

TransPerfect supports client

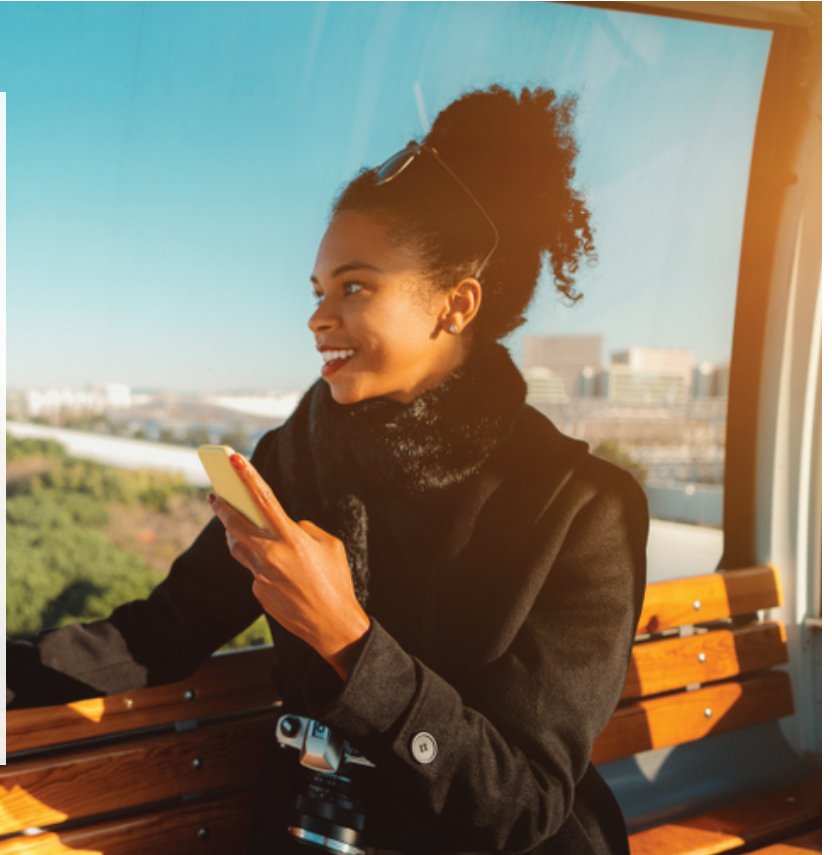
with large-scale, multilingual voice data collection to streamline their app localization process.



TRANSPERFECT
DATAFORCE

“We’re impressed with the success of the Korean and Mandarin real-world collection that TransPerfect has undertaken on our behalf.”

—Manager, Data Engineering



THE CHALLENGE

There is an increasing demand for solutions that voice-controlled smart products can deliver. To meet the end user's expectations, voice identification must be accurate, regardless of background noises or language and voice parameters.

To prevent possible bias, you need a large-scale audio collection of native speakers of the target languages covering various demographics and environments. Our client, an industry leader in the far-field speech and voice recognition market, did not have the necessary internal resources to execute such a complex data collection. Therefore, they requested TransPerfect's support in the Korean and Chinese markets.

THE SOLUTION

To expand the coverage of the machine-learning solution in Mandarin and Korean, TransPerfect resourced more than 500 participants from various demographic groups for each language. The participants were asked to complete ten recording sessions using a TransPerfect app on their mobile phone. These sessions were conducted at different locations and times of day to capture various background noises and voice parameters.

TransPerfect delivered the project in less than eight weeks. The audio data sets from the participants' recordings enabled the improvement of the client's audio/voice-recognition solutions.