

## **Gen AI Assurance**

Test Strategy for TIAA

### May 2024

# Table of contents

[1.0 Background 3](#_Toc160821319)

[2.0 Project Objective 4](#_Toc160821320)

[3.0 Scope of Assurance Solutions 4](#_Toc160821322)

[5.0 Test Approach 4](#_Toc160821325)

[6.0 Benefits 5](#_Toc160821326)

[7.0 Test Tools 6](#_Toc160821328)

[11.0 Assumptions, Dependencies and Constraints 6](#_Toc160821337)

[12 Risks 6](#_Toc160821338)

[13 Change Log 7](#_Toc160821339)

# Document Control

* Change Record:

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Author | Version | Summary of changes |
| 28/05/2024 | Nancy | 1.0 | First version of the document |
|  |  |  |  |
|  |  |  |  |

* Reviewer / Approver:

|  |  |  |
| --- | --- | --- |
| Name | Role | Status |
| Arunadevi | Associate Director |  |

* Reference Documents:

|  |  |
| --- | --- |
| Document name | Location |
|  |  |

# 1.0 Background

In TIAA, when interviewing candidates for open roles all candidates are being interviewed with the same interview guide. The current guides are not specific enough to each job and are often not used or not applicable. Creating interview guides specific to each job description that incorporates, the needed skills while still adhering to a standard format and layout would provide consistency help to streamline interview process and help TIAA improve the quality of new hires.

# 2.0 Project Objective

Given a job description that has been customised for specific open role, LLM generates an interview guide along with the rubric tailored to that job description and the corresponding skills.

Gen AI Assurance Resource:

|  |  |  |
| --- | --- | --- |
| Name | QA Role | Email |
| Shankar Bharanidharan | Gen AI Assurance Tester |  |

# 3.0 Gen AI Assurance Areas (In scope & out of scope)

|  |  |
| --- | --- |
| In Gen AI Assurance Scope | Out of QA Scope |
| Assessment of Prompts built | Functional Testing |
| Assessment of existing workflow | Cross Browser Testing |
| Plan for use case validation & metrics selection | UI Testing |
| Developing assurance component with selective Metrics needed | Performance Testing |
| Working with SME to enable additional insights needed for assurance | Security Testing |
| Execution of tests | Automation Testing |
| Review results with SME | CI CD Integration of Automation Scripts |
| Submission of execution report | Evaluation of Training dataset/Fine tuning models |

*Table 1: Scope*

# 4.0 Test Approach

**Phases of Assurance approach:**

The automation approach spans across 4 phases.

1. Analysis
2. Test Design
3. Execution
4. Review O/P

*A blue squares with white text

Description automatically generated*

Figure 1: Phases

# 5.0 Benefits

|  |  |  |
| --- | --- | --- |
| Benefits |  | Brief |
| Diverse evaluations |  | Solution follows a multi-faceted approach ranging from use case-based Metrics, standard metrics evaluations, to gather better insights |
| Enriched Datasets for  your testing needs. |  | Generate diversified set of prompts to evaluate the response being generated in all angles. |
| Precision Checks |  | Innovative techniques, generate scores and provides reason for every evaluation performed. |
| Easy Visualization |  | Filter results below threshold for further SME evaluation. |

*Table 2: Benefits*

## **6.0** **Test Tools**

### **6.1 Equipment/Hardware/Environment**

Test Execution for an AI solution would demand a minimal configuration and Software needs, shown below:

* Installed RAM: 12 GB RAM
* Disk Space: 150 - 200 GB of disk space is required for package installation.

### **6.2 Software**

|  |  |  |
| --- | --- | --- |
| Tool | Version | Comments |
| Pycharm | community version 2023.1.1 |  |
| POSTMAN | Latest Version |  |
| Browsers (Chrome, Firefox and IE8) | Latest version |  |

*Table 3: Software Required*

# 7.0 Assumptions, Dependencies and Constraints

1. A stable version of the Gen AI Assurance application should be available in Test environment before the test activities begin.
2. All privileges/ model access provided to development team, would be enabled for Gen AI Assurance team as well.
3. Upon any new version release / changes made to the ongoing use case, TIAA team will notify Gen AI Assurance tester immediately.
4. Complete knowledge on AI Implementation performed for specific use cases would be provided to Gen AI Assurance Team
5. In case of any firewall related issues, TIAA team will take the complete responsibility to get it resolved with any internal system communications.

# 8 Risks

|  |  |  |  |
| --- | --- | --- | --- |
| Risk | Priority | Impact | Mitigation Plan |
| Availability of required model access/ hardware/software resources on time | High | * Schedule Slippage | * The project team ensures that all the necessary environment/ model access /hardware /software resource are available before the implementation begins |
| Necessary Security Approvals | High | * Schedule Slippage | * Facilitation by TIAA team to ensure Cognizant Gen AI Assurance tester gets access to the Servers to perform necessary installations/implementations/monitoring if required |
| Delay in Iteration due to issues in connectivity, environment, or infrastructure | High | * Schedule Slippage | * The project team to ensure the connectivity networks between servers are stable and no restrictions imposed |
| There may be an impact on Assurance Implementation timeline if the requirement keeps changing | High | * Schedule Slippage | * It has been assumed that requirement shared in the initial phases once frozen will not go through any further changes. Upon any change Cognizant Gen AI Assurance tester, it’s not possible to provide any support on the impacts caused by the change. |
| New Usecase added onto the application which requires the above process to be repeated starting from Analysis | High | * Schedule Slippage | * Cognizant Gen AI Assurance tester to discuss with TIAA to understand the additional functionality and come up with a feasibility check performed |

*Table 4: Risks*

# 9 Change Log

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Version Number | Changes made | | | |
| V1.0 | *<First version>* | | | |
| V1.1 | *<If the change details are not explicitly documented in the table below, reference should be provided here>* | | | |
| Page no | Changed by | Effective date | Changes effected |
|  |  |  |  |



Thank you!