, existing solutions are often costly, complex, or designed for hospital structures, leaving a gap for independent practitioners.

Consequently, the objective of the work presented in this paper is to develop a secure data transfer solution tailored to professionals operating outside the confines of institutional infrastructures. This solution prioritizes a security-focused design approach, integrating advanced mechanisms such as End-to-End Encryption (E2EE), robust authentication methods, and comprehensive traceability of exchanges. Concurrently, the user experience remains a central tenet of our development process, ensuring that the solution is not only effective but also straightforward to adopt, even by users with limited expertise in cybersecurity.

The paper is structured as follows. We first provide an overview of the legal framework and regulatory constraints governing the management and exchange of healthcare data, along with the key criteria derived from these constraints. Then, we present the results of a comparative analysis of existing solutions that could potentially fulfill the identified requirements. Subsequently, we introduce U-HealthSec, the architecture we propose to address these challenges through a user-centric and regulation-compliant design. Finally, we conclude with a forward-looking discussion on the future enhancements and research directions of U-HealthSec to further strengthen the protection of healthcare data.

## 2 Legal framework

The protection and security of healthcare data is subject to a strict and complex legal framework established to guarantee the confidentiality, integrity, and traceability of sensitive information. In this section, we present the main relevant legal and normative obligations, with a particular focus on the regulatory constraints governing the management and exchange of healthcare data, especially regarding confidentiality, retention, and non-repudiation.

The General Data Protection Regulation (GDPR) [10, 30], a foundational law in Europe, imposes several crucial principles on data controllers, including lawfulness, fairness, and transparency in data processing; the minimization of data collection; and the implementation of appropriate technical and organizational measures for data security (GDPR, Articles 5 [9] and 32 [8]). Additionally, the GDPR stipulates that organizations must ensure the effective exercise of individuals' rights, including the right of access, rectification, deletion, and data portability (Articles 15-20 of the GDPR [1-5, 7]). A critical regulatory requirement is the prompt notification (within 72

hours) to the competent authority in case of a data breach (GDPR, Article 33 [6]).

"France's Code de la Santé Publique" (CSP) [28] further elaborates this general framework by introducing specific obligations tailored to health data. The fundamental principle of medical secrecy, as delineated in Article L1110-4 [25], is further reinforced. Additionally, the CSP stipulates that health data must be stored by a certified Health Data Host (HDS) (Article L1111-8 [26]). The mandatory retention of medical records is specified as a minimum of 20 years after the last treatment (Article R1112-7 [27]).

Furthermore, the "Commission Nationale de l'Informatique et des Libertés" (CNIL) [12] periodically disseminates specific guidelines pertaining to the management of sensitive data. These guidelines underscore the significance of data encryption (both at rest and in transit), robust authentication (MFA), and the implementation of pseudonymization or anonymization, contingent upon the prevailing circumstances. Furthermore, the CNIL underscores the necessity for comprehensive logging of accesses to ensure unassailable traceability and non-repudiation.

The "Politique Générale de Sécurité des Systèmes d'Information de Santé" (PGSSI-S) [15–17, 19] standard further delineates operational security requirements. These latter include the implementation of robust user authentication protocols, the mandatory use of end-to-end encryption (E2EE) for all sensitive communications, and the conducting of regular audits to verify ongoing compliance with security requirements. The PGSSI-S emphasizes the necessity of comprehensive logging of all accesses and systematic verification of integrity and non-repudiation for all actions performed on healthcare data.

The regulatory requirement for explicit, revocable consent from patients for the processing and sharing of their health data is paramount. In conjunction with the requirement for a continuous privacy impact assessment (DPIA) [11], this ensures that the risks associated with the processing of personal data are fully taken into account.

### 3 User-centric, security and organizational requirements derived from legal constraints

In this section, we present the key technical and organisational requirements derived from the legal constraints to design an effective and security solution tailored to the medical context.

In light of the various regulatory and normative constraints, several key criteria emerge for the design and evaluation of an effective security solution:

- **Transparency and lawfulness of processing:** Data processing activities must be transparent, clearly communicated, and lawful, complying with the principles of fairness and transparency defined by GDPR (Articles 5 and 32).
- **Data minimization:** Only data that is strictly necessary should be collected and processed, aligning directly with the GDPR principle of data minimization (Article 5).
- **Robust encryption (End-to-End Encryption - E2EE):** Data must be encrypted directly on the sender's device and decrypted only on the recipient's device. This ensures that data remains permanently encrypted during transmission,

storage, and processing, thereby preventing any intermediary from accessing unencrypted information. This method protects both confidentiality and integrity, in accordance with the CNIL's recommendations on the encryption of sensitive data and the mandatory requirements defined by the PGSSI-S standards.

- **Mandatory Multi-factor Authentication (MFA):** Access to sensitive health data must require robust authentication methods (at least two independent authentication factors), in compliance with CNIL and PGSSI-S recommendations for secure user authentication protocols.
- **HDS-certified hosting:** Health data must be hosted by service providers certified under the French HDS framework, as required by Article L1111-8 of the CSP. This certification ensures that the provider meets strict national requirements for handling sensitive medical data, including the implementation of an ISO/IEC 27001-compliant information security management system. In addition, HDS-certified providers are expected to comply with ISO/IEC 27018, which specifically addresses the protection of personal information in cloud environments. Together, these standards ensure robust confidentiality, integrity, availability, and traceability controls that include technical safeguards such as encryption, access controls, and secure backups, as well as organizational measures to ensure full accountability and regulatory compliance.
- **Full traceability and non-repudiation of actions:** All actions performed on health data must be meticulously logged, ensuring complete traceability and non-repudiation. The execution of operations by users cannot be denied, as this provides indisputable evidence of each action through reliable identification and timestamping. This approach is consistent with the guidelines established by the CNIL and the requirements stipulated within the PGSSI-S framework for comprehensive logging of accesses and actions. The objective of this approach is to reinforce accountability, facilitate audits, and enhance overall data integrity.
- **Compliance with legal retention periods:** Medical records must be retained for a minimum of 20 years following the last treatment, adhering explicitly to the CSP (Article R1112-7).
- **Clear and ergonomic management of user consent:** Systems must enable explicit, informed, and easily revocable user consent for data processing, aligning with GDPR requirements for explicit consent and continuous DPIA evaluations.
- **ISO 27001 certification:** Implementation of this certification is recommended to proactively manage and continuously improve information security, providing compliance assurance with GDPR and PGSSI-S operational security mandates.

These criteria will then be used as the basis for a detailed analysis of existing solutions, such as ProtonMail, Signal, or IDOMED, to highlight their strengths and limitations in meeting the specific needs of healthcare professionals.

## 4 Existing solutions and limitations

The secure exchange of healthcare data is a critical issue that necessitates an in-depth analysis of existing solutions. This analysis should include an assessment of their compliance with the regulatory and technical criteria defined above. In this regard, two predominant categories are emerging: secure messaging solutions for the general public and solutions dedicated to healthcare professionals.

Secure messaging solutions for the general public, including ProtonMail, Signal, Session, and Threema, offer advanced communication encryption guarantees, often based on proven protocols such as OpenPGP or Signal Protocol. **ProtonMail** [29], for instance, employs end-to-end encryption and a zero message access policy for its operators. However, this solution does not fully satisfy the regulatory requirements of the medical sector. While adhering to the stringent encryption standard (E2EE), ProtonMail conspicuously lacks mechanisms that ensure the traceability and non-repudiation of exchanges, a crucial component in safeguarding the integrity of medical records. Additionally, ProtonMail's metadata remains exposed, compromising the confidentiality of communications, and there is no systematic provision of clear management of user consent or multi-factor authentication. Additionally, ProtonMail does not offer HDS-certified hosting, nor does it comply with ISO 27001 or ISO 27018 standards.

Similarly, **Signal** [21, 22] is a robust alternative for instant messaging, thanks to its end-to-end encryption and automatic deletion of messages on servers after delivery, thus partially meeting the data minimization criterion. However, its implementation remains ill-suited to the medical sector due to the absence of a robust authentication mechanism and exchange logging, which hinders the ability to provide formal proof of information transmission. Additionally, the absence of HDS-certified hosting, clear management of user consent, and ISO certifications poses limitations in terms of meeting regulatory obligations within the healthcare sector.

In addressing these concerns, **Session** [24, 31] proposes a decentralized solution via Lokinet, ensuring high confidentiality by eliminating any trace of communications. While this model satisfies data minimization criteria and offers robust encryption, it does not ensure complete traceability or non-repudiation of exchanges, making it difficult to adopt in a context requiring structured archiving of medical data. The absence of mandatory multi-factor authentication, HDS-certified hosting, and ISO certification poses a significant challenge in meeting the rigorous standards demanded by the medical sector.

In the context of digital privacy and data security, **Threema** [32] stands out due to its commitment to zero-metadata harvesting, robust end-to-end encryption, and user anonymization. However, it does not meet the requirements of traceability, non-repudiation, mandatory multi-factor authentication, user consent management, and it does not offer HDS-certified hosting or ISO certifications. Consequently, Threema's incompatibility with healthcare regulatory frameworks hinders its practical application in contexts necessitating strict legal compliance.

Conversely, certain solutions have been developed with healthcare professionals in mind, taking into account the regulatory constraints and security requirements particular to this sector.

**MSSanté** [18, 20], a French government platform, offers a secure framework thanks to a trusted infrastructure and encrypted communications via SMTP over TLS, partially complying with traceability requirements. However, this solution is constrained by its restricted access to healthcare professionals registered in official databases such as the Shared Directory of Professionals Involved in the Healthcare System RPPS (Répertoire Partagé des Professionnels intervenant dans le système de Santé, the French shared directory of healthcare professionals), thus excluding certain non-registered professions. Additionally, the exchange platform imposes a maximum file size of 10 MB, a limitation that significantly restricts the sharing of voluminous medical documents such as MRIs (Magnetic Resonance Imaging) and scans. The absence of definitive deletion and the lack of transparency regarding internal mechanisms also impede full compliance with the stated criteria.

**IDOMED** [23] is a private alternative that offers advanced teleconsultation and medical document sharing functionalities. However, its lack of transparency regarding the encryption mechanisms used raises concerns about the reliability of its security measures. This, in turn, hinders the ability to fully validate its compliance with the transparency and lawfulness criterion. Additionally, while IDOMED utilizes HDS-certified hosting, it does not provide a definitive method for the secure and permanent removal of exchanged documents. Furthermore, it does not ensure comprehensive traceability or non-repudiation mechanisms.

**FAST-Échanges** [13], a secure file transfer system, is notable for its exchange traceability and guaranteed integrity via secure protocols such as SFTP, FTPS, and HTTPS. A notable advantage of FAST-Échanges is the absence of limitations on the size of transferred files. However, its suitability for medical use is constrained by several factors. Firstly, the absence of strong authentication hinders the implementation of strict access management protocols. Secondly, the lack of integration with specific healthcare data requirements, such as HDS certification and compliance with the principles of the French Public Health Code, is a notable limitation.

**Doctolib** [14], a widely adopted solution in France, offers a comprehensive platform encompassing teleconsultation, appointment management, and secure exchange of medical documents. It features HDS-certified hosting, mandatory multi-factor authentication, and compliance with legal retention periods. However, the platform exhibits certain limitations, including imperfect traceability and non-repudiation of user actions, and at times, opaque management of user consent. Furthermore, access is primarily restricted to practitioners registered with professional orders, often excluding alternative medicine practitioners and certain self-employed professionals.

A thorough examination of these solutions reveals several significant limitations with respect to the regulatory and security requirements of the medical sector. Despite their high level of encryption, general-purpose solutions suffer from the absence of traceability and non-repudiation mechanisms, as well as from a lack of strong compliance features such as HDS-certified hosting or clear user consent management. Conversely, while healthcare-specific solutions offer enhanced regulatory alignment, they frequently exhibit technical limitations, including restricted file size, inadequate deletion capabilities, and poor accessibility, particularly for independent

**Table 1: Comparison of existing solutions**

| Solution/<br>Requirements                              | ProtonMail | Signal  | Session | Threema | MSSanté | IDOMED  | FAST-Échange | Doctolib | U-HealthSec |
|--|------------|---------|---------|---------|---------|---------|--------------|----------|-------------|
| Transparency and<br>lawfulness of processing           | No         | Yes     | Yes     | Yes     | Yes     | Partial | Yes          | Partial  | Yes         |
| Data<br>minimization                                   | Partial    | Partial | Yes     | Yes     | Partial | Partial | Partial      | Partial  | Yes         |
| Robust encryption<br>(E2EE)                            | Yes        | Yes     | Yes     | Yes     | Partial | Partial | Partial      | Partial  | Yes         |
| Mandatory<br>multi-factor<br>authentication            | Partial    | No      | No      | No      | Partial | Yes     | No           | Yes      | Yes         |
| HDS-certified<br>hosting                               | No         | No      | No      | No      | Yes     | Yes     | No           | Yes      | Yes         |
| Full traceability and<br>non-repudiation<br>of actions | No         | No      | No      | No      | Partial | Partial | Yes          | Partial  | Yes         |
| Compliance with legal<br>retention periods             | No         | No      | No      | No      | Yes     | Yes     | Partial      | Yes      | Yes         |
| Clear and ergonomic<br>management of<br>user consent   | No         | No      | No      | No      | Partial | Partial | Partial      | Partial  | Yes         |
| ISO 27001  | Yes        | No      | No      | Yes     | No      | No      | No           | Yes      | No          |

practitioners or those working in alternative medicine. These findings underscore the necessity of a comprehensive solution that integrates all mandatory criteria: robust end-to-end encryption, mandatory multi-factor authentication, complete traceability and non-repudiation, clear and ergonomic consent management, HDS-certified hosting, ISO compliance, and inclusive access to all categories of healthcare professionals.

The Table 1 summarizes the results of this comparative analysis according to the criteria identified in section 3, providing a synthetic view of the strengths and weaknesses of each approach. Given that no existing solution fully satisfies the technical and organizational requirements necessary for legal compliance and practical ease of use, we present U-HealthSec, a dedicated architecture designed to meet these needs. The following section details this solution, explaining the underlying technological choices and how they align with both regulatory obligations and actual usage needs in the healthcare context.

## 5 U-HealthSec: Proposed Architecture

The proposed architecture, U-HealthSec (Figure 1), is specifically designed to address regulatory and technical constraints by integrating distinct functionalities, each explicitly linked to key compliance and security criteria:

- Local client-side encryption (AES-256-GCM):** Files are encrypted on the user's device using a unique AES-256 key generated for each file. This is essential for robust encryption (E2EE) and data minimization. This client-side encryption guarantees that neither the files nor their encryption keys are exposed or transmitted unencrypted, thus meeting confidentiality, integrity, and minimal data exposure regulatory criteria.
- Asymmetric encryption via Elliptic Curve Cryptography (ECC):** The AES key, which is generated for each file, is then encrypted using the recipient's ECC public key, thereby creating a "secure envelope." This ensures that only the intended recipient with the corresponding private key can decrypt the content, thus addressing criteria for robust encryption (E2EE) and strict access control.
- Secure File Transfer to HDS-Certified AWS S3 Storage:** Encrypted files are securely transmitted directly to AWS S3 via pre-signed URLs dynamically generated by the backend. This method ensures compliance with mandatory HDS-certified hosting regulations, simultaneously preventing exposure of sensitive data in clear text during transit and storage.
- Metadata management in secure PostgreSQL database:** Metadata, including encrypted AES keys, user identities, and comprehensive access logs, are securely stored in PostgreSQL. This configuration enables comprehensive traceability and systematic non-repudiation, ensuring that every critical action (e.g., connection, upload, download, sharing, deletion) is recorded in immutable, timestamped audit logs. This, in turn, facilitates compliance with transparency, lawful processing, and legal retention periods.
- Robust Authentication with OAuth2 and Mandatory Multi-Factor Authentication (MFA):** The proposed solution employs OAuth2 for authentication, in conjunction with mandatory MFA, further reinforced by encrypted PASETO tokens for secure session management. This robust authentication mechanism satisfies the stringent regulatory requirements for managing sensitive medical data, effectively

preventing unauthorized access and ensuring accountability.

- **An ergonomic interface for managing user rights and consent:** The user interface has been designed to be intuitive, thereby facilitating the management of consent preferences and data access permissions. Users have the option of revoking their consent or requesting permanent file deletions, which in turn trigger the complete deletion of encrypted files and associated metadata in AWS S3 and PostgreSQL. This protocol enhances transparency, guarantees lawful processing, and ensures compliance with the General Data Protection Regulation (GDPR)'s right to be forgotten.
- **Attribute-Based Keyword Search (ABKS) for secure data search:** ABKS technology facilitates secure searches within encrypted documents, allowing authorized users to search according to predefined attributes without server-side decryption. This mechanism directly aligns with regulatory criteria for data minimization and strict privacy protection, providing an efficient and secure user experience.

Having presented the various components and functionalities integrated into the U-HealthSec architecture, it is essential to illustrate their interactions using typical use cases. The sequence diagrams delineate the dynamic interactions between the various elements of the system during the primary phases: authentication (Figure 2), secure upload (Figure 3) and secure download (Figure 4).

Specifically, the diagrams illustrates that authentication via OAuth2 and MFA occurs prior to the issuance of a secure token (PASETO), which subsequently conditions all subsequent actions. These actions include the generation of AES-256-GCM encryption keys on the client side, their protection via ECC, and the use of pre-signed links to transfer encrypted files to object storage. Furthermore, it highlights how file access requests are verified and audited by the backend using metadata stored in relational database, ensuring full traceability and strict access control in accordance with regulatory requirements.

U-HealthSec integrates a suite of advanced and proven technologies, including AES-256-GCM, ECC, OAuth2, MFA, PASETO, and ABKS, to deliver a robust and comprehensive solution that is fully compliant with the stringent regulatory requirements of the medical sector. In addition, while ISO 27001 certification is not mandatory, the proposed architecture has been designed to align with their core principles. Its structured approach to access control, audit logging, risk minimization, and data protection by design lays the groundwork for a potential certification process. As such, U-HealthSec is well-positioned for future compliance with these standards, should formal certification be pursued. This solution is particularly well-suited for users without specialized IT security expertise. It is imperative to acknowledge that the architecture delineated is presently in the implementation phase.

## 6 Conclusion

The security and confidentiality of healthcare data remain central challenges in the evolving digital landscape. Through an in-depth analysis of legal requirements and existing solutions, this work has

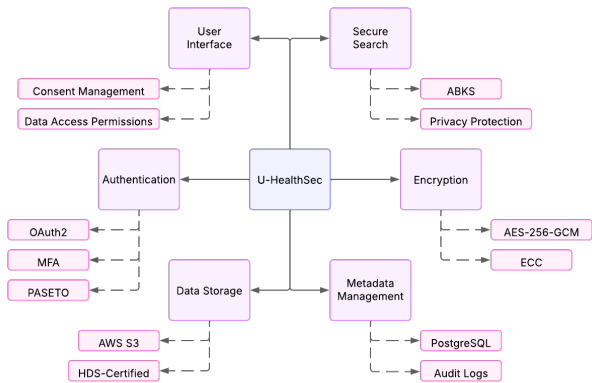


Figure 1: U-HealthSec Architecture

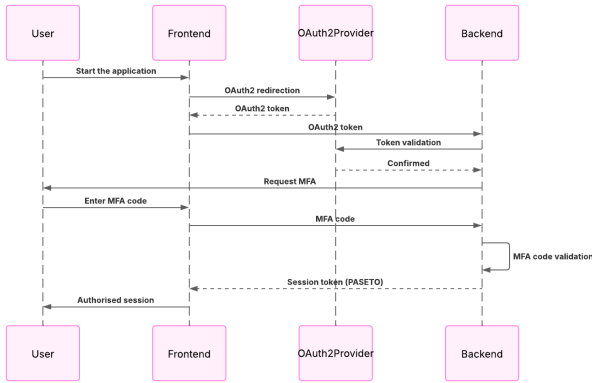


Figure 2: U-HealthSec Sequence Diagram – Authentication

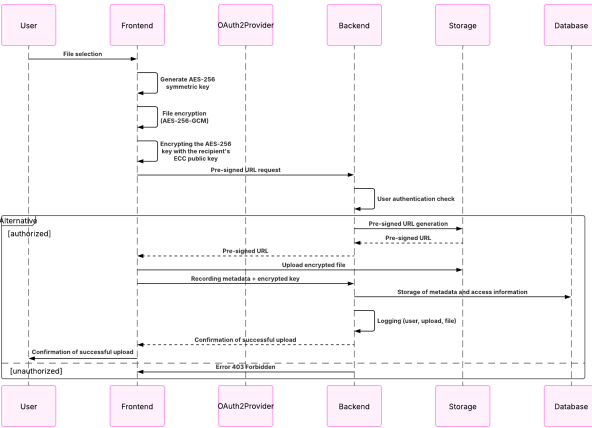


Figure 3: U-HealthSec Sequence Diagram – Secure Upload

demonstrated the pressing need for a comprehensive architecture



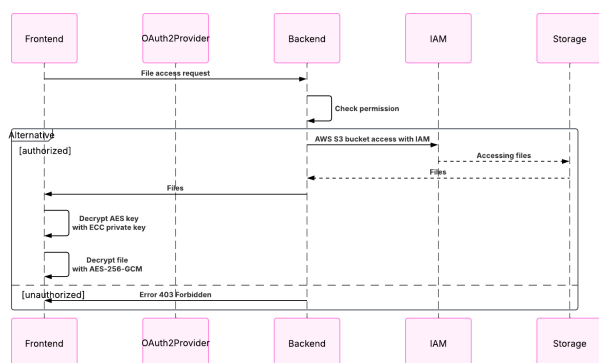


Figure 4: U-HealthSec Sequence Diagram – Secure Download

that meets both regulatory obligations and practical constraints specific to healthcare professionals, especially those operating outside institutional infrastructures.

To address these needs, a robust, user-centric solution has been introduced: U-HealthSec, currently undergoing implementation. This incorporation encompasses essential functionalities, ensuring full alignment with the primary compliance criteria of the RGPD, CSP, CNIL, and PGSSI-S, while maintaining accessibility for users lacking specialized cybersecurity expertise.

In parallel, we are currently exploring the integration of advanced cryptographic techniques such as homomorphic encryption, Searchable Symmetric Encryption (SSE) and Public Key Encryption with Keyword Search (PEKS) to further strengthen U-HealthSec. These technologies offer promising ways to provide secure data processing and privacy-preserving search capabilities without compromising confidentiality. However, their inclusion requires careful evaluation of performance, scalability, and implementation complexity in the healthcare context.

Beyond cryptographic enhancements, future improvements may include the development of a secure mobile application to increase accessibility, the implementation of automated compliance dashboards to support non-expert users, and the integration of anonymization workflows to enable secondary use of data for research while preserving privacy. These evolutions aim to create a trusted ecosystem for the exchange and management of sensitive healthcare data, fostering long-term trust among professionals and patients alike.

## References

- [1] CNIL. 2018. *Article 16 - Droit de rectification*. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre3#Article16> Consulté en mai 2025.
- [2] CNIL. 2018. *Article 17 - Droit à l'effacement ("droit à l'oubli")*. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre3#Article17> Consulté en mai 2025.
- [3] CNIL. 2018. *Article 18 - Droit à la limitation du traitement*. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre3#Article18> Consulté en mai 2025.
- [4] CNIL. 2018. *Article 19 - Obligation de notification en cas de rectification ou d'effacement de données à caractère personnel ou de limitation du traitement*. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre3#Article19> Consulté en mai 2025.
- [5] CNIL. 2018. *Article 20 - Droit à la portabilité des données*. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre3#Article20> Consulté en mai 2025.
- [6] CNIL. 2018. *Article 33 - Notification à l'autorité de contrôle d'une violation de données à caractère personnel*. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre4#Article33> Consulté en mai 2025.
- [7] CNIL. 2023. *Article 15 : Droit d'accès de la personne concernée*. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre3#Article15> Consulté en mai 2025.
- [8] CNIL. 2023. *Article 32 : Sécurité du traitement*. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre4#Article32> Consulté le 9 mai 2025.
- [9] CNIL. 2023. *Article 5 : Principes relatifs au traitement des données à caractère personnel*. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre2#Article5> Consulté en mai 2025.
- [10] CNIL. 2023. *Le règlement européen sur la protection des données*. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees> Consulté le 9 mai 2025.
- [11] CNIL. 2025. *Analyse d'impact relative à la protection des données (AIPD)*. <https://www.cnil.fr/fr/ce-quel-faut-savoir-sur-lanalyse-dimpact-relative-la-protection-des-donnees-aipd> Consulté en mai 2025.
- [12] CNIL. 2025. *Site officiel de la CNIL*. <https://www.cnil.fr/fr> Consulté en mai 2025.
- [13] Docaposte. 2025. *FAST-Échanges – Solution d'échange sécurisée de données*. <https://www.fast-docaposte.fr/nos-solutions/fast-echanges/> Consulté en mai 2025.
- [14] Doctolib. 2025. *Confidentialité des données de santé chez Doctolib*. <https://www.doctolib.fr/sante/confidentialite/> Consulté en mai 2025.
- [15] Agence du numérique en santé. 2018. *PGSSI-S – Politique Générale de Sécurité des Systèmes d'Information de Santé (document du 28 mai 2018)*. [https://esante.gouv.fr/sites/default/files/media\\_entity/documents/180528\\_PGSSI-S\\_0.pdf](https://esante.gouv.fr/sites/default/files/media_entity/documents/180528_PGSSI-S_0.pdf) Consulté en mai 2025.
- [16] Agence du numérique en santé. 2023. *Fiche pratique : Référentiel d'Identité Électronique (v4)*. [https://esante.gouv.fr/sites/default/files/media\\_entity/documents/231002\\_rie-fiche-pratique\\_v4-%281%29.pdf](https://esante.gouv.fr/sites/default/files/media_entity/documents/231002_rie-fiche-pratique_v4-%281%29.pdf) Consulté en mai 2025.
- [17] Agence du numérique en santé. 2025. *Corpus documentaire PGSSI-S*. <https://esante.gouv.fr/produits-services/pgssi-s/corpus-documentaire> Consulté en mai 2025.
- [18] Agence du numérique en santé. 2025. *MSSanté – Messagerie Sécurisée de Santé*. <https://esante.gouv.fr/produits-services/mssante> Consulté en mai 2025.
- [19] Agence du numérique en santé. 2025. *PGSSI-S – Politique Générale de Sécurité des Systèmes d'Information de Santé*. <https://esante.gouv.fr/produits-services/pgssi-s> Consulté en mai 2025.
- [20] Agence du numérique en santé. 2025. *Référentiel socle MSSanté v2*. [https://esante.gouv.fr/espace\\_documentation/mssante-clients-de-messageries-securisees-de-sante/referentiel-socle-mssante-2](https://esante.gouv.fr/espace_documentation/mssante-clients-de-messageries-securisees-de-sante/referentiel-socle-mssante-2) Consulté en mai 2025.
- [21] Signal Foundation. 2025. *Signal - Messagerie privée et sécurisée*. <https://signal.org/fr/> Consulté en mai 2025.
- [22] Signal Foundation. 2025. *Signal Protocol Documentation*. <https://signal.org/docs/> Consulté en mai 2025.
- [23] IDOMED. 2025. *Présentation de la solution IDOMED*. <https://www.idomed.fr/solution/> Consulté en mai 2025.
- [24] Kee Jefferys, Maxim Shishmarev, and Simon Harman. 2020. *Session: End-To-End Encrypted Conversations With Minimal Metadata Leakage*. <https://arxiv.org/pdf/2002.04609>. Consulté en mai 2025.
- [25] Légifrance. 2025. *Article L1110-4 du Code de la santé publique*. [https://www.legifrance.gouv.fr/codes/article\\_lc/LEGIARTI000043895798](https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000043895798) Consulté en mai 2025.
- [26] Légifrance. 2025. *Article L1111-8 du Code de la santé publique*. [https://www.legifrance.gouv.fr/codes/article\\_lc/LEGIARTI000033862549](https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000033862549) Consulté en mai 2025.
- [27] Légifrance. 2025. *Article R1112-7 du Code de la santé publique*. [https://www.legifrance.gouv.fr/codes/article\\_lc/LEGIARTI000036658351/](https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000036658351/) Consulté en mai 2025.
- [28] Légifrance. 2025. *Code de la santé publique*. [https://www.legifrance.gouv.fr/codes/texte\\_lc/LEGITEXT000006072665/](https://www.legifrance.gouv.fr/codes/texte_lc/LEGITEXT000006072665/) Consulté en mai 2025.
- [29] Proton Mail. 2025. *Sécurité de Proton Mail*. <https://proton.me/fr/mail/security> Consulté en mai 2025.
- [30] Ministère de l'Économie. 2023. *Règlement général sur la protection des données (RGPD)*. <https://www.economie.gouv.fr/entreprises/reglement-general-protection-donnees-rgpd> Consulté en mai 2025.
- [31] Session. 2025. *Messagerie anonyme, sécurisée et décentralisée*. <https://getsession.org/> Consulté en mai 2025.
- [32] Threema. 2025. *Security and Privacy*. <https://threema.ch/en/security> Consulté en mai 2025.

---

# Data Fusion of Observability Signals to Detect Anomalies and Failures in Microservices

Adalberto F. P. Chaves  
Federal University of Ceará (UFC)  
Crateús, Ceará, Brazil  
felipe.pinheiro.c27@gmail.com

Francisco A. A. Gomes  
Federal University of Ceará (UFC)  
Crateús, Ceará, Brazil  
almada@crateus.ufc.br

Paulo A. L. Rego  
Federal University of Ceará (UFC)  
Fortaleza, Ceará, Brazil  
paulo@dc.ufc.br

Fernando A. M. Trinta  
Federal University of Ceará (UFC)  
Fortaleza, Ceará, Brazil  
fernando.trinta@dc.ufc.br

José N. de Souza  
Federal University of Ceará (UFC)  
Fortaleza, Ceará, Brazil  
neuman@ufc.br

## ABSTRACT

Microservices have become a reality, with large companies adopting this architecture due to its advantages, such as development agility, decoupling, scalability, resilience, and operational efficiency. However, this approach has drawbacks, such as the need for complex cloud infrastructure and communication between microservices through protocols. In organizations with hundreds of microservices, any failure in one of them can trigger cascading effects, resulting in more failures. Identifying failures in microservices is not a trivial task. The proposed work focuses on data fusion for detecting failures and anomalies in microservices, combining two pillars of observability: distributed tracing and metrics, with the help of OpenTelemetry. This fusion aims to provide faster and more efficient failure detection, as well as the identification of anomalies in metrics. This allows problems to be detected and addressed more quickly, improving the overall stability and reliability of the system.

## CCS CONCEPTS

• **General and reference** → *Performance; Evaluation*; • **Computer systems organization** → *Cloud computing*.

## KEYWORDS

microservices, failures, anomalies, observability, data fusion

## 1 INTRODUCTION

A well-designed software architecture can help ensure that a system meets key requirements in areas such as performance, reliability, portability, scalability, and interoperability [8]. As the size and complexity of software systems increase, the design problem extends beyond computing algorithms and data structures: designing and specifying the overall structure of the system emerges as a new type of challenge [9]. Due to the complexity of modern systems, the software industry has begun adopting a new architectural standard, where the system is modularized into services [6].

Microservices architecture has become a reality in the industrial domain. Microservices architectures assist project managers and developers by providing guidelines for the design and implementation of distributed applications [7]. Such architectures are present in various types of applications, including web, mobile, and e-commerce systems, bringing positive factors such as scalability, flexibility, and portability. In this architecture, services are small

applications that can be deployed, scaled, and tested independently, each with a single responsibility, i.e., a single business rule [21].

Due to their highly distributed nature, microservices are more challenging to operate. Therefore, they require virtualization and infrastructure technologies (based on containers), such as Docker and Kubernetes, which provide continuous updates, automated scaling, and rebalancing in the event of node failure [10]. Given the operational complexity, monitoring them becomes a difficult task.

Conducting extensive validation testing with microservices before each deployment is not feasible due to the high frequency of their releases. Instead, the quality assurance of microservices is often complemented or even replaced by refined monitoring techniques in production environments exposed to real-world workloads [16]. Large systems may contain hundreds of services communicating with one another. Failures in some of these services can lead to failures in others, and quickly identifying these failures can prevent cascading effects, resulting in a more stable system. Furthermore, the detection of anomalies is crucial in this context [18]. Anomalies in metrics, such as unexpected spikes in resource usage or unusual response times, can signal potential issues before they escalate into failures.

Traditional monitoring solutions often lack the capability to provide comprehensive visibility and to proactively detect deviations from expected behavior before they impact services. Addressing these gaps requires a new level of monitoring known as observability [13]. Observability is defined as the ability to deduce the state of a complex system from its outputs [11]. It also refers to a system's property that allows its state to be observed and compared with expected conditions throughout the software development lifecycle [12]. The three observability signals are metrics, traces, and logs. Due to its critical importance, a range of tools have been proposed to support observability. In recent years, OpenTelemetry (OTel) has become the most widely adopted tool in the industry and is considered the gold standard for observability [14]. OTel is an open-source observability framework that provides a suite of standardized, vendor-agnostic SDKs, APIs, and tools for data ingestion, transformation, and transmission to an observability backend [3].

Data fusion allows for the integration of information from different sensors, systems, and sources [22]. This can provide a more comprehensive and accurate view of the system's state, helping to detect anomalies and failures that might be difficult to detect

using a single data source. Additionally, by combining data from multiple sources, it is possible to identify anomalous patterns that may indicate an impending failure.

This study aims to perform data fusion of observability signals to enhance the effectiveness of failure and anomaly detection in microservices architectures. The solution developed, named **FOCUS** (data Fusion of Observability signals to detect anomalies and failures in microservices), uses OTel as the observability framework. This research will leverage observability techniques—specifically metrics and traces—to develop a more robust system that is less susceptible to failures. By integrating and analyzing these diverse observability signals, the study aims to detect and diagnosing issues, ultimately contributing to a more stable and resilient microservices environment. As a result, through a load testing experiment, it was found that the solution is able to detect failure and anomaly of the application. In the context presented, this paper aims to address one initial research question (RQ):

**RQ** *How does data fusion from different observability signals facilitate the detection of anomalies and faults in microservices architectures?*

The contributions of this paper include the data fusion of observability signals by combining metrics and distributed tracing, implemented through instrumentation with OTel. This approach enables the association of metrics with traces, allowing distributed tracing to identify faulty services by detecting anomalous values in the configured metrics. A library was created to facilitate this setup, which is integrated with the microservices application. Additionally, it is possible to select the services to be monitored, bringing this data into a dashboard that classifies services into risk levels based on real-time metrics collected during execution. Finally, a reference application called Spring Petclinic was used for load testing experiment, to verify the effectiveness of the solution. The implemented codes were not made available due to conference restrictions about *Do NOT include hyperlinks in the document*.

The remainder of the paper is organized as follows. Section 2 presents the background. Section 3 discusses related work. The integration of OTel, Observability to Detect Failures, and applied solutions are explained in Section 4. Section 5 details the experiments and results. Finally, Section 6 concludes our study and provides suggestions for future research.

## 2 BACKGROUND

### 2.1 Microservices

Microservices is an architectural style that emphasizes the decomposition of a system into small, lightweight services that are deliberately designed to perform highly cohesive business functions. It represents a traditional evolution of Service-Oriented Architecture (SOA) [17]. SOA is a software architecture pattern characterized by low coupling, which makes components reusable by employing services with a common language over a network. These aspects lead to various benefits for microservices, including increased agility, decoupling of functionalities, and ease of deployment and maintenance [19]. Consequently, microservices have been widely adopted, with notable examples including companies such as *Netflix* and *Amazon*.

As microservices are small applications that can be deployed independently, each service is ideally designed to implement a specific business rule or system functionality. Furthermore, each service should possess its own logic and avoid dependencies on other microservices in the system, meaning its responsibility should be singular [21]. These services are built around business resources and deployed independently by fully automated systems. There is minimal centralized management of these services; each service is a separate codebase that can be managed by a development team and can be written in different programming languages and use various database technologies [24]. Consequently, this architecture necessitates a DevOps culture responsible for implementing continuous delivery and continuous integration (CI/CD) pipelines and orchestrating microservices in a cloud environment.

Due to these characteristics, this approach also has some drawbacks. With multiple applications, all must be monitored individually rather than just one, creating a need for distributed monitoring infrastructure and centralized visualization [5]. Integration testing becomes more challenging, as a test environment where all components are operational and communicating is required. When interactions in a microservices-based application are designed as Application Programming Interface (API) calls, all processes necessary for API management must be implemented. This includes various components that need to be developed alongside the microservices architecture, such as message queues for communication, a new deployment pipeline, and an API gateway [1].

### 2.2 Observability

Observability provides high-level overviews of system health and granular insights into implicit failure modes of the system. An observable system offers extensive context about its internal operations, enabling the discovery of deeper systemic failures [23]. Observability empowers those who build and operate distributed systems and applications to understand the behavior of their code during production [3]. Observability can also be understood as a superset of monitoring, in the sense that if a system is observable, it can be monitored. Observability provides information that aids in monitoring, as it allows navigation from effects to causes in a production system, understanding what, where, and why things happen. The observability of a distributed system relies on three core pillars (signals): logs, metrics, and traces.

Logs are files that record events, warnings, and errors as they occur in a software environment. Most logs include contextual information, such as the time an event occurred and which user or endpoint was associated with it [14]. For example, a web server log file may include when the server was started, client requests, and how the server responded to those requests. It records information about each successful transaction, as well as errors and failures in client connections.

Measuring application and system performance through metrics collection is a common practice in software development. These data are converted into graphs to generate meaningful visualizations for those responsible for monitoring system health. In certain environments, metrics are used to automate workflows in response to changes in the system, such as scaling up the number of application instances or rolling back a faulty deployment [14]. Metrics

are low-level data that can indicate issues with underlying computational, storage, or network equipment. Careful monitoring of these metrics can highlight performance degradation and detect anomalies, which can prevent system-wide failures. This makes them crucial for decision-making regarding system improvements.

Tracing applications means having the capability to execute the application code and ensure it performs as expected [14]. Achieving this, given the number of services in larger projects, is a challenging task, necessitating the use of mechanisms that perform this tracing. One approach in this context is distributed tracing. Distributed tracing is a diagnostic technique that helps engineers locate failures and performance issues in applications, especially those that may be distributed across multiple computers or processes. With distributed tracing, it is possible to determine whether any communication steps failed, how long each step took, and potentially log messages produced by each step during execution. It observes requests following the steps taken in the system. A unique tracing context (trace ID) is inserted into the header of each request, and mechanisms are implemented to ensure that the tracing context is propagated throughout the request's path. In other words, to enable identification, an identifier (ID) is created in the first request and passed through all subsequent requests via the request header.

In system monitoring, metrics and traces are crucial observability signals that provide different but complementary insights. Metrics offer quantitative measurements of system performance, such as response times or error rates. Traces, on the other hand, track the flow and interactions of requests across various system components. By fusing metrics and traces, engineers can achieve a deeper understanding of system behavior and enhance fault detection. For instance, if a metric indicates a spike in response times, traces can help pinpoint which specific service or component is causing the delay by showing the path and timing of requests through the system. This combined view allows for more precise diagnosis of performance issues.

### 2.3 OpenTelemetry

OpenTelemetry (OTel) is a framework and toolkit for observability designed to create and manage telemetry data, such as traces, metrics, and logs [3]. OTel is vendor- and tool-agnostic, meaning it can be used with a wide variety of observability backends. It originated from the merger of OpenTracing and OpenCensus in 2019. OpenTracing is an open standard for distributed tracing in applications and open-source software (OSS) packages, allowing developers of application code, OSS packages, and OSS services to instrument their own code without being tied to any specific tracing vendor. OpenCensus is a set of libraries that enables the collection of application metrics and distributed traces, and subsequently transfers the data to a backend of choice in real-time. This merger brought two main improvements: (i) the unification of standards for both instrumentation and data collection, and (ii) a larger and more robust ecosystem.

To be useful, telemetry data must be exported to a data storage system where storage and analysis can occur [3]. To achieve this, each implementation provides a range of mechanisms for generating, processing, and transmitting telemetry. These processes are referred to as pipelines. Pipelines are executed at the start of

implementation to ensure no data is lost, and after this stage, the pipeline's role is to collect and send data, characterized by three processes: receivers, processors, and exporters.

Receivers are used to provide access to the application code or an entity used to generate telemetry data, and can be configured at the outset without the need for any telemetry data to be generated. Processors perform data aggregation, filtering, sampling, and other processing logic on the collected telemetry data, given the potential for changes in the data that was previously processed. Processors can be chained to produce complex processing logic [20]. Exporters handle the emission of telemetry data to one or more destinations in various formats and protocols, including examples such as Prometheus, Jaeger, and Zipkin [20].

## 3 RELATED WORK

Bento et al. [2] conducted a study focused on finding anomalous regions through distributed tracing. They claim that manual queries to the distributed tracing tool are necessary, considering the time and annotations involved. Therefore, an automated approach is emphasized. The authors developed an Open Trace Processor (OTP) to extract metrics from traces and feed them into a data analyzer, which identifies anomalies in time series data such as the number of received calls, made calls, and response times. Based on existing traces and the metrics derived from them—such as the number of received and sent requests, response times, or service error codes—the goal was to identify threats to software resilience. A drawback of this work is its reliance on OpenTracing, an older standard for observability.

Casse et al. [4] focused on exploring OpenTelemetry trace data to detect inefficient communications within a cloud application. Their goal was to expose inefficient resource allocations made by Kubernetes in a geographically distributed cluster. To obtain the data, Kubernetes was extended with a service mesh in each microservice, aimed at observing network calls between services. The Linkerd service mesh was chosen due to its compatibility with the OpenCensus format, ensuring compatibility with OpenTelemetry binaries. In a Kubernetes environment, Linkerd automatically injects and configures HTTP proxies between each microservice instance to enable communication observation. The next step was to create a property graph, where all traces were considered as a single graph, decomposing trace data into multiple vertices and edges merged with previously observed data, thus establishing correlations across multiple traces. A drawback of this work is its reliance solely on traces for detecting resource inefficiencies, whereas data fusion, incorporating metrics as well, could contribute to more effective detection.

Marie et al. [15] conducted a study demonstrating the feasibility of observability-oriented monitoring using a set of customized tools. The proposed architecture includes an observability layer that extracts data from cloud-native applications, using this data to inform DevOps teams. The observability framework was implemented in the environment of the company Lectra SA, with the constraint that all microservices depend on Microsoft Azure infrastructure. This framework provides functionalities such as collection, storage, and management of metrics, utilizing Prometheus and Grafana. The microservices are deployed by an orchestrator based on a manifest

file containing all necessary parameters for the deployment and execution of instances. The research offers insights into concepts, features, and prototypes related to observability in cloud-native applications. A drawback of this work is its reliance only on metrics for detecting the need for auto-scaling resources.

Tzanettis et al. [22] tackle the challenge of integrating data from orchestration platforms, distributed tracing, and logging tools to simplify the performance analysis of distributed applications. Their central contribution is the specification of a data binding scheme supporting the collection and integration of container resource data, performance metrics, distributed tracing, and logs. The proposed scheme is capable of merging and correlating data from different types of signals, implemented through open-source tools. These tools utilize the Prometheus monitoring engine (supported by Kubernetes), the Zipkin distributed tracing tool, the Fluentd logging software, and the Prometheus Python instrumentation library for defining instances in the source code. A drawback of this work is that, although it performs data fusion of all signals, it only validates the data scheme and does not conduct an experiment to demonstrate that scalability actually improves with the proposed solution.

None of the evaluated studies use data fusion through observability signals to detect anomalies and failures in microservices using the modern observability standard of OTel.

## 4 METHODOLOGY USED

This section describes the methodology employed in our study. Section 4.1 discusses the implementation of the FOCUS library for data fusion, while Section 4.2 presents the Petclinic application. Section 4.3 describes how the Petclinic project was integrated with the library, Section 4.4 covers the creation of a service that manages FOCUS information and detects anomalies and failures. Finally, Section 4.5 details the creation of a dashboard service for visualizing the configured services.

### 4.1 FOCUS Library for Data Fusion

The first step of the study was to develop the FOCUS library, which is based on Java and OTel, for data collection and fusion, focusing on metrics and traces of observability signals. In data fusion, each trace span includes, in addition to standard values such as *traceId* and *spanId*, metrics associated with the trace. The library is accessed through a Java Annotation called `@ObservabilityParam`, where parameters for metrics with maximum values to be monitored are passed (more details in Section 4.3) to identify anomalous metrics, which are those exceeding the configured maximum values. The parameters configured by the developer include CPU usage, memory, response time, and throughput. Thus, both infrastructure and application metrics are evaluated.

To collect observability signals, as previously described, OTel operates with the Collector, which has specific components for data reception and export. The **OTLP Receiver** was used to receive application observability signals in OTLP format (OTel's data protocol). Metrics are sent to the **Prometheus Exporter**, which exports metrics to a Prometheus server via scraping. To retrieve infrastructure metrics, the **HostMetrics Receiver** was used, which generates a wide variety of host system metrics, including CPU

usage, memory, disk I/O, among others. Additionally, the **OTLP Exporter** was used to export data in OTLP format via gRPC. This allows retrieval of response times for each request. Throughput calculation was done by summing bytes received and transmitted through the network interface provided by the HostMetrics, divided by the response time of the request. Finally, a component called **Logging Exporter** was used, which sends data to the console via `zap.Logger` and supports traces, metrics, and logs in the pipeline. Thus, although not part of the data fusion in this study, logs can be retrieved and analyzed.

To retrieve the values of the collected metrics and compare them with the values configured by the developer, an Aspect-Oriented Programming (AOP) implementation in Java (AspectJ) was necessary. In this approach, an advice can be defined, which are code snippets executed at join points in the program. Thus, code can be included before, after, or around the execution of methods. This made it possible to retrieve CPU and memory usage metrics during the execution of any method annotated with `@ObservabilityParam`. In the subsequent step, these metrics had to be incorporated into the OTel span to perform data fusion. For this, the `setAttribute` method of OTel was used on the current span of the request to assign the collected metrics. Thus, each span contains CPU, memory, response time, and throughput metrics. As a result, the fusion of metric and trace data was completed, allowing both to be viewed in a single span in distributed tracing systems. Jaeger, a distributed tracing system created by Uber that offers adaptive sampling and has native support for OTLP, was used in the proposed solution.

### 4.2 Overview of Application

Petclinic is a reference application created to demonstrate practices of the Spring Framework in microservices, simulating the management system of a veterinary clinic. Its main functionalities include client management, where it is possible to register and manage the information of pet owners; pet management, which allows the registration of pets associated with an owner, and the viewing of these pets' medical history; and veterinarian management, with the ability to register and list veterinarians and their specializations. Developed in 2021 and with ongoing modifications, it currently consists of 4 microservices related to business logic and uses Java. This application was used to validate the proposed solution and conduct the experiments for this work.

### 4.3 Integration Application and Library

To integrate the Petclinic application with the FOCUS library, it was necessary to add the OTel agent to the Docker image and include the library dependency in each service project. Figure 1 presents an overview of the architecture of the developed solution, which operates through docker-compose, a tool for defining and running multiple Docker containers. In addition to the OTel Collector components explained earlier, data was exported to observability backends, which, as illustrated in Figure 1, are Prometheus and Jaeger.

After configuring OTel and the application, it was necessary to configure the usage of the `@ObservabilityParam` Annotation on the methods of the services that have the endpoints (controllers) and define the metric values that should not be exceeded. CPU

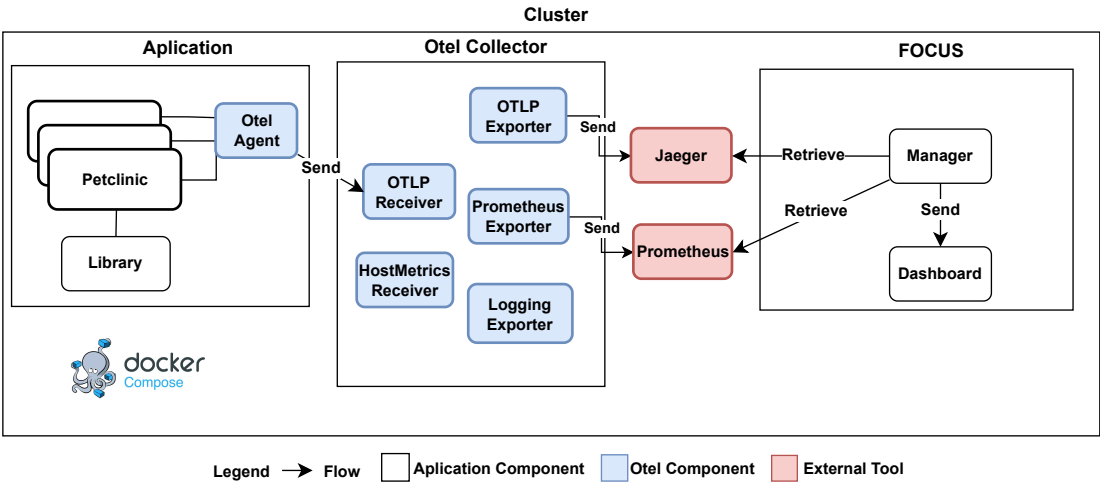


Figure 1: Proposed Architecture

usage is given as a percentage, memory in MB, response time in milliseconds (ms), and throughput in bytes per second. For example, the annotation has the following parameters:

Listing 1: Example of Annotation

```
@ObservabilityParam(params = {
    @Param(key = "cpuUsage", value = "20"),
    @Param(key = "memory", value = "30"),
    @Param(key = "responseTime", value = "200"),
    @Param(key = "throughput", value = "800")
})
```

From Listing 1, it can be seen that the CPU usage is set not to exceed 20%, memory usage is limited to 30 MB, response time should not be more than 200 ms, and maximum throughput is set to 800 Bytes/ms. The metric configurations are saved by the library in the span and sent to Jaeger. Figure 2 shows how this data is stored. This data is subsequently retrieved by the Manager service of FOCUS, which is responsible for analyzing the metrics to determine if they are anomalous, as well as identifying any failures in the requests, and performing data export. More details are presented in the next section.

4.4 FOCUS Manager

After integrating with the Petclinic project, setting up the library, and sending data to Jaeger, the next step was to create the FOCUS Manager service to export data and perform metric analysis to detect anomalies and failures in the requests, as well as to provide information to the dashboard service that will be shown later. The Manager service has four main endpoints: `/api/metrics/system-info`, which returns general information about the application, in this case, Petclinic, including the number of requests, number of errors, and requests per second; `/api/metrics/all`, which returns the values received at runtime from the methods where the Annotation was applied; `/api/metrics/system-metrics`, which returns metrics from

|                           |                    |
|---------------------------|--------------------|
| cpuUsage                  | 20                 |
| cpuUsageReceived          | 2.0000000000000018 |
| http.request.method       | GET                |
| http.response.status_code | 200                |
| http.route                | /owners            |
| internal.span.format      | otlp               |
| isObservabilitySpan       | true               |
| memory                    | 30                 |
| memoryUsageReceived       | 62357504           |
| responseTime              | 200                |
| responseTimeReceived      | 1000               |
| throughput                | 800                |
| throughputReceived        | 500                |

Figure 2: Data stored in Jaeger

the Prometheus API about the system’s health at that moment, including information on CPU usage, memory usage, and throughput. Finally, `/api/analysis` returns the results of anomaly detection in the metrics, as well as requests with failures.



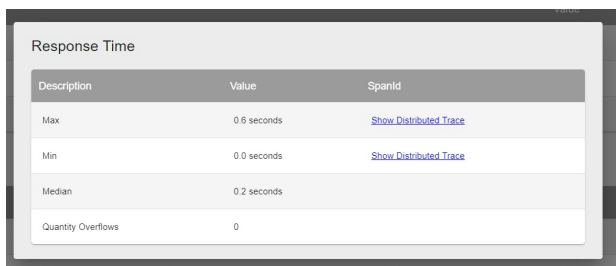
For anomaly detection in metrics, a risk classification of the application's integrity is performed. The algorithm determines if a method is at low risk if it is below 50% of the value set as a parameter in the Annotation. For example, if it was specified that the `getAllUsers` method should not exceed 20% CPU usage, results below 10% are classified as low risk. Medium risk is assigned to methods with values between 80% and 99% of the value defined in the Annotation. Finally, high risk is assigned to methods that exceed the defined values. This classification was determined by the authors of the study. Thus, all metrics classified as high risk are considered anomalous, and the rest are normal. This risk classification is used in the Dashboard service to indicate to system operators the status of the metrics for the annotated methods. In FOCUS, request traces have one of two states: normal or failure. Any annotated method with a status code of 4xx or 5xx is classified as a failure state. Therefore, FOCUS can also detect and report requests with failures

#### 4.5 Focus Dashboard

After creating the Manager service to expose data fusion and analyze metrics and traces, a dashboard service was developed to facilitate visualization. The Dashboard has two main flows: on its home screen, all methods with the Annotation are listed, along with information on whether the metrics are anomalous and if the traces have failures.

The first flow shows the classification according to the risk to the application's integrity. The classifications, as shown earlier, come from the Manager. The second flow corresponds to the details screen, which provides insights into the metrics of the annotated method, including general information such as the number of requests, errors, and average data throughput.

Figure 3 provides information about the maximum, minimum, average, and quantity overflows of the metrics, including request traces. By clicking on the trace, you are directed to Jaeger to analyze the distributed trace, which facilitates identifying why the metrics received those values at that moment. This highlights the importance of data fusion. Metrics allow us to see the health of the system, and distributed tracing helps us find what caused our system to have metrics that exceeded the defined values.



| Description        | Value       | SpanId                                 |
|--------------------|-------------|--|
| Max                | 0.6 seconds | <a href="#">Show Distributed Trace</a> |
| Min                | 0.0 seconds | <a href="#">Show Distributed Trace</a> |
| Median             | 0.2 seconds |  |
| Quantity Overflows | 0           |  |

Figure 3: Specific Metric Values

## 5 EXPERIMENTS AND RESULTS

We conducted experiments to evaluate our solution in identifying methods that exceed pre-defined metrics in the Petclinic application, as well as in detecting failures.

### 5.1 Experiment Configuration

We deployed the application on a cluster using Docker Compose in a local infrastructure composed of an Intel® Core™ i7-11390H 11th generation processor with 4 cores, 2918 MHz, 16 GB RAM, and Windows. We ran two scenarios in Petclinic, where we modified one of the services to include a specific method that, when triggered, significantly increases memory usage and response time. The execution of this method occurs randomly, generated by a random value from 1 to 100 at service startup. If the value is between 1 and 8, the anomalous method is triggered, resulting in approximately an 8% chance of occurrence.

This strategy was designed to simulate conditions where unexpected failures may occur, providing a practical way to test the system's resilience and its ability to detect anomalies. The project scenarios were as follows: (i) the original version of Petclinic, with the mentioned modifications, and (ii) the modified version, including the library dependency and the use of the annotation. The goal was to assess whether triggering the method that overloads the metrics would enable the proposed solution to facilitate problem identification. For this test, JMeter was used with 50 users and a total of 4 executions.

### 5.2 Results and Discussions

The results obtained from the experiments are presented below. In the load test, it was observed that in scenario (i), without our solution, it was very difficult to identify the anomalous service using only Jaeger with the call listing, and the method ended up blending in with the many calls that occurred. In scenario (ii), with our solution, the endpoint was configured with the Annotation specifying the values that should not be exceeded, as shown in Listing 2

#### Listing 2: Annotation Called

```
@GetMapping
@ObservabilityParam(params = {
    @Param(key = "cpuUsage", value = "15"),
    @Param(key = "memory", value = "25"),
    @Param(key = "responseTime", value = "550"),
    @Param(key = "throughput", value = "940")
})
public List<Owner> findAll(){ ... }
```

In addition to being set in this endpoint, the Annotation was applied to other services with different maximum values, as these values are chosen by the developer. Upon conducting the load test and checking the dashboard of our solution, we observed that five services with our annotation were called, as shown in Figure 4. We decided to examine the details of the highlighted endpoint, as it was flagged with a critical risk and we knew it could invoke our method that increases memory usage.

5 Services Called

Select by Priority ▾

|                                     |           |
|-------------------------------------|-----------|
| findOwner: /owners/{ownerId}        | low risk  |
| findAll: /owners                    | high risk |
| read: /owners/*/pets/{petId}/visits | low risk  |
| findPet: /owners/*/pets/{petId}     | low risk  |
| read: /pets/visits                  | low risk  |

### Figure 4: List of Called Services

Upon examining the service metrics, it was found that memory consumption was at a critical risk level. Figure 5 displays the memory usage metrics collected from that service, including maximum, minimum, average, and quantity overflows, along with the corresponding trace.

| Description        | Value        | SpanId                                 |
|--------------------|--------------|--|
| Max                | 96.546875 MB | <a href="#">Show Distributed Trace</a> |
| Min                | 3 MB         | <a href="#">Show Distributed Trace</a> |
| Median             | 51.1 MB      |  |
| Quantity Overflows | 23           |  |

### Figure 5: Service Details

By examining both traces, it is possible to compare them through distributed tracing. As shown in Figure 6, one trace included a call to the *causeMemoryLeak* method (the method responsible for the anomaly), while the other trace, as evidenced in Figure 7, did not include this call. Thus, with the fusion of metrics and distributed tracing, detecting failures or anomalies in microservices becomes more straightforward.

In addition to the anomaly tests, we also conducted tests to determine the percentage of failures that our system can identify. For this, an endpoint with the Annotation was configured to call a non-existent service, resulting in an exception. The result of this test, illustrated in Figure 8, shows that the proposed solution was able to identify 100% of the failures. Since the data is extracted from the Jaeger API, the solution retrieves the status of the traces, and all failed requests were correctly displayed on the dashboard.

### 5.3 Lessons Learned

Technology professionals can use this solution to optimize the maintenance of their microservices-based applications, reducing the occurrence of failures and improving the efficiency of anomaly detection. As demonstrated by the results, the integration of observability techniques, combining metrics and distributed tracing, proves to be highly effective. This fusion allows both to work together, providing valuable insights that help identify issues that, due to the size of this architecture, might otherwise remain hidden.

[illegible]

Figure 6: Trace with Anomalous Method

[illegible]

**Figure 7: Trace without Anomalous Method**

| GENERAL INFO         |       |
|----------------------|-------|
| Description          | Value |
| Requests             | 85    |
| Errors               | 85    |
| Response Time Median | 0.0s  |

### Figure 8: Failure Rate

It helps in identifying anomalous services and enhancing the overall stability and performance of the application. In this way, the research question is answered.

## 6 CONCLUSION AND FUTURE WORK

This document presented the FOCUS solution, which was developed using an observability tool (OTel) to perform data fusion of distributed tracing with metrics for anomaly and failure detection. This fusion provided a broader view of the microservices application, as it combined metrics with distributed tracing to easily



locate traces in case of metric exceedances. This results in quicker analysis and reduced chances of failures, as maximum values are set for the methods studied, and any values exceeding these thresholds are displayed on the dashboard by FOCUS. A limitation of this solution is that it is specific to Java-based microservices projects because it relies on Annotations, a feature unique to this language. Additionally, another limitation is the need for project team synchronization to select the most appropriate parameter values to avoid false positives in anomaly detection.

For future work, we plan to expand the FOCUS library to support other programming languages beyond Java and seek more efficient methods for anomaly detection, as well as apply machine learning algorithms using the data collected by OTel to recommend better metric configurations.

## REFERENCES

- [1] Omar Al-Debagy and Peter Martinek. 2018. A comparative review of microservices and monolithic architectures. In *2018 IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI)*. IEEE, 000149–000154.
- [2] Andre Bento, Jaime Correia, Ricardo Filipe, Filipe Araujo, and Jorge Cardoso. 2021. Automated analysis of distributed tracing: Challenges and research directions. *Journal of Grid Computing* 19, 1 (2021), 9.
- [3] A. Boten and C. Majors. 2022. *Cloud-Native Observability with OpenTelemetry: Learn to gain visibility into systems by combining tracing, metrics, and logging with OpenTelemetry*. Packt Publishing.
- [4] Clément Cassé, Pascal Berthou, Philippe Owezarski, and Sébastien Josset. 2021. Using distributed tracing to identify inefficient resources composition in cloud applications. In *2021 IEEE 10th International Conference on Cloud Networking (CloudNet)*. IEEE, 40–47.
- [5] Ramaswamy Chandramouli. 2019. Microservices-based application systems. *NIST Special Publication* 800, 204 (2019), 800–204.
- [6] Nicola Dragoni, Saverio Giallorenzo, Alberto Lluch Lafuente, Manuel Mazzara, Fabrizio Montesi, Ruslan Mustafin, and Larisa Safina. 2017. Microservices: yesterday, today, and tomorrow. *Present and ulterior software engineering* (2017), 195–216.
- [7] Nicola Dragoni, Ivan Lanese, Stephan Thordal Larsen, Manuel Mazzara, Ruslan Mustafin, and Larisa Safina. 2017. Microservices: How to make your application scale. In *International Andrei Ershov Memorial Conference on Perspectives of System Informatics*. Springer, 95–104.
- [8] David Garlan. 2008. Software architecture. (2008).
- [9] David Garlan and Mary Shaw. 1993. An introduction to software architecture. In *Advances in software engineering and knowledge engineering*. World Scientific, 1–39.
- [10] Robert Heinrich, André Van Hoorn, Holger Knoche, Fei Li, Lucy Ellen Lwakatare, Claus Pahl, Stefan Schulte, and Johannes Wettinger. 2017. Performance engineering for microservices: research challenges and directions. In *Proceedings of the 8th ACM/SPEC on international conference on performance engineering companion*. 223–226.
- [11] Rudolf E Kalman. 1960. On the general theory of control systems. In *Proceedings First International Conference on Automatic Control, Moscow, USSR*. 481–492.
- [12] Suman Karumuri, Franco Solleza, Stan Zdonik, and Nesime Tatbul. 2021. Towards observability data management at scale. *ACM SIGMOD Record* 49, 4 (2021), 18–23.
- [13] Joanna Kosińska, Bartosz Baliś, Marek Konieczny, Maciej Malawski, and Sławomir Zieliński. 2023. Towards the Observability of Cloud-native applications: The Overview of the State-of-the-Art. *IEEE Access* (2023).
- [14] C. Majors, L. Fong-Jones, and G. Miranda. 2022. *Observability Engineering: Achieving Production Excellence*. O'Reilly Media, Incorporated.
- [15] Nicolas Marie-Magdelaine, Toufik Ahmed, and Gauthier Astruc-Amato. 2019. Demonstration of an observability framework for cloud native microservices. In *2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*. IEEE, 722–724.
- [16] Nabor C Mendonça, Pooyan Jamshidi, David Garlan, and Claus Pahl. 2019. Developing self-adaptive microservice systems: Challenges and directions. *IEEE Software* 38, 2 (2019), 70–79.
- [17] Cesare Pautasso, Olaf Zimmermann, Mike Amundsen, James Lewis, and Nicolai Josuttis. 2017. Microservices in practice, part 1: Reality check and service design. *IEEE software* 34, 01 (2017), 91–98.
- [18] Mahsa Raeeszadeh, Amin Ebrahimzadeh, Ahsan Saleem, Roch H Glitho, Johan Eker, and Raquel AF Mini. 2023. Real-Time Anomaly Detection Using Distributed Tracing in Microservice Cloud Applications. In *2023 IEEE 12th International Conference on Cloud Networking (CloudNet)*. IEEE, 36–44.
- [19] Andy Singleton. 2016. The economics of microservices. *IEEE Cloud Computing* 3, 5 (2016), 16–20.
- [20] Aadi Thakur and MB Chandak. 2022. A review on opentelemetry and HTTP implementation. *International journal of health sciences* 6 (2022), 15013–15023.
- [21] Johannes Thönes. 2015. Microservices. *IEEE software* 32, 1 (2015), 116–116.
- [22] Ioannis Tzanettis, Christina-Maria Androna, Anastasios Zafeiropoulos, Eleni Fotopoulou, and Symeon Papavassiliou. 2022. Data Fusion of Observability Signals for Assisting Orchestration of Distributed Applications. *Sensors* 22, 5 (2022), 2061.
- [23] Muhammad Usman, Simone Ferlin, Anna Brunstrom, and Javid Taheri. 2022. A Survey on Observability of Distributed Edge & Container-based Microservices. *IEEE Access* (2022).
- [24] Muhammad Waseem, Peng Liang, and Mojtaba Shahin. 2020. A systematic mapping study on microservices architecture in devops. *Journal of Systems and Software* 170 (2020), 110798.

---

# Rapid Review of Digital Sovereignty and Open-Source Software

José A. P. Martins

jamp@alu.ufc.br

Federal University of Ceará - UFC

Fortaleza, Ceará, Brazil

José A. F. de Macêdo

jose.macedo@insightlab.ufc.br

Federal University of Ceará - UFC

Fortaleza, Ceará, Brazil

Paulo A. L. Rego

paulo@dc.ufc.br

Federal University of Ceará - UFC

Fortaleza, Ceará, Brazil

Vinícius L. R. da Costa

vinicius.1232@abin.gov.br

Research and Development Center for Communication

Security - CEPESC

Brasília, Distrito Federal, Brasil

## Abstract

The growing reliance on proprietary digital technologies and the global dominance of large technology providers have raised concerns about national autonomy, prompting the emergence of digital sovereignty as a strategic policy objective. This concept highlights the importance of states maintaining control over digital infrastructures, data governance, and technological capabilities. Despite its relevance, the relationship between digital sovereignty and open-source software (OSS) remains insufficiently mapped. Although OSS is often cited as a means to reduce vendor dependency and enhance transparency, the literature lacks a consolidated assessment of how OSS is adopted across domains, the challenges involved, and its actual contribution to sovereignty goals. This paper addresses this gap by conducting a structured Rapid Review of scholarly publications, grey literature, and case studies from 2020 to 2025. It offers a comparative synthesis of OSS alternatives to proprietary software across critical categories, including operating systems, cloud infrastructure, productivity tools, and communication tools. The review identifies strategic opportunities for sovereignty through OSS, such as improved auditability and support for local capacity. It also outlines persistent challenges related to sustainability, legal frameworks, and geopolitical tensions. The results suggest that realizing digital sovereignty through OSS requires integrated governance, long-term investment in technical expertise, and alignment with international regulatory efforts.

## Keywords

Digital Sovereignty, Open-Source Software, Government Policy, Cybersecurity, Software Governance, Vendor Lock-in, Technological Autonomy, Software Supply Chain, Rapid Review, ICT Infrastructure

## 1 Introduction

Digital sovereignty refers to the capacity of governments, institutions, or communities to retain control over digital infrastructure, data, and technology choices while minimizing external dependency [12, 18]. This topic has gained prominence in policy and technical agendas, particularly in response to the dominance of multinational technology providers and the strategic implications of relying on proprietary systems [4, 6, 15].

Within this context, open-source software (OSS)—software with publicly accessible source code—has emerged as a relevant tool for

supporting sovereignty objectives [22, 25]. OSS is associated with benefits such as auditability, transparency, and reduced vendor lock-in, and has been explored by various governments and institutions aiming to strengthen control over digital infrastructure [4, 8].

Despite these advantages, OSS adoption for sovereignty purposes introduces specific challenges. These include project sustainability, security vulnerabilities in supply chains, limited local expertise, and legal uncertainties [13, 14]. In some cases, reliance on OSS ecosystems dominated by large global companies may reproduce forms of dependency that sovereignty strategies aim to reduce [1, 22].

Recent efforts in emerging economies have broadened the scope of digital sovereignty beyond state-centered approaches. The BRICS countries, for instance, have articulated diverse strategies that include community-led infrastructures and open digital ecosystems. Belli and Jiang argue that digital sovereignty encompasses state power and grassroots initiatives where empowered communities use, develop, and control their own technological infrastructures [3]. Examples include community networks in Brazil, India, and South Africa, where OSS is foundational in enabling network self-determination and resisting corporate and governmental digital centralization [3]. These initiatives highlight how open-source software supports autonomy in design and implementation and localized innovation aligned with social and cultural needs.

In the European context, OSS has also been recognized as critical infrastructure for public administration. The FOSSEPS report, commissioned by the European Commission, identifies various open-source components essential to digital services across EU member states, including LibreOffice and libXML2. The study highlights key risks associated with the sustainability and governance of such projects, such as limited funding, contributor burnout, and lack of institutional coordination [7]. These findings reinforce the strategic value of OSS in supporting long-term digital autonomy and the importance of structured policies to maintain and evolve these assets.

Complementing this perspective, the European Commission launched the *EU Open Source Solutions Catalogue*, a centralized repository designed to promote reuse, transparency, and sovereignty across digital public services. This initiative, part of the broader Interoperable Europe policy framework, enables public administrations to discover and adopt OSS tools vetted and shared by peers across the Union [11]. The catalogue exemplifies how OSS ecosystems can serve as pillars of digital self-determination at scale

by fostering cross-border collaboration and reducing reliance on proprietary vendors.

This article offers a structured review of how OSS has supported digital sovereignty efforts. It applies a Rapid Review methodology to examine academic publications, policy reports, and grey literature produced between 2020 and 2025. The objective is to identify key OSS alternatives to proprietary solutions, analyze their adoption in different institutional settings, and discuss their use's main technical, organizational, and strategic implications.

The review does not aim for exhaustiveness but synthesizes recent contributions. It highlights recurring patterns in OSS-based sovereignty initiatives and identifies areas that require further institutional or policy development.

The remainder of this paper is organized as follows. Section 2 introduces key concepts underpinning digital sovereignty. Section 3 reviews recent research that connects open-source software to sovereignty agendas. Section 4 presents the study's design, objectives, and methodological approach. Section 5 synthesizes the main findings, emphasizing how OSS has been positioned within sovereignty strategies. Section 6 discusses structural and institutional challenges that affect the effectiveness of OSS in this context. Finally, Section 7 offers concluding remarks and outlines potential directions for further investigation.

## 2 Fundamental Concepts of Digital Sovereignty

Digital sovereignty is a multifaceted concept broadly defined as the capacity of a state, organization, or community to autonomously control, supervise, and manage its digital infrastructure, data, and technologies without undue reliance on external actors or foreign corporations [12, 20]. While the notion is often discussed regarding state power, it also encompasses individual and organizational dimensions of autonomy in digital spaces, including control over personal data, access to open technologies, and governance of on-line services [18, 25].

Historically, the emergence of digital sovereignty as a policy imperative reflects increasing concerns about the global dominance of large technology companies and the inadequacy of traditional territorial sovereignty frameworks in addressing transnational digital infrastructures [12, 15]. The EU, in particular, has integrated the concept into its broader digital strategy, promoting the development of regulatory tools to reduce dependence on foreign technologies while preserving openness and competition. Instruments such as the General Data Protection Regulation (GDPR), the Digital Services Act (DSA), and the Digital Markets Act (DMA) exemplify efforts to assert European control over digital governance [22, 25].

Elsewhere, national approaches differ. The United States does not explicitly invoke the term "digital sovereignty" but has pursued security-oriented policies to limit technological reliance on strategic rivals, as evidenced by bans on certain foreign suppliers in telecommunications and infrastructure [22]. Conversely, China has promoted an assertive model of "cyber sovereignty" focused on centralized state control, data localization, content regulation, and the strategic advancement of domestic platforms. Legal frameworks such as the Cybersecurity Law underpin a vision of technological independence aligned with national priorities [1, 22].

Digital sovereignty entails navigating a constellation of legal, technical, and economic dependencies. Legal risks include exposure to extraterritorial legislation, such as the U.S. CLOUD Act, which can compel access to data stored by American firms abroad [10]. Technical dependency arises when critical systems rely on proprietary, opaque platforms that limit auditability or adaptability. Economically, sovereignty is constrained when dominant external providers marginalize domestic innovation [18, 22].

To counter these forms of dependency, governments have explored regulatory reform, national digital infrastructure initiatives, and the strategic deployment of OSS. It has become a tool aligned with sovereignty goals, enabling greater transparency, cost control, and adaptability. However, its successful deployment also depends on local capacity to sustainably maintain, govern, and evolve open systems. As such, sovereignty through OSS cannot be understood as purely technological—it must also be institutional and participatory [8, 22].

## 3 Related Work

This section presents a comparative overview of selected studies that address digital sovereignty from different perspectives—technical, regulatory, and institutional—while integrating OSS-related strategies. These works represent a diverse set of methodological approaches and thematic contributions, providing important foundations and contrast points for positioning this review. Table 1 synthesizes the scope and methods of each study.

The report by Blind et al. [5] commissioned by the European Commission provides a comprehensive assessment of the strategic relevance of OSS and open hardware in promoting EU technological autonomy. It connects OSS adoption to innovation ecosystems and long-term competitiveness. In parallel, Tan et al. [23] deliver a systematic review focused on self-sovereign identity, discussing governance and interoperability—issues indirectly relevant to OSS but highly pertinent to digital autonomy.

On the technical side, Biström et al. [4] analyze OSS deployment in educational institutions, identifying sociotechnical barriers and institutional inertia. Osborne [21] examine the funding and governance of the scikit-learn project, emphasizing sustainability challenges in volunteer-based development models. These works offer insights into domain-specific OSS adoption but are not comparative in nature.

From a policy and governance angle, Tridgell [25] explore regulatory alignment between OSS and EU cybersecurity frameworks. Bechara and Lechner [2] reflect on governance models for public infrastructure, proposing that OSS must be embedded in broader institutional frameworks to achieve sovereignty outcomes. Dreyer [9], through empirical data from DACH institutions, highlights the operational complexities and adoption rates of OSS within academic and government settings.

While these studies offer valuable insights into specific aspects of digital sovereignty and OSS, this review contributes in two principal ways. First, it applies a structured *Rapid Review* methodology to consolidate findings from technical, regulatory, and institutional perspectives. Second, it provides a cross-domain synthesis of OSS alternatives to proprietary solutions, offering a comparative view

**Table 1: Selected comparative studies on digital sovereignty and OSS.**

| Author(s)                  | Ref. | Scope                               | Methodology                            |
|----------------------------|------|-------------------------------------|--|
| Blind et al. (2021)        | [5]  | EU policy, OSS/hardware adoption    | Technical Report                       |
| Tan et al. (2024)          | [23] | Digital identity, interoperability  | Systematic Review                      |
| Biström et al. (2024)      | [4]  | OSS adoption in education           | Technical Case Study                   |
| Osborne (2024)             | [21] | OSS sustainability and governance   | Case Study (Qualitative)               |
| Tridgell (2025)            | [25] | EU cybersecurity and OSS regulation | Regulatory/Policy Analysis             |
| Bechara and Lechner (2024) | [2]  | OSS governance in public policy     | Conceptual/Policy Paper                |
| Dreyer (2025)              | [9]  | Institutional OSS usage in DACH     | Survey-Based Empirical Report          |
| This Work (2025)           | —    | Cross-domain OSS alternatives       | Rapid Review and Comparative Synthesis |

across key infrastructure layers. Rather than exhausting the research agenda, this work seeks to establish an analytically grounded entry point for future investigations. By mapping existing knowledge and identifying persistent gaps, it provides a structured foundation for subsequent empirical, comparative, and policy-oriented studies on digital sovereignty and open-source adoption.

#### 4 Research Design and Methodology

This study adopts a Rapid Review approach, which follows systematic principles but allows for flexibility in scope and time. This method is particularly suited for emerging topics that span multiple fields, such as digital sovereignty and open-source software [24].

The research was guided by three primary questions:

**RQ1:** How is OSS adopted to support digital sovereignty goals in practice?

**RQ2:** What are the main benefits and challenges of using OSS to reduce technological dependency?

**RQ3:** Which categories of OSS are most relevant in sovereignty-oriented contexts, and how do they compare to proprietary alternatives?

**RQ4:** How do the identified OSS adoption strategies relate to the specific challenges in each scenario, according to software category?

These questions shaped the scope of the review and the selection criteria for relevant documents.

We conducted searches across academic databases, including ACM Digital Library, IEEE Xplore, and Google Scholar, as well as grey literature from institutional sources like the European Commission and OSPO Alliance. Additionally, policy reports and technical documentation available on Google Scholar were incorporated into the review. The time frame for the review was from 2020 to early 2025.

Documents were included based on specific criteria:

- The document discusses the adoption or role of OSS in digital sovereignty;
- It focuses on public sector, government, or institutional settings;
- It presents empirical evidence, technical analysis, or policy arguments; and
- It was published in English during the specified period.

The screening process involved reviewing titles and abstracts to identify relevant texts, which were then read in full. Data extraction was performed based on predefined categories, such as publication

type, domain, OSS categories, and relevance to sovereignty (autonomy, transparency, governance). The classified documents were then synthesized to highlight common themes. Figure 1 provides a summary of the review process, including the definition of the research scope, document search and selection, data extraction, and synthesis.

As a Rapid Review, this study followed a streamlined approach with clearly defined scope and time constraints. Although efforts were made to ensure comprehensive coverage and relevance, some documents may have been overlooked. Grey literature presented challenges regarding consistency and methodological transparency. To mitigate these issues, we established clear inclusion criteria and cross-validated sources wherever possible.

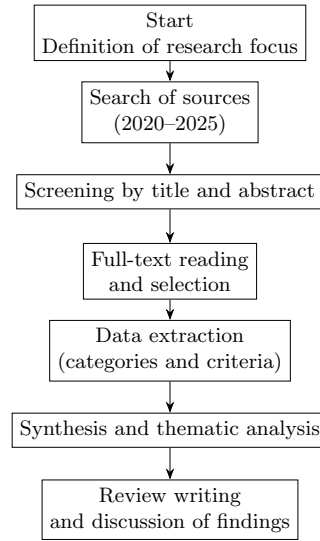
#### 5 Findings in Digital Sovereignty and Open-Source Software

This section presents the main findings of the review, based on academic literature, policy documents, and case studies. It highlights recurring themes and contrasting approaches related to the role of open-source software (OSS) in advancing digital sovereignty. Rather than summarizing each source individually, the discussion draws connections across documents, emphasizing patterns in adoption strategies, institutional contexts, and software categories.

Attention is given to convergences and divergences across national initiatives, as well as to practical and policy-relevant implications. Examples and comparative insights are used to illustrate how OSS is positioned as both a technical and strategic component within broader efforts to reduce digital dependency and promote technological autonomy.

##### 5.1 RQ1: OSS adoption strategies for digital sovereignty

Several studies highlight government and institutional efforts to adopt OSS as a means of achieving technological autonomy. Case studies from the EU, Germany, and China show divergent strategies shaped by distinct regulatory cultures and geopolitical interests. For example, Klare et al.[22] describe China’s promotion of HarmonyOS and Unified Operating System (UOS) as part of a broader national strategy to reduce dependency on U.S. technology and assert state control over digital ecosystems. This strategy is reinforced by strong state-led coordination and legal instruments such as the Cybersecurity Law, which frames OSS as a vehicle for domestic technological development[1].



**Figure 1: Steps followed in the Rapid Review methodology.**

In the European Union, the approach is anchored in regulatory frameworks that seek to balance openness, innovation, and sovereignty. Tridgell [25] and Moerel and Timmers [19] emphasize the EU’s use of instruments like the Cyber Resilience Act and Gaia-X to integrate OSS into efforts aimed at strengthening control over infrastructure while preserving interoperability and competition. These initiatives represent a governance model of digital sovereignty, using OSS not only for cost efficiency but also as a strategic asset aligned with democratic values and transparency.

Germany represents an intermediate case. As discussed by Dreyer [9], public institutions across the DACH region have adopted OSS with varying levels of success. While municipalities and universities increasingly rely on open platforms for service delivery, adoption often depends on local leadership and internal expertise. The lack of consistent national policy frameworks has led to fragmented implementations, revealing the importance of institutional capacity and coordination.

Together, these examples illustrate that OSS adoption for sovereignty goals is shaped not only by technical criteria but also by political models, regulatory environments, and administrative traditions. The role of the state—centralized, federated, or regulatory—plays a key role in determining the coherence and impact of such strategies.

## 5.2 RQ2: Benefits and challenges of OSS for reducing dependency

In our analysis, we identified a set of recurring challenges that organizations face when adopting OSS. Table 2 summarizes these obstacles, grouped by category. Studies consistently emphasize that OSS can reduce technological dependency by offering transparent, adaptable, and cost-effective alternatives to proprietary systems. Key advantages include inspecting and modifying source code, avoiding vendor lock-in, and aligning technological infrastructure

with public values such as transparency and interoperability [8, 14, 19]. OSS also fosters local innovation by allowing institutions to customize solutions to their specific needs, which can strengthen domestic expertise and reduce reliance on external vendors [16].

OSS is increasingly framed as a critical infrastructure component from a policy perspective. In the EU, regulatory instruments such as the Open Source Software Strategy (2020–2023) and the Interoperable Europe Act explicitly promote OSS adoption to build resilient digital ecosystems [25]. These initiatives are based on the premise that sovereignty requires technological independence and transparent and accountable systems.

However, these benefits are often offset by structural and operational challenges. One of the most frequently cited limitations is the lack of long-term sustainability in OSS projects, particularly those maintained by small teams or volunteers. As illustrated by the Log4Shell vulnerability, critical OSS components can become weak points in national infrastructure when underfunded or inadequately governed [13, 14].

Security in open ecosystems also poses specific risks. While openness enables auditability, it does not guarantee that audits occur. Dependency hijacking, supply chain attacks, and delayed patching are recurrent concerns. As noted by Osborne [21], even mature projects like scikit-learn face funding shortages that limit their capacity to respond to emerging threats.

Human capital remains another persistent constraint. Effective OSS adoption requires teams with the ability to deploy, maintain, and contribute to open technologies. Public institutions often lack this internal capacity, particularly in sectors where proprietary vendors have dominated for decades [8, 9].

Finally, the perception of OSS as a “free” alternative can lead to underestimation of the investment required for governance, integration, and lifecycle maintenance. Several studies warn that substituting proprietary tools with OSS without accompanying

Table 2: Structural Challenges for OSS Adoption in Sovereignty Strategies.

| Category               | Key Challenge                                    | Critical Considerations  |
|------------------------|--|--|
| Sustainability         | Long-term maintenance of OSS projects            | Dependency on small, underfunded contributor bases can lead to critical vulnerabilities (e.g., Log4Shell), requiring state-backed maintenance policies [14]. |
| Security               | Supply chain and code integrity risks            | Requires audit mechanisms, SBOM policies, and cryptographic code validation to ensure infrastructure resilience [13].  |
| Human Capital          | Limited technical expertise in the public sector | Successful use of OSS requires strategic investment in technical training and career support [8, 17].  |
| Legacy Systems         | Complexity in migration from proprietary systems | Transitioning from embedded proprietary environments requires phased planning and change management [17, 22].  |
| Corporate Influence    | Dominance of OSS projects by tech firms          | Risk of reproducing dependency unless governments fund and participate in OSS governance structures [1, 22].   |
| Geopolitical Pressures | Reactions to sovereignty-driven procurement      | Sovereignty decisions may generate diplomatic tensions; must be managed through multilateral cooperation [13, 22].   |
| Legal Frameworks       | Complexity of OSS licenses                       | Procurement and compliance policies must adapt to license models like the GPL [8, 14].   |

organizational change may result in superficial gains and hidden dependencies [2, 22].

The evidence shows that OSS can reduce dependency and support sovereignty efforts, but only when adopted as part of a coordinated institutional strategy. This includes financial support, training, governance structures, and integration with broader public digital policies.

5.3 RQ3: OSS categories relevant to sovereignty and comparisons

The reviewed literature identifies fundamental categories of software where OSS adoption is strongly associated with sovereignty objectives. These include desktop and server operating systems, office productivity tools, cloud platforms, mobile systems, and database management systems.

Table 3 presents a comparative synthesis of open-source and proprietary solutions across these categories. It draws on findings from recent policy and technical analyses [1, 8, 14, 22], highlighting the strategic advantages of OSS in terms of transparency, cost control, and customizability, while also acknowledging recurring adoption challenges such as a lack of internal expertise and interoperability issues.

This comparison illustrates that OSS solutions are more mature and widely adopted in infrastructure-related domains such as servers and databases. In contrast, user-facing categories such as mobile platforms, office productivity suites, and communication tools still face notable adoption hurdles, particularly regarding ecosystem maturity, usability, and user experience.

In the specific case of communication tools, using platforms like Matrix/Element and Jitsi Meet underscores the strategic importance of local data hosting, transparency, and interoperability—key aspects for digital sovereignty. However, proprietary services such as Zoom and Microsoft Teams remain dominant due to their ease of use, polished interfaces, and widespread adoption. These factors present considerable challenges for OSS alternatives. Addressing these obstacles requires technical improvements and sustained institutional commitment, user training, and investments aimed at improving performance and scalability [19, 25]. The effective

implementation of open-source communication platforms can significantly enhance sovereignty outcomes by reducing dependency on foreign providers and mitigating surveillance risks.

Overall, these findings suggest that technical readiness alone is not sufficient. The adoption of OSS in sovereignty-relevant categories depends equally on coordinated institutional support, political mandates, and long-term investments in training and infrastructure support.

5.4 RQ4: Integrated View of Strategies and Challenges

Table 4 shows that for *Infrastructure*, partnerships, and open-source audits are most cited to address legacy compatibility and data security. In *Enterprise Applications*, workshops and pilot cases help overcome resistance. In *Development Tools*, training and academic collaborations mitigate skill gaps. These correlations confirm the relevance of strategies and challenges.

The strategy versus challenge mapping reveals that no single solution fits all contexts. Each OSS category requires specific approaches. For example, governments that form collaborative initiatives between universities and local companies tend to address skill deficits faster. This integrated perspective helps tailor implementation plans to real-world obstacles.

Beyond category-specific strategies, the table reveals patterns in the types of challenges that recur across multiple OSS categories. For instance, “lack of skills” appears not only under *Development Tools* but also in *Enterprise Applications*, indicating that skill deficits can hinder both tool usage and broader organizational adoption. In such cases, a combined approach—pairing internal training programs with pilot demonstrations—can create a positive feedback loop: as staff become more proficient, they serve as advocates for OSS within their departments. This dual emphasis on capacity building and hands-on exposure accelerates cultural acceptance and reduces the time required to scale solutions beyond initial pilots.

Moreover, the table highlights that some challenges demand multi-layered strategies. Take “data security,” which is cited under *Infrastructure* but can also affect *Enterprise Applications* and *Development Tools*. Addressing data security purely through code

**Table 3: Comparative Analysis of Open-Source Software vs. Proprietary Software.**

| Software Category             | Open-Source Examples                                | Proprietary Examples                      | Key Advantages of OSS   | Adoption Challenges for OSS                                     |
|-------------------------------|---|---|---|---|
| Desktop Operating Systems     | Ubuntu, Debian, UOS                                 | Windows, macOS                            | Security, flexibility, cost-efficiency                              | User familiarity, compatibility issues                          |
| Server Operating Systems      | Ubuntu Server, Red Hat Linux                        | Windows Server, IBM Unix                  | Scalability, transparency, lower TCO                                | Requires skilled personnel, maintenance complexity              |
| Office Productivity Suites    | LibreOffice, Apache OpenOffice                      | Microsoft Office 365, Google Workspace    | Cost-saving, open standards, autonomy                               | Document compatibility, user acceptance                         |
| Cloud and Collaboration Tools | OpenStack, Nextcloud                                | AWS, Azure, Microsoft Teams               | Data control, lower costs, reduced surveillance risk                | Need for internal expertise, security management                |
| Mobile Operating Systems      | LineageOS, HarmonyOS                                | Google Android, Apple iOS                 | Data autonomy, regulatory compliance                                | Ecosystem maturity, security updates                            |
| Database Management Systems   | PostgreSQL, MariaDB, MongoDB                        | Oracle DB, Microsoft SQL Server           | Transparency, customization, reduced costs                          | Migration complexity, internal expertise required               |
| Communication Tools           | Matrix/Element, Jitsi Meet, Mattermost, Rocket.Chat | Zoom, Microsoft Teams, Slack, Google Meet | Local data hosting, auditability, interoperability, privacy control | User interface polish, performance scalability, network effects |

audits may not suffice if organizational policies remain misaligned. Therefore, pairing open-source audits with clear governance frameworks—such as formal security guidelines or compliance checklists—bridges the gap between technical and institutional levels. In practice, this means that decision makers should treat certain challenges as systemic, requiring synchronized actions: technical verification through audits, policy alignment via governance, and continuous monitoring through community or third-party channels.

## 6 Challenges to Digital Sovereignty in OSS

Adopting OSS to support digital sovereignty entails a range of strategic benefits, but it also introduces complex structural and institutional challenges that must be addressed to ensure sustainable and autonomous outcomes [8, 14, 25]. This section discusses key obstacles based on recent analyses and practical experiences, and synthesizes them into critical domains.

A major concern is the long-term sustainability of OSS projects. Many critical components of digital infrastructure depend on a few contributors or underfunded maintainers. Incidents such as the Log4Shell<sup>1</sup> vulnerability have revealed how fragile this model can be [14]. Governments and institutions must consider upstream contributions, support OSPOs, and ensure adequate funding mechanisms to stabilize maintenance cycles.

Security in OSS supply chains is another recurring issue. The openness of code allows for transparency but also introduces risks such as dependency hijacking or undetected vulnerabilities in widely used libraries [13]. Countermeasures include code signing, Software Bills of Materials (SBOM), and independent security audits, all strengthening trust in OSS infrastructure.

Human capital limitations further constrain OSS adoption in public systems. Sustained deployment and customization of open technologies require qualified personnel, who are often lacking in institutional settings [22]. Investments in public-sector training, OSS career pathways, and long-term capacity-building are necessary to support adoption beyond procurement.

<sup>1</sup>Log4Shell is a critical remote code execution vulnerability in the Apache logging library.

Migrating from legacy proprietary systems remains another barrier. These systems are often deeply embedded in workflows, with limited documentation and vendor entrenchment [17]. Effective transition strategies should be gradual, risk-aware, and supported by end-user engagement and technical coordination.

A less visible but significant challenge is the increasing concentration of influence in OSS ecosystems. While OSS is intended to decentralize control, many core projects are dominated by a small set of global technology firms [1, 22]. Governments must engage proactively in governance processes and fund the development of independent alternatives.

Legal and geopolitical dimensions also shape OSS adoption. Complexities in software licensing—particularly in copyleft models—require adaptations in procurement and legal frameworks [8, 14]. At the same time, shifting away from proprietary foreign vendors may trigger diplomatic sensitivities, underscoring the importance of multilateral coordination and digital diplomacy [13]. Table 2 summarizes the aforementioned challenges.

In short, OSS adoption can reinforce digital sovereignty, but only if paired with long-term institutional alignment. These challenges highlight that sovereignty is not achieved through technical substitution alone, but through coordinated strategies involving governance, infrastructure, and human capability.

Addressing the challenges outlined in this section is essential to unlock the full potential of OSS in digital sovereignty strategies. The issues discussed—ranging from sustainability and security to institutional capacity and geopolitical alignment—highlight that sovereignty cannot be achieved through technical substitution alone. Rather, it must be supported by long-term governance mechanisms, coordinated public investment, and adaptive legal frameworks. These findings reinforce the view that digital sovereignty is not merely a matter of control over software infrastructure but a broader political and institutional project that requires systemic commitment and strategic foresight.

Table 4: Matrix of OSS Adoption Strategies × Challenges, by Software Category

| Software Category       | Challenge                         | Adoption Strategy   |
|-------------------------|-----------------------------------|---|
| Infrastructure          | Compatibility with legacy systems | Partnership with external support providers; gradual migration using test environments. |
| Infrastructure          | Data security                     | Open-source code audits; collaboration with security-focused communities.               |
| Enterprise Applications | Organizational resistance         | Awareness workshops; pilot demonstrations of successful cases.                          |
| Development Tools       | Lack of skills                    | Internal training programs; collaboration with local universities.                      |

7 Conclusion and Future Work

This article examined how OSS contributes to digital sovereignty strategies, with emphasis on institutional adoption, technical alternatives, and structural challenges. Using a Rapid Review methodology, the study mapped recent academic, policy, and technical literature to assess the relevance of OSS in different national and organizational contexts.

The findings show that OSS is often regarded as a way to strengthen autonomy, transparency, and control over digital infrastructure. Its use is growing in key domains such as server environments, office suites, and cloud services. However, the analysis also identified persistent limitations. These include lack of skilled personnel, difficulties in migrating from legacy systems, legal complexity, and concentration of influence in certain OSS ecosystems. Together, these factors indicate that OSS adoption requires coordinated planning and long-term support.

While the review offers valuable insights, it is important to note that the Rapid Review methodology emphasizes thematic breadth rather than exhaustive coverage, which may have resulted in the omission of some relevant experiences. Additionally, the dynamic nature of digital policy and OSS development may influence the applicability of the findings over time.

Further work could examine concrete national strategies for OSS adoption, analyze the role of public OSPOs, or assess the influence of specific governance models. There is also space for interdisciplinary research on how legal, institutional, and technical factors interact in sovereignty-oriented initiatives.

Overall, the study highlights that OSS is not a ready-made solution, but a strategic tool whose impact depends on broader institutional and political arrangements.

References

[1] R. Arcesati and M. Meinhardt. 2021. China Bets on Open-Source Technologies to Boost Domestic Innovation. MERICS China Industry Brief, August 2021. Available online: <https://merics.org/en/report/china-bets-open-source-technologies-boost-domestic-innovation>.

[2] John Bechara and Ulrike Lechner. 2024. Digital Sovereignty and Open-Source Software-A Discussion Paper. In *International Conference on Innovations for Community Services*. Springer, UE, 397–407.

[3] Luca Belli and Min Jiang. 2024. Digital Sovereignty from the BRICS: Structuring Self-determination, Cybersecurity, and Control. In *Digital Sovereignty in the BRICS Countries: How the Global South and Emerging Power Alliances Are Reshaping Digital Governance*, Min Jiang and Luca Belli (Eds.). Cambridge University Press. <https://ssrn.com/abstract=4666615> Pre-print version available.

[4] Malte Biström, Kristoffer K. Adolfsson, and Matteo Stocchetti. 2024. Open-Source Software and Digital Sovereignty: A Technical Case Study on Alternatives to Mainstream Tools. In *Smart Technologies in Education 2024*. Vol. 1027. Springer, 106–113. doi:10.1007/978-3-031-61891-8\_10

[5] Knut Blind, Mirko Böhm, Paula Grzegorzewska, Andrew Katz, Sachiko Muto, Sivan Pätsch, and Torben Schubert. 2021. The impact of Open Source Software and Hardware on technological independence, competitiveness and innovation in the EU economy. 330 pages. doi:10.2759/430161

[6] D. Broeders, F. Cristiano, and M. Kaminska. 2023. In Search of Digital Sovereignty and Strategic Autonomy. *Journal of Common Market Studies* 61, 5 (2023), 1261–1280. doi:10.1111/jcms.13448

[7] Deloitte Consulting and Inno3. 2022. *Critical Software Study Report: Identify (and Help Fix) Critical Open Source Software Used by European Public Services*. Technical Report. European Commission, DIGIT. <https://joinup.ec.europa.eu/collection/foseps/Study-for-the-FOSSEPS-Pilot-Project>.

[8] Courmont. 2021. A Primer on Digital Sovereignty & Open Source. <https://www.opensourcerers.org/2021/08/09/a-primer-on-digital-sovereignty/> Accessed: 2025-04-15.

[9] Malte Dreyer. 2025. Digital Sovereignty – Results from the ZKI Survey in DACH. *EPIC series in computing* 105 (2025), 180–171. <https://easychair.org/publications/paper/bfg7/download>

[10] OpenForum Europe. 2022. Cloud Services and the EU Digital Autonomy: For Digital Sovereignty and EU Security, Neutralise the US CLOUD Act. Illuminem article, 2022. Available online at [urlhttps://illuminem.com/illuminemvoices/for-digital-sovereignty-and-eu-security-the-first-step-is-to-neutralise-the-us-cloud-act](https://illuminem.com/illuminemvoices/for-digital-sovereignty-and-eu-security-the-first-step-is-to-neutralise-the-us-cloud-act).

[11] European Commission. 2023. EU Open Source Solutions Catalogue Now Live. <https://interoperable-europe.ec.europa.eu/interoperable-europe/news/eu-open-source-solutions-catalogue-now-live>. Accessed April 2025.

[12] Luciano Floridi. 2020. The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU. *Philosophy & Technology* 33, 3 (2020), 369–378. doi:10.1007/s13347-020-00423-6

[13] D. Ghack. 2022. GitHub Blocks Accounts of Two Russian Banks Amid US Sanctions. HackRead, April 16, 2022.

[14] B. Johnson. 2021. Open-Source Software as a Global Public Good. *Journal of Digital Innovation* 3, 2 (2021), 10–22.

[15] B. C. Larsen. 2022. The Geopolitics of AI and the Rise of Digital Sovereignty. Brookings Institution Report, December 8, 2022.

[16] Johan Linåker and Sachiko Muto. 2024. Software Reuse through Open Source Software in the Public Sector: A Qualitative Survey on Policy and Practice. *RISE Report 2024:7* (2024). <https://ri.se>

[17] J. Miller. 2020. Munich’s Linux Migration and Reverse: A Socio-Technical Analysis. *IEEE Software* 37, 5 (2020), 15–17.

[18] Sanjay Misra, Kousik Barik, and Petter Kvalvik. 2025. Digital Sovereignty in the Era of Industry 5.0: Challenges and Opportunities. *Procedia Computer Science* 254 (2025), 108–117. doi:10.1016/j.procs.2025.02.069

[19] Lokke Moerel and Paul Timmers. 2021. Reflections on Digital Sovereignty. *EU Cyber Direct, Research in Focus series* (2021). Available at SSRN: <https://ssrn.com/abstract=3772777>.

[20] Milton L. Mueller. 2020. Against Sovereignty in Cyberspace. *International Studies Review* 22, 4 (2020), 779–801. doi:10.1093/isr/viz044

[21] Michael Osborne. 2024. Public-Private Funding Models in Open Source Software Development: A Case Study on scikit-learn. *Journal of Open Research Software* (2024). Forthcoming.

[22] Alice Pannier. 2022. Software Power: The Economic and Geopolitical Implications of Open Source Software. IFRI Report, December 2022. Available online.

[23] Chunlei Tan, Jun Zhang, Xiaodong Cheng, Yutao Ma, and Qiang Yang. 2024. Survey on Digital Sovereignty and Identity: From Digitization to Digitalization. *Comput. Surveys* 57, 1 (2024), 1–38. doi:10.1145/3616400

[24] Andrea C. Tricco et al. 2017. Rapid reviews to strengthen health policy and systems: a practical guide. *World Health Organization* (2017).

[25] Jennifer Tridgell. 2025. Open or Closing Doors? The Influence of Digital Sovereignty in the EU’s Cybersecurity Strategy on Open-Source Software Security. *Computer Law & Security Review* 56 (2025), 106078. doi:10.1016/j.clsr.2024.106078



---

# From male dominated-Telecommunication, Network and Cyber jobs to more win-win jobs and companies

Corinne Houssain

Département R&T - IUT,

Aix-Marseille Université

- INSPE, ADEF

Marseille, France

corinne.houssain@univ-amu.fr

Karima Boudaoud

Université Côte d'Azur -

Laboratoire I3S-CNRS

Sophia Antipolis, France

karima.boudaoud@univ-

cotedazur.fr

Sabrina Marques

Docaposte

Sophia Antipolis, France

sabrina.marques@docaposte.fr

## Abstract

In France, tech professions, and in particular those linked to cybersecurity, do lack a pool of experts. It is well known that women desert this sector in which men prevail. Several studies have been conducted to better understand this phenomenon in the general tech field. However, only few have focused on networks and telecommunications sectors. Furthermore, studies regarding the lack of women following cybersecurity curricula or choosing jobs in this sector are sparse. The aim of this paper is to present some results of a study carried out to measure the reasons and factors responsible for the lack of women in this field in France, reviewing the situation both in the preprofessional and professional periods to better understand the issue. More specifically, we have focused on university students (in a three-year-postbaccalaureate) following Networks and Telecommunications path at the Technology University Institute (IUT - Institut Universitaire de Technologie), and alumni of this curriculum who are currently working.

## CCS Concepts

• **Human-centered computing** → **User studies**; • **Applied computing** → **Education**; • **Social and professional topics** → **Computing profession**; **Gender**.

## Keywords

Gender, Telecommunications, Networks, Cybersecurity, Women in tech, Higher education

## 1 Introduction

Today, in France, only 27% of tech jobs are held by women [17] and only 16% in technical professions [15]. Internationally, only 22% of AI professionals are women [19] and 5% of tech startups are founded by women [13]. Moreover, 50% of women leave tech by the age of 35 [21] compared with 20% in other sectors. This under-representation of women has major economic, social, and technological consequences in this sector, where there is a shortage of expertise.

The **November 2023 report by the French High Council for Equality** points to the persistence of gender stereotypes in career choices and sexism in many tech environments [21]. It is undeniable that the development of computer programs and the democratization of AI via ChatGBT, for example, are helping to transform our societies and the place of women, for better or worse. Algorithms and AI in fact reproduce the biases of their designers. At the second edition of the National conference on the feminization of digital professions and sectors organized by *femmes@numérique* in February 2024 [17], it was recalled that the major challenge is to eliminate bias in technology, particularly in the construction of algorithms and tools using AI. As a digital scientist and entrepreneur Aurélie Jean points out that “Artificial intelligences possesses algorithmic biases (leading to what we call technological discrimination) because of our own cognitive biases” [1].

The documentary ‘Coded Bias’, made in 2020, clearly presents the fact that facial recognition performed on women works less well than on men and lifts the veil on the sexist and racist biases of algorithms. So, for instance, biases were pointed in facial recognition which was less effective for women and particularly for women of colour where 34% errors were found out [24] or in recruitment tools. For example, Amazon abandoned its algorithms in 2018, which had biases that penalized female CVs. Some negative impacts have certainly been noted and corrected since then. Google tried to reduce the biases of its translation algorithms by 50% by integrating more women into its teams in 2021; however, it is still perpetuating sexist stereotypes [23]. These gender biases are mainly due to biases in the data collected and used to construct the learning and AI models.

It is clear that if new digital technologies are created solely by men, gender bias will be introduced into these technologies. Having diverse teams of engineers is therefore essential if we want to use unbiased tech systems.

In addition, there is one technical field more obscure to the public than AI, where the under-representation of women is even more glaring which is the cybersecurity sector. In

France, even if the computer science field has seen a significant increase in the number of young women as attitudes are changing, it is surprising to find that technology in general remains a field neglected by young women, particularly those involved in cybersecurity. This is even more striking in the poorly identified digital field of Networks and Telecommunications, as we are witnessing the emergence of exciting new research and work opportunities in which women could and should have their place.

The aim of this paper is to present the results of a study we have conducted to measure the impact of external influence on the under-representation of young women in the R&T (Réseaux & Télécoms) specialization of the Institut Universitaire Technologique (IUT) focusing on the first three years of university education. The ultimate goal of this study is to suggest avenues for improvement in the light of the few elements highlighted by the data collected.

The paper is structured as follows. We first give an overview about existing related studies. Then, we present the methodology followed to conduct the study. Afterwards, we discuss the results based on the collected data. Finally, we propose some actions that can be done to improve the situation and increase the number of women in tech fields.

## 2 Related work

The barriers to digital equality between the sexes are varied, and very often rooted in long-standing social structures and ideas that privilege men over women. Studies such as those carried out by Accenture[2], which focuses on cloud computing, digital and security, activities that now account for around 70% of its revenues, show that this lack of role models in the digital professions can perpetuate future under-representation, as the absence of role models prevents young women from imagining themselves as successful computer scientists or engineers

Surprisingly enough, the gender gap is often greater in economically more developed countries, where well-paid and more stable jobs can further attract women [22]. AI is being denounced as the “new engineering of power” [20]. From design to use, gender norms circulate in productions, traces, discourses, and practices.

Both men and women mistakenly believe that men in computing have a better average. This may be explained by the fact that girls tend to have less computer experience [3]. They thus reproduce, at the level of schooling, the unequal gender relations that organize society as a whole. Nicole Mosconi [25] questions the responsibility of teaching in this process. In her study, she shows that 2/3 of a teacher’s time is devoted to boys and 1/3 to girls. It is therefore at the educational level that this imbalance is initiated, but also at the level of the family and societal environment, as gift theory is taken up by

students and shared by parents [4]. Relationships between students also play a part in girls’ censorship of scientific studies. Female minority definitely creates gendered differences in confidence in succeeding which prevent those who could have been attracted to choose this field or to specialize in it [9]. In the digital professions, women have always been in the minority, but never to the point where today men represent 90% of them.

Society, too, recognizes the importance of diversity in the working environment to achieve higher quality results [6]. There is an undeniable effect of diversity on the performance of science companies [28].

There is a perception that anyone with the right skills, knowledge and experience can pursue certain technological careers. And yet, the IT sector is still considered a “man’s job” by society [14].

There is no doubt, women as a minority in IT, specifically in Cybersecurity, is an inextricable case as concretely barely anything for the past 30 years has been changing, their number in the technical side of IT being still very low [5]. This low number, that is confirmed both in the tech university curriculum and job market, has motivated the study that we conducted to better understand the blocking factors.

## 3 Methodology

To conduct our study, we used a quantitative method (an online survey), the use of questionnaires being the preferred method, although answers to open-ended questions can shed more qualitative light on the results. Initially, the target was put on the first three years of university education focusing on the local scale (i.e. local students at Aix-Marseille R&T département), then on a national one (i.e. students of the other R&T départements in France). Then, the data collection continued with alumni who had obtained a two-year technology diploma (DUT - Diplôme Universitaire Technologique) or a 3rd year Professional Bachelor diploma at the Aix-Marseille département.

The online survey was designed to collect information from five graduating classes locally, in Aix-Marseilles, making the total sample covered of 167 students, among whom 15 women. The survey included closed-ended multiple-choice questions, open-ended questions, and rating questions using a Likert scale.

To broaden the sample and reinforce the results obtained, in winter 2023, the questionnaire was sent to colleagues in the R&T community, i.e., the 28 other departments, with the same message and the same questionnaire. A few members of the teaching team passed it on to students. Students from 8 RT départements (from 8 different cities) responded, including 1st and 2nd year students, as well as students from the professional bachelor’s program. The percentage of students, by

department, who responded was as follows : **Blagnac** (25.2%), **Colmar** (19.6%), **Nice** (18.9%), **Valence** (15.8%), **Mont de Marsan** (7.7%), **Châtellerauld** (3.5%), **St Malo** (3.4%) and **Lannion** (0.3%). 5.8% of respondents did not specify their geographical area.). As responses depended on colleagues relaying information to students, those aware of gender issues were undoubtedly more responsive, which probably explains why the response rate from young women was higher than the one obtained locally. 285 responses were collected, i.e., around 10% of the R&T community, including 43 women, i.e., 15% of the total participants.

As said previously, for this study, we targeted not only graduating students but also former students to collect feedback about what job they ended up with. Data was collected via a questionnaire and 195 responses were collected, including 12 women (corresponding to 6.1% of responses, in line with the proportion of female students in R&T courses), in January 2023, with a very wide range of graduating classes from 2006 to 2021. Methodologically, the various data were collected between March 2021 and May 2023. All respondents' answers to this data collection are confidential.

The details about the preliminary analysis, and the online surveys are described hereafter.

## 4 Results

In this section, we present the results of the collected data for both the graduating and former students.

### 4.1 Graduating Students

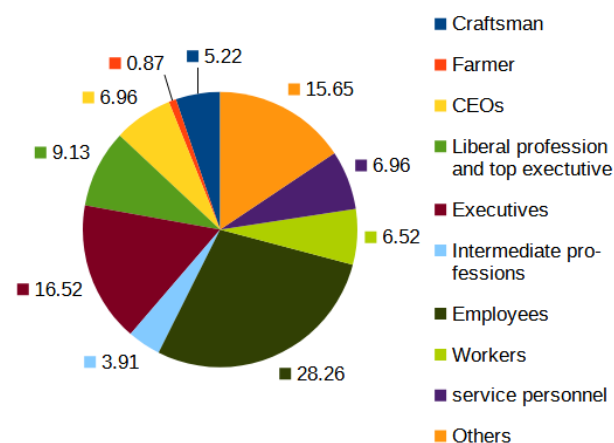
The questionnaire was sent three times to 5 different local classes ( 3 times for the 1st year class, once for 2nd year class and once for the 3rd year class of R&T Aix-Marseille) and once nationwide (i.e. to the other R&T departments). Locally, the response rate to this questionnaire varies between 55% and 69% (for the 3 years). It is similar for the response rate for women, which varies between 6% and 10.7%, corresponding to the high end of their representation in percents in R&T departments. The R&T department we sampled is in the south of Marseille, in the 9th district, a rather privileged area. The local students come from all the districts of Marseille, but also from the PACA region and other French departments. The national questionnaire, on the other hand, was carried out in geographically distant departments.

The questionnaire consisted in 14 questions, related to the following topics:

- Socio-demographic background
- Career path choice
- Know-How and Soft skills
- Soft skills
- Self-confidence
- Mixity benefits

**4.1.1 Socio-demographic background.** The question: **What is(are) the socio-professional category of your parents?** relates to the parents' professions and serves to better understand the socio-demographic background of the respondents.

According to the data collected, the students appear to be issued from all social categories, mostly employees as shown in Figure 1



**Figure 1: What is(are) the socio-professional category of your parents? Out of 230 responses**

**4.1.2 External influencing factors.** To understand which external factors impacted the students' choice for higher education study path, in the digital sector, we have asked two questions:

- To pursue academic studies, which has been the most important influencing factor?
- To choose your curriculum, which has been the most determining factor?

The most decisive factors in students' choice are their interests (59,5% at the local level 55,1% at national level) but also the opportunities available (around 29% at local and national level). Results vary little between the 5 graduating classes and are broadly similar across all geographic regions as shown in table. 1 (**What was the most decisive factor in your choice of higher education?** -Only one answer possible)

Thus, these results underline that when embracing a career, young people not only follow what they may appreciate but also the future job opportunities they have.

Moreover, parents (around 11% at the local and national level) and friends (8,6% at the local level and 16% at the national one) have also an impact on their choice. In addition, we would like to emphasize that the results show the crucial role of the educational staff and environment (25% at the

**Table 1: Results of student questionnaires**

| Q1 Choice R&T | Local Questionnaire | National Questionnaire |
|---------------|---------------------|------------------------|
| Interest      | 59,5%               | 55,1%                  |
| Outlet        | 26,6%               | 29,8%                  |
| Cost          | 1,2%                | 0,4%                   |
| Localisation  | 2,1%                | 4,5%                   |
| Other         | 4,9%                | 2,1%                   |
| Study Level   | 4,7%                | 6,3%                   |
| Parents Job   | 1%                  | 1,8%                   |

local level and 35% at the national one) in the choice of the career path and studies.

**4.1.3 Career path choice.** The results of the question **"What career path(s) would you like to pursue?"** show that a quarter of respondents want to become network and systems administrators, a second quarter network engineers, and a third one cybersecurity engineers. The other options are sales, telecoms, developer, IT and "no idea". Responses to the questionnaire sent during the three years are similar.

Among the responses received, it is worth highlighting the fact that young people know why they made this choice of career path. Indeed, they were not mistaken, as the network field, was among the other topics, the preferred subject of most respondents (57,7% at the local level and 49,8% at the national level).

Moreover, 1/4 of the students chose R&T training from various resources (mainly Internet resources), and 1/4 from the educational and school environment. We therefore can see again the substantial role played by the educational environment in guiding students, in addition to the Internet and ONISEP resources (that provides useful information to help students to choose their study path)[27].

**4.1.4 Know-How and soft skills.** To understand their feelings about the know-how and soft skills they need to master in R&T, we asked two questions:

- In your opinion, what are the main "soft skills" (e.g., interpersonal skills) you need to master to enter the digital professions?
- In your opinion, what is the main "know-how" (e.g. analytical skills) you need to master to enter the digital professions?

The responses to the second question **"In your opinion, what are the main "soft skills" (e.g., interpersonal skills) you need to master to enter the digital professions?"** show that communication, interpersonal skills and team spirit are the most frequently cited. Then there is curiosity and patience, followed by rigor, precision, meticulousness, adaptability, and professionalism. Finally, there is passion, listening, teaching and autonomy.

The responses to the question **"In your opinion, what is the main "know-how" (e.g. analytical skills) you need to master to enter the digital professions?"** show that the main skills identified can be classified into two groups, and are identical for the 5 graduating classes (i.e. during the years and 3 educational levels). First, curiosity, analysis, logic, adaptation, and understanding are most often cited. Then there is autonomy, passion, knowing how to use IT tools and how to document themselves. The skills required in the digital professions seem to favour men, as they are more accustomed to using a computer and its associated functions. They seem to master the technical side much better than women, because they have practiced more often. They know this and let the girls know, judging them from the outset, which can have an impact on their self-confidence.

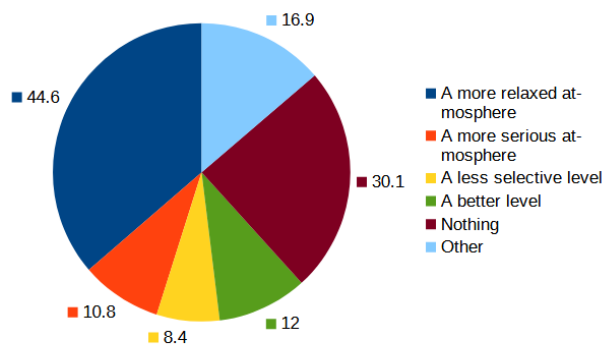
**4.1.5 Self-confidence.** In this study, we were also interested in the feelings of the students regarding their self-confidence level (i.e. how they feel in certain situations). To do so, we took a pre-existing measurement scale as a basis, i.e. the items used to measure general social confidence and information processing confidence from Peter Wright [29].

This question was divided into 10 affirmations related to the students' feelings, to assess their degree of self-confidence. Respondents could answer strongly agree, agree, neutral, disagree or strongly disagree (Likert scale). The results show that firstly, both local and national students are undisturbed integrating into a group that has already been formed. Secondly, they are, again, relatively at ease communicating even with people they don't know. Overall, they do not seem to suffer from a lack of self-confidence in the presence of strangers. Indeed, more than half of them feel comfortable in most social situations, and the majority are confident in their ability to judge mass media messages.

The last affirmation regarding **"I am afraid not to succeed in my studies (Likert scale)"**, points out that, on average, almost half the respondents (females and males) are afraid of failing in their studies (47,2% at the local level and 46,3% at the national level), which seems normal for their age. Yet, young men's and women's apprehension regarding digital technology appears to differ. In fact, the initial results

show that fear of the future career prospects in the digital sector is more marked among girls than boys. Indeed, not a single female student is confident that she will succeed in her studies. All the girls seem to have doubts about their abilities, which is not the case for the young men, and this doubt does not necessarily seem to be linked to the way others look at them. However, the data collected indicate that, despite their apparent self-confidence, young women are not necessarily confident in their ability to succeed in R&T. Even though they have gone through the various obstacles to being in minority in their studies, women still wonder if they are good enough to achieve success in embracing a profession in this field. This is also true in the digital professions, where women are more likely to change career paths.

**4.1.6 Mixity benefits.** In addition to self-confidence, we wanted to check if the students were seeing the benefits of having mix classrooms. We therefore asked the question: "**What would more women bring in the courses?**". The responses to this question show that it is first a more relaxed atmosphere (44,6%) followed by nothing (30,1%) (See Figure 2). Thus, we clearly notice that the under-representation of women in their training do not affect the respondents much, except for a more pleasant climate, the majority of whom being young men. In fact, in view of the results, more gender diversity seems of little importance to the students, as it would bring 'nothing'. They do not really see the point of increasing the gender mix in R&T, except sometimes for the social side, for "the atmosphere" or for "the physical attraction" as indicated in "other" response. The serious side is not really addressed, as they seem to think that in terms of skills, young women are not to be feared.



**Figure 2: what could make your training more gender-balanced? Multiple answers possible- Out of 352 answers**

## 4.2 Former students

The data collected from the former classes of Aix-Marseille department, show interesting results regarding :

- Path choice after the Bachelor,
- Positions in companies,
- Digital professional landscape.

### Path choice after the Bachelor

- In a range of 15 years (between 2006 and 2021), 34% (a third of them) entered the job market after obtaining their diploma (the 2nd year diploma or the 3rd year, i.e Professional degree, which is now part of the Bachelor R&T curriculum) and work in small, medium, and large companies in the PACA region (Provence Alpes Côte d'Azur), the Paris region, other regions or abroad.
- 65% have chosen to follow their studies as there is a wide range of specializations after the R&T Bachelor. A third followed a Master's degree (29%), almost a third went to an engineering school (28%) and only a few went on to private IT schools (8%)
- two out of 193 (1%) switched to journalism or business.

We can see, then, that the field appeals, and that the career opportunities are considerable. In fact, the career fields chosen by the former students cover computer science, networks, and telecoms, which is in line with the R&T curriculum.

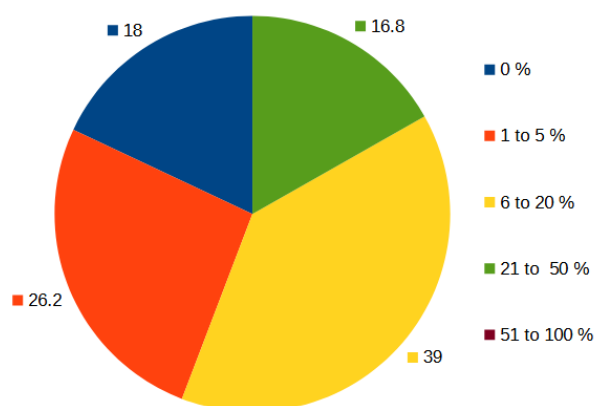
### Positions in companies

We notice that a significant number of graduated students are in project management positions (50%) regardless the gender. Other positions include IT director, systems engineer, presales engineer, service centre manager. For women, they can be found as business manager, technician, presales manager, team or operational manager.

### Digital professional landscape

The data collected regarding women in technical positions is striking. In 44,2% of French companies, the proportion of women in technical roles is between 0 and 5%, or 1 woman per team of 20 people or more. Almost one in 5 French companies (18%) have no woman in their digital-technical related teams. 39% of companies have a female workforce that fluctuates between 10 and 20%. Finally, 83,2% of companies have between 0 and 20% of women staff in technical roles which is not that surprising regarding the low number of women in the tech curriculums. And a priori, even when our graduate students change companies, the result remains almost identical, namely: women are almost non-existent in technical positions in digital firms in France as shown in Figure 3.

A clear parallel can be drawn between the low number of academic studies in R&T and the absence of women in the associated jobs. What remains worrying is that there seems



**Figure 3: The number of women professionals in technical positions in the alumni firms**

to be a lack of consciousness, if not of concrete actions, to find solutions to get more mixity in this crucial sector.

## 5 Discussion: A collective imperative

Considering the issue of self-confidence, the issue is rooted in lesser education, such as middle and high school, where prejudices and discrimination takes place as an ambiance, or misplaced remarks and jokes. The first solution would be to crack down on such bullying, applying more severe punishment and raising awareness, be it for the student and professoral body

For the lack of perceived benefit to gender equality, another solution would be to make the biases known and explain the issue it would cause in the long run inciting companies to devote resources to encourage women and hire them.

After presenting the results of our study, we underline some actions that can be considered at three levels: Education, Research and Industry. Some of these actions have already been implemented as a step towards a slow but constant change.

### 5.1 Education level - Schools and universities

At the education level, several actions could be considered. First, for the issues to be tackled efficiently it becomes urgent to consider the early age of our education and society system where gender differentiation is rooted. Indeed, if we don't address the flaws at the very start, inequalities in this sector are likely to prevail. Second, considering the issue of self-confidence, the issue is anchored in lesser education, such as middle and high school, where prejudices and discrimination take place as an atmosphere, or misplaced remarks

and jokes, as shown recently by the researcher in sociology Clémence Perronnet [7]. A solution would be to crack down on such bullying, applying more severe punishment and raising awareness, be it for the student and educational staff. Third, science courses among girls should be promoted by encouraging them to take up STEM subjects (only 29.2% in engineering, and this figure has been falling steadily for several years). It is also important to set up scholarships dedicated to girls studying in tech. Offering mentoring to girls/women by inspirational women/role models (e.g. the Tech4Elles application from the WHAT06 association[31]) or creating partnerships with specialized associations (such as Femmes et Sciences in France[30] and WHAT06) could be other actions to undertake.

## 5.2 Research

In France, only 16.6% of computer science researchers are women (9). The possible actions could be to raise researchers' awareness of algorithmic biases through compulsory courses in algorithmic ethics, to initiate projects such as "Les Décodeuses du Numérique" [8] to help break down stereotypes that discourage women from pursuing research in the digital sciences or to create specific scholarships for women involved in STEM research projects.

## 5.3 Companies

The first action that could be done in companies to avoid bias in computer programs and AI would be to integrate women into R&D teams which would help reducing bias by 40% [24]. In this context, the European commission recommends gendered ethical audits for AI systems[11].

Secondly, the workforce should be doubled to meet the sector's needs as the digital sector suffers from a shortage of talent (recruitment tensions 4 to 8 times higher), with 500,000 unfilled positions in Europe in the digital sector (specifically big data and analytic) and 300,000 in cybersecurity sector, by 2025 [10](European Commission). Therefore, bringing in more women would enable us to fill job vacancies, strengthen corporate competitiveness and reduce job insecurity among women.

Thirdly, projects adapted to everyone's needs should be created as tech products designed by male-dominated teams can ignore women's needs. For instance, 70% of health applications ignore women's specific needs (Journal of Medical Internet Research, 2022) or in autonomous cars they are 73% more dangerous for women because crash-test dummies are male [26]

Fourthly, operational and financial efficiency should be improved with greater gender diversity to reach better results. Indeed, there is a 21% increase in profitability for tech companies with gender-balanced management [12] and a



45% increase in innovation in mixed-gender teams. It has also been proved that mixed teams solve problems 30% faster.

## 6 Conclusion

In this paper we have given an overview of the reasons why women are deserting cybersecurity sector. Throughout our survey, we were able to determine that there is no significant disparity between genders, in socio-economical background and motivation for their study path. Yet, It clearly appeared that the main disparity was the lack of self-confidence with women ( 100 per cent of them are afraid when thinking about their future career against 40 per cent of their male colleagues).

Furthermore, inequality is not perceived as a problem at all by the students, which highlights the need for better communication and awareness campaigns to make this issue taken more seriously into consideration. The feminization of tech is not just a question of equality, but a lever for performance, innovation, and social equity.

If we don't reverse the trend soon, this means that gender inequalities will not be reduced in the years to come, but will on the contrary multiply, as digital and new technology professions will soon be essential and will therefore be professions of power.

An active inclusive recruitment policy has to be implemented drafting inclusive job descriptions, encouraging more women to apply (a woman will apply if she meets 80% of the criteria, whereas a man will apply if he meets only 50-60% of the criteria), anonymizing CVs (EPFL in Switzerland increased hiring of women by 30% by masking names in 2020), correcting pay gaps to achieve equal pay for men and women in equivalent positions or ensuring women's representation in management positions (Rixain law: 30% women in management bodies by 2026 [16]).

To transform this situation, we need to attack the discriminatory system at root working to deconstruct systems of representation of knowledge and professions. But do the people who are in charge really want to do it?" asked Isabelle Collet [18].

## Acknowledgments

We would like to thank the R&T community for their contribution to the collection of learner data. We wish to express our gratitude to all respondents.

## References

- [1] Jean A. 2019. *De l'autre côté de la machine*. L'observatoire.
- [2] Accenture. 2018. *Cracking the Gender Code*.
- [3] Perrault J. Beyer S., Rynes K. and al. 2023. Gender differences in computer science students. In *SIGCSE '03: Proceedings of the 34th SIGCSE technical symposium on Computer science education*. 49–53. doi:10.1145/611892.611930
- [4] Pierrel A. Blanchard M., Orange S. 2016. Filles + Sciences = une équation insoluble? (2016).
- [5] Houssain C. 2023. RT et les genres, l'atonie dans le changement ?. In *Workshop Pédagogique Réseaux et Télécom 2023, Saint-Pierre, La Réunion*.
- [6] Medeiros C. 2019. Engaging women in computer science - Past, present and future. *CLEI electronic journal* 22, 2 (2019). doi:10.19153/cleiej.22.2.1
- [7] Perronnet C. 2024. *Les filles sont l'avenir des mathématiques*. CNRS Edition.
- [8] Thiebault L. Castor L., Esnault C. and al. 2021. Les décodeuses du numériques - The codebreakers of the digital world. <https://www.calameo.com/read/00729227958a8778661a4?page=1>
- [9] Silbey S. Cech E., Rubineau B. and Seron C. 2011. Professional role confidence and gendered persistence in engineering. (2011).
- [10] European Commission. 2019. . <https://digital-strategy.ec.europa.eu/en/library/digital-skills-all-europeans-brochure> Last visit 16 april 2025.
- [11] European Commission. 2024. EU AI Act. <https://artificialintelligenceact.eu/the-act/> last visit 16 april 2025.
- [12] LivingInstitute Diversity Intelligence Consultancy. 2024. The Impact of DEI on Business Performance and Profitability. <https://hai.stanford.edu/ai-index/2023-ai-index-report> Last visit 16 april 2025.
- [13] Crunchbase. 2022. *Startups Founded in 2022*. <https://www.crunchbase.com/hub/startups-founded-in-2022> Last visit 16 april 2025.
- [14] Peacock D. and Irons A. 2017. Gender inequalities in Cybersecurity: Exploring the gender gap in opportunities and progression,. *International Journal of Gender, Science and Technology* Vol.9, 1 (2017).
- [15] Laboratoire de l'Egalité. 2024. La Femme Invisible dans le numérique - le cercle vicieux du sexisme. [https://www.haut-conseil-egalite.gouv.fr/IMG/pdf/hce-parster-rapport-la\\_femme\\_invisible\\_dans\\_le\\_numerique-vdef.pdf](https://www.haut-conseil-egalite.gouv.fr/IMG/pdf/hce-parster-rapport-la_femme_invisible_dans_le_numerique-vdef.pdf)
- [16] Ministère du travail de la santé des solidarités et des familles. 2022. *La loi Rixain : Accélérer la participation des femmes à la vie économique et professionnelle*. <https://travail-emploi.gouv.fr/la-loi-rixain-acceler-la-participation-des-femmes-la-vie-economique-et-professionnelle> Last visit 16 april 2025.
- [17] Moreno E. 2024. 2ème édition des Assises Nationales de la féminisation des métiers et filières du numérique. <https://www.assises-feminisation-metiers-numerique.fr/programmation-2024/> Last visit 16 april 2025.
- [18] Collet I. 2019. *Les oubliées du numérique*. Le Passeur.
- [19] Stanford AI Index. 2023. Artificial Intelligence Index Report 2023. <https://hai.stanford.edu/ai-index/2023-ai-index-report> Last visit 16 april 2025.
- [20] Crawford K. 2021. *tlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. A. Yale University Press, New Haven.
- [21] Haut Conseil à l'égalité entre les Femmes et les Hommes. 2023. *La Femme Invisible dans le numérique – le cercle vicieux du sexisme*. [https://www.haut-conseil-egalite.gouv.fr/IMG/pdf/hce-parster-rapport-la\\_femme\\_invisible\\_dans\\_le\\_numerique-vdef.pdf](https://www.haut-conseil-egalite.gouv.fr/IMG/pdf/hce-parster-rapport-la_femme_invisible_dans_le_numerique-vdef.pdf) Last visit 16 april 2025.
- [22] Charles M. and Bradley K. 2009. Indulging Our Gendered Selves? Sex Segregation by Field of Study in 44 Countries. *Amer. J. Sociology* 114, 4 (2009), 924–976.
- [23] Medel L. M. 2021. Gender bias in machine translation: an analysis of Google Translate. *Academia Letters* (07 2021). doi:10.20935/AL2288
- [24] MIT. 2018. *Study finds gender and skin-type bias in commercial artificial-intelligence systems*. <https://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212> Last visit 16 april 2025.
- [25] Mosconi N. 2002. Comment les pratiques enseignantes fabriquent-elles de l'inégalité entre les sexes ? *Les dossiers des sciences de l'éducation* 5

- (2002), pp. 97–109.
- [26] Medical College of Wisconsin. 2024. *Study Finds Women More Likely Than Men to Die or Be Seriously Injured in Car Crashes*. <https://www.wiareport.com/2024/03/study-finds-women-more-likely-than-men-to-die-or-be-seriously-injured-in-car-crashes/> Last visit 16 april 2025.
- [27] ONISEP. 2025. *Office national d’information sur les enseignements et les professions*. <https://www.onisep.fr/> Last visit 16 april 2025.
- [28] Timmermans B. & Kristinson K. Ostergaard C.R. 2011. Does a different view create something new? The effect of employee diversity on innovation. *Research Policy* 40, 3 (2011), 500–509. doi:10.1016/j.respol.2010.11.004
- [29] Wright P. R. 2006. Drama Education and Development of Self: Myth or Reality? *Soc Psychol Educ* 9 (2006), 43–65. <https://doi.org/10.1007/s11218-005-4791-y>
- [30] Femmes & Sciences. 2025. *Association Femmes & Sciences*. <https://www.femmesetsciences.fr/who-are-we> Last visit 16 april 2025.
- [31] WHATO6. 2025. *Women Hackers Action Tank 06*. <https://what06.fr/> Last visit 16 april 2025.



---

# Data Beyond Control

Loïc Pantano  
Polytech Nice-Sophia  
Biot, France  
loic.pantano@gmail.com

Karima Boudaoud  
Université Côte d'Azur - Laboratoire  
I3S-CNRS  
Sophia Antipolis, France  
karima.boudaoud@univ-cotedazur.fr

Emmy Holveck  
Université Côte d'Azur - Faculté de  
Droit et de Science Politique  
Nice, France  
emmy.holveck@etu.univ-  
cotedazur.fr

Elena Ecerf  
Université Côte d'Azur - Faculté de  
Droit et de Science Politique  
Nice, France  
elena.lecerf@etu.univ-cotedazur.fr

Alicia Matingou  
Université Côte d'Azur - Faculté de  
Droit et de Science Politique  
Nice, France  
alicia.matingou@etu.univ-  
cotedazur.fr

Jean-Sylvestre Bergé  
Université Côte d'Azur - Faculté de  
Droit et de Science Politique  
Nice, France  
Jean-Sylvestre.berge@univ-  
cotedazur.fr

## Abstract

In a context of increasing digitalization, controlling data flows constitutes a major challenge for organizations, particularly when it comes to sensitive data. This study examines the technical and legal challenges related to secure data flow management, based on the concrete case of the medical analysis laboratory MedLab and its facial recognition system implementation project. The study proposes an integrated approach combining Zero-Trust architecture, multi-factor authentication mechanisms, and advanced encryption solutions, while ensuring compliance with GDPR and other applicable regulations. Special attention is paid to subcontractor management and international data transfers. The results demonstrate that a holistic approach, combining innovative technical solutions and robust governance, can effectively meet security and compliance requirements. The study also emphasizes the importance of continuous adaptation to technological and regulatory developments, offering concrete recommendations for organizations wishing to strengthen control of their data flows.

## CCS Concepts

• Security and privacy → Intrusion detection systems; Security protocols.

## Keywords

Data flows, GDPR, Compliance

## 1 Introduction

In an increasingly interconnected world, data flows constitute a fundamental element of modern information systems. These flows, whether continuous or intermittent, transport often sensitive information that proves essential to the functioning of organizations. The multiplication of these exchanges, however, raises a major issue: how to guarantee rigorous control of these flows while respecting the growing requirements of performance, security, and regulatory compliance? In an increasingly interconnected world, data flows constitute a fundamental element of modern information systems. These flows, whether continuous or intermittent, transport often sensitive information that proves essential to the functioning of organizations. The multiplication of these exchanges, however,

raises a major issue: how to guarantee rigorous control of these flows while respecting the growing requirements of performance, security, and regulatory compliance? To concretely illustrate these challenges, the case of the medical analysis laboratory MedLab proves particularly enlightening. Following several security incidents involving unauthorized access to its facilities, this laboratory is considering implementing a facial recognition system to secure access to its sensitive areas, particularly the storage spaces for biological samples and servers hosting patients' medical data. This project, seemingly simple, crystallizes the complexity of current challenges in controlling data flows. Indeed, its implementation involves not only collecting and processing biometric data of employees but also securely managing multiple flows of sensitive data. The situation becomes more complex with the intervention of two subcontractors: BiometricsCorp, specializing in biometric systems, and DataSecure Systems, an American company being considered for software development. This configuration raises crucial questions both technically and legally: how to ensure the security of biometric data while respecting the fundamental rights of employees? How to guarantee the traceability and confidentiality of data during their processing by different actors, particularly in a context of international transfers? The current context is marked by rapid evolution of digital technologies and continuous strengthening of the regulatory framework. The emergence of new technical solutions such as artificial intelligence, biometric systems, and the Internet of Things offers unprecedented opportunities to improve the control of data flows. At the same time, the entry into force of the General Data Protection Regulation (GDPR) imposes strict standards for the protection of personal data, particularly for sensitive data such as health or biometric data. This regulatory evolution is accompanied by increased awareness of the issues of privacy protection and fundamental rights of individuals. The medical world, in particular, is at the heart of these issues. Healthcare establishments must manage growing flows of sensitive data while guaranteeing their security and confidentiality. The complexity of these issues lies in their multidimensional nature. On one hand, technical aspects require a deep mastery of security technologies, system architectures, and communication protocols. On the other hand, legal considerations require a fine understanding of the regulatory framework and the obligations that derive from it. This technical and legal duality

requires an integrated approach, capable of providing coherent answers to the different challenges posed. The current context is marked by rapid evolution of digital technologies and continuous strengthening of the regulatory framework. The emergence of new technical solutions such as artificial intelligence, biometric systems, and the Internet of Things offers unprecedented opportunities to improve the control of data flows. At the same time, the entry into force of the General Data Protection Regulation (GDPR) imposes strict standards for the protection of personal data, particularly for sensitive data such as health or biometric data. This regulatory evolution is accompanied by increased awareness of the issues of privacy protection and fundamental rights of individuals. The medical world, in particular, is at the heart of these issues. Health-care establishments must manage growing flows of sensitive data while guaranteeing their security and confidentiality. The case of MedLab laboratory perfectly illustrates this problem. Its project to implement a facial recognition system to secure access to its facilities raises complex questions at the intersection of technology and law. How to ensure the security of employees' biometric data while respecting their fundamental rights? How to manage the outsourcing of technical development while maintaining control over sensitive data? These questions reflect the challenges many organizations face in their digital transformation. The complexity of these issues lies in their multidimensional nature. On one hand, technical aspects require a deep mastery of security technologies, system architectures, and communication protocols. On the other hand, legal considerations require a fine understanding of the regulatory framework and the obligations that derive from it. This technical and legal duality requires an integrated approach, capable of providing coherent answers to the different challenges posed. Our study aims to analyze these issues from both technical and legal angles. The objective is to identify solutions that guarantee the security and traceability of data while ensuring their compliance with the current legal framework. This analysis revolves around four critical points in the data journey: their entry into a system, their processing, their storage, and their output. At each of these stages, technical vulnerabilities can appear and legal obligations must be respected. Identifying and understanding these critical points is an essential prerequisite for implementing effective solutions. The methodology adopted combines an in-depth analysis of available technical solutions with a detailed examination of legal requirements. This two-dimensional approach allows us to understand the complexity of situations encountered in the field. The technical solutions studied cover a wide spectrum, ranging from authentication and encryption mechanisms to distributed system architectures. The legal analysis, for its part, focuses on deciphering regulatory obligations and identifying contractual levers to secure data exchanges. Particular attention is paid to practical implications through the study of the MedLab case, which serves as a common thread to illustrate concrete challenges and solutions. This case study allows us to anchor our analysis in the operational reality of organizations. The stake is twofold: it is not only about protecting data against growing cyber threats but also about guaranteeing respect for the fundamental rights of the people concerned. The protection of personal data can no longer be considered as a simple technical or legal constraint but must be integrated as a central element of organizations' strategy. This study is addressed to technical

and legal managers facing these challenges, offering them a global approach and concrete recommendations to maintain effective control over their data flows. The growing complexity of information systems and the constant evolution of threats require increased vigilance and continuous adaptation of practices. Our analysis aims to provide a reference framework allowing organizations to develop a structured approach to controlling data flows. The first part of our study will present a comprehensive state of the art, covering both technical and legal aspects of data flow control. This review will identify existing best practices and emerging trends in the field. The second part will detail our methodology for analyzing critical points, proposing a systematic approach for identifying and managing risks related to data flows. The third part will deal with legal and organizational aspects, proposing a governance framework adapted to current challenges. The fourth part will present the proposed technical solutions, emphasizing their practical applicability and effectiveness in different contexts. This section will particularly address questions of system architecture, securing exchanges, and data traceability. The fifth part will propose a concrete technical architecture of these technologies. Finally, we will discuss the results obtained and formulate practical recommendations for organizations wishing to strengthen the control of their data flows. This discussion will put into perspective the different elements analyzed and propose an integrated vision of the problem of controlling data flows. The recommendations will take into account the operational constraints of organizations while guaranteeing a level of security and compliance adapted to current challenges.

## 2 State of the art

### 2.1 Evolution of data flow control architectures

The history of data flow control dates back to the first operating systems and their memory protection mechanisms [6]. Since then, the evolution towards distributed architectures has considerably complicated this issue, requiring new approaches to maintain control over data in transit [11]. One of these approaches is Zero-Trust architectures, conceptualized by Kindervag in 2010, which marks a major turning point by establishing the principle of "never trust, always verify" [18]. The emergence of cloud computing has also introduced new challenges, particularly the loss of visibility on the physical storage of data [3]. The work of Gupta et al. [14] shows that more than 22 billion datasets were leaked in the year 2021 alone, a number that is theorized to increase over the years. To address these challenges, new hybrid architectures have emerged, combining local and cloud controls [19]. Finally, microservices have also transformed data flow management. Research by Hannousse, Abdelhakim, Yahiouche and Salima [15] shows that breaking down applications into autonomous services multiplies data exchange points, requiring adapted control mechanisms. Service mesh solutions, as analyzed by Li et al. [20], can offer an abstraction layer for centralized management of these flows.

### 2.2 Data Protection Mechanisms in Transit

Encryption of data in transit constitutes the first line of defense. Modern protocols like TLS 1.3 offer robust guarantees against interception [25]. However, as Archimbaud [4] points out, encryption alone is not enough: key management and certificate validation

remain critical points. Additionally, recent work on homomorphic encryption opens new perspectives. Gentry's research [12] allows computations to be performed on encrypted data, maintaining their confidentiality throughout processing. Although performance remains a major challenge [29], recent optimizations by Dewangan et al. [9] have reduced the computational overhead from 10 to 15%. Finally, secure routing protocols constitute another crucial aspect. For this, the work of Sarkar et al. [27] proposes a new routing approach based on security properties, allowing dynamic optimization of data paths according to detected threats.

### 2.3 Data Flow Traceability and Auditing

Data traceability constitutes a major challenge, particularly in regulated environments. The work of Wang, Haiyan and Zhang, Jiawei [32] proposes a blockchain-based approach to guarantee the integrity of audit logs, a solution that meets GDPR requirements for traceability [23]. AI-based anomaly detection systems have also evolved considerably. El Hatib, Souad [10] demonstrates the effectiveness of machine learning techniques for identifying suspicious software. Furthermore, recent research by Coelho et al. [8] introduces federated learning models enabling distributed detection while preserving the confidentiality of training data.

### 2.4 Access Rights Management and Granular Control

Access control models have evolved toward finer granularity. The ABAC model allows decisions based on context attributes [5]. Masood, Rahat et al. [21] show that this approach offers increased flexibility compared to role-based models (RBAC), allowing more precise management of authorizations. Another option consisting of using identity as a security perimeter is emerging as a new paradigm [13]. This approach allows more dynamic access control in hybrid cloud environments. Recent work by Rostami, Kamyar [26] introduces a continuous authentication framework based on behavioral factors. Finally, decentralized identity management (DID) also represents a major evolution, with research by Hao, Jiakun et al. [16] demonstrating the effectiveness of self-sovereign identity systems for access control in distributed environments.

### 2.5 Emerging Technologies and New Approaches

Artificial intelligence is revolutionizing anomaly detection in data streams. The work of Sendera, Marcin et al. [28] demonstrates the effectiveness of deep neural networks, while the research of Zhang, Zheyu [33] explores the use of reinforcement learning for the dynamic adaptation of security policies. Quantum computing represents both a threat and an opportunity. Indeed, quantum computing poses a threat as it could compromise current encryption systems due to its ability to quickly factor large numbers using Shor's algorithm, making cryptographic protections like RSA vulnerable. However, it also offers significant opportunities with quantum cryptography enabling theoretically unbreakable communications thanks to the properties of entanglement and Heisenberg's uncertainty principle. This duality drives the development of post-quantum algorithms and hybrid architectures to prepare for data security in the quantum era. Buchmann, Johannes et al. [7] explore

these post-quantum algorithms, while recent work by Sykot, Arman et al. [31] proposes a hybrid architecture combining classical and quantum algorithms.

### 2.6 Compliance and Security Standards

Standards and regulations play a crucial role in harmonizing practices. ISO/IEC 27701 [17] provides a framework for information privacy management, while the article by Smartpoint [30] demonstrates the significant impact of poor data governance, with 90% of Big Data projects failing for this reason, resulting in an average cost of 6 million dollars per data breach. The lack of initial governance can lead to a fivefold increase in costs, not to mention GDPR fines that can reach up to 20 million euros.

### 2.7 Current Challenges and Limitations

Despite these advancements, several challenges persist. The increasing complexity of architectures multiplies points of vulnerability. The latency introduced by access control mechanisms impacts performance, particularly in real-time applications. Data management in multi-cloud environments, as well as the difficulties in maintaining consistent control across different cloud providers, are also significant issues. The standardization of best practices through ISO 27017:2023 [1] emerges as a potential solution to facilitate the implementation of security measures. Finally, among the challenges related to the scalability of control solutions in hyperconnected environments, the search for a balance between security and performance remains a major concern.

### 2.8 Future Perspectives

Current research opens up new perspectives, such as the use of artificial intelligence for the self-adaptation of control mechanisms. Notably, the research by John, Grace [24] on self-healing systems promises better resilience against attacks.

## 3 METHODOLOGY FOR DATA FLOW ANALYSIS

Our methodological approach to data flow analysis is based on a systematic process that combines technical and legal perspectives. This section presents the foundations of our methodology and its application in the context of the MedLab laboratory.

### 3.1 Identification of Critical Points

Data flow analysis requires a precise identification of critical points where control must be particularly rigorous. Our methodology relies on a comprehensive mapping of the data lifecycle, from creation or collection to deletion or archiving. The first critical point concerns the entry of data into the system. In the case of MedLab, this notably involves the capture of biometric data during employee registration. This initial phase requires particular attention as it conditions the quality and security of all subsequent processing. The validation of incoming data must incorporate technical controls, such as verifying the integrity and format of the data, as well as legal checks, particularly regarding the legal basis for processing and obtaining consent. The second critical point is at the data processing level. This phase involves transforming raw data into

usable information, such as converting facial images into biometric templates. Our methodology imposes complete traceability of operations performed on the data, allowing for the reconstruction of the history of transformations and verifying their compliance with regulatory requirements.

### 3.2 Risk Assessment

Risk assessment is a crucial component of our methodology, relying on a multi-criteria analysis that evaluates five key dimensions: technical risks (compromise of biometric data, privacy violations), operational risks (impact on business continuity), legal risks (GDPR compliance), risks related to outsourcing (security of data transfers), and geopolitical risks (international transfers). In the context of MedLab, this approach allows for the identification and prioritization of critical vulnerability areas, particularly in collaboration with BiometricsCorp, a biometric systems design company, and DataSecure Systems, which is responsible for software development and based in the United States, while ensuring compliance with regulatory requirements.

### 3.3 Design of Control Mechanisms

The design of control mechanisms is based on the results of the risk assessment. Our methodology favors a defense-in-depth approach, combining multiple layers of protection. For MedLab, this translates into the implementation of multi-level access controls, advanced encryption mechanisms, and intrusion detection systems. The proposed technical architecture incorporates the principles of Zero Trust, imposing systematic verification of access regardless of its origin. This approach is particularly relevant in the context of MedLab, where the sensitivity of biometric data requires a maximum level of protection.

### 3.4 Validation and Testing

The validation of control mechanisms is a crucial step in our methodology. It relies on test scenarios that replicate real usage conditions, including unauthorized access attempts and technical failure situations. These tests allow for the verification of the effectiveness of the implemented controls and the identification of potential security vulnerabilities. Validation also includes legal aspects, particularly the verification of compliance of processing activities with GDPR requirements, such as the establishment of a processing activities register, conducting impact assessments for biometric data, and verifying the mechanisms for exercising the rights of data subjects. Special attention is given to documenting technical and organizational measures, including procedures for notifying data breaches and privacy by design mechanisms. Regular audits are planned to ensure ongoing compliance over time.

### 3.5 Documentation and Training

Our methodology places particular importance on documenting processes and training users. Control procedures are formalized in reference documents, which are regularly updated to reflect technical and regulatory developments. The training program developed for MedLab covers both the technical and legal aspects of data flow control. It aims to raise employee awareness of security issues and

to train them in best practices for data protection. This comprehensive and structured methodology allows for a systematic approach to the challenges of data flow control. Its application to the MedLab case demonstrates its relevance for environments handling sensitive data and subject to strict regulatory constraints.

## 4 LEGAL AND ORGANIZATIONAL ASPECTS

### 4.1 Applicable Regulatory Framework

The processing of biometric data in the context of MedLab is subject to a strict regulatory framework. The GDPR classifies biometric data among the special categories of data requiring enhanced protection. Article 9 of the regulation establishes a general principle of prohibition on the processing of such data, with strictly regulated exceptions. The processing of biometric data for access control purposes can be justified based on Article 9.2(b) of the GDPR, relating to obligations in labor law. However, this legal basis must be supplemented by appropriate safeguards for the fundamental rights of the data subjects. In the case of MedLab, these safeguards include the implementation of an alternative authentication solution for employees who do not wish to use facial recognition.

### 4.2 Organization of Outsourcing

Collaboration with BiometricsCorp and DataSecure Systems requires rigorous legal oversight. The outsourcing contract must meet the requirements of Article 28 of the GDPR, particularly in terms of technical and organizational safeguards. The contractual clauses must specify the respective obligations of the parties regarding security, confidentiality, and data protection. The potential transfer of data to DataSecure Systems in the United States raises particular issues following the invalidation of the Privacy Shield. The use of Standard Contractual Clauses (SCCs) must be accompanied by additional technical measures, in accordance with the recommendations of the EDPB (European Data Protection Board).

### 4.3 Management of Data Subject Rights

A comprehensive system for managing the rights of data subjects has been established. Employees must be clearly informed about the processing of their biometric data, in accordance with Articles 13 and 14 of the GDPR. This information includes the purposes of processing, the retention period of the data, and the recipients of the data. The system allows for the effective exercise of the rights of access, rectification, and erasure of data. Specific procedures have been defined for managing requests to exercise these rights, with response times compliant with regulatory requirements.

### 4.4 Data Governance

Data governance is based on a clearly defined organization, with specific roles and responsibilities. A Data Protection Officer (DPO) has been appointed to oversee compliance with processing activities. A steering committee, comprising representatives from various relevant functions, ensures regular monitoring of data protection issues. Formalized policies and procedures govern all data processing operations. These documents specifically define the rules for data classification, access modalities, and applicable security procedures.

## 4.5 Documentation and Compliance

Compliance documentation includes all elements required by the GDPR, notably the register of processing activities (Article 30) and the data protection impact assessment (Article 35). The latter has been particularly thorough given the sensitive nature of biometric data. Regular audits are scheduled to verify ongoing compliance over time. These audits cover both technical and organizational aspects, with particular attention paid to security measures and the management of subcontractors. This comprehensive approach to compliance and governance ensures that data processing respects the regulatory framework while meeting the operational needs of the organization. 1

## 5 PROPOSED TECHNICAL SOLUTIONS

This section details the technical solutions developed to ensure effective control of data flows, using the MedLab laboratory case as a reference.

### 5.1 Secure Technical Architecture

The proposed architecture is based on a Zero Trust approach that integrates multiple levels of security. The biometric capture system uses high-resolution sensors equipped with liveness detection technologies to prevent fraud attempts. These sensors are positioned at strategic access points in the laboratory, particularly at the entrance to sensitive areas containing samples and medical data servers. The processing of biometric data initially occurs locally on secure processors embedded in the sensors. This edge computing approach minimizes the risks associated with the transmission of raw data. The extracted biometric templates are immediately encrypted before any transmission to the central system. The network infrastructure is segmented into distinct security zones, with strict filtering rules between each zone. Communications between the various components of the system use secure protocols, with systematic mutual authentication of endpoints.

### 5.2 Access Control Mechanisms

The access control system implements multifactor authentication combining several elements: facial biometrics, a secure physical token, and a PIN code. This approach maintains a high level of security while providing a smooth user experience. Access rights management is based on a Role-Based Access Control (RBAC) model enhanced by contextual attributes (Attribute-Based Access Control - ABAC). Authorization decisions take into account not only the user's role but also factors such as the time of day, location, or detected risk level.

### 5.3 Data Protection in Transit and at Rest

Data encryption is a central element of the solution. Data in transit is protected by TLS 1.3 protocols with robust cryptographic suites such as TLS\_AES\_256\_GCM\_SHA384, as recommended by ANSSI [2]. Data at rest is encrypted with AES-256 in GCM mode, with key management based on a dedicated PKI infrastructure. For particularly sensitive data, such as biometric templates, additional encryption is applied at the application level. The encryption keys are stored in HSMs (Hardware Security Modules) compliant with FIPS 140-2 Level 3 standards [22].

## 5.4 Traceability and Audit Systems

The solution integrates an advanced logging system that allows for tracking all operations performed on the data. Logs are timestamped and electronically signed to ensure their integrity. A cryptographic chaining mechanism ensures that traces cannot be modified retroactively. The SIEM (Security Information and Event Management) collects and analyzes security events in real-time. Sophisticated correlation rules enable the detection of suspicious behaviors and intrusion attempts. Behavioral analysis relies on machine learning algorithms to identify anomalies.

## 5.5 Incident Management and Business Continuity

An incident detection and response system (SIEM/SOC) continuously monitors the infrastructure. In the event of an anomaly, automated response procedures are triggered, which may include the temporary isolation of certain system components. Business continuity is ensured by a redundant architecture, with synchronous replication of critical data to a backup site. Automatic failover procedures allow for service maintenance in the event of a failure at the primary site.

## 5.6 Integrated Emerging Technologies

The solution integrates several emerging technologies to enhance security. Artificial intelligence is used for anomaly detection and predictive threat analysis. Post-quantum encryption algorithms are implemented for data requiring long-term protection. A private blockchain is utilized for the immutable recording of critical events, such as changes in access rights or transfers of sensitive data. This approach ensures undeniable traceability of operations. These technical solutions form a coherent and comprehensive set that ensures effective control of data flows. Their implementation at MedLab demonstrates their ability to meet the strictest requirements for security and regulatory compliance. 2

## 6 PROPOSED TECHNICAL ARCHITECTURE

### 6.1 Fundamental Infrastructure

The infrastructure is based on a highly available architecture deployed on OpenStack, with redundant availability zones. The Docker Enterprise containerization solution provides the necessary isolation of components, orchestrated by Kubernetes for workload management. Persistent storage is ensured by Ceph Storage, offering synchronous replication between different availability zones.

The network layer uses Calico as the Kubernetes Container Network Interface (CNI) plugin, ensuring traffic segmentation and the application of security policies at the container level. Communications between different services are managed by Istio as a service mesh, ensuring mutual TLS encryption and service authentication.

### 6.2 Authentication and Access Control System

The authentication system relies on Keycloak as the identity server, integrated with Microsoft Windows Server 2022 Active Directory for user management. The biometric solution uses HID Signo sensors equipped with secure processors for local processing of biometric data. Access control is implemented via Open Policy Agent

**Table 1: MedLab System Requirements**

| Category                      | Regulatory Requirement                 | Associated Technical Requirement |
|-------------------------------|--|----------------------------------|
| Data Protection               | Confidentiality of sensitive data      | Securing storage                 |
|                               | Legal basis for processing             | Protection of transmissions      |
|                               | Purpose limitation                     | Isolation of sensitive data      |
| Traceability and Transparency | Documentation of processing activities | Recording of operations          |
|                               | Register of processing activities      | Integrity of logs                |
|                               | Ability to demonstrate compliance      | Access traceability              |
| Management of Subcontractors  | Contractual framework                  | Control of external access       |
|                               | Appropriate safeguards                 | Separation of environments       |
|                               | Supervision of processing activities   | Monitoring of flows              |
| Continuity and Resilience     | Availability guarantee                 | High availability                |
|                               | Protection against incidents           | Data backup                      |
|                               | Recovery plan                          | Anomaly detection                |

(OPA), allowing for fine-grained access policy management based on attributes. Authentication tokens are JWTs signed by a public key infrastructure (PKI) based on HashiCorp Vault.

### 6.3 Data Storage and Protection System

The storage of sensitive data is managed by MongoDB in a cluster configuration, with transparent data encryption at rest using AES-256-GCM. The encryption keys are managed by HashiCorp Vault in auto-unseal mode, utilizing nCipher nShield Connect XC HSMs for hardware protection of the keys. Data backup is managed by Veeam Backup & Replication 12, with backup encryption and replication to a remote site. Data is archived on a MinIO object storage system configured for high availability.

### 6.4 Monitoring and Logging System

System monitoring is provided by Prometheus for metric collection, with Grafana for visualization. Logs are centralized via an Elasticsearch cluster, collected by Filebeat agents, and visualized through Kibana. The cryptographic chaining of traces is implemented via Chainpoint, ensuring the integrity and immutability of event logs by anchoring them in the Ethereum blockchain. The electronic signatures of the logs use X.509 certificates managed by a private OpenSSL PKI. A private Hyperledger Fabric blockchain is deployed for recording critical events, with custom smart contracts for the automatic validation of sensitive operations. Consensus nodes are distributed across different availability zones to ensure high availability. Anomaly detection is performed by Elastic Security, using machine learning models to identify suspicious behaviors. Alerts are managed by Prometheus AlertManager, with PagerDuty integration for team notifications.

### 6.5 Security Infrastructure

The network perimeter is protected by Palo Alto Networks PA-7080 firewalls in an active/passive cluster, with SSL/TLS inspection of incoming traffic. Web requests are filtered by an F5 Advanced WAF, deployed in transparent bridge mode. Incident detection and

response are managed by a Splunk Enterprise Security SIEM, integrated with Suricata probes for network traffic analysis. Vulnerabilities are continuously detected by Tenable.io, with automated scans of containers and dependencies.

### 6.6 DevSecOps Pipeline

Continuous integration and deployment are managed by GitLab CI/CD, with automated security testing via GitLab Security Scanner. Container images are analyzed by Aqua Security Scanner before deployment. Code is versioned in GitLab Enterprise, with commits signed via GPG. Secrets are managed by HashiCorp Vault, with automatic credential rotation. Deployments follow a Blue/Green approach via ArgoCD to minimize service interruptions.

### 6.7 Resilience and Business Continuity

High availability is ensured by HAProxy in an active/passive configuration, with custom health checks for each service. Business continuity relies on a secondary site synchronized via Veeam, with a recovery time objective (RTO) of 4 hours and a recovery point objective (RPO) of 15 minutes. Backups are automatically tested each week through partial restorations in an isolated environment. Business continuity is maintained even in the event of complete data center loss due to synchronous replication of critical data.

## 7 DISCUSSION AND RECOMMENDATIONS

### 7.1 Lessons Learned from the Case Study

The in-depth analysis of the MedLab case has highlighted several fundamental lessons regarding the control of sensitive data flows in a modern medical environment. The experience demonstrates that a purely technical approach is insufficient to ensure effective data control. The success of the project relies on the harmonious integration of technological solutions, legal frameworks, and organizational aspects. The implementation of a Zero Trust architecture, while technically sophisticated, could only yield results through user adherence and clear data governance. The integration of emerging technologies such as artificial intelligence for anomaly detection and blockchain for traceability has also underscored the

**Table 2: Solutions to MedLab’s Needs**

| Technical Requirement       | Technologies and Solutions   |
|-----------------------------|--|
| Storage Security            | AES-256 encryption in GCM mode<br>FIPS 140-2 Level 3 HSM Modules<br>Zero Trust Architecture  |
| Transmission Protection     | TLS 1.3 protocols with TLS_AES_256_GCM_SHA384 cryptographic suite<br>Immediate encryption of biometric templates<br>Mutual authentication of endpoints |
| Isolation of Sensitive Data | Network segmentation into security zones<br>Edge computing for local processing<br>Embedded secure processors  |
| Operation Recording         | Advanced logging system<br>Timestamping and electronic signing of logs<br>Cryptographic chaining of traces   |
| Log Integrity               | Private blockchain for critical events<br>Cryptographic chaining mechanisms<br>Electronic signatures   |
| Access Traceability         | SIEM/SOC for real-time monitoring<br>Behavioral analysis using AI<br>Event correlation   |
| External Access Control     | Multifactor authentication (biometrics, token, PIN)<br>RBAC model enhanced with ABAC<br>Contextual validation of access                                |
| Environment Separation      | Distinct security zones<br>Inter-zone filtering rules<br>Segmented network infrastructure  |
| Flow Monitoring             | SIEM for real-time analysis<br>Anomaly detection using AI<br>Sophisticated correlation rules   |
| High Availability           | Redundant architecture (Kubernetes/Docker Swarm)<br>Synchronous replication<br>Automatic failover procedures   |
| Data Backup                 | Synchronous replication to backup site<br>Backup encryption<br>Integrity validation  |
| Anomaly Detection           | Machine learning for behavioral analysis<br>SIEM/SOC<br>Anomaly detection algorithms   |

importance of continuous technological monitoring to maintain the effectiveness of data control.

## 7.2 Ongoing Challenges and Proposed Solutions

Despite significant advancements, some challenges persist and warrant particular attention. Performance management is a major issue, as control mechanisms can introduce latency that impacts user experience. To address this issue, the study recommends adopting a graduated approach to control, adapting the level of verification

to the context and sensitivity of the data. The question of international data transfers, particularly in the context of collaboration with DataSecure Systems in the United States, also raises complex challenges. The proposed solution relies on a combination of enhanced contractual safeguards and complementary technical measures, such as end-to-end encryption and tokenization of sensitive data. The constantly evolving security threats represent another major challenge. In light of this reality, the study advocates for the development of a rapid adaptation capability for control mechanisms. This involves using artificial intelligence technologies

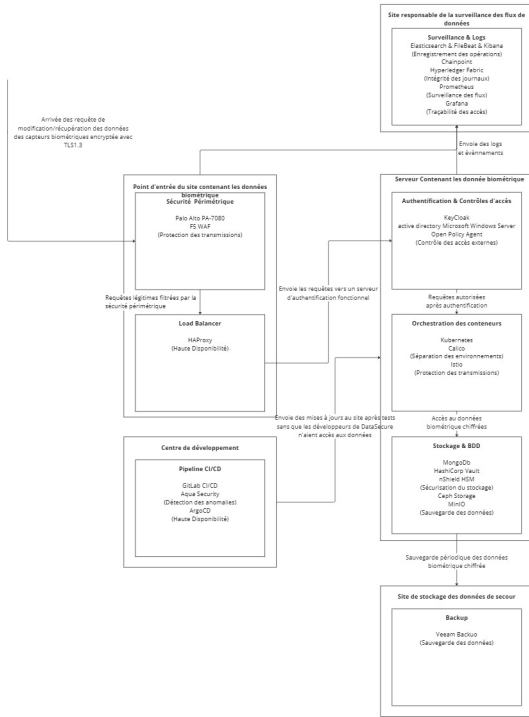


Figure 1: Proposed Architecture

for early threat detection and automating security responses, as well as establishing a dedicated security monitoring team.

### 7.3 Recommendations for Successful Implementation

Based on the lessons learned and the identified challenges, several concrete recommendations can be made for organizations looking to strengthen their control over data flows. First, it is essential to adopt a gradual approach in implementing solutions. This involves starting with a detailed mapping of data flows and a precise risk assessment before deploying control mechanisms. This preparatory phase should include a thorough data protection impact assessment, particularly for sensitive data such as biometric data. User training is another fundamental pillar of success. The training program should cover not only the technical aspects of the implemented solutions but also raise awareness of data protection issues and individual responsibilities. Regular knowledge update sessions should be organized to maintain a high level of vigilance.

### 7.4 Future Evolution Perspectives and Development Areas

The future evolution of data flow control revolves around several promising areas. The emergence of quantum computing, while posing a potential threat to current encryption systems, also paves the way for new approaches to data security. The development of post-quantum algorithms and their gradual integration into existing control systems is a priority development area. Homomorphic

encryption, despite its current limitations in terms of performance, represents a significant advancement that allows for the processing of encrypted data without prior decryption, thus opening new perspectives for the secure processing of sensitive data in cloud environments. Artificial intelligence is also expected to play an increasing role in optimizing control mechanisms. Machine learning systems will enable a more nuanced adaptation of controls to the usage context, thereby reducing their impact on performance while maintaining a high level of security.

## 8 Conclusion

At the end of this in-depth study on data flow control, several major conclusions emerge. The analysis of the MedLab case has highlighted the complexity of the technical and legal issues related to the management of sensitive data, particularly in the context of biometric systems.

The technical solutions developed demonstrate that it is possible to reconcile security and usability in data flow control. The Zero-Trust architecture, combined with multifactor authentication mechanisms and advanced encryption, offers a level of protection suitable for current requirements. The integration of emerging technologies such as artificial intelligence and blockchain opens new perspectives for further strengthening this control.

From a legal standpoint, the study emphasizes the importance of a comprehensive approach to compliance, incorporating not only the requirements of the GDPR but also sector-specific considerations. The management of subcontracting and international data transfers remains a major challenge, requiring particular vigilance in defining contractual and technical guarantees.

The experience of MedLab also demonstrates the crucial importance of the organizational aspect. The success of a data flow control project relies as much on the quality of processes and user training as on the robustness of technical solutions. Establishing clear governance and a continuous training program are essential elements of the framework.

The evolving landscape of the field presents new challenges, particularly related to the emergence of quantum computing and the increasing complexity of the regulatory environment. The ability to anticipate these developments and adapt solutions accordingly will be critical for maintaining effective data flow control.

The challenge for organizations now lies in their ability to develop an agile approach to data control, capable of quickly adapting to new threats and regulatory changes while maintaining a balance between security and operational efficiency. This agility must be based on a judicious combination of innovative technical solutions, robust organizational processes, and continuous monitoring of developments in the field.

Ultimately, data flow control appears as a major strategic issue for organizations, requiring a holistic and evolving approach. The solutions presented in this study provide a solid foundation to address these challenges while underscoring the need for continuous adaptation to the evolving technological and regulatory context.

## References

- [1] 2015. ISO/IEC 27017:2023 Technologies de l'information — Techniques de sécurité — Code de bonnes pratiques pour les contrôles de sécurité de l'information fondés



- sur l'ISO/IEC 27002 pour les services du nuage. <https://www.iso.org/fr/standard/56572.html>
- [2] Agence nationale de la sécurité des systèmes d'information (ANSSI). [n. d.]. *Recommandations de sécurité relatives à TLS*. [https://cyber.gouv.fr/sites/default/files/2017/07/anssi-guide-recommandations\\_de\\_securite\\_relatives\\_a\\_tls-v1.2.pdf](https://cyber.gouv.fr/sites/default/files/2017/07/anssi-guide-recommandations_de_securite_relatives_a_tls-v1.2.pdf)
  - [3] Cloud Security Alliance. 2010. Top threats to cloud computing v1. 0. *White Paper* 23 (2010).
  - [4] Jean-Luc Archimbaud. 2003. Les principes techniques des certificats électroniques. *Les Cahiers du numérique* 4, 3 (2003), 101–110.
  - [5] Hamed Arshad, Christian Johansen, and Olaf Owe. 2022. Semantic attribute-based access control: A review on current status and future perspectives. *Journal of Systems Architecture* 129 (2022), 102625.
  - [6] Paul M Beach, Logan O Mailloux, Brent T Langhals, and Robert F Mills. 2022. Analysis of systems security engineering design principles for the development of secure and resilient systems. In *Handbook of Scholarly Publications from the Air Force Institute of Technology (AFIT), Volume 1, 2000-2020*. CRC Press, 33–63.
  - [7] Johannes A Buchmann, Denis Butin, Florian Göpfert, and Albrecht Petzoldt. 2016. Post-quantum cryptography: state of the art. *The New Codebreakers: Essays Dedicated to David Kahn on the Occasion of His 85th Birthday* (2016), 88–108.
  - [8] Kristopher K Coelho, Michele Nogueira, Alex B Vieira, Edelberto F Silva, and José Augusto M Nacif. 2023. A survey on federated learning for security and privacy in healthcare applications. *Computer Communications* 207 (2023), 113–127.
  - [9] Revati Raman Dewangan, Sunita Soni, and Ashish Mishal. 2024. Optimized Homomorphic Encryption (OHE) algorithms for protecting sensitive image data in the cloud computing environment. *International Journal of Information Technology* (2024), 1–11.
  - [10] Souad El Hatib. 2020. Une approche sémantique de détection de maliciel Android basée sur la vérification de modèles et l'apprentissage automatique. (2020).
  - [11] Philip Harrison Enslow. 1978. What is a "distributed" data processing system? *Computer* 11, 1 (1978), 13–21.
  - [12] Craig Gentry. 2009. Fully homomorphic encryption using ideal lattices. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*. 169–178.
  - [13] A Gunuganti. 2023. Identity Based-Zero Trust. *JAIMLD* 1 (2023), 492–497.
  - [14] Ishu Gupta, Ashutosh Kumar Singh, Chung-Nan Lee, and Rajkumar Buyya. 2022. Secure data storage and sharing techniques for data protection in cloud environments: A systematic review, analysis, and future directions. *IEEE Access* 10 (2022), 71247–71277.
  - [15] Abdelhakim Hannousse and Salima Yahiouche. 2021. Securing microservices and microservice architectures: A systematic mapping study. *Computer Science Review* 41 (2021), 100415.
  - [16] Jiakun Hao, Jianbo Gao, Peng Xiang, Jiashuo Zhang, Ziming Chen, Hao Hu, and Zhong Chen. 2023. TDID: Transparent and Efficient Decentralized Identity Management with Blockchain. In *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 1752–1759.
  - [17] ISO/IEC 27701:2019. Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines. International Organization for Standardization [n. d.]. Geneva, Switzerland, 2019.
  - [18] John Kindervag et al. 2010. Build security into your network's dna: The zero trust network architecture. *Forrester Research Inc* 27 (2010), 1–16.
  - [19] Jonne Koskinen. 2023. Cloud Security Architecture. (2023).
  - [20] Wubin Li, Yves Lemieux, Jing Gao, Zhuofeng Zhao, and Yanbo Han. 2019. Service mesh: Challenges, state of the art, and future research opportunities. In *2019 IEEE International Conference on Service-Oriented System Engineering (SOSE)*. IEEE, 122–1225.
  - [21] Rahat Masood, Muhammad Awais Shibli, et al. 2012. Comparative analysis of access control systems on cloud. In *2012 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*. IEEE, 41–46.
  - [22] National Institute of Standards and Technology. [n. d.]. *Security Requirements for Cryptographic Modules*. Technical Report FIPS PUB 140-2. U.S. Department of Commerce. <https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.140-2.pdf>
  - [23] OJ L 119, 4.5.2016, p. 1–88 [n. d.]. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
  - [24] Emmanuel Ok, Grace John, and Praise Chris. 2024. Autonomous Infrastructure & Self-Healing Clouds. (2024).
  - [25] Eric Rescorla. 2018. *The transport layer security (TLS) protocol version 1.3*. Technical Report.
  - [26] Kamyar Rostami. 2024. *Authentification continue par le biais de justificatifs vérifiables sur les blockchains*. Ph.D. Dissertation. École de technologie supérieure.
  - [27] Tiyas Sarkar, Ravi Kumar, Mellachervu Sathwik Kumar, Sanchit Aggarwal, Achyuta Sandhya, and Anand Mohan Shukla. 2023. Review Paper of Performance Analysis in Wireless Sensor Networks. *Kilby* 100 (2023), 7th.
  - [28] Marcin Sendera, Marek Śmieja, Łukasz Maziarka, Łukasz Struski, Przemysław Spurek, and Jacek Tabor. 2021. Flow-based SVDD for anomaly detection. *arXiv preprint arXiv:2108.04907* (2021).
  - [29] Vasily Sidorov, Ethan Yi Fan Wei, and Wee Keong Ng. 2022. Comprehensive performance analysis of homomorphic cryptosystems for practical data processing. *arXiv preprint arXiv:2202.02960* (2022).
  - [30] Smartpoint. 2024. Gouvernance des données et architecture IT, conformité et sécurité. *Smartpoint* (2024). <https://www.smartpoint.fr/gouvernance-donnees-architecture-data/>
  - [31] Arman Sykot, Md Shawmoon Azad, Wahida Rahman Tanha, BM Morshed, Syed Emad Uddin Shubha, and MRC Mahdy. 2024. Multi-Layered Security System: Integrating Quantum Key Distribution with Classical Cryptography to Enhance Steganographic Security. *arXiv preprint arXiv:2408.06964* (2024).
  - [32] Haiyan Wang and Jiawei Zhang. 2019. Blockchain Based Data Integrity Verification for Large-Scale IoT Data. *IEEE Access* PP (11 2019), 1–1. doi:10.1109/ACCESS.2019.2952635
  - [33] Zheyu Zhang. 2024. Reinforcement Learning-Based Approaches for Enhancing Security and Resilience in Smart Control: A Survey on Attack and Defense Methods. *arXiv preprint arXiv:2402.15617* (2024).

---

# Modeling the Complexity of the Georgian Verbal System through Linked Data and Machine Learning

Elizbarashvili Archil  
Tbilisi State University  
Tbilisi, Georgia  
archil.elizbarashvili@tsu.ge

Ducassé Mireille  
Université de Rennes, IRISA  
Rennes, France  
mireille.ducasse@irisa.fr

Tsintsadze Magda  
Tbilisi State University  
Tbilisi, Georgia  
magda.tsintsadze@tsu.ge

Khachidze Manana  
Tbilisi State University  
Tbilisi, Georgia  
manana.khachidze@tsu.ge

## Abstract

The Georgian verbal system poses significant challenges due to its agglutinative and inflectional nature, lack of infinitives, rich morphological structure, and irregular patterns. To assist Georgian language learners and support lexicographic work, we are developing KARTUVERBS, a comprehensive linked-data platform for Georgian verb forms. This system integrates structured lexical data, semantic web technologies, and machine learning to infer missing lemmas (verbal nouns). Current results show a prediction accuracy of 98% using decision trees, with cross-validation against the Georgian National Corpus. This paper outlines our methodology, the unique linguistic challenges of Georgian, and future directions for integrating crowdsourcing and deep learning into the tool's development.

## Keywords

Linked data, machine learning, Georgian verbs, lexicography, verbal noun, decision tree

## 1 Introduction

The Georgian language's verbal morphology is notably complex: it is agglutinative, inflectional, and lacks a traditional infinitive form. Dictionaries typically use verbal nouns, third-person singular present or future forms, or roots as lemma entries. This ambiguity poses significant challenges for learners and computational models. For example, the form *mekirava* (მეჭირავა, "I were renting") can correspond to the verbal noun *kiraoba* (ჭირაობა), the present third person singular *kiraobs* (ჭირაობს), or simply the root *kira* (ჭირა).

Existing Georgian linguistic resources and dictionaries, such as the Comprehensive Georgian-English Dictionary Rayfield et al. [10] and Georgian-German dictionary Tschenkéli et al. [11], provide valuable but limited help for learners because of their varying lemma strategies and incomplete morphological information. Digital resources, such as the Georgian National Corpus and various online translation platforms, offer broader access but still fall short in consistently addressing lemma disambiguation and providing extensive morphological details.

To bridge these gaps, we are developing KARTUVERBS, an advanced linked-data-based semantic database containing over 5 million Georgian verb forms. This database leverages semantic web technology

and supervised machine learning to accurately predict and reconstruct verbal nouns and facilitate intuitive navigation among inflected forms and their morphological characteristics.

## 2 Literature Review

Research on Georgian verb morphology underscores its complexity and the challenges faced in lexicographic representation. Georgian verbs exhibit unique morphological features, such as polypersonal agreement, version markers, and preverbs, which significantly alter meaning and grammatical relationships Cherchi [2]; Tutite [12]; Makharoblidze [7]. The absence of an infinitive form further complicates dictionary compilation and learner comprehension Gippert [5]; Margalitadze [8].

Previous works, like those by Daraselia and Sharoff [3], initiated corpus-based lexicographic methodologies but lacked robust morphological linkage and lemma provision. Clarino INESS provides structured Georgian morphological data but primarily targets linguistic researchers rather than language learners or computational lexicography Meurer [9].

Advancements in linked open data (LOD) methodologies as crucial for future lexicographic developments and semantic web technologies have been suggested by Gracia et al. [6] and Ferré [4], enabling interoperability and user-friendly data navigation. Additionally, supervised machine learning methods, particularly decision trees, have shown promise in linguistic data prediction tasks due to their low complexity and robustness in handling missing values in multiclass classification problems, Bansal et al. [1].

KARTUVERBS integrates these methodological advancements into a comprehensive and learner-focused platform. The proposed system not only enhances dictionary usability through automatic lemma inference but also sets a foundation for future computational linguistic developments in Georgian and other morphologically rich languages.

## 3 Linguistic Peculiarities of Georgian Verbs

The Georgian verbal system represents one of the most intricate morphological architectures among world languages, characterized by numerous linguistic peculiarities that significantly complicate computational modelling, lexicography, and language learning.

### 3.1 Complex Tense and Aspect System

Unlike Indo-European languages, Georgian lacks a conventional infinitive form. This absence poses considerable challenges for dictionary compilation and language learners, as verbs are instead represented through multiple possible entry points such as verbal nouns, third-person singular forms, or roots. For instance, the English infinitive "to write" is typically presented as a verbal noun *ts'era* (წერა, literally "writing") or through third-person singular present forms such as *ts'ers* (წერს, "he/she writes"). Each choice complicates verb lookup and comprehension due to the considerable morphological differences between lemmas and their conjugated forms.

Georgian verbs exhibit polypersonal agreement, simultaneously encoding the grammatical roles of the subject, direct object, and frequently the indirect object. This polypersonal characteristic significantly expands conjugational possibilities. For example, the single verb form *mogts'era* (მოგწერა, "he/she wrote it to you") encapsulates agreement with three grammatical persons simultaneously, encoding subject, direct object, and indirect object within a compact morphological form. Georgian verbs follow a complex system of tense-aspect categories distributed across three morphologically distinct series:

Series I (Present and Future): Represents ongoing, habitual, or future actions (e.g., *ts'er* წერ, "you write" / *dats'er* დაწერ, "you will write").

Series II (Aorist): Denotes completed actions in the past with no current relevance, often associated with perfective aspects (e.g., *dats'ere* დაწერე, "you wrote").

Series III (Perfect): Marks actions completed in the past but with present relevance, similar to English perfect tenses (e.g., *dagits'eria* დაგიწერია, "you have written (at some point)").

Each series introduces specific morphological changes affecting verb roots, preverbs, and endings, significantly increasing morphological complexity.

### 3.2 Preverbs and Version Markers

Georgian preverbs and version markers alter verb meanings substantially, introducing semantic nuances related to motion, directionality, benefaction, and other semantic dimensions. Preverbs such as *gada-* (გადა-, "across, again"), *mo-* (მო-, "towards the speaker"), and *ts'a-* (წა-, "away from the speaker") interact with verb stems to produce new lexical meanings and aspectual distinctions. Similarly, version markers (e.g., subjective, objective, and neutral) encode nuances concerning participants' roles and interests within an event, adding complexity to verb conjugation and interpretation.

### 3.3 Verbs

Georgian verb morphology includes numerous irregular forms and instances of suppletion (where different morphological forms are entirely unrelated). Common verbs such as *q'opna* (ყოყნა, "to be") demonstrate extensive irregularity across tense and person conjugations, significantly complicating both learning and computational modeling.

Verbs are traditionally categorized into morphological classes based on conjugation patterns (e.g., Class I: transitive, active verbs;

Class II: intransitive, stative verbs; Class III: mediopassive verbs; and Class IV: indirect verbs). Each class follows distinct morphological rules, further complicating systematization and lemma generation.

Georgian morphology is highly agglutinative, frequently resulting in verb forms composed of multiple morphological units (preverbs, root, version markers, thematic suffixes, mood-tense markers, and person-number markers). This extensive concatenation of morphemes within verb forms leads to extremely high form variety and necessitates meticulous morphological parsing for computational modeling.

Understanding these linguistic peculiarities is crucial for computational and lexicographic projects, such as KARTUVERBS, as they directly impact the design of databases, lemma prediction algorithms, and user interfaces. By explicitly modeling and accommodating these complexities, our approach aims to provide a robust computational tool capable of supporting both learners and researchers in navigating the rich and challenging Georgian verbal system.

## 4 Methodology

To effectively manage and utilize the complex morphological data inherent to Georgian verbs, our approach combines structured data extraction, transformation into linked data, and supervised machine learning. As illustrated by Figure 1, the process to transfer data from Clarino to KARTUVERBS consists of three main blocks. The first stage extracts data from Clarino web pages and generates an intermediate CSV file. The second stage applies several normalization and enrichment processes to improve the quality and usability of the data, addressing linguistic and structural inconsistencies, and includes machine learning techniques to reconstruct missing lemmas. The final stage transforms the cleaned CSV data into RDF triples and deploys a SPARQL endpoint to support semantic querying and interoperability.

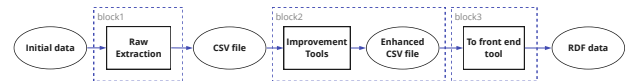


Figure 1: Global structure of the transformation process

### 4.1 Data Transformation Pipeline

The KARTUVERBS database relies initially on morphological data obtained from Clarino, an online resource primarily targeting linguistic researchers. To convert this linguistic data into a structured, semantic linked-data resource tailored to learners and computational applications, we implemented a multi-stage data transformation pipeline detailed as follows:

The original data, available as web pages from the Clarino INESS repository, was scraped and parsed using script. Due to Clarino's hierarchical data structure (organized by common roots, verbs, and inflected forms), we first converted the scraped JSON data into a flat CSV format, resulting in a dataset of approximately 22 million verb form entries.

The raw JSON data extracted was initially complex and included redundant hierarchical structures. We flattened these structures by

extracting each inflected form along with associated morphological fields such as tense, person, number, root, preverb, and morphological markers. The dataset was filtered to retain only 14 key morphological attributes crucial for our lexicographic goals. Each verb form was represented by a single, easily manageable record, facilitating later transformations and analyses.

The Clarino dataset utilized linguistic shorthand and symbolic notations, necessitating normalization and enrichment. Special Python scripts were developed to decode symbolic notations (such as Latin placeholders indicating optional or obligatory morphemes) and reconstruct full verbal nouns by combining root lemmas and preverbs. For example, Clarino’s shorthand verbal noun “*\*adami-aneba*” (\*ადამიანება) was programmatically expanded to the fully reconstructed verbal noun “*gaadamianeba*” (გაადამიანება, “to humanize somebody”).

This normalization step included automatic handling of variant morphemes, removal of redundant alternatives, and the correction of encoding irregularities present in Clarino’s linguistic annotation.

To enable semantic querying and interoperability, the normalized CSV dataset was transformed into RDF (Resource Description Framework) triples. Each CSV entry was converted into multiple RDF triples using scripts designed specifically for this transformation. The resulting RDF dataset (in Turtle N-Triples format) exceeded 80 million links, enabling rich semantic querying capabilities through SPARQL endpoints.

An Apache Jena Fuseki server was employed to manage RDF indexing, supporting efficient real-time queries through the Sparklis<sup>1</sup> linked-data frontend. This allowed learners and lexicographers to intuitively navigate between verb forms, lemmas, and their morphological details without requiring technical query syntax knowledge. In addition to the integrated KARTUVERBS endpoint, users can specify a remote SPARQL endpoint address to query external resources. Having a locally running endpoint is particularly convenient when working with frequent minor revisions or experimental dataset versions.

## 4.2 Machine Learning for Missing Lemmas

Although extensive, the Clarino-derived dataset initially contained approximately 600,000 entries lacking essential verbal noun lemmas, posing a significant challenge. To address this, we employed supervised machine learning to predict and reconstruct missing lemmas from morphological features present in other dataset entries.

We utilized a subset of around 300,000 fully annotated verb forms to train a supervised decision tree classifier. Fourteen carefully selected morphological features, such as tense, person, number, preverb, stem formants, and morphological class, were extracted and encoded numerically. To represent Georgian textual features numerically, we devised an encoding approach based on UTF-8 character representation, ensuring that each verbal noun uniquely mapped to a distinct numerical identifier, essential for accurate classification and later decoding.

A decision tree classifier was selected due to its robustness in handling multi-class classification tasks and inherent interpretability, which significantly aids linguistic validation. Decision trees efficiently manage datasets with missing values, an essential requirement given the irregular nature of morphological data. The dataset was randomly split into training (80%) and testing (20%) subsets, a standard practice to evaluate generalization performance reliably.

The trained decision tree model exhibited an overall classification accuracy ranging from 98% to 99% in reconstructing missing verbal noun lemmas. Precision, recall, and F1-score metrics further confirmed robust model performance. This high accuracy suggests that morphological properties encoded within Georgian verb forms provide sufficient predictive power for automatic lemma reconstruction.

Individual class performance varied, reflecting inherent dataset imbalance and morphological irregularities specific to certain verbal noun groups. Decision trees can sometimes overfit or struggle with specific patterns in the data. To address this, we experimented with another robust model, Support Vector Machine (SVM). However, SVM yielded significantly worse performance on our dataset.

To further validate our predictions, we have cross-checked reconstructed lemmas against lemmas provided by the Georgian National Corpus (GNC). Approximately 92% of our predicted lemmas matched GNC lemmas, confirming their reliability. A minor discrepancy (approximately 6.5%) pointed toward either genuine linguistic variability or minor errors, motivating further expert-driven verification and crowdsourcing initiatives.

To address residual ambiguities and discrepancies identified in cross-validation, we initiated a crowdsourcing campaign through the Headwork<sup>2</sup>, an open-source academic platform for the crowdsourcing of complex tasks. Experts and native Georgian speakers can collaboratively review and validate verb lemma predictions flagged as uncertain by our automated verification process. Contributors evaluate morphological consistency across multiple verb forms and make linguistic judgments on predicted verbal nouns, enabling continual improvement of dataset accuracy.

## 5 Conclusion

KARTUVERBS leverages semantic web and machine learning to model the Georgian verbal system. It contains approximately 5 million inflected forms associated with over 16,000 verbs, featuring more than 80 million link. For learners, it serves as a digital lexicographic tool that identifies verbal nouns from any inflected form and provides an API for translation through electronic dictionaries or machine translation applications, such as Google Translate. For Georgian lexicographers, KARTUVERBS offers a comprehensive digital system that consolidates nearly all Georgian verbs and their inflected forms, supporting both pedagogical and scholarly endeavors given the language’s intricate and exception-rich verb system. For lexicographers of other languages, KARTUVERBS demonstrates how Semantic Web technologies combined with linked data, can create a machine-readable, interoperable database for linguistic resources. Our linked-data database platform achieves acceptable response times even when hosted on a private machine, validating the feasibility of the approach.

<sup>1</sup><http://www.irisa.fr/LIS/ferre/sparklis>

<sup>2</sup><https://headwork.irisa.fr/headwork/>

All scripts produced during this study, including those for experimental validation, are openly available online<sup>3</sup>. These resources can be adapted to other projects, enabling the transformation of data from one purpose to another effectively.

## References

- [1] Malti Bansal, Apoorva Goyal, and Apoorva Choudhary. 2022. A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning. *Decision Analytics Journal* 3 (2022), 100071. doi:10.1016/j.dajour.2022.100071
- [2] Marcello Cherchi. 1997. *Modern Georgian Morphosyntax: A grammatico-categorial hierarchy-based analysis with special reference to indirect verbs and passives of state*. Harrassowitz Verlag · Wiesbaden.
- [3] Sophiko Daraselia and Serge Sharoff. 2016. Enriching Georgian Dictionary Entries with Frequency Information. In *Proceedings of the 17th EURALEX International Congress* (6-10), Tinatin Margalitadze and George Meladze (Eds.). Ivane Javakhishvili Tbilisi University Press, Tbilisi, Georgia, 321–327.
- [4] Sébastien Ferré. 2017. Sparklis: An Expressive Query Builder for SPARQL Endpoints with Guidance in Natural Language. *Semantic Web: Interoperability, Usability, Applicability* 8, 3 (2017), 405–418. doi:10.3233/SW-150208
- [5] Jost Gippert. 2016. Complex Morphology and its Impact on Lexicology: the Kartvelian Case. In *Proceedings of the 17th EURALEX International Congress* (6-10), Tinatin Margalitadze and George Meladze (Eds.). Ivane Javakhishvili Tbilisi University Press, Tbilisi, Georgia, 16–36.
- [6] Jorge Gracia, Ilan Kernerman, and Julia Bosque-Gil. 2017. Toward Linked Data-native Dictionaries. In *Proceedings of the eLex 2017: Lexicography from Scratch* (19-21). Netherlands, 550–559. <https://elex.link/elex2017/wp-content/uploads/2017/09/paper33.pdf>
- [7] Tamar Makharoblidze. 2012. *The Georgian Verb*. LINCOM GmbH.
- [8] Tinatin Margalitadze. 2020. The Lexicography of Georgian. In *International Handbook of Modern Lexis and Lexicography*, Patrick Hanks and Gilles-Maurice de Schryver (Eds.). Springer Berlin Heidelberg, 1–24. doi:10.1007/978-3-642-45369-4\_103-1
- [9] Paul Meurer. 2007. A computational grammar for Georgian. In *International Tbilisi Symposium on Logic, Language, and Computation*. Springer, 1–15.
- [10] Donald Rayfield, Shukia Apridonze, Ariane Chanturia, Rusudan Amirejibi, Laurence Broers, Levan Chkhaidze, and Tina Margalitadze (Eds.). 2006. *A Comprehensive Georgian-English Dictionary*. Garnett Press.
- [11] Kita Tschenkéli, Yolanda Marchev, and Lea Flury. 1965. *Georgisch-deutsches Wörterbuch*. Vol. 2. Amirani-Verlag Zürich.
- [12] Kevin Tuite. 1998. *Kartvelian morphosyntax: Number agreement and morphosyntactic orientation in the South Caucasian languages*. Lincom Europa Munich.

<sup>3</sup><https://github.com/aelizbarashvili/KartuVerbs>

# Smart Cities Using Intelligent IoT and WSNs: Technologies, Concepts, and Challenges

Tanya Garg\*\*  
tanya.garg@thapar.edu  
Thapar Institute of Engineering and  
Technology  
Patiala, Punjab, India

Jose Neuman de Souza  
Federal University of Ceará  
Ceara, Brazil  
neuman@ufc.br

Vishnu Batra  
Thapar Institute of Engineering and  
Technology  
Patiala, India  
vbatra\_be21@thapar.edu

## Abstract

The Wireless Sensor Networks and IoT are evolving technologies in the paradigm of Smart Cities. Interconnection among smart devices is to be done to provide the ease of communication using these technologies. Sensors are deployed in smart devices and embedded systems which are further integrated with Internet of Things to impart advanced facilities for developing smart cities. This paper explores the transformative role of Artificial Intelligence (AI) in enhancing Internet of Things (IoT) and Wireless Sensor Networks (WSNs) for smart city development. By leveraging AI-driven data analytics, predictive modeling, and automation, IoT and WSNs can optimize urban resource management, improve infrastructure efficiency, and support sustainable growth in smart cities. The Internet of Things are capable of bridging the physical objects by using actuator networks, sensors and then they control these components. It is the most crucial area of research in the field of wireless sensor networks that can be combined with IoT to work in developing Smart Cities. The main focus of writing this paper is on how smart devices can be deployed in various locations to collect the required information, how this collected data is to be analyzed and how it helps the Smart Cities. In this paper, our main objective is to discuss the technologies, concepts, various issues and challenges in developing Smart Cities using these intelligent IoT and WSNs.

## Keywords

Internet of Things (IoT), Wireless Sensor Networks (WSNs), Smart Cities, Physical Systems, Embedded Systems

## ACM Reference Format:

Tanya Garg, Jose Neuman de Souza, and Vishnu Batra. 2025. Smart Cities Using Intelligent IoT and WSNs: Technologies, Concepts, and Challenges. In *Proceedings of International Workshop on ADVANCES in ICT Infrastructures and Services (ADVANCE'2025)*. ACM, New York, NY, USA, 5 pages. <https://doi.org/XXXXXXX.XXXXXXX>

\* All the authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

ADVANCE'2025, Sophia-Antipolis, France

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-1-4503-XXXX-X/2018/06  
<https://doi.org/XXXXXXX.XXXXXXX>

## 1 Introduction

A Smart city is a system, made out of Information and Communication Technologies (ICT), to create, convey, and elevate supportable advancement practices to address developing urbanization challenges. A major piece of this ICT structure is basically a canny organization of associated articles and machines (otherwise called an advanced city) that communicate information utilizing remote innovation and the cloud. Cloud-based IoT applications get, examine, and oversee information progressively to help regions, undertakings, and residents settle on better choices that improve personal satisfaction. Residents draw in with brilliant city biological systems in different manners utilizing cell phones and cell phones and associated vehicles and homes. Blending gadgets and information with a city's actual foundation and administrations can reduce expenses and improve manageability. Networks can improve energy the executives, Agriculture Automation, Embedded frameworks, Building the board, Telemedicine and Healthcare, Everyday things, Vehicle observing, Security and reconnaissance, decline gridlock, and even improve air quality with assistance from the IoT [1]. The Figure 1 below shows the role of IoT in various applications of smart cities. Artificial Intelligence (AI) complements IoT and Wireless Sensor

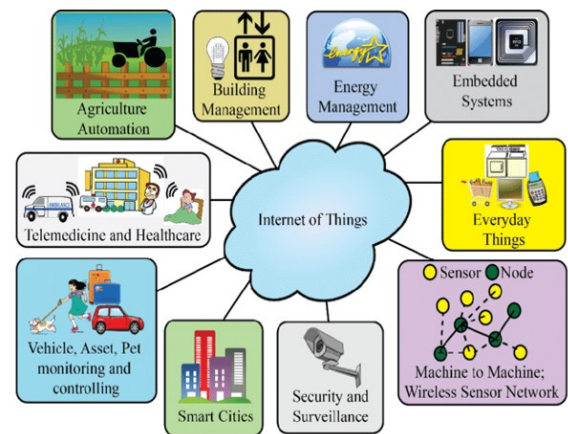


Figure 1: IoT in Smart Cities [1]

Networks (WSNs) by transforming raw sensor data into actionable insights. Through machine learning algorithms and advanced analytics, AI enables predictive maintenance, traffic optimization, and energy-efficient solutions, which are critical for smart city sustainability. The integration of AI with IoT and WSNs enhances the



decision-making capabilities of smart city systems. AI algorithms can identify patterns, forecast urban demands, and automate responses in areas such as traffic control, waste management, and public safety, thereby improving overall urban resilience [13]. Issues identified with wellbeing, traffic, contamination, shortage of assets, squander the executives and helpless framework emerge and subsequently advancement of city falls apart. This has set off the use of innovation as an answer of every one of these issues and to address them in a more astute manner. Thus the idea of Smart Cities is coming up. Smart Cities guarantee economic climate with the assistance of Big Data and Internet of Things. By "Smart", we imply that the city is more economical, decent and productive. The brilliant city market is assessed to arrive at a yearly expenditure of about 16 billion dollars continuously 2020 (Pike Research 2011). Brilliant Cities, with a legitimate guide, will serve individuals everywhere and will unquestionably help in diminishing labor as long as possible. However, the idea of brilliant city is as yet hazy in the personalities of individuals, scarcely any investigations have endeavored to address the inquiries with respect to the advancement of urban areas [4]. In the wake of playing out a broad examination in different spaces like public administration, data innovation, e-administration, we have recognized six basic factors that structure column for the advancement of a Smart city. In this work, various aspects related to the concept of Smart City and how IoT can be used for developing smart cities has been discussed in detail. The Scenario without IoT has also been discussed which forces us to develop the concept of Smart City.

### 1.1 AI driven IoT in Smart Cities

The internet is an essential mechanism that is used for communicating and gathering of information. IoT is prompting an adjustment of the way of life as countless gadgets, actuators, sensors, and different components are connected to one another. A network of enormous gadgets that have been modified to gather information brought about a totally new administration and highlights the structure of the premise of some significant ideas such as "Smart Cities. The use of IoT for smart cities can bring immense market openings and will make the lives of people smarter. AI-driven IoT plays a transformative role in the development of smart cities by enabling real-time data processing, predictive analytics, and intelligent automation. By integrating machine learning algorithms with IoT devices and wireless sensor networks (WSN), smart cities can efficiently manage urban resources, optimize traffic flow, enhance energy distribution, and improve public safety. AI analyzes the vast amounts of data generated by IoT sensors to identify patterns, predict potential problems, and enable proactive decision-making. This synergy between AI and IoT not only enhances operational efficiency but also supports sustainable urban growth, creating adaptive, resilient, and citizen-centric environments. Today, the gadgets around us are gradually becoming wiser. In addition, these advances will undoubtedly change our conduct and the manner in which we use them. We are in a period where we are attempting to find new freedoms that we get by new technologies that intend to exploit the progression of new close-to-home and worldwide information. Urban areas are probably going to contribute about



Figure 2: AI driven IoT in Smart Cities

41 trillion dollars in IoT advances in the next 20 years. To make urban communities more brilliant, legislatures have begun advancing some new companies and different companies to take advantage of advances in IoT with the goal of being implemented in a few circles of metropolitan living [10].

## 2 Problems related with cities

The smart city environment requires the technology as a fundamental need, just as to consideration of social and human resources viewpoints. Smart cities, while offering enhanced efficiency and connectivity, face several critical challenges. One major issue is data privacy, as the vast amount of personal information collected through sensors and devices can be vulnerable to misuse or breaches. Cybersecurity risks are also prominent, with interconnected systems presenting potential targets for hackers. High infrastructure costs pose another barrier, especially for developing regions, requiring substantial investment in technology and maintenance. Additionally, managing and analyzing the enormous volumes of data generated can overwhelm existing systems, leading to inefficiencies or misinformed decisions. These challenges must be carefully addressed to ensure the sustainable growth of smart cities [2].

These days, there are many smart city arrangements dependent on custom frameworks and arrangements; however, these are not generally appropriate to different urban areas all throughout the planet, and, in some cases, just a subset of the different viewpoints should be thought of. In the accompanying, the principle hindrances, issues, and open difficulties recognized in the writing for savvy urban communities are summed up. More in detail, we recognized that these difficulties can be isolated inside four principle measurements: residents, versatility, administration, and climate. This is remarkably significant as residents would fear the presentation of

imaginative advancements or might see it as meddling. The summary of various problems associated with smart cities are as shown in Figure 3 below:



Figure 3: Problems associated with Smart Cities

Multiple strategies must be implemented to address the challenges that smart cities face. Advanced encryption together with regular system audits and real-time threat detection systems should be implemented to enhance cybersecurity protection. Data privacy needs to be protected through strict data governance policies and anonymization techniques and complete transparency about data usage. Governments should implement public-private partnerships (PPPs) and promote open-source platforms and deploy infrastructure in phases to minimize financial burden [11]. Real-time decision-making and analysis can be achieved through the integration of AI and machine learning tools for effective data management. The implementation of smart city technology requires active citizen participation and educational programs to build public trust and promote responsible technology use.

### 3 Smart City Challenges

The convergence of the Internet with radio frequency identification devices, sensors, and other smart devices is termed the Internet of Things. As per reports [10], nearly 70 percent of the world population will live in cities by 2050. This inflow of people has ignited the transformation of urban areas into smart cities. There are a variety of urban issues that need better, sustainable, and efficient solutions and management. Such social issues comprise waste management, energy management, water management, public transportation, healthcare, security, and education, to name a few. In case of a pandemic, it is even more important to digitalize everything in order to facilitate people living in quarantine zones. To address all such issues, technology comes into the picture in the form of smart cities. Various components of a smart city have been displayed in Figure 4 below [2]. In smart cities, various devices are connected to each other through sensor networks in order to facilitate data collection which is further utilized to assist people in their daily routine. Be it any domain; things are getting electrified and connected to

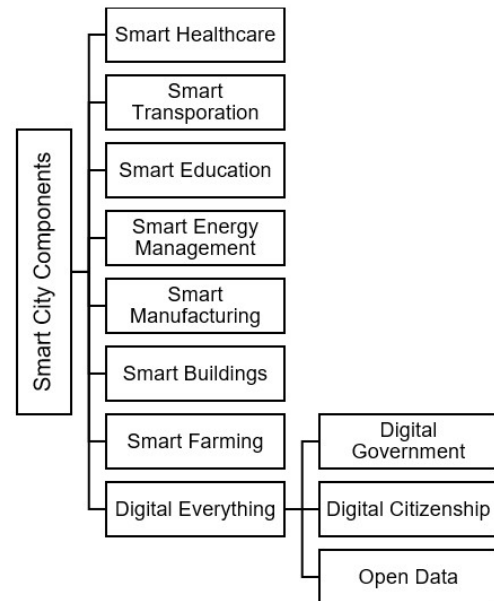


Figure 4: Components of Smart Cities

revolutionize the idea of smart city. But in order to lead the transformation, numerous challenges need to be addressed which have been discussed in this section.

#### 3.1 Infrastructure

The foremost challenge associated with the development of smart cities is infrastructure and investments. The idea of a smart city can be conceptualized very well but its implementation lacks due to the non-availability or affordability of requisite infrastructure. For instance, various sensors in a smart home would need to communicate and pass on information from one device to another; mimicking an algorithm of people's daily routine. Development of underlying infrastructure needs to be done by considering affordable, clean and sustainable technological solutions using green energy. The up gradation of existing infrastructure into 'smart' infrastructure and integration with legacy systems require efficient planning and investments [14].

#### 3.2 Security and Privacy

With the proliferation of sensors and connected devices, security issues arise. Cyber threats are not only limited to computers and internet platforms now a days. Stuxnet [3] is an example of how real time systems can be attacked and controlled by the hackers. Therefore, there is a dire need of developing secure networks and systems to prevent such attacks and safeguard the systems from hackers. On the other hand, quality of life is often associated with loss of privacy such as CCTV cameras inside homes may help with security but invade your privacy if fall into wrong hands. Therefore, ensuring privacy of data keeping into mind the trust of the people in the security solutions is the need of the hour.



### 3.3 Data Management

'Big data' is the term that refers to the huge amount of data collected from different sensors and computational devices. Since smart city is a supposed to be a storehouse of data; efficient data management systems need to be developed to store, process and analyse data from various sources.

Traditional database management systems may not be able to store such variety of data coming from different sources and having multiple dimensions; therefore there is a need of big data storage and processing tools such as HBase, Hadoop, Mahout, Cassandra etc. The availability of data needs to be ensured by using redundancy and replication principles. The disaster management plan should also include safeguarding the data resources of the individuals [7].

### 3.4 Sustainability

With the rise in global warming, climate change, development of sustainable smart city solutions in terms of social, economic and ecological is the need of the hour. Since technology without energy is useless; energy without leveraging clean, green and renewable energy is even more useless. Therefore, 'smart energy' [8] aspect must be taken well into consideration while planning a smart city. Green energies [16] include solar, wind, hydroelectric, biogas and biomass; having huge potential for the future world. Therefore, technologies should be created by harnessing the renewable sources around us to avoid damage to the nature and make them more accessible and affordable.

### 3.5 Social and Economic Challenges

Apart from the techniques, platforms and technology related challenges, there are social and economic barriers as well which may affect the implementation of a smart city if not taken into consideration while preparing the implementation plans. One major social challenge is covering all the spheres of the population. With growth of population; inclusive applications need to be developed which ensure equity among the citizens. Lack of trust by people into the new systems and reluctance in change from traditional approaches makes it difficult for the smart city idea to emerge and rise. Security and privacy are other social concerns which need to be addressed and citizens must be well acquainted with these. On the other hand, economic challenges include insufficient funding, lack of business models, economic security, interoperable services and sustainability [15].

## 4 Applications of AI driven IoT & WSN in Smart Cities

Wireless sensor network is basically a mesh of sensors which connect with each other wirelessly and form a network of communication. With the rise in internet of things, wireless sensor networks have become a part and parcel of life. Since urbanization has led to the emergence of the concept of smart cities, therefore wireless sensor networks have become an important technology to be leveraged [5]. AI-driven IoT and Wireless Sensor Networks (WSN) have a wide range of applications in smart cities, enhancing urban living through intelligent systems. In transportation, AI algorithms optimize traffic flow, reduce congestion, and enable autonomous vehicle integration. In energy management, IoT sensors, combined

with AI, facilitate smart grids that monitor and adjust energy usage in real-time, improving efficiency and sustainability. AI-driven IoT is also instrumental in waste management by optimizing collection routes and schedules, while WSNs monitor environmental conditions such as air quality and pollution levels. Additionally, AI-powered surveillance and security systems ensure public safety by detecting anomalies and responding to incidents in real-time. In a smart city, data can be collected from everywhere such as environment, IT infrastructure, sensors, computational devices, businesses etc. Various kinds of smart city services and applications make use of sensors to collect data for analysis and decision making [9]. In such a scenario, IoTs and WSNs have a significant part in building the smart cities. Integration of IoT and WSN assists in handling complex computational tasks in the areas of disaster management, agriculture modernization, e-health, weather forecasting and privacy [6]. Various IoT and wireless sensor networks applications have been shown in summarized form in the Figure 5 below:

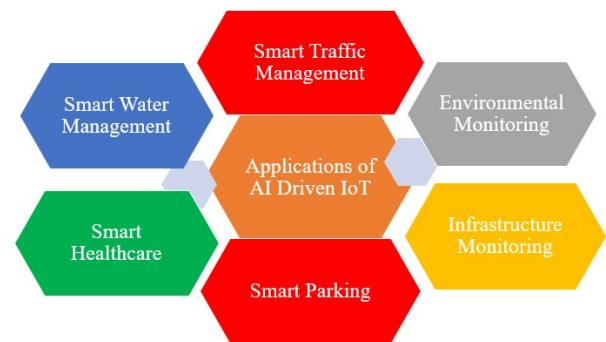


Figure 5: Applications of AI driven IoT

The implementation of smart cities using IoT and WSN involves sensors, devices, applications, and people. Therefore, security becomes a critical aspect to be looked into. These networks are prone to a large number of security threats and therefore require strong authentication, authorization, and privacy and integrity measures. The trust of people in the smart city applications can only be ensured by implementing integrated security solutions. An instance of such a solution is the video surveillance systems, which can provide real-time threat analysis, detection, and generation of alarms in suspicious situations using big data analytics and data mining techniques. There are numerous artificial intelligence-based IoT and WSN architectures that can be utilized in smart city scenarios to ensure security and privacy. The WSN protocols and IoT-based networking architectures can be integrated with artificial intelligence, machine learning, and deep learning-based solutions in order to provide security at each layer. Blockchain is one important technology that has been proven to solve security issues of the smart cities by ensuring decentralization, transparency, immutability, and auditability [17]. There are numerous other technologies and platforms, such as cyber-physical systems, unmanned aerial vehicles, cloud computing, edge computing, biometrics, data centers, geographic information systems, global positioning systems,

vehicular ad hoc networks, etc., that may be explored to prepare robust, safe, and reliable architectures for smart cities.

## 5 Conclusion and Future Research Directions

There are multiple infrastructures, services, and processes that need to be managed in a smart city where the Internet of Things and wireless sensor networks together play a significant role in carrying out the interconnection and networking. The implementation of a smart city has numerous challenges which need to be addressed while designing the architecture of the city. Wireless sensor networks form a backbone of the networking architecture of smart cities and assist in a number of applications, and the Internet of Things facilitates cost-effective, secure, optimal, and sustainable solutions [12]. AI-driven IoT serves as the backbone of smart city development by transforming data into actionable insights and enabling intelligent automation. The integration of AI with IoT and wireless sensor networks enhances urban efficiency through real-time monitoring, predictive analytics, and adaptive decision-making. This synergy not only optimizes resource management and infrastructure performance but also fosters sustainable, resilient, and citizen-centric urban environments, paving the way for the future of smart cities." There are various emerging state-of-the-art techniques, technologies, and other research directions that have huge potential to make cities smarter than ever.

## Acknowledgments

We would like to thank the Department of Computer Science and Engineering, Thapar Institute, for their continuous support and encouragement. We also acknowledge the insightful feedback from reviewers, which greatly contributed to enhancing this work.

## References

- [1] Kashif Ahmad, Majdi Maabreh, Mohamed Ghaly, Khalil Khan, Junaid Qadir, and Ala Al-Fuqaha. 2020. Developing future human-centered smart cities: Critical analysis of smart city security, interpretability, and ethical challenges. *arXiv preprint arXiv:2012.09110* (2020).
- [2] Karan Belsare, Manwinder Singh, Anudeep Gandam, Praveen Kumar Malik, Ritu Agarwal, and Anita Gehlot. 2024. An integrated approach of IoT and WSN using wavelet transform and machine learning for the solid waste image classification in smart cities. *Transactions on Emerging Telecommunications Technologies* 35, 4 (2024), e4857.
- [3] Gaurav Dhiman and Norah Saleh Alghamdi. 2024. Smose: Artificial intelligence-based smart city framework using multi-objective and iot approach for consumer electronics application. *IEEE Transactions on Consumer Electronics* 70, 1 (2024), 3848–3855.
- [4] Nikhita Reddy Gade, Nishanth Reddy Gade, and GU Reddy. 2016. Internet of things (IoT) for smart cities-The future technology revolution. *Global Journal of Computer Science and Technology: E Network, Web & Security* 16, 1 (2016), 28–33.
- [5] Taher M Ghazal, Mohammad Kamrul Hasan, Haitham M Alzoubi, Muhammad Alshurideh, Munir Ahmad, and Syed Shehryar Akbar. 2023. Internet of things connected wireless sensor networks for smart cities. In *The effect of information technology on business and marketing intelligence systems*. Springer, 1953–1968.
- [6] Kun Guo, Yueming Lu, Hui Gao, and Ruohan Cao. 2018. Artificial intelligence-based semantic internet of things in a user-centric smart city. *Sensors* 18, 5 (2018), 1341.
- [7] Essam H Houssein, Mahmoud A Othman, Waleed M Mohamed, and Mina Younan. 2024. Internet of Things in Smart Cities: Comprehensive Review, Open Issues and Challenges. *IEEE Internet of Things Journal* (2024).
- [8] Hiroshi Ito. 2025. Energy Transition in Japan's SDGs Future Cities: Toyama, Sapporo, and Kumamoto. In *Smart Cities and Japan's Energy Transition*. Routledge, 165–187.
- [9] Santosh Kumar Jha, S Jennathu Beevi, MN Babitha, Somu Chinnusamy, Sampath Boopathi, et al. 2024. Artificial Intelligence-Infused Urban Connectivity for Smart Cities and the Evolution of IoT Communication Networks. In *Blockchain-Based Solutions for Accessibility in Smart Cities*. IGI Global, 113–146.
- [10] Sujata Joshi, Saksham Saxena, Tanvi Godbole, et al. 2016. Developing smart cities: An integrated framework. *Procedia Computer Science* 93 (2016), 902–909.
- [11] KM Karthick Raghunath, Manjula Sanjay Koti, R Sivakami, V Vinoth Kumar, Grande Nagajyothi, and V Muthukumaran. 2024. Utilization of IoT-assisted computational strategies in wireless sensor networks for smart infrastructure management. *International Journal of System Assurance Engineering and Management* 15, 1 (2024), 28–34.
- [12] Nahid Neoaz. 2025. Internet of Things (IoT) and Smart Cities Examine how IoT technologies can improve urban living and infrastructure management. *Author Nahid Neoaz* (2025).
- [13] Emmanuel Utochukwu Ogbodo, Adnan M Abu-Mahfouz, and Anish M Kurien. 2022. A survey on 5G and LPWAN-IoT for improved smart cities and remote area applications: From the aspect of architecture and security. *Sensors* 22, 16 (2022), 6313.
- [14] Shailendra Pratap Singh, Naween Kumar, Norah Saleh Alghamdi, Gaurav Dhiman, Wattana Viriyasitavat, and Assadaporn Sapsomboon. 2024. Next-Gen WSN Enabled IoT for Consumer Electronics in Smart City: Elevating Quality of Service Through Reinforcement Learning-Enhanced Multi-Objective Strategies. *IEEE Transactions on Consumer Electronics* (2024).
- [15] Amin Ullah, Syed Myhammad Anwar, Jianqiang Li, Lubna Nadeem, Tariq Mahmood, Amjad Rehman, and Tanzila Saba. 2024. Smart cities: The role of Internet of Things and machine learning in realizing a data-centric smart environment. *Complex & Intelligent Systems* 10, 1 (2024), 1607–1637.
- [16] Or Yatzkan, Reuven Cohen, Eyal Yaniv, and Orit Rotem-Mindali. 2025. Urban Energy Transitions: A Systematic Review. *Land* 14, 3 (2025), 566.
- [17] Fan Zeng, Chuan Pang, and Huajun Tang. 2024. Sensors on internet of things systems for the sustainable development of smart cities: a systematic literature review. *Sensors (Basel, Switzerland)* 24, 7 (2024), 2074.

Received 16 February 2025; revised 22 May 2025; accepted 25 May 2025

# User-Centric Challenges in Digital Identity Wallets: Insights from Industry Experimentation

Louis RAFFIN  
I3S / CNRS & Docaposte  
Université Côte d'Azur  
Sophia Antipolis, FR, EU  
louis.raffin@i3s.univ-cotedazur.fr

Karima BOUDAUD  
I3S / CNRS  
Université Côte d'Azur  
Sophia Antipolis, FR, EU  
karima.boudaoud@i3s.univ-cotedazur.fr

Yves ROUDIER  
I3S / CNRS  
Université Côte d'Azur  
Sophia Antipolis, FR, EU  
yves.roudier@i3s.univ-cotedazur.fr

## Abstract

This article explores the human constraints associated with the use of digital identity wallets, emphasizing the importance of a user-centered approach. We have examined the regulatory constraints, such as the eIDAS regulation, which impact the use cases of digital identity wallets. By developing an innovative solution that constructs documents from identity attributes stored in digital wallets, we aim to identify and address users' privacy concerns and strengthen user control. Our findings provide valuable insights into the future development of digital identity technologies, highlighting the need for security, privacy, and user autonomy.

## CCS Concepts

• **Security and privacy** → *Multi-factor authentication; Access control; Digital rights management; Privacy-preserving protocols; Pseudonymity, anonymity and untraceability; Authorization; Social aspects of security and privacy; Privacy protections; Usability in security and privacy.*

## Keywords

Personal Data Injection, Digital Identity Management, Dynamic Document Security, Access Control Framework, DRM (Digital Rights Management), ODRL (Open Digital Rights Language), Usage Rights Management, Data Privacy Protection, Information Security in Documents, Digital Document Permissions, Integration of Technologies, Cross-Technology Integration, Compliance Standards, Data Protection Laws, Security Compliance, Privacy Compliance

## ACM Reference Format:

Louis RAFFIN, Karima BOUDAUD, and Yves ROUDIER. 2025. User-Centric Challenges in Digital Identity Wallets: Insights from Industry Experimentation. In *Proceedings of Advance 2025 (ADVANCE'2025)*. ACM, New York, NY, USA, 5 pages. <https://doi.org/XXXXXXX.XXXXXXX>

## 1 Introduction

### 1.1 Context of research

Digital identity wallets are essential tools for the secure and private management of identity information. Their importance is growing, especially with the evolving regulations and technologies. These

wallets enable users to store, manage, and share their identity attributes securely, offering advanced functionalities such as electronic signatures and identity proof [13].

Today, there is a significant challenge surrounding digital identity wallets. These technologies are rapidly expanding, but mass adoption cannot occur without considering the users. **This paper aims to address and identify the issues and concerns of users. Our experimentation serves to identify initial user-centric constraints and will form the basis for future work.**

We have developed an innovative solution that constructs a document containing identity attributes sourced from the wallet, while also allowing users to configure permissions for the information they share. Our demo, conducted as a trusted eIDAS electronic signature service provider, aims to exceed the European Commission's expectations, which primarily view the wallet as a means of identity proof or signature. This approach addresses increasing privacy and security concerns.

A user-centered approach is crucial in this context, as it helps understand and meet the real needs of users. By focusing on the human and technical challenges encountered, we can develop solutions that not only comply with regulations but also enhance user experience and build trust in digital identity wallets. The results of our experimentation provide valuable insights for the future development of these technologies, emphasizing the importance of security, privacy, and user autonomy.

### 1.2 A quick reminder about Regulatory Constraints and Definitions

As detailed in our previous article, "Impacts of EU regulations on Digital Identity Wallets," digital identity wallets are essential tools for securely storing and managing identity attributes, enabling users to share their information for identification and authentication purposes. These wallets, which incorporate verifiable credentials (VCs) and verifiable presentations (VPs), are influenced by several European regulations. The GDPR imposes constraints such as data anonymization, maintaining a register of data processing operations, and ensuring data security. The eIDAS regulation establishes a harmonized framework for legally binding identification and introduces the European Digital Identity Wallet (EUDIW) by 2026, reinforcing the role of Qualified Trust Service Providers (QTSP). Unlike other regulations, eIDAS specifically defines trust services and sovereign identity wallets for EU member states, ensuring a high level of security and interoperability for these national digital identity solutions. The NIS2 directive aims

This research project is financed by Docaposte with the support of the Association Nationale de la Recherche et de la Technologie (ANRT).

ADVANCE'2025, Sophia Antipolis, FR

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM  
<https://doi.org/XXXXXXX.XXXXXXX>

to standardize security for network and information systems, enhancing the resilience of critical infrastructures. Additionally, the MiCA regulation provides a framework for crypto-assets, impacting decentralized identities on blockchains and accelerating the adoption of self-sovereign identity. These regulations collectively ensure that digital identity wallets comply with privacy, security, and interoperability standards, addressing the evolving needs of users and the market. Digital identities refer to the information individuals share online, including personal data, social media activity, and online behavior. The eIDAS regulation defines three levels of identity assurance (Low, Substantial, High) and introduces the concept of qualified identity, verified by trusted third parties. The selective disclosure feature of digital identity wallets allows users to limit the disclosure of their data to only necessary attributes, enhancing privacy. Self-sovereign identity (SSI) mechanisms on blockchains offer decentralized identity management, giving individuals complete ownership and control over their identities. These regulations aim to protect personal data, ensure cybersecurity, and support innovation in digital identity technologies.

## 2 Description of our experiment

### 2.1 Origin of the experiment needs and objectives

Our experimentation aims to explore the human constraints associated with digital identity wallets through a user-centered approach. This research is part of a thesis, funded by the industrial sector, and aligns with the standardization efforts of the eIDAS sovereign identity wallet. This approach addresses privacy concerns, the primary objective is to develop an innovative solution that constructs a document containing identity attributes sourced from the wallet, while allowing users to configure permissions for the information they share.

Our solution is particularly innovative as it enables the construction of documents directly from the attributes stored in digital identity wallets. This functionality not only enhances the usability and flexibility of digital identity solutions but also ensures that users have greater control over their personal information. By allowing users to selectively disclose only the necessary attributes, our solution significantly improves privacy and security.

By focusing on the real needs and challenges faced by users, we aim to enhance user experience, build trust, and ensure compliance with regulatory standards. Specifically, our experimentation seeks to address the needs for greater privacy control, user autonomy, and seamless integration of digital identity solutions into everyday activities. Additionally, our research aims to drive industrial innovation by developing cutting-edge technologies that meet the evolving demands of the digital identity landscape. Our findings will provide valuable insights into the future development of digital identity technologies, emphasizing the importance of security, privacy, and user autonomy [12].

### 2.2 Technical overview

Our solution does not yet rely on a signature platform, as we are not aiming to create an industrial product, and integrating such a platform would be overly complex. For practical reasons, we

have implemented a demonstrator based on open-source tools. Furthermore in this initial demonstration, we are not focusing on the signature aspect. Instead, we needed tools to manage document storage, user management, sharing, and permissions. We chose Nextcloud and OnlyOffice for these purposes. Additionally, we use Keycloak with WaltID to simplify identity management and avoid redeveloping this component multiple times.

Our solution allows users to create a contract template and associate it with a DRM file, using ODRL (Open Digital Rights Language) technology to manage digital rights. The user utilizes their digital identity wallet to share personal information and select the permissions they wish to associate with the document. They can choose with whom they want to share the document, and the final document can be constructed in multiple iterations. This approach ensures flexibility, user control, and enhanced privacy, aligning with our goal of addressing the real needs and challenges faced by users in the digital identity landscape.

Our tool communicates with Nextcloud via its REST APIs, ensuring seamless integration for document management. Keycloak, on the other hand, communicates with Nextcloud and our solution using OpenID Connect, facilitating secure and efficient identity management.

Refer to the schema below for a detailed description of the architecture of our solution:

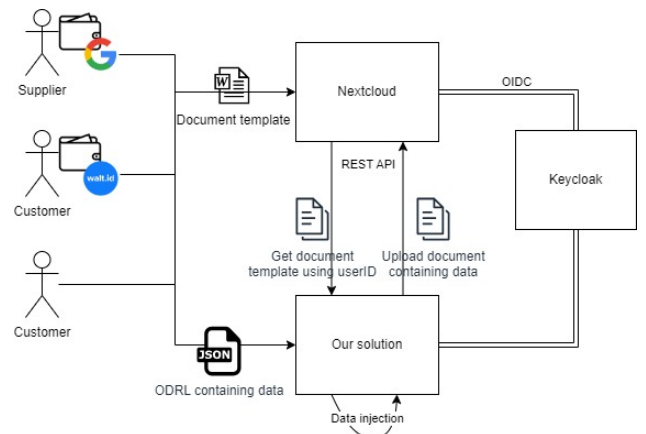
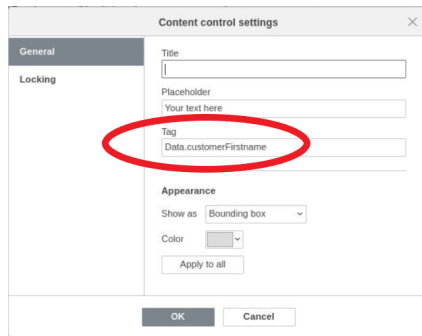


Figure 1: Solution architecture overview

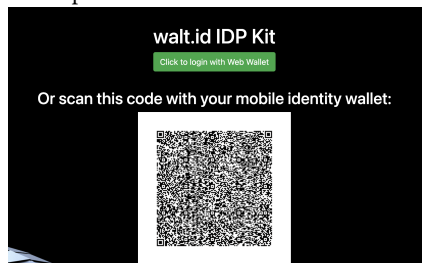
### 2.3 Use case scenario

Imagine a university that needs to facilitate the signing of internship agreements involving three parties: the university, the student, and the company offering the internship.

- (1) The university administrator, Alice, creates a Word document template for the internship agreement, including content control tags that specify the required information, such as the student's name, surname, email address, phone number, and similar details from the company. Alice uploads this template to Nextcloud:



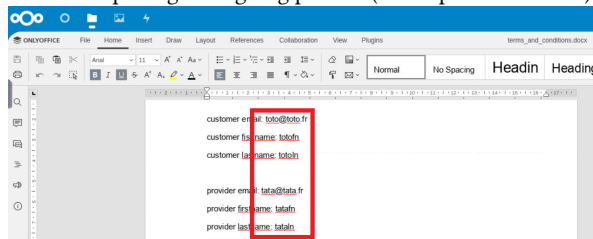
- (2) The student, Bob, and the company representative, Charlie, receive notifications to complete their parts of the agreement. They use their digital identity wallets to share the required personal information, with Keycloak & WaltID managing the identity authentication process. Bob and Charlie use their wallets to present their attributes:



- (3) They select the permissions they wish to associate with the document, ensuring that only necessary information is disclosed and for what purpose:



- (4) The final document is constructed iteratively, with each party contributing their information and reviewing the agreement before completing the signing process (not implemented here):



This approach enhances privacy, user control, and efficiency, demonstrating how our solution leverages digital identity wallets to streamline the signing of agreements. Moreover, our solution can

be applied to various other types of contracts or documents, such as real estate leases, vehicle rentals, and employment contracts, showcasing its versatility and broad applicability.

### 3 Human Constraints and Barriers to Adoption

#### 3.1 Currently identified constraints

These following constraints highlight the need for careful consideration and strategic planning to overcome the obstacles to the effective adoption and use of our digital identity wallet solution. In future work, we will need to confront these identified constraints with actual users to confirm and refine them.

- **Privacy Concerns:** Users may be worried about how their personal data will be used and whether it will be adequately protected. Managing permissions and ensuring transparency about data handling practices are essential to build trust. The process of setting up and using digital identity wallets can be complex and intimidating for some users, leading to lower adoption rates. Simplifying the user experience is crucial to encourage usage.
- **Technical Barriers:** Ensuring that our solution can seamlessly interact with various systems and services is a significant challenge. Standardizing data specifications and exchange norms is necessary to facilitate integration. Additionally, the user experience can vary depending on the digital identity wallet chosen, which may add complexity. As we introduce more granular permissions in the future, configuring these permissions precisely may also become challenging for users.
- **Adoption and Usage:** Users may be resistant to adopting new technologies, preferring traditional methods. Change management efforts and awareness campaigns are necessary to overcome this resistance. Adequate training and continuous support are needed to ensure users are comfortable and proficient with the new technology. This includes understanding wallet functionalities and managing permissions. Furthermore, users might find the complexity of DRM technologies frustrating, such as restrictions on sharing or downloading documents, and the requirement to authenticate to access documents. While these measures enhance privacy, they can be a barrier to adoption.
- **Regulatory Compliance:** Navigating various regulatory requirements, such as GDPR and eIDAS, can be complex and resource-intensive. It is essential to ensure that our solution complies with all security and privacy standards. Depending on the use case, it may be mandatory to archive documents over time and provide users with the ability to delete their data. These requirements, while crucial for regulatory compliance, can add layers of complexity that may be perceived as cumbersome by users.

#### 3.2 Future work and new constraints

To validate our theories and further develop our solution, we propose several future work items. First, we will conduct extensive user testing sessions and surveys to gather feedback on the usability and effectiveness of our solution, analyzing user interactions to identify pain points and areas for improvement. Additionally, we plan to test our solution with various digital identity wallets to ensure compatibility and interoperability, evaluating the user experience across different wallets to identify specific challenges. Developing

more granular permission settings will provide users with finer control over their data, and we will assess the impact of these settings on user experience and adoption rates. Implementing features for long-term data archiving and deletion will ensure compliance with regulatory requirements, while enhanced privacy and security measures will protect user data. Scalability and performance testing will help us optimize our solution to handle large volumes of data and users, identifying any technical constraints and ensuring efficient operation. Continuous monitoring and updating of our solution will ensure compliance with evolving regulatory requirements, and user education and training programs will facilitate adoption. Collaborating with industry partners and standardization bodies will align our solution with emerging standards, and exploring new use cases will demonstrate its versatility.

In our privacy-focused research, we have identified that many policy management technologies are not well-suited for digital identity wallets. Typically, these technologies involve consenting and choosing the information to share, similar to our solution. However, once the data is transmitted, users lose visibility over it. We have identified a need to address this issue using Information Rights Management technologies, digital identity wallets, and blockchain such as [3]. Specifically, we aim to leverage these technologies, including smart contracts, to track shared attributes and verify that permissions are correctly applied.

Also a significant technical constraint closely tied to human factors is the responsibility users have for the security of their data in self-sovereign identity (SSI) systems. In SSI, users store their identity attributes in their digital wallets, potentially on local devices, and must ensure the security of this data. This includes safeguarding against unauthorized access and being cautious about where and with whom they share their information. The risk of phishing attacks and other security threats is likely to increase. Users must be vigilant and adopt best practices to protect their digital identities, highlighting the need for robust security measures and user education to mitigate these risks.

Our choice for the use case for electronic signatures is not solely tied to our industrial solutions. As previously mentioned, there are legal constraints governing this activity, particularly in the definition of trust services under eIDAS. Our future solution must consider the level of signature required and cannot always be 100% zero-knowledge, especially in cases of qualified signatures that require strong signer identification using an eIDAS sovereign wallet and impose strict archiving requirements. One potential solution we are exploring is to move away from traditional certificate-based signatures and instead use Verifiable Credentials provided by the signer. We envision a hybrid solution where only parts of the document are signed, and not all users have access to the entire document. This approach aims to balance the need for strong identification and compliance with privacy and usability considerations.

These efforts will help us refine our solution and address potential constraints, ensuring it meets the evolving needs of users and the market, in the continuity of this work: [5].

## 4 Related work

In preparation for our research, we reviewed several studies related to our work.

The field of electronic signatures has seen significant research, particularly in the areas of identification mechanisms and methods of authentication. Surveys have been conducted to understand the various technologies and their adoption rates [1]. One comprehensive survey explores the different mechanisms of identification used in electronic signatures, highlighting the importance of public key infrastructure (PKI) and digital certificates issued by trusted certificate authorities (CAs) [9] [8]. Another state-of-the-art review delves into the cryptographic techniques employed in digital signatures, such as hashing and asymmetric encryption, which ensure the authenticity and integrity of digital documents [4]. These studies provide a foundational understanding of the technologies that underpin electronic signatures and their role in secure digital transactions.

Research on digital identity wallets is rapidly expanding, reflecting their growing importance in managing personal identity information securely [2] [10]. A systematic literature review (SLR) on digital identity wallets discusses their usage and the various attributes they manage [11]. This review categorizes the features of digital identity wallets and examines their diffusion across different regions. Another study investigates user readiness to adopt digital identity wallets [7], revealing insights into user perceptions and potential barriers to adoption [14]. Additionally, a survey focuses on the choices users make when selecting digital identity wallets, providing valuable data on user preferences and the factors influencing their decisions [6]. These works collectively highlight the critical aspects of digital identity wallets, from technological features to user adoption challenges.

## 5 Conclusion

In this article, we have explored the human constraints and technical challenges associated with the adoption of digital identity wallets, particularly in the context of our innovative solution. Our research aligns with the standardization efforts of the eIDAS sovereign identity wallet and aims to exceed the European Commission's expectations. By developing a solution that constructs documents from identity attributes stored in digital wallets and allows users to configure permissions, we address critical privacy concerns and enhance user control, proposing a new use case with identity wallets.

Our experimentation has highlighted several key barriers, including privacy concerns, technical complexities, user adoption resistance, and regulatory compliance challenges. These constraints underscore the need for careful consideration and strategic planning to ensure the effective adoption and use of digital identity wallet solutions. Future work will involve confronting these identified constraints with actual users to confirm and refine our findings, ensuring that our solution meets the evolving needs of the digital identity landscape.

The insights gained from this research provide valuable guidance for the future development of digital identity technologies, emphasizing the importance of security, privacy, and user autonomy. By addressing the real needs and challenges faced by users, we aim to build trust and drive industrial innovation in the field of digital identity management.



## Acknowledgments

We would like to express our gratitude to M. Jérémie BLANC (Docaposte) for his expertise, insights, feedback and invaluable support and guidance throughout this research work.

The authors used AI tools to revise the text and to correct any typos, grammatical errors, and awkward phrasing.

## References

- [1] Aki. 1983. Digital signatures: A tutorial survey. *Computer* 16, 2 (Feb. 1983), 15–24. doi:10.1109/MC.1983.1654294 Conference Name: Computer.
- [2] Ali Abdullah S. AlQahtani, Zakaria El-Awadi, and Manki Min. 2021. A Survey on User Authentication Factors. In *2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. 0323–0328. doi:10.1109/IEMCON53756.2021.9623159 ISSN: 2644-3163.
- [3] Eranga Bandara, Xueping Liang, Peter Foytik, Sachin Shetty, and Kasun De Zoysa. 2021. A Blockchain and Self-Sovereign Identity Empowered Digital Identity Platform. In *2021 International Conference on Computer Communications and Networks (ICCCN)*. 1–7. doi:10.1109/ICCCN52240.2021.9522184 ISSN: 2637-9430.
- [4] Harmandeep Kaur and Munish Kumar. 2023. Signature identification and verification techniques: state-of-the-art work. *Journal of Ambient Intelligence and Humanized Computing* 14, 2 (Feb. 2023), 1027–1045. doi:10.1007/s12652-021-03356-w
- [5] Maina Korir, Simon Parkin, and Paul Dunphy. 2022. An Empirical Study of a Decentralized Identity Wallet: Usability, Security, and Perspectives on User Control. 195–211. <https://www.usenix.org/conference/soups2022/presentation/korir>
- [6] Sandra Kostic. 2024. Who is the Better Operator of an Identity Wallet Prioritised by the User? - A Quantitative Survey Between State and Company. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, 1–7. doi:10.1145/3613905.3647961
- [7] Sandra Kostic, Fraunhofer Aisec, Maija Poikela, and Fraunhofer Aisec. 2022. Do Users Want To Use Digital Identities? A Study Of A Concept Of An Identity Wallet. (2022).
- [8] Ashish Kumar, Rahul Saha, Mauro Conti, Gulshan Kumar, William J. Buchanan, and Tai Hoon Kim. 2022. A comprehensive survey of authentication methods in Internet-of-Things and its conjunctions. *Journal of Network and Computer Applications* 204 (Aug. 2022), 103414. doi:10.1016/j.jnca.2022.103414
- [9] Qais Abd Al-Salam Mohammed, Mohammed Joudah, and Hamed Mohammed. 2024. A survey on digital signature schemes. *AIP Conference Proceedings* 3232, 1 (Oct. 2024), 020057. doi:10.1063/5.0236576
- [10] Doruntina Murtezaj. 2023. Unlocking Digital Trust: A Study of User Trust and Usability in a Digital Identity Wallet Concept. (Nov. 2023).
- [11] Blaž Podgorelec, Lukas Alber, and Thomas Zefferer. 2022. What is a (Digital) Identity Wallet? A Systematic Literature Review. In *2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)*. 809–818. doi:10.1109/COMPSAC54236.2022.00131 ISSN: 0730-3157.
- [12] Alexander Rieger, Tamara Roth, Johannes Sedlmeir, Linda Weigl, and Gilbert Fridgen. 2022. Not yet another digital identity. *Nature Human Behaviour* 6, 1 (Jan. 2022), 3–3. doi:10.1038/s41562-021-01243-0 Publisher: Nature Publishing Group.
- [13] Johannes Sedlmeir, Reilly Smethurst, Alexander Rieger, and Gilbert Fridgen. 2021. Digital Identities and Verifiable Credentials. *Business & Information Systems Engineering* 63, 5 (Oct. 2021), 603–613. doi:10.1007/s12599-021-00722-y
- [14] Fatemeh Stodt and Christoph Reich. 2023. A Review on Digital Wallets and Federated Service for Future of Cloud Services Identity Management. (2023).

# Next Step Towards Intent-based Scheduling for Performance and Security in Edge-to-Cloud Networks

Karim Boutih

*Computer and Network Systems Msc  
University of Evry Paris-Saclay  
Evry, France  
karim.boutih@etud.univ-evry.fr*

Mazigh Djelouah

*Computer and Network Systems Msc  
University of Evry Paris Saclay  
Evry, France  
mazigh.djelouah@etud.univ-evry.fr*

Nazim Agoulmine

*IBISC Laboratory  
University of Evry Paris Saclay  
Evry, France  
nazim.agoulmine@univ-evry.fr*

**Abstract**—This paper presents a framework for intent-based scheduling in Kubernetes (K8s) by advancing the design of a dynamic Segmentation Custom Resource Definition (CRD). The focus is on the interplay between security and performance in multi-tenant environments, proposing a framework for intent-driven scheduling without requiring manual node-level management. By establishing the architectural foundation for declarative segmentation, this research sets the stage for future implementation and validation of this concept in the Kubernetes environment.

**Index Terms**—Kubernetes, Security, Cloud Computing, Containers, Intent-based Scheduling

## I. INTRODUCTION

The introduction of 5G has opened up new possibilities for ultra-reliable low-latency communication applications with stringent performance requirements, such as Virtual/Augmented Reality (VR/AR), autonomous vehicles, and the Internet of Things (IoT). These applications demand ultra-high reliability and low latency, necessitating the adoption of edge computing solutions in mobile networks to provide the required computational and communication resources. Cloud native computing models, which leverage lightweight, portable containers and orchestration frameworks such as Kubernetes (K8s), offer greater efficiency than traditional, resource-heavy virtual machine-based solutions. However, adopting these models introduces additional complexities, particularly in edge computing environments that support 5G networks.

In telco-managed clusters, where the infrastructure for telecommunication services is hosted and managed, a strict separation of responsibilities exists between Cluster Administrators (CAs) and application administrators (AAs). CAs oversee the physical or virtual machines that form the cluster, focusing on resource allocation, hardware utilization, tenant isolation, and security. In contrast, AAs manage Kubernetes Pods — the smallest deployable units containing containers—prioritizing workload performance and resource efficiency. This division of duties often results in complex configuration inconsistencies between multiple layers of the system. For example, while application administrators may

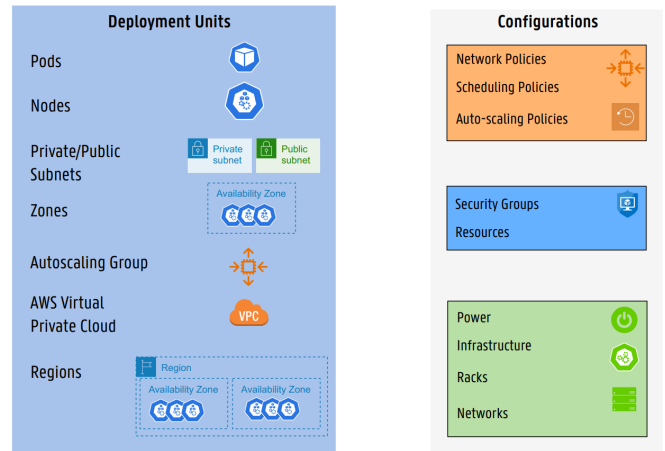


Fig. 1. The layered relationship in a Kubernetes cluster hosted on Amazon Web Services.

define network policies to optimize Pod-level performance, CAs prioritize security and tenant isolation at the VM or node level. Misalignment between these layers can lead to conflicts, such as network policies at the Pod level that conflict with security group rules at the VM level, causing performance degradation and security vulnerabilities [1]. Fig. 1 [1] illustrates this layered relationship in a Kubernetes cluster hosted on Amazon Web Services (AWS), highlighting the coordination required between deployment units and network configurations.

Reconciling the conflicting priorities of CAs and application administrators while ensuring interlayer coordination is a significant challenge in the deployment and management of Kubernetes clusters and represents the objectives of this paper. The goal of this paper is therefore to advance the contributions of authors presented in [1][5] by fully designing and implementing Segmentation CRD and enhancing dynamic enforcement of segmentation policies in Kubernetes. Although the original study introduced the concept of intent-based scheduling for multi-tenant environments, it stopped short of



automating the segmentation process.

## II. DEFINITIONS

### A. Intent-Based Scheduling

Intent-based scheduling is a paradigm that shifts focus from manual resource allocation to high-level declarations of desired outcomes. It allows orchestration platforms to dynamically adapt their behavior to align with goals, reducing complexity and enhancing scalability in multi-tenant environments. Modern cloud platforms have embraced microservice architectures, enhancing scalability, resilience, and development agility. Kubernetes is the de facto standard for container orchestration, providing a comprehensive suite of tools for deploying, scaling, and managing microservice-based applications.

### B. Microservices in Modern Cloud Applications

Modern cloud platforms use microservice architectures for scalability, resilience, and development agility, but robust orchestration frameworks are needed for efficient deployment, scaling, and communication.

### C. Kubernetes Overview

Kubernetes is the standard for container orchestration, offering tools for deploying, scaling, and managing microservice-based applications, enabling seamless communication and automation in dynamic and distributed infrastructure [2].

### D. Kubernetes Cluster Architecture

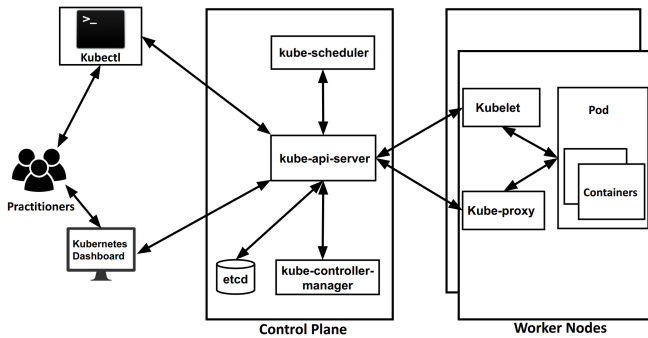


Fig. 2. The Kubernetes Cluster Architecture.

A Kubernetes cluster [3] is composed of two primary types of nodes: control plane nodes and worker nodes. These nodes collaborate to deliver the core functionality of Kubernetes, from resource orchestration to efficient workload management.

The control plane serves as the brain of the cluster, making global decisions and managing overall operations. It comprises several critical components that are: (1) Kube-API Server, The central access point for the Kubernetes cluster, (2) Etcd: A distributed, fault-tolerant key-value store that holds the cluster's configuration data and state information, (3) Kube-Controller-Manager: A monitor point of the cluster's state. It is responsible for maintaining the desired configuration of the cluster by managing controllers, (4) Kube-Scheduler: that

determines the optimal placement of Pods on worker nodes based on resource availability and defined constraints.

The Worker nodes are responsible for running the applications managed by Kubernetes. Each worker node contains the following components: (1) Kube-Proxy: that manages network communication, ensuring proper routing of traffic between Pods and services across nodes, (2) Kubelet: An agent that ensures containers within Pods are running as intended, maintaining communication with the control plane for status updates and lifecycle management and (3) Pod: The smallest deployable unit in Kubernetes. Pods encapsulate one or more tightly coupled containers, as well as shared storage and network resources. Containers within a Pod share the same network namespace, allowing seamless communication within it.

The Kubernetes scheduler is crucial for managing Pod placement within a cluster, affecting application performance and security. It evaluates nodes and assigns Pods based on resource availability, constraints, and workload requirements. However, it cannot address complex issues like interlayer misconfigurations or advanced network security vulnerabilities.

## III. RELATED WORKS

In this research, we have thoroughly explored various sources of academic literature, utilizing platforms such as IEEE Xplore, ResearchGate, and arXiv to gather relevant articles. Among the many articles we reviewed, the majority came from IEEE Xplore, covering a diverse range of topics in the fields of Kubernetes security and Intent-Based Networking (IBN). In particular, we identified several studies on Kubernetes security [8][9] and others focused on IBN [10][11]. Furthermore, we found 3 articles exploring the intersection of Kubernetes and IBN [1][5].

### A. Related Work on Network-Aware Scheduling and Diktyo

The work [7] reveals the limitations of current network-aware scheduling techniques, highlighting the need for a novel solution. Traditional methods, like ILP, are computationally intensive and unsuitable for real-time Kubernetes clusters. Diktyo, an open-source tool, integrates network topology and microservice dependencies, allowing efficient, dynamic placement of latency-sensitive and bandwidth-intensive microservices across Kubernetes clusters.

### B. Related Work on Intent-Based Scheduling and Segmentation

The enhanced version of Diktyo, integrated with GrassHopper (GH), extends beyond performance optimization to address security-oriented scheduling, as discussed in [1]. Existing microsegmentation platforms, such as Cisco ACI [4] and VMware NSX [6], use fine-grained firewall rules to limit lateral movement within containerized environments. However, these platforms face key limitations:

- **Fragmented Policies:** VM-level and container-level security policies operate independently, leading to misconfigurations and administrative overhead.

- **Manual Conflict Resolution:** Reconciliation of security policies across domains relies on API-driven checks, which are error-prone and tedious.
- **Delays in Policy Convergence:** Sequential enforcement steps, like those in Illumio's platform [7], introduce delays that can significantly hinder Pod readiness and performance.

Security-oriented schedulers have been proposed to address isolated aspects of container security. For example, syscall-aware schedulers reduce attack surfaces by limiting unnecessary system calls. However, such approaches lack a holistic view, failing to dynamically adapt scheduling decisions to align with both security policies and network topology.

The Diktyo and GrassHopper integrate performance metrics with dynamic security enforcement, reducing administrative complexity and aligning with security objectives through intent-based scheduling, though implementation is incomplete.

This work aims to close this gap by fully realizing the Segmentation CRD, enabling dynamic segmentation, and extending Diktyo's capabilities to deliver a truly intent-driven scheduling framework.

### C. Synthesis

After careful consideration of all discovered related works, we found that articles [1] and [7] are the most appropriate for our research, as they are compelling treatments of emerging challenges and practical solutions and align with the objectives of our research work. In the following section, we will dive into the presented work and solutions and provide a critical analysis of the proposed methodologies, with particular emphasis on the identified challenges.

## IV. ADVANCED ANALYSIS OF MOST RELEVANT CONTRIBUTIONS

### A. Analysis of article [7]

Article [7] introduces Diktyo, a network-aware scheduling framework for Kubernetes, designed to optimize pod placement based on both microservice dependencies and the underlying network topology. The challenges are threefold: (1) lack of network awareness in K8s: The default scheduler does not account for network latency or bandwidth during scheduling decisions, (2) lack of latency awareness: Applications with tightly coupled microservices, such as distributed databases and web services, suffer from high latency when placed on distant nodes, (3) lack of awareness about the dynamicity of the environment: Scalable scheduling in large-scale, distributed clusters remains a significant challenge in Kubernetes.

The objectives of the presented Diktyo framework are the following: (1) minimize end-to-end application latency by optimizing pod placement based on network performance metrics. (2) maximize bandwidth utilization for data-intensive applications to ensure efficient communication. (3) enable scalability through a modular framework that adapts to changes in the state of the cluster.

1) **Diktyo Framework:** As highlighted in Fig. 3 [7], Diktyo integrates with Kubernetes using a plugin-based architecture, enhancing the scheduling process with three key plugins:

- **TopologicalSort Plugin:** Orders pods based on dependency graphs to prioritize scheduling of latency-sensitive pods.
- **NodeNetworkCostFit Plugin:** Filters nodes that do not meet the bandwidth or latency constraints defined in the AG and NT CRDs.
- **NetworkMinCost Plugin:** Scores nodes based on aggregated network costs, favoring placements that minimize latency and maximize bandwidth for critical pod dependencies.

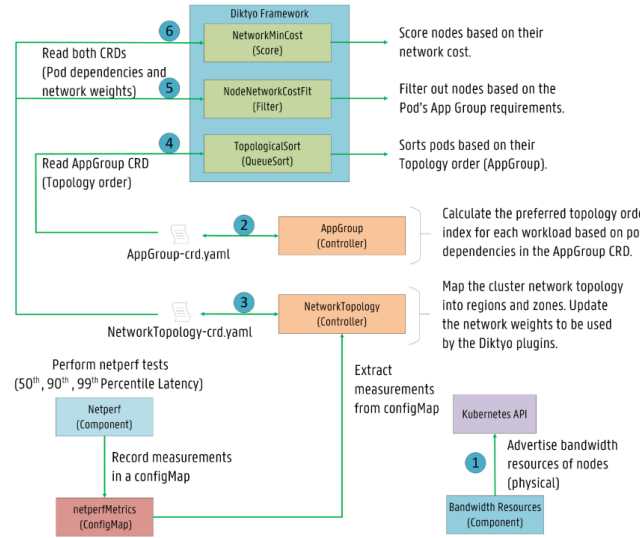


Fig. 3. Illustration of the Diktyo framework and the Kubernetes scheduling workflow.

Diktyo framework also introduces two Custom Resource Definitions (CRDs) to model application dependencies and the network state of the cluster. These two CRDs (AppGroup and NetworkTopology) are introduced in the following paragraphs:

#### a) AppGroup (AG) CRD:

- **Dependency Modeling:** Captures the logical relationships between microservices, specifying latency and bandwidth requirements between dependent pods.
- **Declarative Constraints:**
  - **MinBandwidth:** Sets a minimum threshold for bandwidth between communicating pods.
  - **MaxNetworkCost:** Specifies the upper limit of acceptable network latency.
- **Prioritized Scheduling:** Incorporates dependency graphs and sorting algorithms to ensure that pods with tighter constraints are scheduled first, preventing bottlenecks.

#### b) NetworkTopology (NT) CRD:

- **Cluster Network Representation:** Maps the topology of the Kubernetes cluster, capturing metrics such as latency and bandwidth across regions and zones.



ration efforts. It also highlighted the necessity of segmentation-aware scheduling to reconcile performance and security requirements in multi-tenant Kubernetes environments.

## V. ADVANCING THE INTENT-BASED SCHEDULING VISION

In this paper, we introduce an intent-based interface for Cluster Administrators (CAs) to allow them to define high-level segmentation requirements without managing low-level configurations. To achieve this, we propose novel mechanisms to enable dynamic segmentation by translating high-level intents into real-time segmentation actions to optimize performance and security in Kubernetes clusters.

The solution will provide the following features:

- **High-Level Intent Specifications:** through a web interface, CAs can define intents in terms of security isolation objectives, low-latency communication objectives, and optimal resource allocation. It also allows for the automatic translation of these intents into actionable configurations in Kubernetes cluster.
- **Dynamic Network Segmentation:** The proposed solution allows for dynamic network configuration, such as dynamically adapt network segments based on real-time metrics (e.g., latency, bandwidth). It uses Segmentation Custom Resource Definition (CRD) to manage these segments.
- **Automated Pod Placement:** With the use of Diktyo, our solution aims to schedule the placement of Pods according to the defined intents taking into account the topology of the network topology and the security constraints.
- **Security Enforcement:** Integrate Grasshopper for dynamic security policy enforcement, ensuring tenant isolation and least-privilege access.

### A. Key Components of the System

1) **Intent-Based Interface:** As previously introduced, this interface allows CAs to specify their intents to the infrastructure. Intents are related to workload isolation to ensure security and/or to pod placements to ensure low-latency communication and, therefore, performance. The intents specified through the web interface are then automatically translated into configuration actions in the Kubernetes environment (e.g., segment creation, pod scheduling).

2) **Dynamic Segmentation via Segmentation CRD:** The fulfillment of intents is achieved through continuous monitoring of the infrastructure to adjust network segments in real time based on workload needs and performance metrics. In addition, nodes are dynamically assigned to segments following specified intents. Finally, segments are updated on the basis of the evolving requirements such as new workloads or new introduced tenants.

3) **Interface with Diktyo Scheduler to enforce performance:** Our solution interfaces with Diktyo Schedule through the Diktyo API. Our solution allows for filtering and scoring the worker nodes against the metrics specified in the intents (e.g., bandwidth, latency). It will then ask Diktyo to perform

the placement, ensuring tenant isolation and compliance with performance.

4) **Interface with Grasshopper to enforce Security:** Our solution interfaces with Grasshopper to enforce the required network isolation and communication rules based on defined intents. It dynamically updates security policies with segmentation changes.

## VI. CONCLUSION

This paper introduces a scheduling interface for Kubernetes, allowing Cluster Administrators to define intents for performance optimization and security enforcement. The underlying system automates Pod placement and aligns it with network performance metrics and security as defined in the intents, and reconciles application-level performance demands with cluster administrator (CA) priorities. This process is aimed to be transparent to the Cluster Administrator as well as to the end-users. The next step of this work is to implement the system and test it in terms of scalability and conflict resolution between intents.

## REFERENCES

- [1] J. Santos, E. Truyen, C. Baumann, F. De Turck, G. Budigiri, and W. Joosen, "Towards intent-based scheduling for performance and security in edge-to-cloud networks," in *Proc. 27th Int. Conf. Innov. Clouds, Internet, Netw. (ICIN)*, pp. 222–229, IEEE, 2024.
- [2] G. Budigiri, C. Baumann, J. T. Mühlberg, E. Truyen, and W. Joosen, "Network policies in Kubernetes: Performance evaluation and security analysis," in *Proc. Joint Eur. Conf. Netw. Commun. & 6G Summit (EuCNC/6G Summit)*, pp. 407–416, IEEE, 2021.
- [3] M. S. I. Shamim, "Mitigating security attacks in Kubernetes manifests for security best practices violation," in *Proc. 29th ACM Joint Eur. Softw. Eng. Conf. Symp. Found. Softw. Eng. (ESEC/FSE)*, pp. 2, ACM, Aug. 2021.
- [4] Cisco, "Cisco ACI and Kubernetes integration," [Online]. Available: [https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b\\_Kubernetes\\_Integration\\_with\\_ACI.html](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b_Kubernetes_Integration_with_ACI.html).
- [5] J. Santos, C. Wang, T. Wauters, and F. De Turck, "Diktyo: Network-aware scheduling in container-based clouds," *IEEE Trans. Netw. Serv. Manag.*, vol. 20, no. 4, pp. 4461–4475, Dec. 2023.
- [6] VMware, "NSX Container Plugin for Kubernetes and Tanzu Application Service - Installation and Administration Guide," [Online]. Available: [https://docs.vmware.com/en/VMware-NSX-Container-Plugin/4.0/ncp\\_40\\_kubernetes.pdf](https://docs.vmware.com/en/VMware-NSX-Container-Plugin/4.0/ncp_40_kubernetes.pdf).
- [7] Illumio, "Illumio Core for Kubernetes and OpenShift," [Online]. Available: [https://docs.illumio.com/core/22.4/Content/Resources/PDF/Illumio\\_Core\\_for\\_Kubernetes\\_and\\_OpenShift\\_21.5.18.pdf](https://docs.illumio.com/core/22.4/Content/Resources/PDF/Illumio_Core_for_Kubernetes_and_OpenShift_21.5.18.pdf).
- [8] D. Brighenti, R. Sisto, and F. Valenza, "Security automation for multi-cluster orchestration in Kubernetes," in *Proc. IEEE NetSoft 2023 - 5th Int. Workshop Cyber-Security Threats, Trust, Privacy Manage. Emerging Software-defined Virtualized Infrastructures*, pp. 480–487, IEEE, 2023.
- [9] Md. S. I. Shamim, F. A. Bhuiyan, and A. Rahman, "XI commandments of Kubernetes security: A systematization of knowledge," in *Proc. IEEE SecDev 2020*, pp. 21–28, IEEE, 2020.
- [10] V. Gowtham, F. Schreiner, M.-I. Corici, and T. Magedanz, "NEMI: A standardized approach to intent-based networking," in *Proc. IEEE Conf. Standards Commun. Netw. (CSCN)*, pp. 357–362, IEEE, 2023.
- [11] A. Rafiq, A. Mehmood, and W.-C. Song, "Intent-based slicing between containers in SDN overlay network," *J. Commun.*, vol. 15, no. 3, pp. 237–244, Mar. 2020.

---

# Modeling the Complexity of the Georgian Verbal System through Linked Data and Machine Learning

Elizbarashvili Archil  
Tbilisi State University  
Tbilisi, Georgia  
archil.elizbarashvili@tsu.ge

Ducassé Mireille  
Université de Rennes, IRISA  
Rennes, France  
mireille.ducasse@irisa.fr

Tsintsadze Magda  
Tbilisi State University  
Tbilisi, Georgia  
magda.tsintsadze@tsu.ge

Khachidze Manana  
Tbilisi State University  
Tbilisi, Georgia  
manana.khachidze@tsu.ge

## Abstract

The Georgian verbal system poses significant challenges due to its agglutinative and inflectional nature, lack of infinitives, rich morphological structure, and irregular patterns. To assist Georgian language learners and support lexicographic work, we are developing KARTUVERBS, a comprehensive linked-data platform for Georgian verb forms. This system integrates structured lexical data, semantic web technologies, and machine learning to infer missing lemmas (verbal nouns). Current results show a prediction accuracy of 98% using decision trees, with cross-validation against the Georgian National Corpus. This paper outlines our methodology, the unique linguistic challenges of Georgian, and future directions for integrating crowdsourcing and deep learning into the tool's development.

## Keywords

Linked data, machine learning, Georgian verbs, lexicography, verbal noun, decision tree

## 1 Introduction

The Georgian language's verbal morphology is notably complex: it is agglutinative, inflectional, and lacks a traditional infinitive form. Dictionaries typically use verbal nouns, third-person singular present or future forms, or roots as lemma entries. This ambiguity poses significant challenges for learners and computational models. For example, the form *mekirava* (მეჭირავა, "I were renting") can correspond to the verbal noun *kiraoba* (ჭირაობა), the present third person singular *kiraobs* (ჭირაობს), or simply the root *kira* (ჭირა).

Existing Georgian linguistic resources and dictionaries, such as the Comprehensive Georgian-English Dictionary Rayfield et al. [10] and Georgian-German dictionary Tschenkeli et al. [11], provide valuable but limited help for learners because of their varying lemma strategies and incomplete morphological information. Digital resources, such as the Georgian National Corpus and various online translation platforms, offer broader access but still fall short in consistently addressing lemma disambiguation and providing extensive morphological details.

To bridge these gaps, we are developing KARTUVERBS, an advanced linked-data-based semantic database containing over 5 million Georgian verb forms. This database leverages semantic web technology

and supervised machine learning to accurately predict and reconstruct verbal nouns and facilitate intuitive navigation among inflected forms and their morphological characteristics.

## 2 Literature Review

Research on Georgian verb morphology underscores its complexity and the challenges faced in lexicographic representation. Georgian verbs exhibit unique morphological features, such as polypersonal agreement, version markers, and preverbs, which significantly alter meaning and grammatical relationships Cherchi [2]; Tutit [12]; Makharobidze [7]. The absence of an infinitive form further complicates dictionary compilation and learner comprehension Gippert [5]; Margalitadze [8].

Previous works, like those by Daraselia and Sharoff [3], initiated corpus-based lexicographic methodologies but lacked robust morphological linkage and lemma provision. Clarino INESS provides structured Georgian morphological data but primarily targets linguistic researchers rather than language learners or computational lexicography Meurer [9].

Advancements in linked open data (LOD) methodologies as crucial for future lexicographic developments and semantic web technologies have been suggested by Gracia et al. [6] and Ferré [4], enabling interoperability and user-friendly data navigation. Additionally, supervised machine learning methods, particularly decision trees, have shown promise in linguistic data prediction tasks due to their low complexity and robustness in handling missing values in multiclass classification problems, Bansal et al. [1].

KARTUVERBS integrates these methodological advancements into a comprehensive and learner-focused platform. The proposed system not only enhances dictionary usability through automatic lemma inference but also sets a foundation for future computational linguistic developments in Georgian and other morphologically rich languages.

## 3 Linguistic Peculiarities of Georgian Verbs

The Georgian verbal system represents one of the most intricate morphological architectures among world languages, characterized by numerous linguistic peculiarities that significantly complicate computational modelling, lexicography, and language learning.



### 3.1 Complex Tense and Aspect System

Unlike Indo-European languages, Georgian lacks a conventional infinitive form. This absence poses considerable challenges for dictionary compilation and language learners, as verbs are instead represented through multiple possible entry points such as verbal nouns, third-person singular forms, or roots. For instance, the English infinitive "to write" is typically presented as a verbal noun *ts'era* (წერა, literally "writing") or through third-person singular present forms such as *ts'ers* (წერს, "he/she writes"). Each choice complicates verb lookup and comprehension due to the considerable morphological differences between lemmas and their conjugated forms.

Georgian verbs exhibit polypersonal agreement, simultaneously encoding the grammatical roles of the subject, direct object, and frequently the indirect object. This polypersonal characteristic significantly expands conjugational possibilities. For example, the single verb form *mogts'era* (მოგწერა, "he/she wrote it to you") encapsulates agreement with three grammatical persons simultaneously, encoding subject, direct object, and indirect object within a compact morphological form. Georgian verbs follow a complex system of tense-aspect categories distributed across three morphologically distinct series:

Series I (Present and Future): Represents ongoing, habitual, or future actions (e.g., *ts'er* წერ, "you write" / *dats'er* დაწერ, "you will write").

Series II (Aorist): Denotes completed actions in the past with no current relevance, often associated with perfective aspects (e.g., *dats'ere* დაწერე, "you wrote").

Series III (Perfect): Marks actions completed in the past but with present relevance, similar to English perfect tenses (e.g., *dagits'eria* დაგიწერია, "you have written (at some point)").

Each series introduces specific morphological changes affecting verb roots, preverbs, and endings, significantly increasing morphological complexity.

### 3.2 Preverbs and Version Markers

Georgian preverbs and version markers alter verb meanings substantially, introducing semantic nuances related to motion, directionality, benefaction, and other semantic dimensions. Preverbs such as *gada-* (გადა-, "across, again"), *mo-* (მო-, "towards the speaker"), and *ts'a-* (წა-, "away from the speaker") interact with verb stems to produce new lexical meanings and aspectual distinctions. Similarly, version markers (e.g., subjective, objective, and neutral) encode nuances concerning participants' roles and interests within an event, adding complexity to verb conjugation and interpretation.

### 3.3 Verbs

Georgian verb morphology includes numerous irregular forms and instances of suppletion (where different morphological forms are entirely unrelated). Common verbs such as *q'opna* (ყოყნა, "to be") demonstrate extensive irregularity across tense and person conjugations, significantly complicating both learning and computational modeling.

Verbs are traditionally categorized into morphological classes based on conjugation patterns (e.g., Class I: transitive, active verbs;

Class II: intransitive, stative verbs; Class III: mediopassive verbs; and Class IV: indirect verbs). Each class follows distinct morphological rules, further complicating systematization and lemma generation.

Georgian morphology is highly agglutinative, frequently resulting in verb forms composed of multiple morphological units (preverbs, root, version markers, thematic suffixes, mood-tense markers, and person-number markers). This extensive concatenation of morphemes within verb forms leads to extremely high form variety and necessitates meticulous morphological parsing for computational modeling.

Understanding these linguistic peculiarities is crucial for computational and lexicographic projects, such as KARTUVERBS, as they directly impact the design of databases, lemma prediction algorithms, and user interfaces. By explicitly modeling and accommodating these complexities, our approach aims to provide a robust computational tool capable of supporting both learners and researchers in navigating the rich and challenging Georgian verbal system.

## 4 Methodology

To effectively manage and utilize the complex morphological data inherent to Georgian verbs, our approach combines structured data extraction, transformation into linked data, and supervised machine learning. As illustrated by Figure 1, the process to transfer data from Clarino to KARTUVERBS consists of three main blocks. The first stage extracts data from Clarino web pages and generates an intermediate CSV file. The second stage applies several normalization and enrichment processes to improve the quality and usability of the data, addressing linguistic and structural inconsistencies, and includes machine learning techniques to reconstruct missing lemmas. The final stage transforms the cleaned CSV data into RDF triples and deploys a SPARQL endpoint to support semantic querying and interoperability.

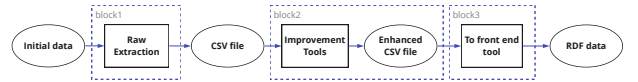


Figure 1: Global structure of the transformation process

### 4.1 Data Transformation Pipeline

The KARTUVERBS database relies initially on morphological data obtained from Clarino, an online resource primarily targeting linguistic researchers. To convert this linguistic data into a structured, semantic linked-data resource tailored to learners and computational applications, we implemented a multi-stage data transformation pipeline detailed as follows:

The original data, available as web pages from the Clarino INESS repository, was scraped and parsed using script. Due to Clarino's hierarchical data structure (organized by common roots, verbs, and inflected forms), we first converted the scraped JSON data into a flat CSV format, resulting in a dataset of approximately 22 million verb form entries.

The raw JSON data extracted was initially complex and included redundant hierarchical structures. We flattened these structures by

extracting each inflected form along with associated morphological fields such as tense, person, number, root, preverb, and morphological markers. The dataset was filtered to retain only 14 key morphological attributes crucial for our lexicographic goals. Each verb form was represented by a single, easily manageable record, facilitating later transformations and analyses.

The Clarino dataset utilized linguistic shorthand and symbolic notations, necessitating normalization and enrichment. Special Python scripts were developed to decode symbolic notations (such as Latin placeholders indicating optional or obligatory morphemes) and reconstruct full verbal nouns by combining root lemmas and preverbs. For example, Clarino’s shorthand verbal noun “*\*adami-aneba*” (\*ადამიანება) was programmatically expanded to the fully reconstructed verbal noun “*gaadamianebe*” (გაადამიანება, “to humanize somebody”).

This normalization step included automatic handling of variant morphemes, removal of redundant alternatives, and the correction of encoding irregularities present in Clarino’s linguistic annotation.

To enable semantic querying and interoperability, the normalized CSV dataset was transformed into RDF (Resource Description Framework) triples. Each CSV entry was converted into multiple RDF triples using scripts designed specifically for this transformation. The resulting RDF dataset (in Turtle N-Triples format) exceeded 80 million links, enabling rich semantic querying capabilities through SPARQL endpoints.

An Apache Jena Fuseki server was employed to manage RDF indexing, supporting efficient real-time queries through the Sparklis<sup>1</sup> linked-data frontend. This allowed learners and lexicographers to intuitively navigate between verb forms, lemmas, and their morphological details without requiring technical query syntax knowledge. In addition to the integrated KARTUVERBS endpoint, users can specify a remote SPARQL endpoint address to query external resources. Having a locally running endpoint is particularly convenient when working with frequent minor revisions or experimental dataset versions.

## 4.2 Machine Learning for Missing Lemmas

Although extensive, the Clarino-derived dataset initially contained approximately 600,000 entries lacking essential verbal noun lemmas, posing a significant challenge. To address this, we employed supervised machine learning to predict and reconstruct missing lemmas from morphological features present in other dataset entries.

We utilized a subset of around 300,000 fully annotated verb forms to train a supervised decision tree classifier. Fourteen carefully selected morphological features, such as tense, person, number, preverb, stem formants, and morphological class, were extracted and encoded numerically. To represent Georgian textual features numerically, we devised an encoding approach based on UTF-8 character representation, ensuring that each verbal noun uniquely mapped to a distinct numerical identifier, essential for accurate classification and later decoding.

A decision tree classifier was selected due to its robustness in handling multi-class classification tasks and inherent interpretability, which significantly aids linguistic validation. Decision trees efficiently manage datasets with missing values, an essential requirement given the irregular nature of morphological data. The dataset was randomly split into training (80%) and testing (20%) subsets, a standard practice to evaluate generalization performance reliably.

The trained decision tree model exhibited an overall classification accuracy ranging from 98% to 99% in reconstructing missing verbal noun lemmas. Precision, recall, and F1-score metrics further confirmed robust model performance. This high accuracy suggests that morphological properties encoded within Georgian verb forms provide sufficient predictive power for automatic lemma reconstruction.

Individual class performance varied, reflecting inherent dataset imbalance and morphological irregularities specific to certain verbal noun groups. Decision trees can sometimes overfit or struggle with specific patterns in the data. To address this, we experimented with another robust model, Support Vector Machine (SVM). However, SVM yielded significantly worse performance on our dataset.

To further validate our predictions, we have cross-checked reconstructed lemmas against lemmas provided by the Georgian National Corpus (GNC). Approximately 92% of our predicted lemmas matched GNC lemmas, confirming their reliability. A minor discrepancy (approximately 6.5%) pointed toward either genuine linguistic variability or minor errors, motivating further expert-driven verification and crowdsourcing initiatives.

To address residual ambiguities and discrepancies identified in cross-validation, we initiated a crowdsourcing campaign through the Headwork<sup>2</sup>, an open-source academic platform for the crowdsourcing of complex tasks. Experts and native Georgian speakers can collaboratively review and validate verb lemma predictions flagged as uncertain by our automated verification process. Contributors evaluate morphological consistency across multiple verb forms and make linguistic judgments on predicted verbal nouns, enabling continual improvement of dataset accuracy.

## 5 Conclusion

KARTUVERBS leverages semantic web and machine learning to model the Georgian verbal system. It contains approximately 5 million inflected forms associated with over 16,000 verbs, featuring more than 80 million links. For learners, it serves as a digital lexicographic tool that identifies verbal nouns from any inflected form and provides an API for translation through electronic dictionaries or machine translation applications, such as Google Translate. For Georgian lexicographers, KARTUVERBS offers a comprehensive digital system that consolidates nearly all Georgian verbs and their inflected forms, supporting both pedagogical and scholarly endeavors given the language’s intricate and exception-rich verb system. For lexicographers of other languages, KARTUVERBS demonstrates how Semantic Web technologies combined with linked data, can create a machine-readable, interoperable database for linguistic resources. Our linked-data database platform achieves acceptable response times even when hosted on a private machine, validating the feasibility of the approach.

<sup>1</sup><http://www.irisa.fr/LIS/ferre/sparklis>

<sup>2</sup><https://headwork.irisa.fr/headwork/>

All scripts produced during this study, including those for experimental validation, are openly available online<sup>3</sup>. These resources can be adapted to other projects, enabling the transformation of data from one purpose to another effectively.

## References

- [1] Malti Bansal, Apoorva Goyal, and Apoorva Choudhary. 2022. A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning. *Decision Analytics Journal* 3 (2022), 100071. doi:10.1016/j.dajour.2022.100071
- [2] Marcello Cherchi. 1997. *Modern Georgian Morphosyntax: A grammatico-categorial hierarchy-based analysis with special reference to indirect verbs and passives of state*. Harrassowitz Verlag · Wiesbaden.
- [3] Sophiko Daraselia and Serge Sharoff. 2016. Enriching Georgian Dictionary Entries with Frequency Information. In *Proceedings of the 17th EURALEX International Congress* (6-10), Tinatin Margalitadze and George Meladze (Eds.). Ivane Javakhishvili Tbilisi University Press, Tbilisi, Georgia, 321–327.
- [4] Sébastien Ferré. 2017. Sparklis: An Expressive Query Builder for SPARQL Endpoints with Guidance in Natural Language. *Semantic Web: Interoperability, Usability, Applicability* 8, 3 (2017), 405–418. doi:10.3233/SW-150208
- [5] Jost Gippert. 2016. Complex Morphology and its Impact on Lexicology: the Kartvelian Case. In *Proceedings of the 17th EURALEX International Congress* (6-10), Tinatin Margalitadze and George Meladze (Eds.). Ivane Javakhishvili Tbilisi University Press, Tbilisi, Georgia, 16–36.
- [6] Jorge Gracia, Ilan Kernerman, and Julia Bosque-Gil. 2017. Toward Linked Data-native Dictionaries. In *Proceedings of the eLex 2017: Lexicography from Scratch* (19-21). Netherlands, 550–559. <https://elex.link/elex2017/wp-content/uploads/2017/09/paper33.pdf>
- [7] Tamar Makharoblidze. 2012. *The Georgian Verb*. LINCOM GmbH.
- [8] Tinatin Margalitadze. 2020. The Lexicography of Georgian. In *International Handbook of Modern Lexis and Lexicography*, Patrick Hanks and Gilles-Maurice de Schryver (Eds.). Springer Berlin Heidelberg, 1–24. doi:10.1007/978-3-642-45369-4\_103-1
- [9] Paul Meurer. 2007. A computational grammar for Georgian. In *International Tbilisi Symposium on Logic, Language, and Computation*. Springer, 1–15.
- [10] Donald Rayfield, Shukia Apridonze, Ariane Chanturia, Rusudan Amirejibi, Laurence Broers, Levan Chkhaidze, and Tina Margalitadze (Eds.). 2006. *A Comprehensive Georgian-English Dictionary*. Garnett Press.
- [11] Kita Tschenkéli, Yolanda Marchev, and Lea Flury. 1965. *Georgisch-deutsches Wörterbuch*. Vol. 2. Amirani-Verlag Zürich.
- [12] Kevin Tuite. 1998. *Kartvelian morphosyntax: Number agreement and morphosyntactic orientation in the South Caucasian languages*. Lincom Europa Munich.

<sup>3</sup><https://github.com/aelizbarashvili/KartuVerbs>



# Author Index

- Agoulmine Nazim, 38–45, 98–102  
Archil Elizbarashvili, 84–87, 103–106
- Barros Nascimento Erick, 21–29  
Batra Vishnu, 88–92  
Bergé Jean-Sylvestre, 75–83  
Bezerra Braga Reinaldo, 14–20  
Billy Maxime, 3–13  
Boudaoud Karima, 3–13, 46–51, 67–83, 93–97  
Boutih Karim, 98–102
- Carvalho Tereza C. M., 38–45  
Chaves Adalberto, 52–59
- De Macêdo José Antonio Fernandes, 60–66  
De Souza José, 52–59  
Delettre Christian, 3–13, 46–51  
Djelouah Mazigh, 98–102  
Ducassé Mireille, 84–87, 103–106
- Ecerf Elena, 75–83
- Flávio De Oliveira Silva, 38–45
- Garg Tanya, 88–92  
Gogorikidze Levan, 30–37  
Gomes Francisco, 52–59
- Holveck Emmy, 75–83  
Houssain Corinne, 67–74
- Journot Ambre, 46–51
- Khachidze Manana, 84–87, 103–106
- Langrota Vinicius, 60–66
- Marques Cavalcante De Oliveira Bruno, 14–20  
Marques Sabrina, 67–74  
Martin Amandine, 3–13  
Martins Joberto S. B., 38–45  
Martins José, 60–66  
Martins Maciel Paulo Romero, 21–29  
Matingou Alicia, 75–83  
Moreira Rodrigo, 38–45
- Neuman De Souza José, 88–92
- Pantano Loic, 75–83
- Raffin Louis, 93–97  
Ramalho Dantas Jamilson, 21–29  
Rego Paulo, 52–66  
Rodrigues Da Silva Gabrielly, 14–20  
Roudier Yves, 93–97
- Svanadze Giorgi, 30–37
- Trinta Fernando, 52–59  
Tsintsadze Magda, 30–37, 84–87, 103–106