The use of AI technologies in the area of conversational interfaces is growing considerably. In domains like Insurance, Banking and Financial Services, and Travel & Transportation, the potential that the conversational interfaces hold is quite immense although the level and extent of integration of AI in all these interfaces is varied. Conversational interfaces are text-based or conversation-based and operate on web or mobile or custom device channels. Though they may all be rolled under the general category of ‘Chatbots’, in reality, they can range from simple domain-specific Chatbots to smart task-oriented Chatbots to Intelligent Agents (IA) with a wide capacity to learn. Conventional testing methodologies and tools fall short if they were to be used without appropriate modification for AI-infused conversational tools such as Chatbots. NIIT Technologies 3-D QA Framework for Testing AI-infused Conversational Interfaces is a comprehensive, end-to-end framework that covers various dimensions like channel integration, conversation flows, and monitoring. The assurance of these dimensions is accelerated by our conversation test accelerators that includes Train-Test-Label approach and frameworks and methodologies for automation, performance testing, security testing, and experience dashboard.
Beyond Traditional Testing for AI-infused Chatbots: Key Considerations

Since the objectives of deploying Chatbots are quite different from those of traditional applications or tools, it is not ‘business as usual’ for the teams charged with testing such conversational interfaces. The considerations that are missed by traditional testing when it is singularly focused on functional testing of Chatbots are:

- Testing for Conversational Flow Assurance in terms of Personality, Behavior and Intelligence
- **Personality:** Testing for Flexibility, ease of conversation, problem solving, pleasant and friendly responses, suggestions even in case of incomplete resolution
- **Behavior:** Testing for ensuring that context is maintained, there is a closure for conversation and there is graceful exit.
- **Intelligence:** Testing for processing of idiomatic and colloquial language, understanding from experience, discrimination between relevant questions and irrelevant questions
- Versatility across multiple channels through Omni-channel integration support
- Performance and avoidance of latency in the case of conversational interfaces as they are highly sensitive to any perceived performance issues.
- Testing for deep localization so that global users can comfortably conduct their conversations. Linguistic and cultural considerations can easily be missed without explicit effort to include them.
- Testing for unpredictable situations and questions which will need special focus on creating diverse test data
- Security focus is critical for unmanned conversational interfaces as they could be easy targets for potential automated social engineering and other cyberattacks.
- Metrics should be an important embedded feature of conversational interfaces as they need to monitor users, conversations, tool performance (goal completion rate, goal completion time) and business results (Revenue growth, self-service rate, customer satisfaction) for learning from experience and future improvements.

It is imperative that these special considerations are integrated into planning and execution of test strategies for AI-infused conversational interfaces so that they can deliver higher rate of business success than traditional tools.
3-D Test and Quality Assurance Framework

When NIIT Technologies found that many of its customers were intent on leveraging AI-infused Chatbots in their operations, it was enthusiastic in applying its niche expertise in AI acquired worldwide but, found that the individual customer journeys were quite different. Some of them would start with simple support Chatbots to answer domain-specific questions and not upgrade them while some others would move through the transition of Support Chatbots to Intelligent Chatbots very quickly. The customer needs and ambitions were quite different just as their capabilities were.

There was a strong need for a scalable framework based on a comprehensive testing methodology that can handle multiple channels, devices, conversational methods, and market leaders such as Amazon Alexa, Google Home and Facebook Messenger. This would allow customers to start with the type of Chatbot they are comfortable with and progress at a desired rate without changing the foundational test platform, tools or methodology at all. They can plug the standard tools into/out of the platform as needed without major changes to tools or skilled resources managing them. In fact, NIIT Technologies found that the lack of standard and scalable test platform was the single biggest reason for project delays and non-realization of desired business results despite investing the required resources.

The 3-Dimensions

The 3 dimensions that were well integrated into the framework are Omnichannel integration, conversational flow and monitoring which can work with the leading conversational chat tools in the market. The framework will work equally well across simple Support Chatbots, Smart Chatbots and Intelligent Chatbots. There are also test planning tools that will ensure that the key considerations of functionality, localization, user interface, channel-specific capabilities, performance and security are addressed adequately for better success rate in the field. A special emphasis is also laid on testing the personality, behavior and intelligence of Chatbots by appropriately quantifying the subjective nature of the tests. A comprehensive dashboard with key test results and trends through advanced analytics will ensure deep insight into the performance of Chatbots.
End-to-End Testing for Development and Release Process

As the systems are tested during development in the Engineering Testing phase and migrated to the NIIT Tech 'Train-Test-Label' Model driven by 'Human-In-The-Loop' Test Automation Frameworks, the feedback is used to strengthen the responsiveness and success of the system by training the AI-infused system to learn from experience in a way that it can be recalled when a response needs to leverage it. The confidence level in the system itself gets a boost by continuously training, testing and labelling the information appropriately as it is captured in various phases.

Another important step in the test and release cycle of an AI-infused application is an orchestrated exposure to a wide variety of real world users, conversations and situations. One of the ways of doing that is to have the AI-infused application like Chatbot go through a ‘Crowd Testing’ phase. Crowd testing, by design is a random mix of users who enroll themselves for testing the application. This will ensure that several language styles, conversations flows and cultural accents are tested ahead of release. This will go a long way in strengthening the application for success in the field.

Conversational Test Accelerators

Conversational flows and personas for functional assurance of conversational interfaces, for both voice and text based, are infinite in number. Even testing a sample could be quite large. Also, traditional non-functional assurance methodologies do not work in many scenarios. Hence, there is a need of conversational test accelerators.

There are a number of key conversational test accelerators that need to be integrated for assurance of the core intelligence and the design of an AI-infused application. An extensive suite of accelerators for automated testing, performance testing, security testing, and experience assessment need to be in place. So that the various conversational test suites can be run repeatedly all through the application life cycle. These will not only speed up the application release process but also ensure the confidence level of conversations every time.
Automated validation of end-to-end conversational flows
Automatically build library of utterances from platform database

Testing the Waters with a Leading APAC Airline

As a hint of things to expect in future, one of NIIT Technologies’ clients, a leading airlines in APAC has released a new set of ‘skills’ for its customers on the move. Its customers can enquire about the flight status and frequent flyer account details with many more skills to be released in due course. Though the airline started with Amazon Alexa to begin with, it has plans to expand to other leading conversational interfaces. This is a tip of an exciting iceberg in terms of ease and convenience that the airline would like to offer to its clients on and off its flights. Future applications are expected to cover more conversational intricacies, personality aspects, language, and cultural aspects. However, the 3-D QA platform can continue to be used as the airline’s needs expand in the domain of AI-infused features of its online presence.

There are several other exciting possibilities in the areas of Banking & Financial services and Insurance as well. There are a number of banking and fintech companies deploying conversational and test-based Intelligent Agents in the areas of customer acquisition and matching appropriate loan products to the customer through easy, friendly and empathetic conversations. AI-infused Chatbots are also appearing in quote and claims areas of Insurance providers. Some of the interesting use cases in various key industries are illustrated below:

**Train-Test-Label Approach**

The key capability of AI-infused systems is its deep ability to learn. Data identified for application training purposes is used to generate the required test conversational data. When the test data is run through the AI-infused application, it produces a clear decision based on both structured data insights and unstructured knowledge reasoning. However, the litmus test applied to each output from the application is the confidence score assigned to it based on validation with training data. The principle of continuous learning is enabled by labeling low confidence results explicitly as pass/fail while high confidence results flow down the process without any additional modification. This incremental approach will help AI-infused applications generate exquisitely better predictions over time.

**Test Automation Framework**

Testing conversational interfaces manually could be very cumbersome and error-prone for multi-platform, voice, and text based chatbots. Therefore, automation of conversational interfaces becomes important, so that test suites can be executed repeatedly throughout the application lifecycle. The automated testing can be executed over the UI, as for voice based chatbots or can be tested using the backend web services. The latter is used where chatbots are built on standard UIs like Facebook, Skype etc. In addition to testing the utterances and their responses, automation frameworks for conversational interfaces, could have the following enhanced capabilities:

- Utterances and conversations library
- Build common test cases for manual and automated testing

---

**Chatbots and Intelligent Assistants**

<table>
<thead>
<tr>
<th>Lex based Shopping Assistant</th>
<th>Enterprise Bot for IT SM</th>
<th>Bot helps with quotes for Auto Insurance</th>
<th>Voice only Bot for Cargo Status Enquiry</th>
</tr>
</thead>
<tbody>
<tr>
<td>A smart chatbot assistant helps a user with shopping for office supplies</td>
<td>An enterprise bot helps user with IT SM tickets</td>
<td>A bot to help in getting quotes for auto Insurance with OCR and NLP</td>
<td>Lex based voice only bot helps cargo handlers in status enquiry</td>
</tr>
<tr>
<td>Smart Assistant</td>
<td>Deployed on Skype for Business</td>
<td>NLP based input</td>
<td>Mobile app based voice only Bot</td>
</tr>
<tr>
<td>Web Based</td>
<td>Developed using Microsoft Bot Framework</td>
<td>OCR for car number plate recognition</td>
<td>Lex based</td>
</tr>
<tr>
<td>Lex based Bot</td>
<td>Raise a service Now Ticket in case it is unable to service a user query</td>
<td>Car image upload</td>
<td>Natural Language input</td>
</tr>
<tr>
<td>Contextual Sharing between web and chat</td>
<td></td>
<td>Available on Skype</td>
<td>Helps senior management in Air cargo to get quick status updates</td>
</tr>
</tbody>
</table>
The NIIT Technologies Thought Board:
3-D QA Framework for Testing AI-infused Conversational Interfaces

What are the defining qualities of AI-infused Conversational Interfaces?

<table>
<thead>
<tr>
<th>Natural conversational ability</th>
<th>Contextual awareness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to narrow down search</td>
<td>Knowledge of external information sources</td>
</tr>
<tr>
<td>Self-learning conversational nuances</td>
<td>Ability to provide constructive pointers</td>
</tr>
<tr>
<td></td>
<td>User Experience</td>
</tr>
</tbody>
</table>

What are the top issues with current solutions?

- Lack of Conversational Flow Assurance for Personality, Behavior and Intelligence
- Missing Omnichannel Versatility
- Lower Performance and Sensitivity to Latency
- Inadequate Localization
- Not enough focus on handling unpredictable situations
- Lack of critical focus on security
- Not enough metrics to cover test performance and business results

How does 3-D Test and QA framework stand out?

- Scalable across channels, devices and conversations
- 3-D: Omnichannel flow and monitoring
- Complete coverage of localization, performance & security beyond functionality
- Integrate feedback from beta phase and field
- Orchestrated exposure to Crowd Testing
- Test Accelerators: Automation, Performance, Security & Experience Dashboard
- Deeper Learning with Train, Test and Label approach

Natural conversational ability
Contextual awareness
Ability to narrow down search
Knowledge of external information sources
Self-learning conversational nuances
Ability to provide constructive pointers
User Experience
Lack of Conversational Flow Assurance for Personality, Behavior and Intelligence
Missing Omnichannel Versatility
Lower Performance and Sensitivity to Latency
Inadequate Localization
Not enough focus on handling unpredictable situations
Lack of critical focus on security
Not enough metrics to cover test performance and business results
Scalable across channels, devices and conversations
3-D: Omnichannel flow and monitoring
Complete coverage of localization, performance & security beyond functionality
Integrate feedback from beta phase and field
Orchestrated exposure to Crowd Testing
Test Accelerators: Automation, Performance, Security & Experience Dashboard
Deeper Learning with Train, Test and Label approach
A Human-Centered Approach to AI Makes a Difference

AI-infused applications are quite unlike traditional systems or applications. There are a number of subjective situations and diverse user base that they need to handle with ease and efficiency. In order to ensure that the sophisticated AI-infused applications deliver their promise in the field with real users and complex situations, testing methodology needs to be evolved to encompass relevant conversational flow assurance and device versatility. This will need to be followed by a platform that concretizes the methodology in a tool form. The established framework needs to handle as many conversational interfaces as the business is willing to add to its scope as it matures. As a testimony to its experience and skill in this area, NIIT Technologies has one such 3-D Quality Assurance Test framework that has been verified by real customers.

However, what sets a winning platform apart from the rest is the human element that has been integrated into an automated test framework that works across the 3 dimensions of Conversation, Omnichannel, and Monitoring. The human element is critical in training the applications to learn from experience and foresee the expected situations. A holistic test framework such as NIIT Technologies’ 3-D QA test framework is one that balances the sophistication of the Artificial Intelligence with the versatility of the critical human element.
NIIT Technologies is a leading global IT solutions organization servicing customers across the Americas, Europe, Asia and Australia. Differentiated on the strength of its industry expertise, NIIT Technologies services clients in travel and transportation, banking and financial services, insurance, manufacturing and media verticals, offering a range of services including Application Development and Maintenance, Infrastructure Management, and Business Process Management. Focused on Digital Services, the Company is helping businesses design agile, scalable and digital operating models. NIIT Technologies adheres to major global benchmarks and standards of quality and Information Security. For further information, please visit www.niit-tech.com