



Design testing for an AI-based Chatbot (multi-industry)

Case Study

research

data

subject

focus

knowledge

KEY CHALLENGE

Unlike requirements-driven software applications, testing an AI-based chatbot has infinite possibilities. Reaching a *go-ahead* for a qualified release is challenging.

Kairos

Context

The customer is an innovative startup that built an AI-based Chatbot primarily serving the recruitment industry. The recruitment processes are largely automated by the bot bringing in efficiencies while preventing fraud. The bot is designed to process audio visual content in addition to the textual conversations. As a design philosophy and for future business, the core design of the bot is industry-agnostic.

Kairos is engaged in providing them with design, implementation, testing and deployment services, and this specific case study deals with how we employed “**design testing**” in order to reduce the defects and make the bot more robust over time (given its machine-learning abilities).

Customer expectations

- To ensure that the bot replicates the activities and expertise of a human recruiter.
- Quantified measurements for the productivity gains achieved (i.e., test workflows to simulate real world volumes).
- Ensure that the design is industry-agnostic and can be scaled into other domains such as retail banking.

Key challenges

- Requirements based testing formed only 25% of the effort and 75% was about testing the *unknown*.
- Need for heterogeneous test data (text, audios and videos) replicating the real world recruiters and job seekers.
- Eliminating false positives in testing with clear justifications.
- Creating a balanced mix of positive and negative paths to test if the bot is indeed able to “learn” from the test inputs.

Test Approach

In order to ensure that we reduce defects in the real world, we created hooks into the bot's core capabilities so that testing can be done without the interface.

From these testing-only entry points, we simulated common patterns of bot interactions covering various aspects of a) Error manages, b) Answering, c) Intelligence, d) Navigation, e) Onboarding, f) Personality, and g) Understanding. We designed voluminous test data covering several textual workflows as well as crafted video/audio files.

We have used Botium to automate bot interactions, both with pre-defined data and with values generated on the fly (based on bot rules).

In essence, we have tested the design of the bot in addition to exercising it with normal functional testing.

Value Delivered

The emergence of a true AI-based (more specifically one that can leverage machine learning over time) Chatbot that serves the recruitment wing for various HR department companies and RPO vendors.