

Automating the testing of an entertainment delivery service driven by voice commands

Digital entertainment services must respond crisply and correctly to every user request, including those via voice commands. This was a considerable challenge for a leading entertainment service given their vast catalog of music, a catalog that ranges from the greatest artists of all time to the most avant-garde, obscure and niche.

Their service must ensure that a simple request to “play Barbra Streisand songs” not only gets interpreted correctly, but that the song playing, the displayed likeness of the artist and album cover, the track name, the song’s credits and lyrics, and any trivia content, all match as expected. This complex problem requires validation of the delivered audio, visual and text.

For one content powerhouse, this had been done on a sampling basis, where a selection of requests was sampled to see and hear what their media player system delivered. But with many, many thousands of entries in the catalog, the chance of a homonym, e.g., Bryan Adams confused with Ryan Adams, delivering the wrong music was too much to risk.

Objectives

The customer’s objectives were to solve this seemingly unbounded problem elegantly and reliably. Their goal was to establish a simple set of building blocks from which every possible request can be spoken to the media player system and every response verified for each part of their vast content inventory.

Fortunately, there were able to collate all possible content requests into simple CSV files. Next they needed to use those collated requests into an automated, scalable testing solution that would validate each request’s delivered audio, visual and text.

Quick Facts

Industry

Media

Region

North America

Applications Under Test

Media Player

The AIQ Solution

The AIQ Services Workbench module proved up to the challenge. It is a remarkable testing platform with over 200 built-in testing functions for even the most difficult quality tasks. They can be combined for tests at massive scale, or for precision testing tasks, and are designed to test the untestable, and to solve the unsolvable. Whether it is reading a .PDF, writing an email, generating test data on the fly, or listening to Barbra Streisand, there is a function for that.

As noted above, the starting point was to collate all possible content requests into a .CSV file. Now, a text-to-speech app reads each content request out loud to the media player. When its UI responds, AIQ's computer vision capability verifies that the visual content matches the request, along with the text content verified as matching the artist, plus matching the audio through sampling.

This is iterated several thousand times so that every piece of content is tested to ensure that the inventory of parts is correctly mapped and properly delivered.

What would take an army of audiophiles several days to complete, now takes a few hours. What was once seemingly impossible, and often unreliable, is now comprehensive, repeatable, and inexpensive.

Key Results



Multimedia results from a huge catalog are reliably validated.



98% reduction in the testing process, from days to hours



Scalable multimedia testing system keeps pace with huge and growing catalog.