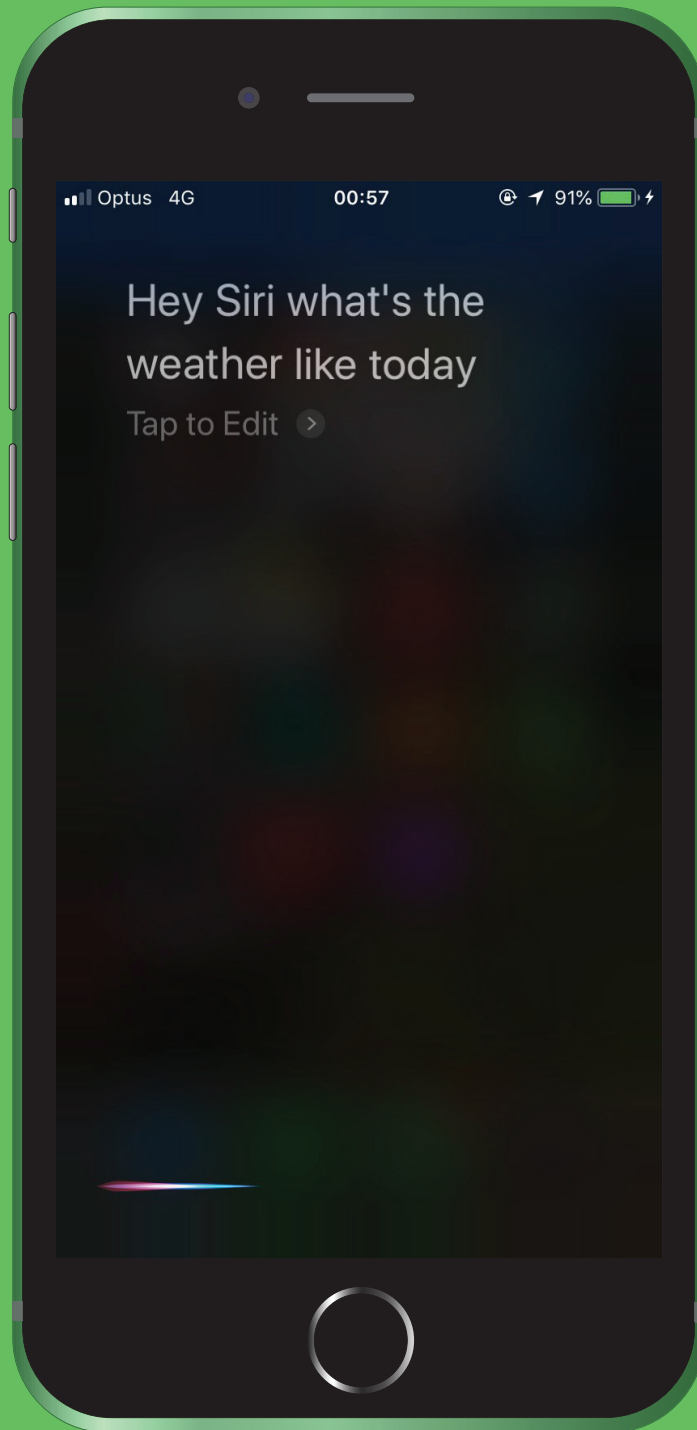


Kim Nguyen

Hey Siri

What's the weather like today?



Hey Siri

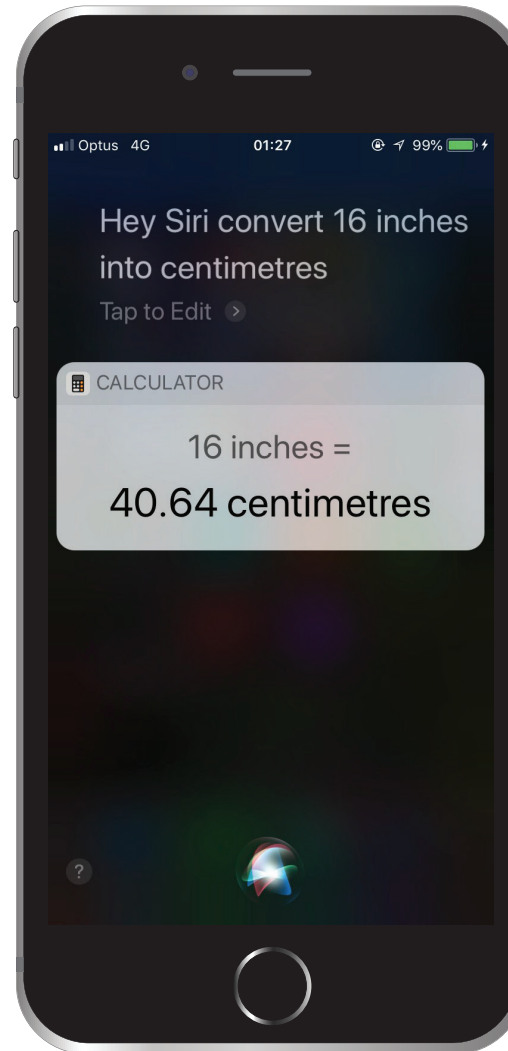
Kim Nguyen

Siri is Apple's voice automated personal assistant. She was born in 2011 with the launch of Apple's iPhone 4S. Siri changed what people thought phones could do, and the way we interact with our phones.

Susan Bennett was the female American voice of Siri from its service launch on October 4th, 2011 until September 18th, 2013 when the iOS 7 update was released. In an interview with CNN on October 4th, 2013 she finally revealed herself as the original voice of Siri.

ScanSoft (now merged with Nuance, a computer software company) found Bennett through GM Voices (a telecommunications company) for a speech construction project. For the entire month of July of 2005, she recorded for over four hours a day, repeating phrases and sentences that made no particular sense.

Bennett only found out what had become of her long lost recordings when a colleague emailed her in October 2011 when they were playing with the newly launched iPhone 4S.



The process behind making Siri can be explained more simply through the following process:

1. A voice talent comes into the recording studio and their flow of speech and phrases they read are captured.
2. The recordings are analysed and pulled apart to get the right sounds from them.
3. After an automation process by the computer, the sounds are segmented, and some sound fragments are hand corrected to make them more appropriate.

At the time this project was carried out, people wanted more expressivity in synthetic voices, so the manual corrections were necessary to help create this unique persona.

Siri is iconic because she is the first concatenated voice that really sounded human, was interactive and had a personality.

Although Siri is not limited to the following capabilities, they are worth mentioning in exploring the diversity of Siri's ability to enhance your quality of life.

1.Siri can perform unit conversions, such as inches into centimetres or pounds into kilograms.

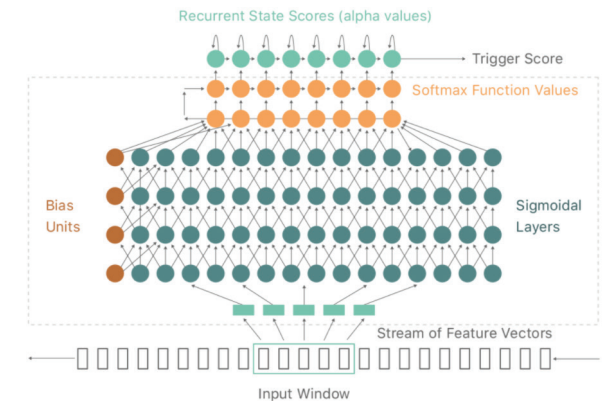
2.You can ask Siri to take you places. For example, saying: "Hey Siri, take me to Western Sydney University Parramatta South Campus" will open maps and navigate you there from your current location.

3.Instead of always saying a full contact name, you can shorten it to their relationship with you. For example, to setup the command "Hey Siri, call my husband", you will need to set a relationship with your husband's contact through Siri. You can do this by saying: "Hey Siri, John Doe is my husband".

4.Siri can set reminders, alarms, appointments and timers.

Our ability to use the "Hey Siri" function handsfree is enabled by a speech recogniser constantly running to listen to those two words. This speech recogniser uses a Deep Neural Network (DNN) to convert your voice pattern into a probability distribution of speech sounds. It then computes a validity score that the phrase you uttered was "Hey Siri". If this score meets the validity threshold then Siri wakes up. Siri then communicates with 'the cloud' (resources on the internet, in this specific case: Apple's servers) to process your speech. Apple's servers run Natural Language Processing (NLP) algorithms to understand the intent of your speech, in which then Siri will respond to your command or query.

Siri wasn't the first speech recognition system, but she is the first to revolutionise smart phone speech recognition. Because of her ability to understand human speech by using context, history and artificial intelligence, Siri remains the powerhouse of Apple's artificial intelligence available for consumers.



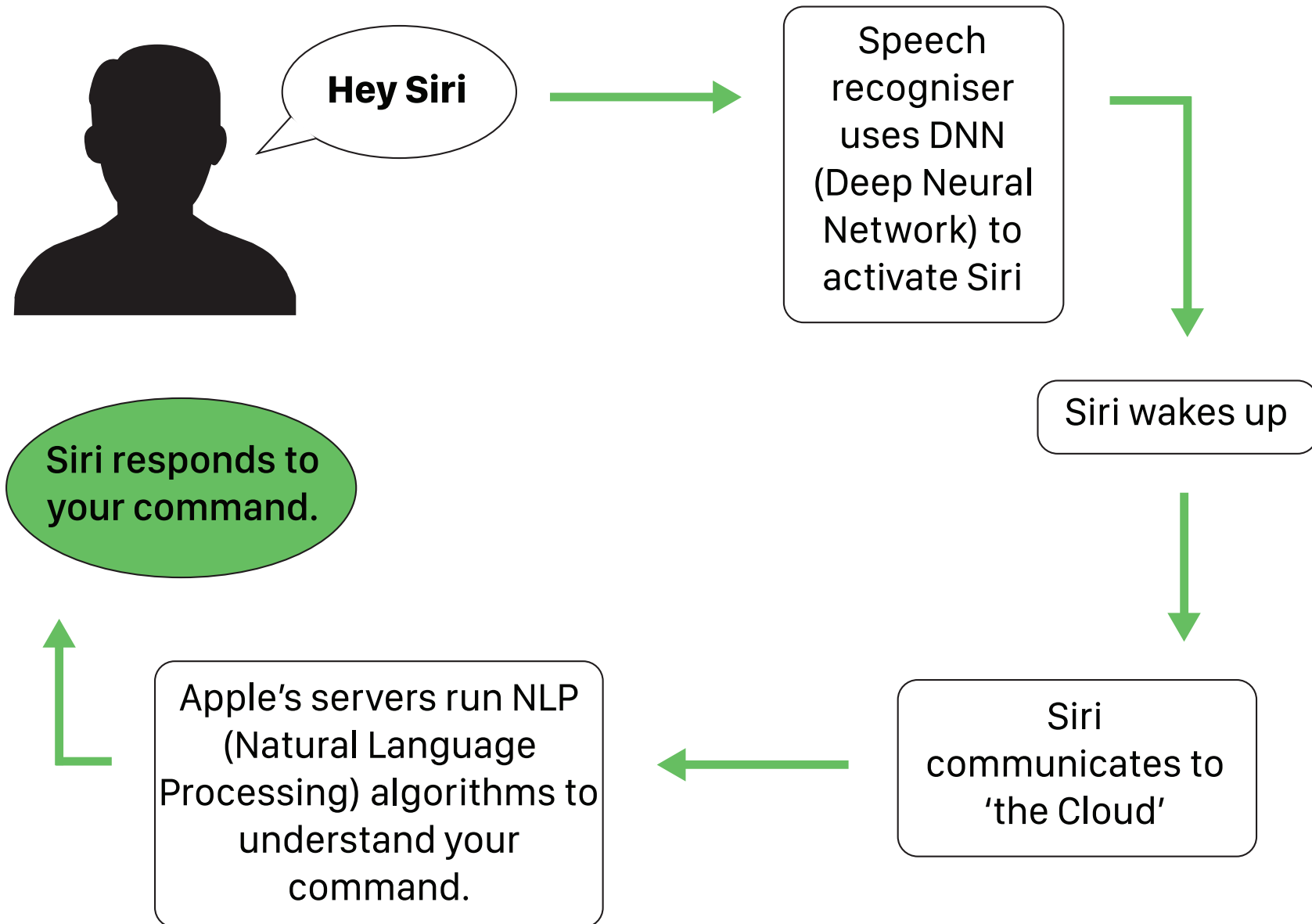
References:

- Barrows, M. (Producer). (2016, November 1). Siri [Audio podcast]. Twenty Thousand Hertz. Retrieved from <https://www.20k.org/episodes/siri>
- Hein, C. (2017, October 18). Apple reveals ai magic behind 'hey Siri'. Cult of Mac. Retrieved from <https://www.cultofmac.com/509163/apple-reveals-ai-magic-behind-hey-siri>
- Hof, R. (n.d.). Retrieved October 22, 2018, from <https://www.technologyreview.com/s/513696/deep-learning>

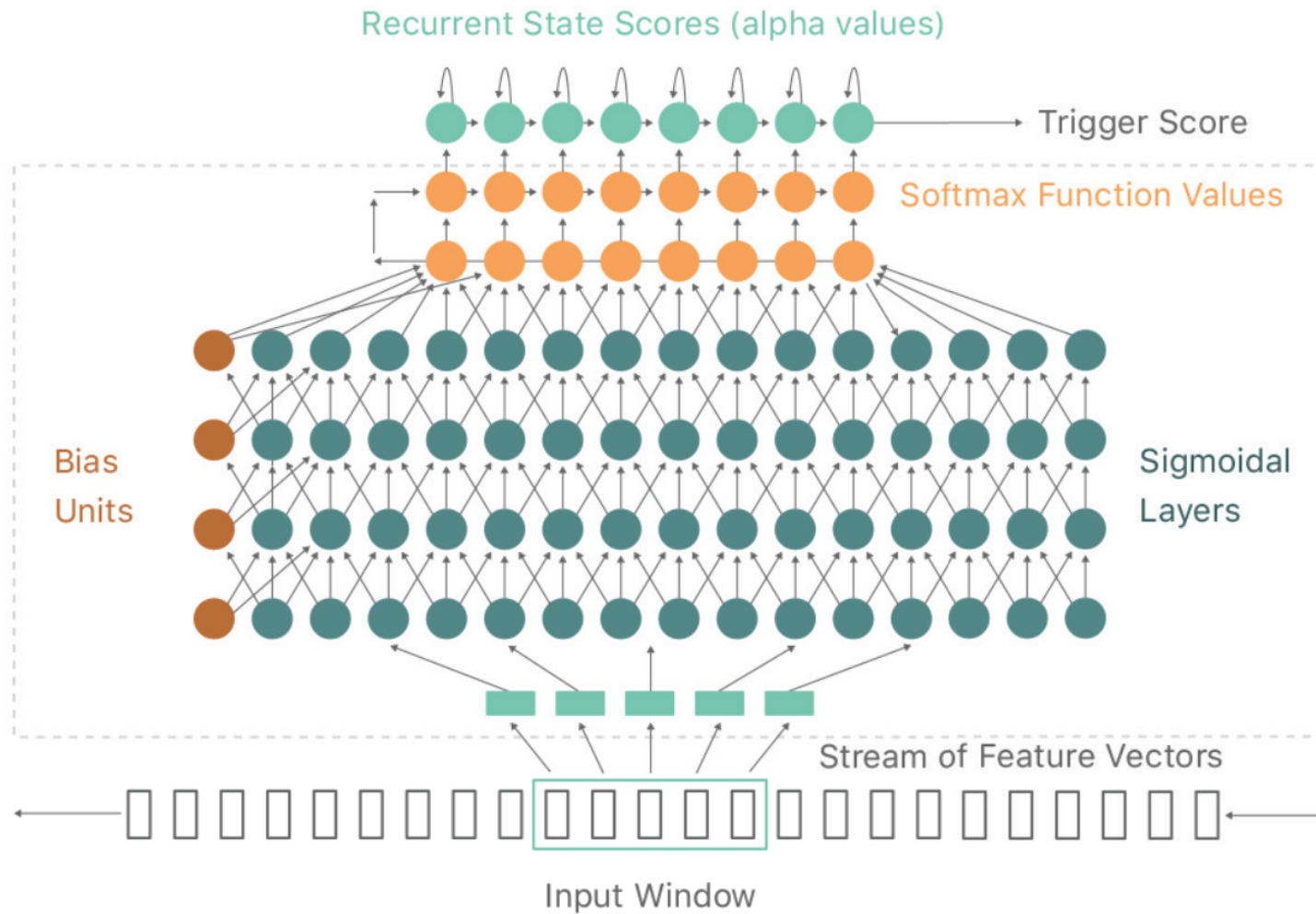
Siri Team (2017). Hey Siri: An on-device dnn-powered voice trigger for Apple's personal assistant. Apple Machine Learning Journal, Vol. 1, Issue 6. <https://machinelearning.apple.com/2017/10/01/hey-siri.html>

Siri Team (2017). Deep learning for Siri's voice: On-device deep mixture density networks for hybrid unit selection synthesis. Apple Machine Learning Journal, Vol. 1, Issue 4. <https://machinelearning.apple.com/2017/08/06/siri-voices.html>

How Siri Works



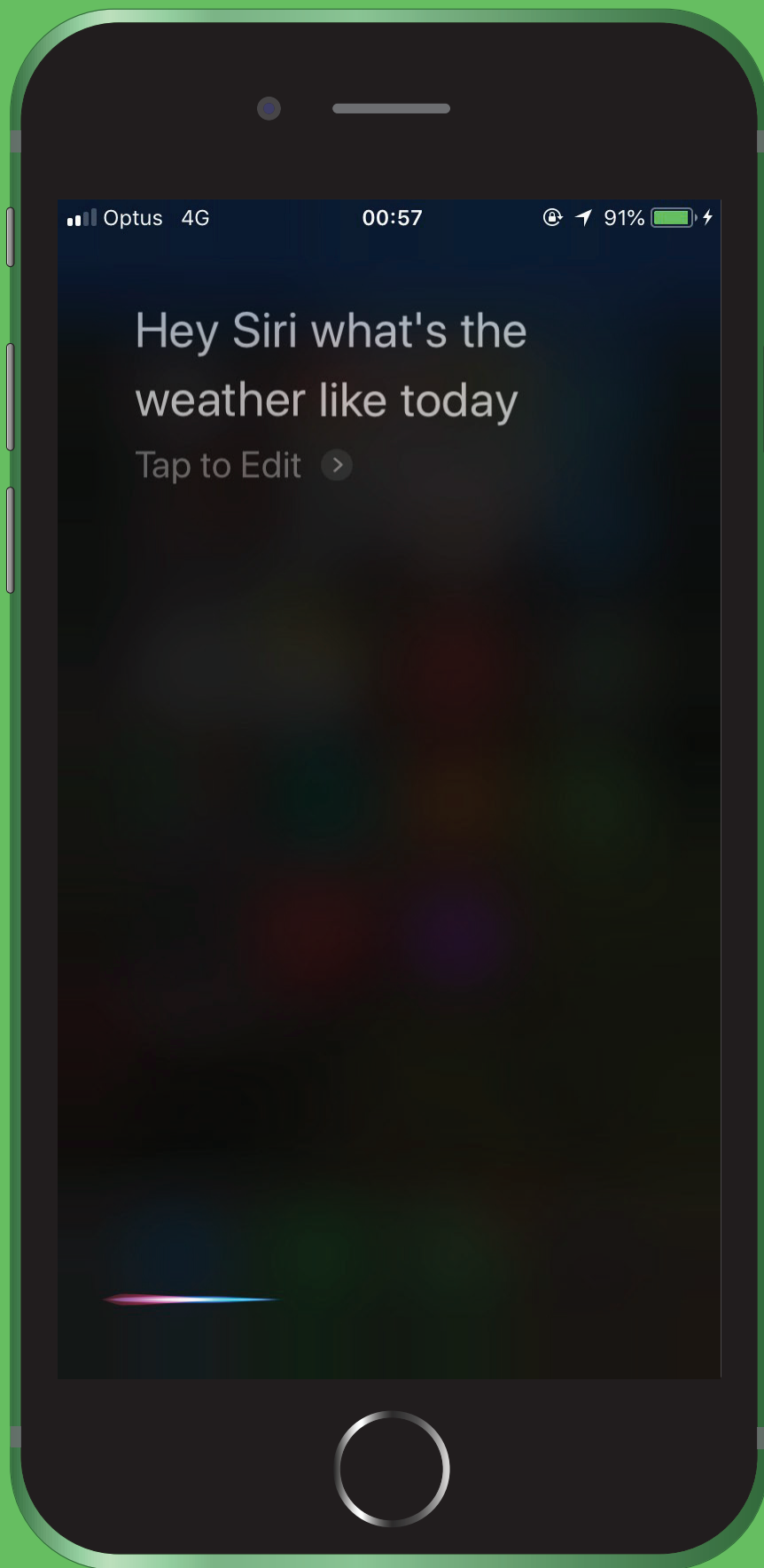
The Deep Neural Network used to detect "Hey Siri." The hidden layers are actually fully connected. The top layer performs temporal integration. The actual DNN is indicated by the dashed box.



Kim Nguyen

Hey Siri

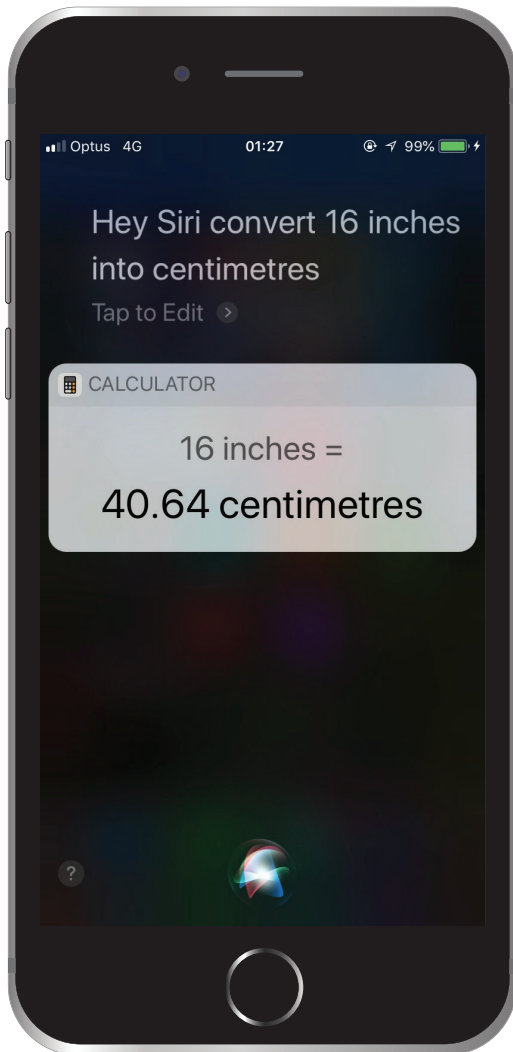
What's the weather like today?



Hey Siri

Kim Nguyen

Siri is Apple's voice automated personal assistant. She was born in 2011 with the launch of Apple's iPhone 4S. Siri changed what people thought phones could do, and the way we interact with our phones.



Susan Bennett was the female American voice of Siri from its service launch on October 4th, 2011 until September 18th, 2013 when the iOS 7 update was released. In an interview with CNN on October 4th, 2013 she finally revealed herself as the original voice of Siri.

ScanSoft (now merged with Nuance, a computer software company) found Bennett through GM Voices (a telecommunications company) for a speech construction project.

For the entire month of July of 2005, she recorded for over four hours a day, repeating phrases and sentences that made no particular sense.

Bennett only found out what had become of her long lost recordings when a colleague emailed her in October 2011 when they were playing with the newly launched iPhone 4S.

The process behind making Siri can be explained more simply through the following process:

1. A voice talent comes into the recording studio and their flow of speech and phrases they read are captured.
2. The recordings are analysed and pulled apart to get the right sounds from them.
3. After an automation process by the computer, the sounds are segmented, and some sound fragments are hand corrected to make them more appropriate.

At the time this project was carried out, people wanted more expressivity in synthetic voices, so the manual corrections were necessary to help create this unique persona.

Siri is iconic because she is the first concatenated voice that really sounded human, was interactive and had a personality.

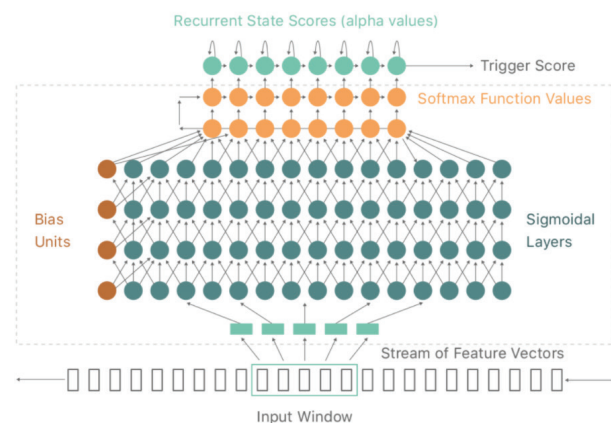
Although Siri is not limited to the following capabilities, they are worth mentioning in exploring the diversity of Siri's ability to enhance your quality of life.

1. Siri can perform unit conversions, such as inches into centimetres or pounds into kilograms.
2. You can ask Siri to take you places. For example, saying: "Hey Siri, take me to Western Sydney University Parramatta South Campus" will open maps and navigate you there from your current location.
3. Instead of always saying a full contact name, you can shorten it to their relationship with you.

For example, to setup the command “Hey Siri, call my husband”, you will need to set a relationship with your husband’s contact through Siri. You can do this by saying: “Hey Siri, John Doe is my husband”.

4.Siri can set reminders, alarms, appointments and timers.

Our ability to use the “Hey Siri” function handsfree is enabled by a speech recogniser constantly running to listen to those two words. This speech recogniser uses a Deep Neural Network (DNN) to convert your voice pattern into a probability distribution of speech sounds. It then computes a validity score that the phrase you uttered was “Hey Siri”. If this score meets the validity threshold then Siri wakes up. Siri then communicates with ‘the cloud’ (resources on the internet, in this specific case: Apple’s servers) to process your speech. Apple’s servers run Natural Language Processing (NLP) algorithms to understand the intent of your speech, in which then Siri will respond to your command or query.

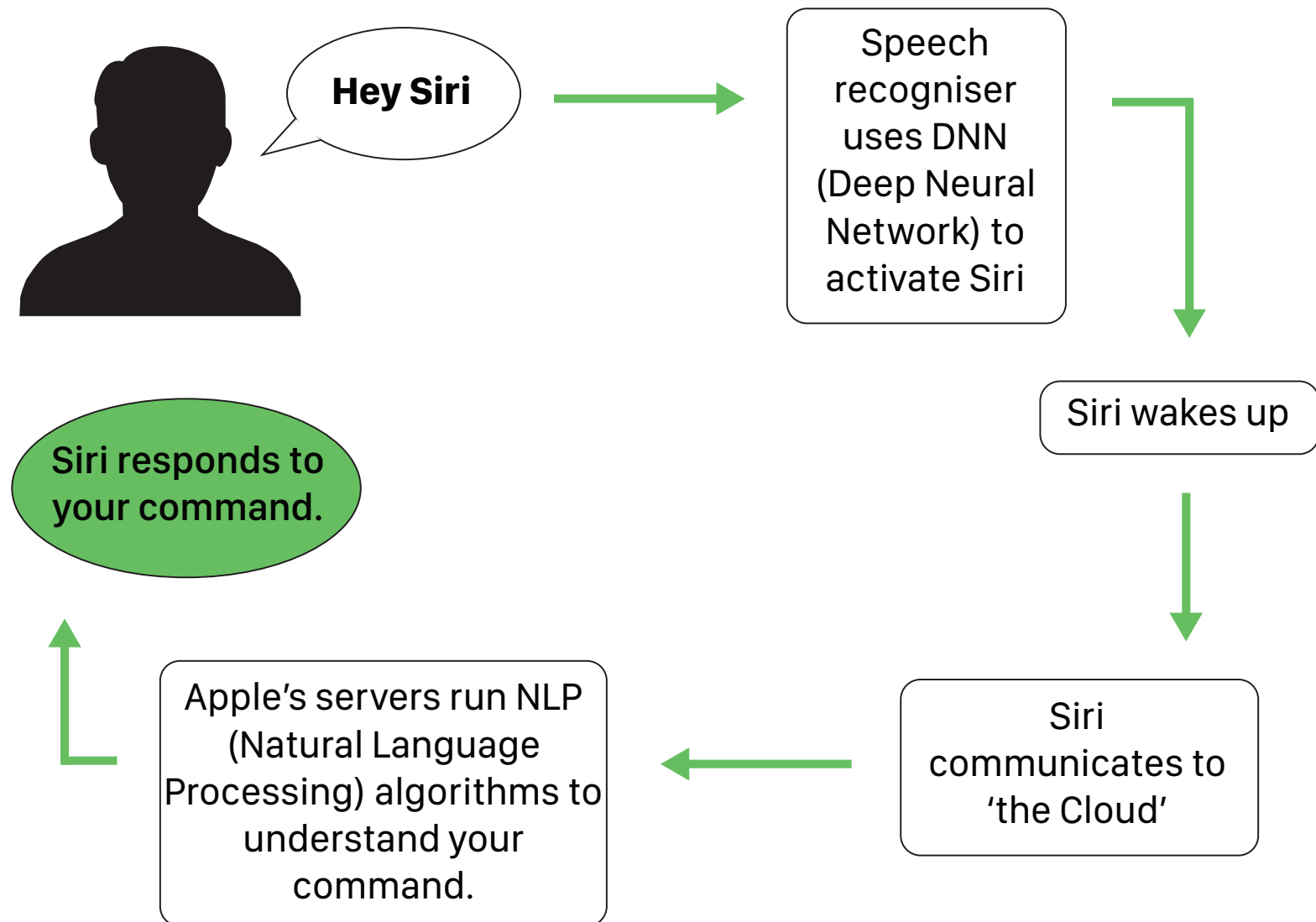


Siri wasn’t the first speech recognition system, but she is the first to revolutionise smart phone speech recognition. Because of her ability to understand human speech by using context, history and artificial intelligence, Siri remains the powerhouse of Apple’s artificial intelligence available for consumers.

References:

- Barrows, M. (Producer). (2016, November 1). Siri [Audio podcast]. Twenty Thousand Hertz. Retrieved from <https://www.20k.org/episodes/siri>
- Hein, C. (2017, October 18). Apple reveals ai magic behind ‘hey Siri’. Cult of Mac. Retrieved from <https://www.cultofmac.com/509163/apple-reveals-ai-magic-behind-hey-siri>
- Hof, R. (n.d.). Retrieved October 22, 2018, from <https://www.technologyreview.com/s/513696/deep-learning>
- Siri Team (2017). Hey Siri: An on-device dnn-powered voice trigger for Apple’s personal assistant. Apple Machine Learning Journal, Vol. 1, Issue 6. <https://machinelearning.apple.com/2017/10/01/hey-siri.html>
- Siri Team (2017). Deep learning for Siri’s voice: On-device deep mixture density networks for hybrid unit selection synthesis. Apple Machine Learning Journal, Vol. 1, Issue 4. <https://machinelearning.apple.com/2017/08/06/siri-voices.html>

How Siri Works



The Deep Neural Network used to detect "Hey Siri." The hidden layers are actually fully connected. The top layer performs temporal integration. The actual DNN is indicated by the dashed box.

