The global crowd tests a chatbot for cars

Testbirds Case Study
The global crowd tests a chatbot for cars

Making the right sound

Arvato SCM Solutions is an innovative, leading international provider of supply chain management and e-commerce services in the automotive, banking, insurance, consumer products, healthcare, high-tech, entertainment, publishing, and telecommunications industries. The company combines the know-how of more than 14,000 employees with the right technologies and suitable business processes, thereby measurably increasing the productivity and performance of its partners. Together with a leading German car manufacturer, Arvato has now developed a solution concept for the digitization of the content of operations manuals. The special thing is that the information is not only available digitally in the form of an app, but can also be reproduced by a virtual assistant, a so-called chatbot, that communicates with the user. In order to find out if the chatbot fulfills the requirements of the target group, Arvato commissioned the Munich-based testing specialist Testbirds to thoroughly assess the chatbot by executing a BugAbility™ test.

The crowd for special tasks

A total of 35 testers were selected for the chatbot test. “In testing, we faced the problem of requiring testers with very specific demographic characteristics,” explains Jan Kittel, Senior IT Project Manager in the department of Corporate Information Management at Arvato SCM Solutions. Testers with either German or US English as their native language were required, and each tester had to have a valid driver’s license and own a car, or regularly use a company car. Another required characteristic of the target group was an annual income of at least 100,000 euros. It was also important for the project that very experienced, as well as inexperienced or barely experienced testers were involved, as testers with different experience use a product in different ways while testing. “By using the Testbirds crowd, we were able to recruit enough people with the profile we needed, and successfully complete the test,” says Kittel.

Testing expertise: German

- No testing experience: 11%
- I test in my free time: 28%
- I test professionally: 61%

Testing expertise: US English

- No testing experience: 6%
- I test in my free time: 24%
- I test professionally: 71%
A test talk with the chatbot

After the target group and the test setup had been configured, and the testers invited, the test phase started. The testing itself focused on the evaluation of the chatbot and its feedback. To test adequately, it was important that the testers use their everyday language, which also included the use of different dialects. Thus, the testers asked the chatbot questions in different language variants, for example with English dialects from Texas, California and Florida, and German dialects from Bavaria, Baden-Württemberg and Cologne. The resulting answers were then reviewed and evaluated. The test itself consisted of ten use cases, in which the testers asked the chatbot five questions each. In order to examine the chatbot with sufficient depth, the use cases were related to the most important and most frequently asked questions on topics such as safety, driver assistance systems, or technical data. While the use cases were pre-empted, the questions needed to be reviewed by the testers themselves, taking into account their personal needs in order to interact with the chatbot as realistically as possible.

Top-notch speech recognition, improvable design

Overall, the chatbot was well received by the German and US English testers, although the probability of using it for their own car was rated rather averagely. The chatbot's ability to recognize and transcribe questions worked very well according to tester feedback (91.1% for German testers, 86.7% for US English testers). There was positive feedback regarding the general use and operation of the app, the structure of the manual, and the fact that the app also included videos. Some testers, however, complained that there was still some inconsistent recognition of keywords when using the chatbot.

As mentioned previously, the testers were generally convinced of the app's abilities. One of the testers put it in a nutshell: "I drive a (sports car) and such an app is a good help, especially if you don't want to have everyday questions answered. An app is definitely faster than browsing through the manual." The feedback provided by the testers also included recommendations that, from their point of view, could help improve the app and the chatbot, as well as the overall user experience, and thus the success of the digital manual. Among other things, the testers recommended modernizing the design of the app, to integrate an onboarding for the app user, and to improve the database further so that ultimately, all the keywords can be referenced in the manual. Jan Kittel was very satisfied with both the test procedure and the test results. He concludes, "Testing a chatbot implicates special requirements that could be met thanks to the use of crowdtesting. Of course, we appreciate the positive feedback from the testers, and their recommendations that are at least as important to us. They allow us to further optimize the app and finally offer our end users a product that has been designed and developed according to their needs."
About Testbirds

Testbirds specialises in the testing of software such as apps, websites and Internet of Things applications by using innovative technologies and solutions. The testing provider investigates software for user-friendliness and functionality issues. With over 300,000 registered testers located in 193 countries, Testbirds is one of the world’s leading crowd-testing providers. The company also utilises cloud based technologies to support customers in optimising their digital products. The combination of the two testing methods delivers a unique portfolio that takes the quality of software to the next level.

Testbirds was founded in 2011 by Philipp Benkler, Georg Hansbauer and Markus Steinhauser. Today, the company has over 100 employees. Other than its headquarter in Munich, there are now offices in Amsterdam, London and Stockholm, franchises in Hungary and Russia and sales partners in Italy and North America.

Would you like to know more?
Visit our website or contact us at info@testbirds.com