

North South University

Department of Electrical and Computer Engineering

School of Engineering and Physical Sciences

CSE323.8

Assignment

# **Title: Siri: The ultimate guide**

**Submitted by:**

|  |  |
| --- | --- |
| **Name** | **ID** |
| Md Sharif Hossain | 1712336642 |

**Submitted to:**

**Faculty:** RMZ1

**Date:** 11th July, 2019

“Everything you need to know about setting up and using Siri, Apple's intelligent personal assistant for iPhone, iPad, Apple TV, Apple Watch, and the Mac. Siri does more than ever, even before you ask”

**Siri** is a virtual assistant that is part of [Apple Inc.](https://en.wikipedia.org/wiki/Apple_Inc.)'s [iOS](https://en.wikipedia.org/wiki/IOS), [iPadOS](https://en.wikipedia.org/wiki/IPadOS), [watchOS](https://en.wikipedia.org/wiki/WatchOS), [macOS](https://en.wikipedia.org/wiki/MacOS), [tvOS](https://en.wikipedia.org/wiki/TvOS) and [audioOS](https://en.wikipedia.org/wiki/AudioOS) [operating systems](https://en.wikipedia.org/wiki/Operating_system). The assistant uses voice queries and a [natural-language user interface](https://en.wikipedia.org/wiki/Natural-language_user_interface) to answer questions, make recommendations, and perform actions by delegating requests to a set of [Internet](https://en.wikipedia.org/wiki/Internet) services. The software adapts to users' individual language usages, searches, and preferences, with continuing use. Returned results are individualized.

Siri is a [spin-off](https://en.wikipedia.org/wiki/Corporate_spin-off) from a project originally developed by the [SRI International](https://en.wikipedia.org/wiki/SRI_International) Artificial Intelligence Center. Its [speech recognition](https://en.wikipedia.org/wiki/Speech_recognition) engine was provided by [Nuance Communications](https://en.wikipedia.org/wiki/Nuance_Communications), and Siri uses advanced [machine learning](https://en.wikipedia.org/wiki/Machine_learning) technologies to function. Its original American, British, and Australian [voice actors](https://en.wikipedia.org/wiki/Voice_acting) recorded their respective voices around 2005, unaware of the recordings' eventual usage in Siri. The voice assistant was released as an app for iOS in February 2010, and it was acquired by Apple two months later. Siri was then integrated into [iPhone 4S](https://en.wikipedia.org/wiki/IPhone_4S) at its release in October 2011. At that time, the separate app was also removed from the iOS [App Store](https://en.wikipedia.org/wiki/App_Store_(iOS)). Siri has since become an integral part of Apple's products, having been adapted into other hardware devices over the years, including newer [iPhone](https://en.wikipedia.org/wiki/IPhone) models, as well as [iPad](https://en.wikipedia.org/wiki/IPad), [iPod Touch](https://en.wikipedia.org/wiki/IPod_Touch), [Mac](https://en.wikipedia.org/wiki/Macintosh), [AirPods](https://en.wikipedia.org/wiki/AirPods), [Apple TV](https://en.wikipedia.org/wiki/Apple_TV), and [HomePod](https://en.wikipedia.org/wiki/HomePod).

Siri supports a wide range of user commands, including performing phone actions, checking basic information, scheduling events and reminders, handling device settings, searching the Internet, navigating areas, finding information on entertainment, and is able to engage with iOS-integrated apps. With the release of [iOS 10](https://en.wikipedia.org/wiki/IOS_10) in 2016, Apple opened up limited third-party access to Siri, including third-party messaging apps, as well as payments, [ride-sharing](https://en.wikipedia.org/wiki/Ride-sharing), and [Internet calling](https://en.wikipedia.org/wiki/Voice_over_IP) apps. With the release of [iOS 11](https://en.wikipedia.org/wiki/IOS_11), Apple updated Siri's voices for more clear, human voices, started supporting follow-up questions and language translation, and additional third-party actions.

Siri's original release on [iPhone 4S](https://en.wikipedia.org/wiki/IPhone_4S) in 2011 received mixed reviews. It received praise for its [voice recognition](https://en.wikipedia.org/wiki/Speech_recognition) and contextual knowledge of user information, including calendar appointments, but was criticized for requiring stiff user commands and having a lack of flexibility. It was also criticized for lacking information on certain nearby places, and for its inability to understand certain [English accents](https://en.wikipedia.org/wiki/English_accents). In 2016 and 2017, a number of media reports have indicated that Siri is lacking in innovation, particularly against new competing voice assistants from other technology companies. The reports concerned Siri's limited set of features, "bad" voice recognition, and undeveloped service integrations as causing trouble for Apple in the field of [artificial intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence) and cloud-based services; the basis for the complaints reportedly due to stifled development, as caused by Apple's prioritization of user [privacy](https://en.wikipedia.org/wiki/Privacy) and executive power struggles within the company.

## **Development:**

Siri is a [spin-out](https://en.wikipedia.org/wiki/Spin-out) from the [SRI International](https://en.wikipedia.org/wiki/SRI_International) Artificial Intelligence Center, and is an offshoot of the US [Defense Advanced Research Projects Agency](https://en.wikipedia.org/wiki/DARPA)'s (DARPA)-funded [CALO](https://en.wikipedia.org/wiki/CALO) project. It was co-founded by Dag Kittlaus, [Tom Gruber](https://en.wikipedia.org/wiki/Tom_Gruber), and [UCLA](https://en.wikipedia.org/wiki/UCLA) alumnus [Adam Cheyer](https://en.wikipedia.org/wiki/Adam_Cheyer).

Siri's [speech recognition](https://en.wikipedia.org/wiki/Speech_recognition) engine was provided by [Nuance Communications](https://en.wikipedia.org/wiki/Nuance_Communications), a speech technology company. This was acknowledged by neither Apple nor Nuance for years, until Nuance CEO Paul Ricci confirmed the information at a 2011 technology conference. The speech recognition system makes use of sophisticated [machine learning](https://en.wikipedia.org/wiki/Machine_learning) techniques, including [convolutional neural networks](https://en.wikipedia.org/wiki/Convolutional_neural_network) and [long short-term memory](https://en.wikipedia.org/wiki/Long_short-term_memory).

The initial Siri prototype was implemented using the Active platform, a joint project between the Artificial Intelligence Center of [SRI International](https://en.wikipedia.org/wiki/SRI_International) and the Vrai Group at [Ecole Polytechnique Fédérale de Lausanne](https://en.wikipedia.org/wiki/Ecole_Polytechnique_F%C3%A9d%C3%A9rale_de_Lausanne). The Active platform was the focus of a PhD thesis led by Didier Guzzoni, who joined Siri as its chief scientist.

Apple's first notion of a digital personal assistant was originally a concept video in 1987, called the [Knowledge Navigator](https://en.wikipedia.org/wiki/Knowledge_Navigator).

### **Voices:**

The original American voice of Siri was provided in July 2005 by [Susan Bennett](https://en.wikipedia.org/wiki/Susan_Bennett), unaware that it would eventually be used for the voice assistant. A report from [*The Verge*](https://en.wikipedia.org/wiki/The_Verge) in September 2013 about voice actors, their work, and machine learning developments, made hints that Allison Dufty was the voice behind Siri, though this was disproven when Dufty wrote on her website that she was "absolutely, positively *not* the voice of Siri." Citing growing pressure, Bennett revealed her role as Siri in October, and her claim was proven by Ed Primeau, an American [audio forensics](https://en.wikipedia.org/wiki/Audio_forensics) expert. Apple has never confirmed the information.

The original British male voice was provided by [Jon Briggs](https://en.wikipedia.org/wiki/Jon_Briggs), a former technology journalist. Having discovered that he was the voice of Siri by watching television, he first spoke only about his role in November 2011, also acknowledging his voice work was done "five or six years ago" without knowing the recordings' final usage form.

The original Australian voice was provided by [Karen Jacobsen](https://en.wikipedia.org/wiki/Karen_Jacobsen), a [voice-over](https://en.wikipedia.org/wiki/Voice-over) artist known in Australia for her work as the [*GPS*](https://en.wikipedia.org/wiki/GPS) *girl*.

As part of an interview between all three voice actors and [*The Guardian*](https://en.wikipedia.org/wiki/The_Guardian), Briggs stated that "the original system was recorded for a US company called Scansoft, who were then bought by Nuance. Apple simply licensed it."

With [iOS 11](https://en.wikipedia.org/wiki/IOS_11), Apple auditioned hundreds of candidates to find a new female voice, then recorded hours of speech, including different personalities and expressions, and built a new [text-to-speech](https://en.wikipedia.org/wiki/Speech_synthesis) voice based on [deep learning](https://en.wikipedia.org/wiki/Deep_learning) technology.

## **Integration:**

Siri was originally released as a [stand-alone application](https://en.wikipedia.org/wiki/Software) for the [iOS](https://en.wikipedia.org/wiki/IOS) operating system in February 2010, and at the time, the developers were also intending to release Siri for [Android](https://en.wikipedia.org/wiki/Android_(operating_system)) and [BlackBerry](https://en.wikipedia.org/wiki/BlackBerry_OS) devices. Two months later, Apple acquired Siri. On October 4, 2011, Apple introduced the [iPhone 4S](https://en.wikipedia.org/wiki/IPhone_4S) with a [beta version](https://en.wikipedia.org/wiki/Beta_version) of Siri. After the announcement, Apple removed the existing standalone Siri app from [App Store](https://en.wikipedia.org/wiki/App_Store_(iOS)). [*TechCrunch*](https://en.wikipedia.org/wiki/TechCrunch) wrote that, despite the Siri app's support for [iPhone 4](https://en.wikipedia.org/wiki/IPhone_4), its removal from App Store might also have had a financial aspect for the company, in providing an incentive for customers to upgrade devices. Third-party developer Steven Troughton-Smith, however, managed to [port](https://en.wikipedia.org/wiki/Porting) Siri to iPhone 4, though without being able to communicate with Apple's servers. A few days later, Troughton-Smith, working with an anonymous person nicknamed "Chpwn", managed to fully hack Siri, enabling its full functionalities on iPhone 4 and [iPod Touch](https://en.wikipedia.org/wiki/IPod_Touch) devices. Additionally, developers were also able to successfully create and distribute legal ports of Siri to any device capable of running [iOS 5](https://en.wikipedia.org/wiki/IOS_5), though a [proxy server](https://en.wikipedia.org/wiki/Proxy_server) was required for Apple server interaction.

[**Siri Remote**](https://en.wikipedia.org/wiki/Siri_Remote)**:**

Over the years, Apple has expanded the line of officially supported products, including newer [iPhone](https://en.wikipedia.org/wiki/IPhone) models, as well as [iPad](https://en.wikipedia.org/wiki/IPad) support in June 2012, iPod Touch support in September 2012, [Apple TV](https://en.wikipedia.org/wiki/Apple_TV) support, and the stand-alone Siri Remote, in September 2015, Mac and AirPods support in September 2016, and [HomePod](https://en.wikipedia.org/wiki/HomePod) support in February 2018.

Venture capital firm Mangrove Capital Partners predicted that Apple will launch a SiriOS at its developer conference in 2020 to further grow the Siri ecosystem. SiriOS could rival something like Amazon’s Alexa Skills platform, which makes it easy for developers to implement Alexa functionality. Apple currently offers SiriKit to developers, but one possibility is that a SiriOS could work across iOS, iPadOS, and macOS with ease.

## **Features and options:**

Apple offers a wide range of voice commands to interact with Siri, including, but not limited to:

* Phone and Text actions, such as "Call Sarah", "Read my new messages", "Set the timer for 10 minutes", and "Send email to mom"
* Check basic information, including "What's the weather like today?" and "How many dollars are in a Euro?"
* Schedule events and reminders, including "Schedule a meeting" and "Remind me to"
* Handle device settings, such as "Take a picture", "Turn off [Wi-Fi](https://en.wikipedia.org/wiki/Wi-Fi)", and "Increase the brightness"
* Search the Internet, including "Define...", "Find pictures of...", and "Search [Twitter](https://en.wikipedia.org/wiki/Twitter) for..."
* Navigation, including "Take me home", "What's traffic like on the way home?", and "Find driving directions to..."
* Translate words and phrases from English to a few languages, such as "How do I say where is the nearest hotel in French"
* Entertainment, such as "What basketball games are on today?", "What are some movies playing near me?", and "What's the synopsis of...?"
* Engage with iOS-integrated apps, including "Pause [Apple Music](https://en.wikipedia.org/wiki/Apple_Music)" and "Like this song"
* Handle payments through Apple Pay, such as "Apple Pay 25 dollars to Mike for concert tickets" or "Send 41 dollars to Ivana."

Siri also offers numerous pre-programmed responses to amusing questions. Such questions include "What is the meaning of life?" to which Siri may reply "All evidence to date suggests it's chocolate"; "Why am I here?", to which it may reply "I don't know. Frankly, I've wondered that myself"; and "Will you marry me?", to which it may respond with "My [End User Licensing Agreement](https://en.wikipedia.org/wiki/End-user_license_agreement) does not cover marriage. My apologies."

Initially limited to female voices, Apple announced in June 2013 that Siri would feature a gender option, adding a male voice counterpart.

In September 2014, Apple added the ability for users to speak "Hey Siri" to enable the assistant without the requirement of physically handling the device.

In September 2015, the "Hey Siri" feature was updated to include individualized voice recognition, a presumed effort to prevent non-owner activation.

With the announcement of [iOS 10](https://en.wikipedia.org/wiki/IOS_10) in June 2016, Apple opened up limited third-party developer access to Siri through a dedicated [application programming interface](https://en.wikipedia.org/wiki/Application_programming_interface) (API). The API restricts usage of Siri to engaging with third-party messaging apps, payment apps, ride-sharing apps, and Internet calling apps.

In [iOS 11](https://en.wikipedia.org/wiki/IOS_11), Siri is able to handle follow-up questions, supports language translation, and opens up to more third-party actions, including task management. Additionally, users are able to type to Siri, and a new, privacy-minded "on-device learning" technique improves Siri's suggestions by privately analyzing personal usage of different iOS applications.

## **Technology behind Siri:**

Siri works on primarily 2 technologies – Speech Recognition and Natural Language Processing. Speech Recognition is the task of converting a human speech into its corresponding textual form. For instance, when you trigger Siri by saying “Hey Siri”, in the back-end, a powerful speech recognition system by Apple kicks off and converts your audio into its corresponding textual form – “Hey Siri.” This is an extremely challenging task simply because we humans have a highly diverse set of tones as well as accents. The accents vary not only across countries, but also across states/cities within a country. Some people speak fast, some speak slowly. Characteristics of male and female voices are also very different.

The engineers at Apple train Machine Learning models on large, transcribed datasets in order to create efficient speech recognition models for Siri. These models are trained with highly diverse datasets that comprise of the voice samples of a large group of people. This way, Siri is able to cater to various accents.

In the recent years, deep learning has proven to produce phenomenal results in speech recognition. The word error rate of speech recognition engines has drastically gone down to less than 10%. This has been possible due to the availability of not only large datasets, but also powerful hardware using speech recognition algorithms that can be trained on the datasets.

Once Siri has understood what you are saying, the converted text is sent to Apple servers for further processing. Apple servers then run Natural Language Processing (NLP) algorithms on this text to understand the intent of what the user is trying to say. For instance, the NLP engines are able to differentiate that when a user is saying “set an alarm for 7AM tomorrow,” the user is asking about setting an alarm and not about making a call. This is challenging because different users speak the same sentence in different ways. For instance, one can say the same thing in the following ways:

* Hey Siri, can you set me an alarm for 7AM tomorrow?
* Siri, can you wake me up tomorrow at 7AM?
* Siri, please set an alarm for tomorrow at 7AM.
* Siri, please wake me up tomorrow at 7AM.

These are just a few right ways of telling Siri to set an alarm. Some people may speak grammatically incorrect sentences – “Siri alarm set me tomorrow at 7AM”. As a result, the intent analysis becomes very challenging. Just like speech recognition, intent analysis also requires a lot of data in order to train Natural Language Processing algorithms. Only when the dataset provided is huge is it the case that Siri is able to generalize and capture the variations of the same sentence that it has never seen. This makes the whole processes an extremely difficult task. In order to accomplish such mammoth tasks, Apple hires top-notch software engineers that have years of experience in Artificial Intelligence, Machine Learning, and Natural Language Processing.

These are just 2 of the most fundamental challenges. Another important technology behind Siri that employs Machine Learning is that of contextual understanding. You can talk to Siri like you are talking to a human:

You: Hey Siri, set an alarm.Siri: What time do you want me to set an alarm?You: 7 AM.

In the last sentence, when you said “7 AM”, Siri was able to understand and correlate that this 7 AM is a continuation of the last message where you asked it to set an alarm.

One final technology that Siri employs in this whole process is that of entity extraction. When you ask Siri to set an alarm for tomorrow at 7AM, Siri not only understands the meaning of your sentence, but also it automatically picks up entities from the sentence – 7AM and tomorrow.

## **Final Words:**

Overall, Siri is based on large-scale Machine Learning systems that employ 2 main aspects of data science – Speech Recognition and Natural Language Processing.

Amazing isn’t it? In our technology-heavy world, we tend to take things around us for granted. However, when we try to uncover them, we realize that there is a lot of technological magic happening behind the scene.

**References:**

1. <https://appleinsider.com/articles/siri>
2. <https://www.apple.com/siri/>
3. <https://en.wikipedia.org/wiki/Siri>
4. <https://www.imore.com/siri>
5. <https://www.lifewire.com/what-is-siri-help-1994303>
6. <https://magoosh.com/data-science/siri-work-science-behind-siri/>