

# Build your own Siri: An open-source digital assistant

March 11 2015

---

An open-source computing system you command with your voice like Apple's Siri is designed to spark a new generation of "intelligent personal assistants" for wearables and other devices. It could also lead to much-needed advancements in the datacenter infrastructure to support them.

Sirius, built by University of Michigan engineering researchers, is similar to Siri, Microsoft Cortana and Google Now—robust applications that accept voice instructions and questions, interpret them and answer in spoken words.

Sirius even uses many of the same "fancy algorithms," said Jason Mars, U-M assistant professor of computer science and engineering and co-director of Clarity Lab where Sirius was developed. But unlike its expensive and locked-down commercial counterparts, Sirius is free and can be customized.

"Now the core technology is out of the bag, and we all have access to it," Mars said. "Instead of making an app to run on the Apple Watch, for example, maybe I could make my own watch. We're very excited to see what the world comes together to build and learn with Sirius as a starting point."

The researchers will introduce their system March 14 at a technology conference in Istanbul. They'll release the software immediately afterwards.

Sirius is an end-to-end voice and vision [personal assistant](#). It bundles [speech recognition](#), image matching, [natural language processing](#) and a question-and-answer system that executes in the cloud. In its initial release, users can enter queries by talking to a device, or with a combination of speech and an image—maybe a photo of a restaurant with the question: "When does this place close?" That's not something current commercial systems can do.

"What we've done with Sirius is pushed the limits of the traditional intelligent personal assistant," said Johann Hauswald, a U-M doctoral student in Clarity Lab. "Not only can you interact with your voice but you can also ask questions about what you're seeing, which is a new way to interact with this type of device."

The demo version of Sirius is a talking Wikipedia. Researchers loaded it with a static version of the site and users can ask it factual questions.

This knowledge base could be swapped for any type of information that researchers or startups deem useful. Developers could make assistants that are experts in particular domains—medicine, cooking or auto repair, for instance. U-M researchers are working with IBM to develop one that could help with academic advising.

Mars describes Sirius as the Linux of intelligent personal assistants. Linux is a free computer operating system—a contemporary of Apple's OS X and Microsoft Windows. While it's used in just a sliver of desktops, it's said to have revolutionized computing. Linux has become the default way to run servers and mainframes and it's the foundation of Google's Android, now the most common operating system for tablets and smartphones.

To make Sirius, the researchers stitched together four established open-source projects that they say rely on techniques and algorithms that

resemble those in commercial systems.

Speech recognition came from Carnegie Mellon University's Sphinx, Microsoft Research's Kaldi and Germany's RWTH Aachen "RASR" project. The question-and-answer system came from OpenEphyra, which laid the foundation for IBM's Jeopardy-winning Watson. Image recognition came from the SURF computer-vision algorithm behind Swiss tech entrepreneur Herbert Bay's company Kooaba, which was recently acquired by Qualcomm.

Mars sees Sirius as an important platform for research into the development of next-generation warehouse computing. It gives researchers a testbed for studying how the data centers that process voice-enabled queries should evolve to keep up with escalating pressure from wearable gadgets. Wearables will rely heavily on voice and image input, and by 2018, sales in the category are projected at 485 million devices each year.

Most of the work to answer these sound-based requests happens in the cloud. Smartphone-based assistants can accept commands or questions and translate them to text. But it's cloud-based software that figures out what the text means, searches for answers, picks the best one and sends it back to the mobile device.

This process, the researchers found, can be more than 100 times more computationally intensive than a simple text web search. They calculated that if voice were to supplant text for web queries, data-center infrastructure would need to grow by 165 times.

"We have to think of new ways to redesign our cloud platforms to support this type of workload," Mars said.

The researchers tested several types of processors datacenter operators

could consider adding to their infrastructure. They found that GPUs—processing units originally developed for graphics but that have proven useful for other applications—could speed up data centers by 10 times. And FPGAs—field-programmable gate arrays that are used for custom applications—could speed them up by 16 times.

The researchers expect their new insights will interest industry, which is also trying to answer these questions.

"Some people ask whether speech or visual-driven computer interaction is just hype or the next big thing, and I truly believe it's the natural trend," said Lingjia Tang, U-M assistant professor of computer science and engineering, and co-director of Clarity Lab. "I think in the future we will communicate with computers more like how we communicate with humans."

The researchers will present their paper titled "Sirius: An Open End-to-End Voice and Vision Personal Assistant and Its Implications for Future Warehouse Scale Computers" March 14 at the International Conference on Architectural Support for Programming Languages and Operating Systems.

**More information:** Sirius: An Open End-to-End Voice and Vision Personal Assistant and Its Implications for Future Warehouse Scale Computers, [jasonmars.org/wp-content/paper ... hauswald15asplos.pdf](https://jasonmars.org/wp-content/paper...hauswald15asplos.pdf)

Provided by University of Michigan

Citation: Build your own Siri: An open-source digital assistant (2015, March 11) retrieved 24 April 2024 from <https://phys.org/news/2015-03-siri-open-source-digital.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.