GALICIA

February 2025

Alberto Stefanini Novareckon

Alberta Bertin (Novareckon)

KPI7 – Terms of Reference of the mid-term Workshop, contents and questionnaires

Project Name: GALICIA	
Report Title: KPI7 – Terms of Reference of the mid-term Workshop, contents and questionnaires	
Authors: Alberto Stefanini and Alberta Bertin (Novareckon)	
Revised by: Alberto Stefanini (Novareckon)	
Date: 28/02/2025	
Version: 1	
Distribution: Public	

TABLE OF CONTENTS

1 INTRODUCTION

The GALICIA project aims to establish a platform for the automated verification of Large Language Model (LLM)-generated code against predefined test cases, with a strong focus on automation and industrial applications in the telecom sector. By encompassing the complete lifecycle—from natural language functional requirements and security specifications to formal verification—GALICIA seeks to bridge the gap between AI-generated code and compliance with user-defined requirements and industry standards. The project's broader ambition is to analyze the reliability of Generative AI in software development and contribute to building trust in AI-driven automation.

This report is structured around two key documents: (1) a Report on a Survey conducted in advance with GALICIA stakeholders, and (2) the Terms of Reference (ToR) for the upcoming workshop. The survey was carried out to gather insights into the challenges and opportunities in verifying AI-generated code, as well as stakeholder perspectives on automated testing and compliance. Its outcomes played a crucial role in shaping and establishing the ToR of the workshop, ensuring that its scope, objectives, and key discussion themes align with the most relevant issues identified. The ToR now serves as a strategic framework to guide stakeholder engagement, facilitating discussions on cybersecurity, AI-driven automation, and compliance verification in a way that directly addresses the survey's findings. The workshop will serve as a platform to refine GALICIA's approach, incorporating feedback from industry and research communities to ensure that the project remains aligned with emerging challenges in digital resilience. With the support of NGI SARGASSO, the event will bring together key partners—including Novareckon, Mind in a Box Inc., and Hal Service—to demonstrate and evaluate the potential of automated verification in real-world applications. This report provides a foundation for these discussions, offering a comprehensive view of the project's current state and its future trajectory.

2 GENERATIVE AI WITH CYBERSECURITY FOR INTERNET APPLICATIONS DEVELOPMENT (GALICIA) – SURVEY

The objective of GALICIA (https://www.galicia-project.eu/) is to provide a platform for automated verification of LLM (Large Language Model) generated code against a set of test cases in automation, encompassing a large case study of industrial relevance in the telecom sector. To this end, GALICIA will assist through and encompass the complete life cycle from the source code automatically generated by LLM, i.e. from the natural language functional requirements and security specifications provided by users to the formal verification of the LLM generated code based on users' natural language specifications. GALICIA's contribution to digital resilience will be to prove in a selected number of industrial case studies how automated testing of LLM generated code can formally prove and ensure compliance against user-defined requirements and given standards. The ambition is to verify source code generated by Generative AI and analyse its limits, thus building trust in Generative AI.

GALICIA project has been selected, among others, for financing support by NGI SARGASSO (https://ngisargasso.eu/) a European Union funded acceleration programme for next-generation internet technologies. GALICIA is a Novareckon project (https://www.novareckon.it/en/) , our project partners are Mind in a Box Inc (https://mindinabox.ai/) a subsidiary of Inmind Technologies, a known Canadian ITC company, and Hal Service (https://www.mywic.it/) that will provide business critical internet application to be used as a test case.

3 SURVEY OBJECTIVES

As part of the GALICIA project, a consultation was launched that involved users of innovative SMEs and Start-ups, Research Centers, Large companies, Industries Associations, and Innovation Hubs, European and Canadian, to learn about their experience and their point of view regarding the innovation prospects of the project.

This report wants to share the results collected until the end of February 2025.

3.1 Survey Methodology

The survey management and collected data are managed through Google Forms.

The link to the survey was sent to direct contact list of the project partners, and it was shared through social media, and Industries association that share it to their associates.

3.2 Survey Results

3.2.1 Geography of respondents and profile of the organisations

The total number of responders are 64.

Most of them are from European countries, in order of numerousness Italy, flowed by Spain, France, Greece, and Belgium, while from Canada responders are 9% of the total.

Within Italy the distribution is among 11 regions over 20.

Respondents are representative of innovative SMEs and Start-ups, Research Centers, Large companies, Industries Associations, and Innovation Hubs. Most respondents are from Small Medium organisations (71.4%), Lage Enterprises are represented by 28.6%.

Source: GALICIA Survey data from Google Forms

![EUROPE pie chart]

EUROPE

Legend:
- FRANCE
- SPAIN
- GREECE
- BELGIUM
- ITALY
- UK

Source: GALICIA Survey data from Google Forms

![Your company dimension pie chart]

Your company dimension

- 11,9%
- 28,6%
- 19%
- 40,5%

Legend:
- fewer than 10 employees
- fewer than 50 employees
- fewer than 250 employees
- more than 250 employee

Source: GALICIA Survey data from Google Forms

3.2.2 Diffusion of generative AI inside organisations

Generative AI is widespread in the organisations and its use varies with respect to the activities for which it is used in companies. The 79.4% of the respondents have used generative AI for their company projects. Most of users experienced the use of generative AI for Software development (60%), followed by Graphic and Video design, Marketing and Sales, and Customer support.

The most used applications are: ChatGPT (96%), Google Gemini (40%), Microsoft Copilot (32%).

In software development, Generative AI is mainly used for automatic generation of code in Python language (65.9%), followed by JavaScript, Java, and C# - C++.
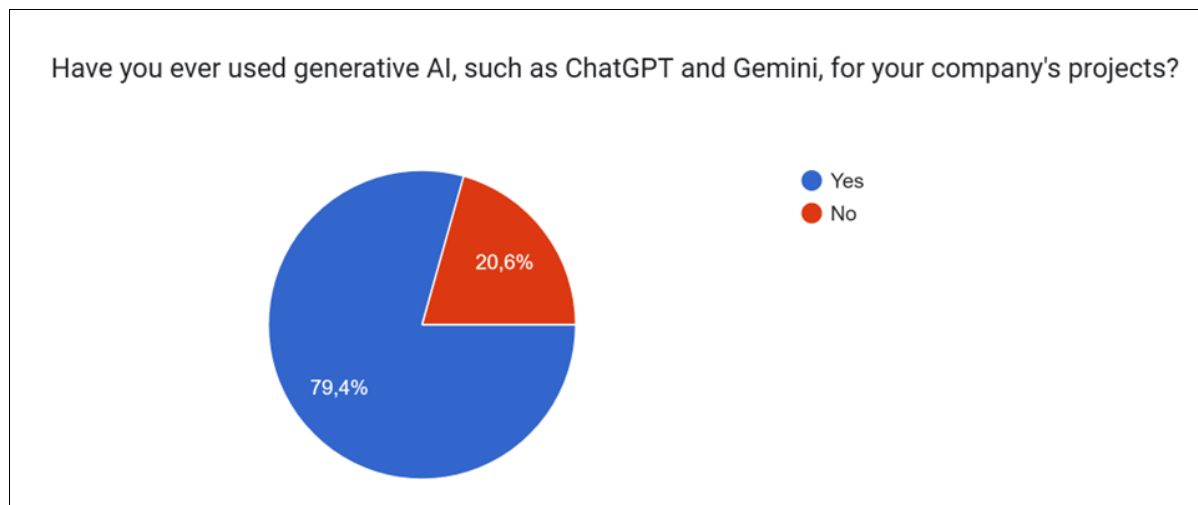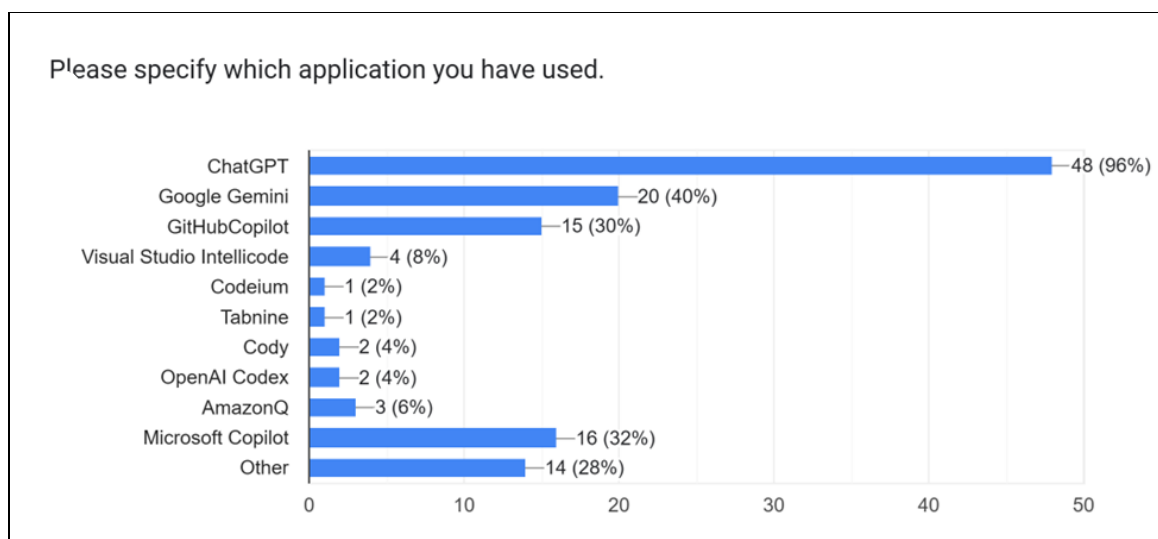
Source: GALICIA Survey data from Google Forms

Source: GALICIA Survey data from Google Forms

For which jobs do you use generative AI applications?
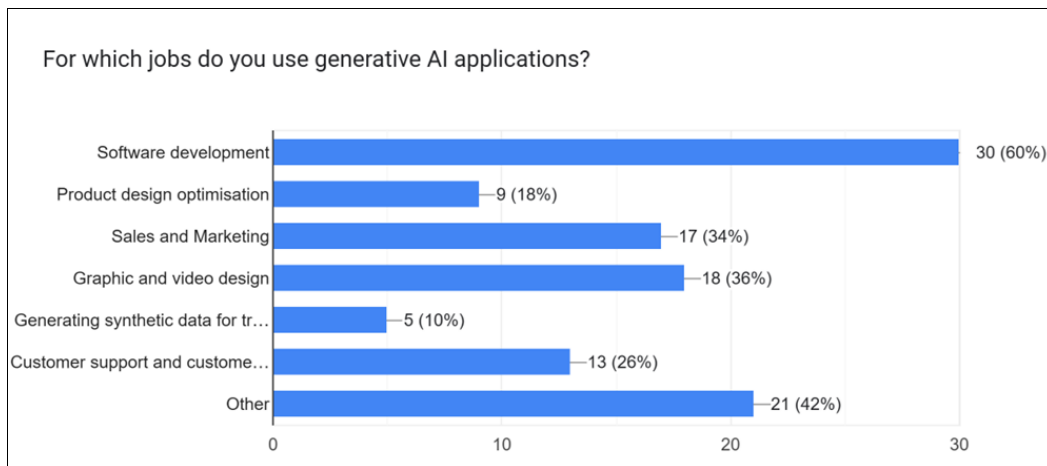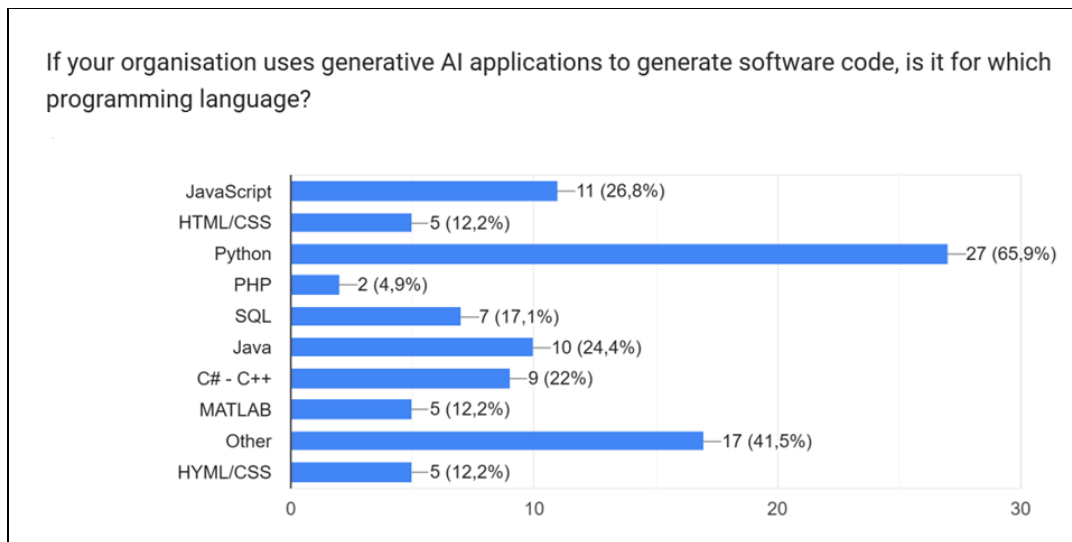
Category	Value
Software development	30 (60%)
Product design optimisation	9 (18%)
Sales and Marketing	17 (34%)
Graphic and video design	18 (36%)
Generating synthetic data for tr…	5 (10%)
Customer support and custome…	13 (26%)
Other	21 (42%)

Source: GALICIA Survey data from Google Forms

If your organisation uses generative AI applications to generate software code, is it for which programming language?

Language	Value
JavaScript	11 (26,8%)
HTML/CSS	5 (12,2%)
Python	27 (65,9%)
PHP	2 (4,9%)
SQL	7 (17,1%)
Java	10 (24,4%)
C# - C++	9 (22%)
MATLAB	5 (12,2%)
Other	17 (41,5%)
HYML/CSS	5 (12,2%)

Source: GALICIA Survey data from Google Forms

3.2.3 Perception of cybersecurity relevance in AI generated software code

Most of respondents are very familiar and somewhat familiar with the concept of verifying source code generated by generative AI (66%). Moreover, for companies it is considered extremely important to analyse and understand the source code generated by AI for ensuring quality and for reducing risks (62%).

Interviewed used were asked how often they notice that the generated code has security flaws or does not match what you requested, and they responded that this happened in their experience very frequently for the 24% of respondents and sometimes for the 42% of them.

How familiar are you with the concept of verifying source code generated by generative AI?

Legend:
- Very familiar
- Somewhat familiar
- Not familiar

34% Not familiar
48% Somewhat familiar
18% Very familiar
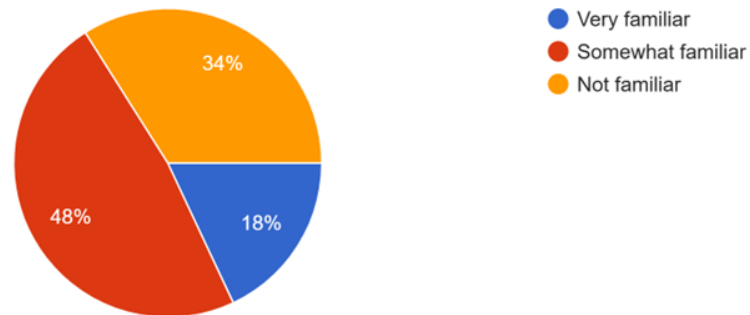
Source: GALICIA Survey data from Google Forms

How important is it for your company to thoroughly analyse and understand the source code generated by generative AI?

Legend:
- Extremely important - it ensures quality and reduces risk
- Moderately important - it can provide valuable insights but may not always be necessary
- Somewhat important - it depends on the use case and level of complexity
- Not important

8%
24%
6%
62%

Source: GALICIA Survey data from Google Forms

How often do you notice that the generated code has security flaws or does not match what you requested?

- Very frequently
- Sometimes
- Rarely
- Never

24%

10%

42%

24%
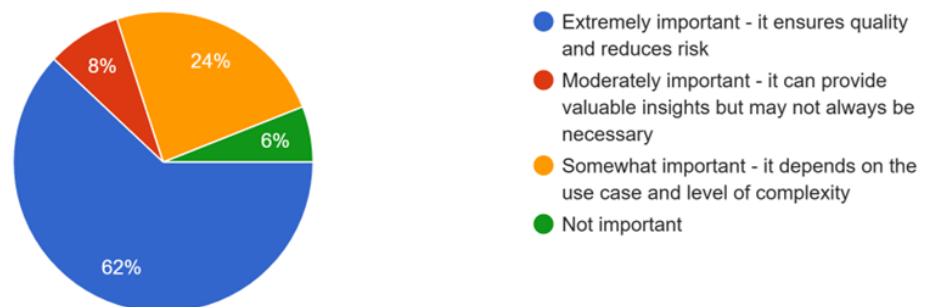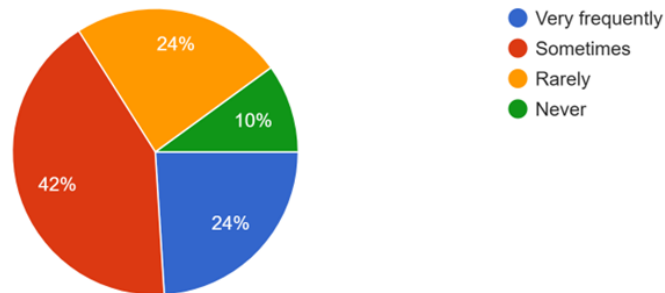
Source: GALICIA Survey data from Google Forms

3.2.4 Need of tools for building trust in Generative AI

Most of respondents that are using generative AI for software development believe that that a tool able to test AI generated code for correctness, security and compliance with user requirements and given standard may found application in their organisation (80%).
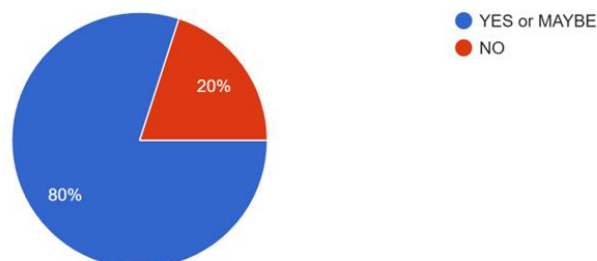
The benefits they expect from such a tool span from time saving (84%), cost reduction (58%), improved security (56%), better customer satisfaction (24%) and revenue growth (10%).
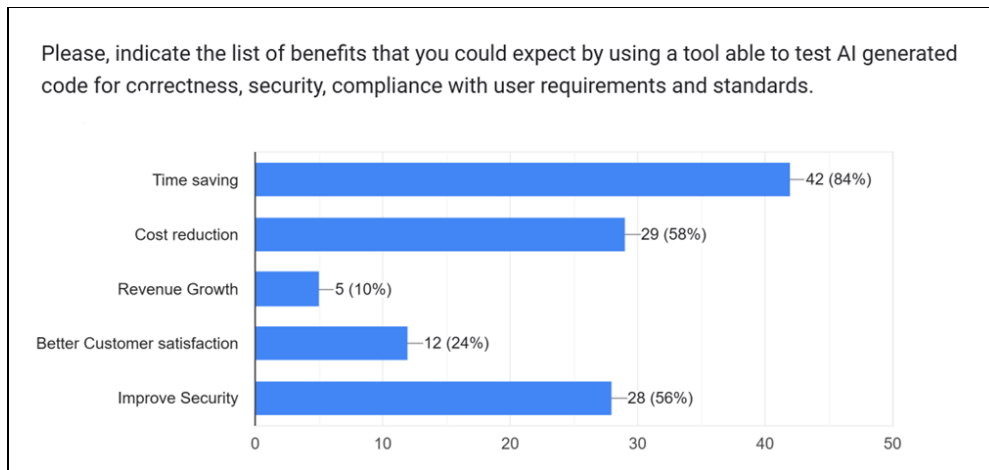
On a scale of 0 to 5, the 64% of respondents rate between 4 and 5 the benefits for their organisations from using a tool able to test AI generated code for correctness and security.

For the 50% of respondents such a new tool, able to assist developers in code generation, verification, and security, should be released as a plugin for integration within existing development environments, the 32.5% would prefer to have it as independent application to be installed in their development environments, the 17.5% see it as a specialized GPT available within existing platforms like OpenAI's GPT.
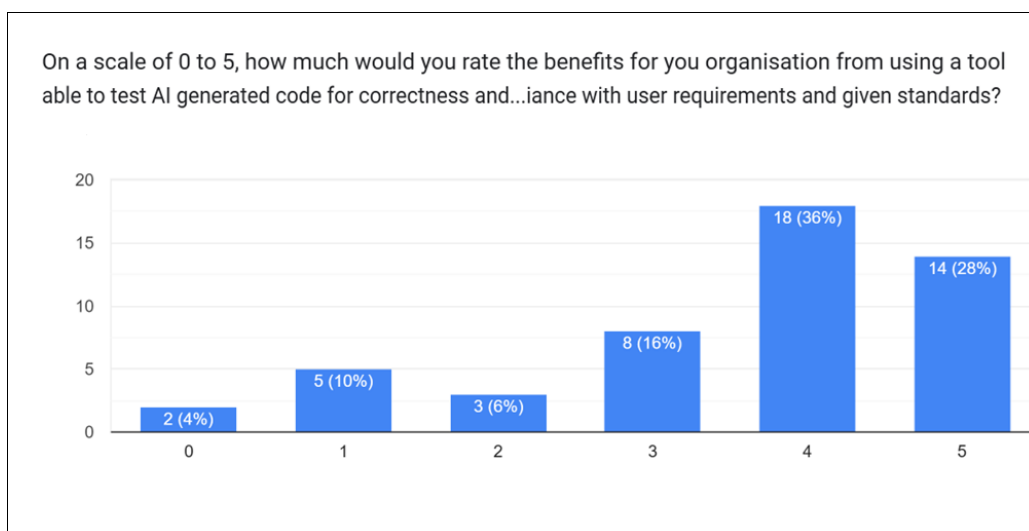

Do you believe that a tool able to test AI generated code for correctness, security and compliance with user requirements and given standard may found application in your organisation?
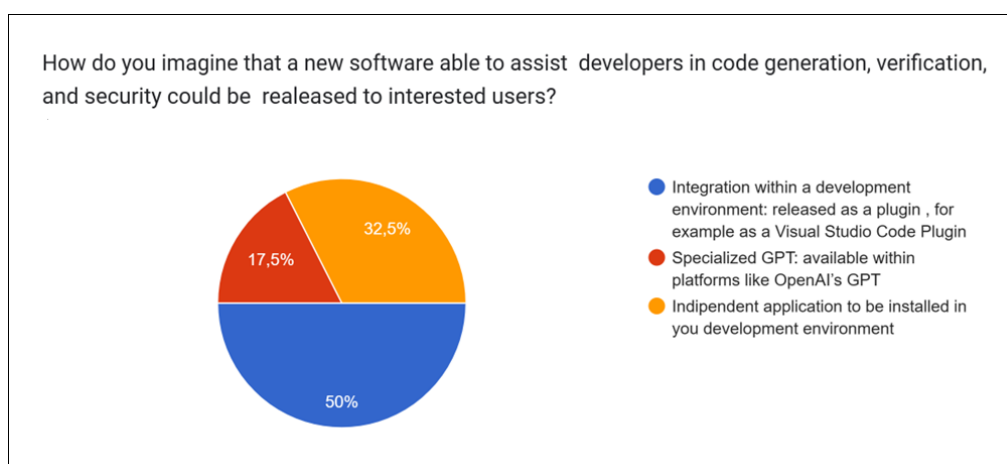
- YES or MAYBE
- NO

20%

80%

Source: GALICIA Survey data from Google Forms

Please, indicate the list of benefits that you could expect by using a tool able to test AI generated code for correctness, security, compliance with user requirements and standards.

Benefit	Value
Time saving	42 (84%)
Cost reduction	29 (58%)
Revenue Growth	5 (10%)
Better Customer satisfaction	12 (24%)
Improve Security	28 (56%)

Source: GALICIA Survey data from Google Forms

On a scale of 0 to 5, how much would you rate the benefits for you organisation from using a tool able to test AI generated code for correctness and...iance with user requirements and given standards?

Rating	Value
0	2 (4%)
1	5 (10%)
2	3 (6%)
3	8 (16%)
4	18 (36%)
5	14 (28%)

Source: GALICIA Survey data from Google Forms

How do you imagine that a new software able to assist developers in code generation, verification, and security could be realeased to interested users?

- 50% Integration within a development environment: released as a plugin, for example as a Visual Studio Code Plugin
- 17,5% Specialized GPT: available within platforms like OpenAI's GPT
- 32,5% Indipendent application to be installed in you development environment

Source: GALICIA Survey data from Google Forms

3.3 CONCLUSIONS

The results of the GALICIA survey highlighted the significant interest and growing reliance on generative AI tools across various sectors, particularly within SMEs, large enterprises, and research centers. Respondents widely recognize the potential of generative AI in streamlining software development processes, with tools like ChatGPT, Google Gemini, and Microsoft Copilot being prominently used. However, this rapid adoption also underscores the pressing need for robust verification and security mechanisms to address the risks associated with AI-generated code.

Key insights from the survey include the following:

- **Adoption of Generative AI**: Most respondents (79.4%) are integrating generative AI into their organizational workflows, particularly for software development (60%), where Python leads as the preferred language.

- **Importance of Cybersecurity**: The survey revealed that security flaws in AI-generated code are a frequent concern, with 66% of respondents emphasizing the need for tools to ensure code quality and compliance.

- **Demand for Validation Tools**: 80% of respondents expressed the necessity for tools that can verify AI-generated code for correctness, security, and compliance, with anticipated benefits such as time savings (84%), cost reduction (58%), and improved security (56%).

These findings affirm the relevance and timeliness of GALICIA's mission to deliver an automated platform for validating AI-generated code. By addressing the identified gaps in trust, compliance, and efficiency, GALICIA is well-positioned to serve a critical role in advancing the secure and reliable adoption of generative AI technologies across industries. As the survey progresses and additional feedback is collected, these insights will further refine GALICIA's development priorities and strengthen its alignment with user needs and expectations.

4 TERMS OF REFERENCE FOR THE GALICIA WORKSHOP

4.1 Advancing GALICIA: Platform Presentation, Case Studies, and Roadmap Development

The upcoming event will serve as a key milestone in the development and refinement of the GALICIA platform, structured around the key outcomes of the stakeholder survey. It will provide an opportunity to consolidate initial findings, present case studies, and outline future directions based on the challenges and priorities identified through the survey. The agenda has been designed to foster technical discussion and stakeholder engagement, ensuring that GALICIA evolves in a way that is both practical and directly aligned with the most pressing issues in cybersecurity, AI-driven automation, and compliance frameworks.

4.2 Presentation and Discussion on the GALICIA Platform

The session will open with a comprehensive presentation of the GALICIA platform, highlighting its key capabilities, architectural foundations, and areas of ongoing refinement. This overview will set

the stage for discussions on how GALICIA can support industries in achieving automated compliance, cybersecurity risk assessment, and software verification through AI-driven approaches. Special attention will be given to recent advancements in integrating formal methods with machine learning techniques, a topic of relevance given the growing demand for robust verification frameworks in mission-critical environments.

4.3 Case Studies: Initial Insights and LLM-Driven Analysis

Following the platform presentation, the discussion will transition into an analysis of several initial case studies that have been conducted using GALICIA's capabilities. These case studies will illustrate practical applications in different domains, ranging from telecommunications and industrial automation to cybersecurity risk evaluation. Each study will highlight specific technical challenges addressed by GALICIA, as well as preliminary outcomes that demonstrate its effectiveness in improving compliance efficiency and security assurance.

Additionally, a separate segment will be dedicated to case studies that leverage a single specific Large Language Model (LLM) to automate certain processes within GALICIA. These cases will provide insights into the feasibility and potential impact of LLM-driven functionalities, such as code analysis, vulnerability detection, and compliance cross-checking. The findings from this exploration will inform discussions on the broader role of AI in automating traditionally complex verification and validation tasks.

4.4 Round Table Discussion: Objectives, Opportunities, and Roadmap

The event will conclude with a round table discussion, where stakeholders will engage in a structured dialogue about the strategic and technical objectives for GALICIA's further development. Key points for discussion will be:

- The identification of specific technical objectives that can drive the next phase of GALICIA's evolution, including areas for refinement and expansion.

- Opportunities for industry adoption and collaboration, particularly in sectors where compliance automation and AI-driven validation can yield significant benefits.

- The initial development of a roadmap that will guide GALICIA's future trajectory. This roadmap will incorporate insights from current stakeholders, ensuring that the platform continues to align with real-world needs and emerging regulatory trends.

By synthesizing these discussions into a clear set of action points, the event will contribute to shaping GALICIA into a more robust and adaptable solution. The engagement of industry experts, research institutions, and SMEs will be instrumental in ensuring that GALICIA remains at the forefront of AI-driven compliance and cybersecurity innovation.

5   CONCLUSIONS

The GALICIA workshop, shaped by the insights gathered through the stakeholder survey, represents a pivotal step in refining the project's objectives and methodology. By structuring discussions around the key findings of the survey, the workshop will ensure that GALICIA remains aligned with industry needs, regulatory developments, and the technical challenges of verifying AI-generated code. The survey not only helped define the scope of the event but also guided the formulation of its Terms of Reference, ensuring that the workshop addresses the most pressing issues in cybersecurity, compliance automation, and AI-driven validation.

The event's agenda, culminating in a structured roundtable discussion, is designed to facilitate meaningful engagement among industry experts, research institutions, and SMEs. The identification of technical objectives, industry adoption strategies, and the initial drafting of a roadmap will provide a concrete foundation for GALICIA's next development phase. The roadmap, enriched by stakeholder contributions, will serve as a guiding instrument to enhance the platform's capabilities while ensuring its adaptability to evolving regulatory and technological landscapes.

Moreover, the workshop will not only consolidate existing research but also act as a catalyst for future collaborations. By bringing together diverse stakeholders, the event aims to strengthen GALICIA's position as a key enabler of trust and reliability in AI-driven software verification. Through continued engagement, the project will refine its methodologies, expand its case studies, and integrate best practices from industry and academia to maximize its impact.

Ultimately, this process will help GALICIA evolve into a robust, industry-ready solution, offering practical and scalable approaches to compliance verification in AI-generated software. The structured dialogue established in this workshop will be instrumental in ensuring that the project remains at the forefront of AI-driven cybersecurity and compliance innovation, fostering a more resilient and trustworthy digital ecosystem.

PROJECT CONSORTIUM: