Speech APIs for Automotive Application

Connecting Your World
One of our goals at AT&T is to expose via APIs the platforms we’ve built so they can be used as broadly as possible to create new, innovative apps and services. Rather than being confined to a single mobile device or operating system, our APIs enable developers to integrate our technologies into almost any Internet-connected equipment. One of our newest APIs is our Speech API for our AT&T Watson℠ speech engine. This software tool enables developers to plug our AT&T Watson℠ speech-to-text and text-to-speech capabilities into their applications without having to build that capability from scratch.

There are multiple use cases for these APIs (imagine being able to issue voice commands to your dishwasher), and one use case we’re experimenting with is for automotive applications. Developers and designers at the AT&T Foundry® in Plano, Texas, along with researchers at AT&T Labs, are currently exploring how these APIs could allow a driver to ask for directions to the nearest coffee shop, listen and respond to text messages, play music and more.

How did the idea hatch?
This project is part of our overall program to experiment with the most innovative ways to deploy our APIs in devices other than smart phones. Eventually we could expand our Speech API capabilities to enable access to a variety of in-car infotainment services, such as web access and video services for passengers. While this project uses the car as the test platform, these Speech API capabilities could eventually be integrated into a variety of devices.

The Future (art of the possible)
While AT&T is getting this project started, we also want to make it accessible for outside developers. Some of the future goals:

- Extend the technology from handheld devices to screens and interfaces built directly into the car.
- Offer real-time analytics and cloud hosting to make it easy to access and share data gathered by connected car apps and services.
- Refine the user interface to be more intuitive and predictive (Project Scrat).

About the Researchers
Connie Wang - After growing up in Beijing, China and graduating from Texas A&M, Connie started her career at AT&T (SBC Services back then) as an application developer. Before moving to the AT&T Foundry® as Principal Technical Architect, she worked on the Strategy Innovation and Prototype team in IT as a development manager for 5 years. Connie has worked on a wide variety of AT&T Foundry® projects prior to Speech APIs for Automotive Applications, including AT&T Toggle, an app which enables one mobile device to contain both personal and enterprise accounts, and the U-verse Interactive Demo for AT&T retail stores.

Michael Johnston, Ph.D. - Michael is a Principal member of technical staff at AT&T Labs. He has more than 21 years of experience in speech and language technology and has worked at the forefront of multimodal interface research for 15 years. He is
currently responsible for AT&T’s research program in advanced multimodal interfaces, holds 14 U.S. patents, has published over 50 technical papers, and currently serves as editor and chair of the W3C EMMA Multimodal standard.