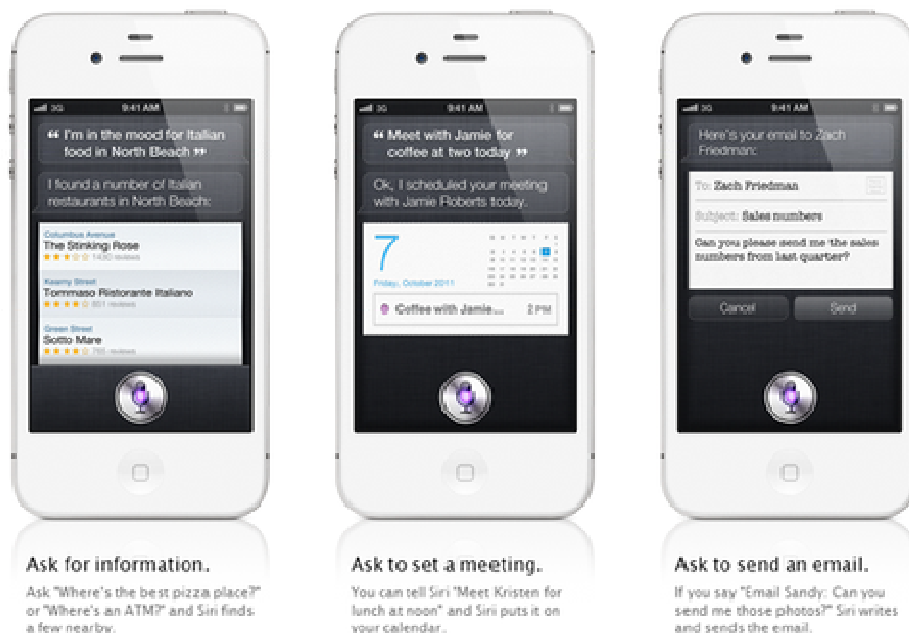


Why Is Siri Important?

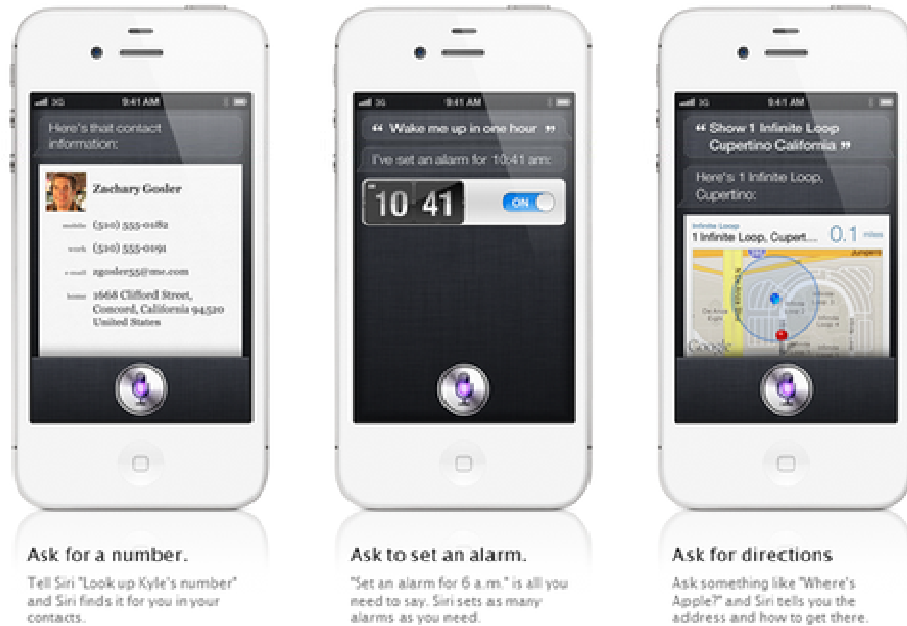
By Brian Roemmele as presented on Quora.com, 10/12/2011

Not Your Dad's Voice Recognition System

It is perhaps easy to discount Siri as just another Voice recognition application, albeit a rather good one. Siri it turns out is far more than this. It is far more than the Artificial Intelligence infrastructures that are dynamically used. It is also far more than the continual learning and contextual awareness systems. Siri is all this and something that could only be held to the definition of true synergy, e.g.: "Two or more things functioning together to produce a result not independently obtainable". None of the individual parts are "new" but the combination Siri created has never really been seen before.



It has been the Holy Grail of computer researchers to one day create a device that could become conversational and intelligent in such a way that it would appear that the dialog is human generated.



We have all pretty much experienced the rather funny hit or miss solutions from most Voice recognition systems. Thus far there has not been the convergence of technology and the synergy it would produce until just the last few years. Siri is a byproduct of this.

DARPA Helps Invent The Internet And Helps Invent Siri

With Siri, Apple is using the results of over 40 years of research funded by DARPA (<http://www.darpa.mil/> Contract numbers FA8750-07-D-0185/0004) via SRI International's Artificial Intelligence Center (<http://www.ai.sri.com/> Siri Inc. was a spin off of SRI Intentional) through the Personalized Assistant That Learns Program (PAL, <https://pal.sri.com>) and Cognitive Agent that Learns and Organizes Program (CALO).

This includes the combined work from research teams from Carnegie Mellon University, the University of Massachusetts, the University of Rochester, the Institute for Human and Machine Cognition, Oregon State University, the University of Southern California, and Stanford University. This technology has come a very long way with dialog and natural language understanding, machine learning, evidential and probabilistic reasoning, ontology and knowledge representation, planning, reasoning and service delegation.

Born In The 1960s

The long history of Siri starts in 1966 when SRI International was tasked by the Defense Department for the "development of computer capabilities for intelligent behavior in complex situations". For decades and up to the present SRI International Artificial Intelligence Center (<http://www.ai.sri.com/timeline/>) has created a path of innovations with a permanent staff of the largest (about 99 computing professionals) and most highly trained (about 55 percent with a Ph.D. or its equivalent) groups of AI professionals in the

world.

It took the amazing vision, imagination and fortitude of [Dag Kittlaus](#) (<http://www.crunchbase.com/person...>), Siri Inc.'s former CEO and co-founder and [Adam Cheyer](#) (<http://adam.cheyer.com/about.html>), Siri Inc.'s former VP Engineering and co-founder, to build Siri, and it was far from easy. I was very remiss in not mentioning them in an early version of this post. They really are the two Fathers of the Siri Apple has today. Both Dag and Adam worked tirelessly to reintroduce voice to the world, this time wrapped up in the ground breaking research and technology from DAPRA. I am rather certain we will hear a great deal more from Dag and Adam.

The Right Timing With The Right Technology

The failure of earlier forms of voice recognition and AI had a number of break points. The primary ones were based on computational power and the workable model for an operable system. Moore's Law, the Internet and Apple has delivered the computer horse power and some 40 years of University research has delivered the other part, Siri. Siri has focused on the 3 important points for this technology:

- Conversational Interface
- Personal Context Awareness
- Service Delegation

The 4th Computer Interface

It is very important to note that Siri is currently just a **1.0 version of the product**, to gain relativity look back on any 1.0 version of a product. Siri will become the 4th and perhaps the most important way to interact with devices. The Mechanical User Interfaces: keyboard, mouse and gestures will always be around and are not going away anytime soon. In fact, I am predicting based on Apple patents a new set of hand gestures and holographic display technology: <http://www.quora.com/How-will-Ap...> However the way humans usually interact is in an even flow of questions and answers most effectively by speaking. There is a huge barrier for most simple questions that is presented once one has to reach for a device and compose a question in a physical manner. The old way of shaping just the right question to get just the right answer in a search field is also not going away anytime soon. But asking a device for a quick answer in just the same manner you would ask a librarian or perhaps a friend will become very, very powerful.

■ Location Context

... near home

... in downtown [San Jose]

■ Task Context

... [serving] burgers

... [described as] romantic

■ Time Context

... this weekend

... tomorrow night

■ Dialog Context

... hotels in Boston

... weather [in Boston]

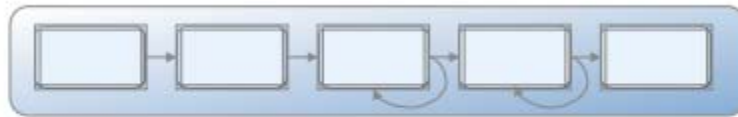
Smaller Needs To Be Smarter

The screen real estate of even the rumored iPhone 5 is limited. Rather than trying to be a search engine, Siri is focusing on mobile use cases (to start with) where models of context like place, time, personal history and limited form factors, magnify the profound power of an Intelligent Assistant. The smaller screen form factor combined with the mobile context and limited mobile bandwidth conspire to make voice the more important interface for most questions. There are a number of benefits by being offered just the right level of detail or being prompted with just the right questions. This interactive process can make the difference between fast task completion or an experience riddled with middle tasks and perhaps endpoint failure.

In a mobile environment, you just don't have time to wade through pages of links and disjointed interfaces and apps to get at simple answers. Just one question can replace 20 tasks by the user. This is the power of Siri.

Task Completion Is The Goal

Using the traditional input systems, the Mechanical User Interfaces, it is hard to see all the tasks that take place. Currently to get at an answer it may require at least a handful of steps to arrive at a satisfying answer.



Find a store

We take all of these steps for granted because there was no other way to do it. With Siri we will be able to eliminate many of the manual tasks to just a simple question. This can be broken out to 3 basic conceptual modes:

- **Does Things For You-** Task completion:

- Multiple Criteria Vertical and Horizontal searches
- On the fly combining of multiple information sources
- Real time editing of information based on dynamic criteria
- Integrated endpoints, like ticket purchases, etc.

- **Gets What You Say-** Conversational intent:

- Location context
- Time context
- Task context
- Dialog context

- **Gets To Know You-** Learns and acts on personal information:

- Who are your friends
- Where do you live
- What is your age
- What do you like

Restaurants



Best sushi near here

Events



What's playing at Carnegie Hall?

Taxis



Book me a taxi

Movies



PG-13 movies this afternoon

Local



Nearest gas station

Weather



What's the temp in Reno?

In the cloud there is quite a bit of heavy lifting working at producing an acceptable result. This encompasses:

- Location Awareness
- Time Awareness
- Task Awareness
- Semantic Data
- Out Bound Cloud API Connections
- Task And Domain Models
- Conversational Interface
- Text To Intent
- Speech To Text
- Text To Speech
- Dialog Flow
- Access To Personal Informations And Demographics
- Social Graph
- Social Data

Of course the A5 dual core processor on the iOS device is also performing quite bit of the front end work. The primary focus is to feed the Awareness data to the Cloud along with the preprocessing of the voice recognition data.

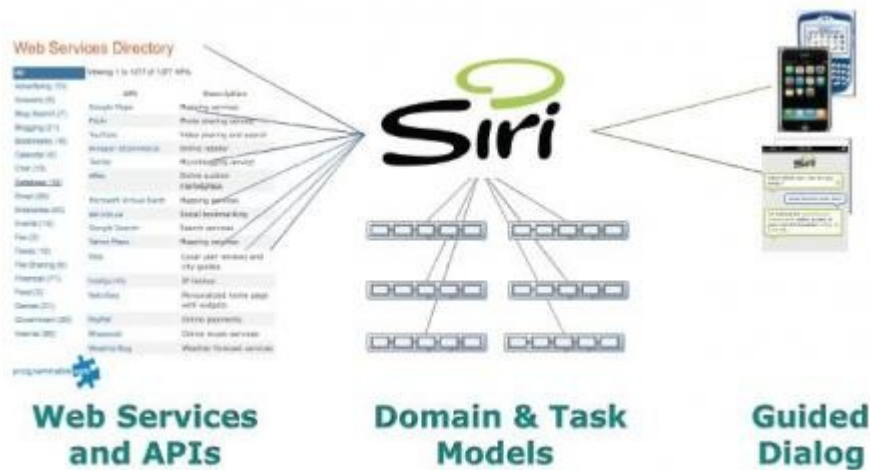
Practical Usage

Siri was demonstrated on October 4th, 2011 using a "Press to ask" system. Siri also has the ability to use the accelerometer and spacial position to use a feature that will be known as "Lift to ask". Siri will also be able maintain active listening mode during a long interaction where no manual activation will be needed. This feature will likely not be available until later versions as a number of noise cancellation algorithms and more refined active listening would be developed. Siri will also be optimized to Bluetooth 4 headsets that will create far more use cases in how it will detect questions from continuous speech. In the future later versions of Siri will be "Active" continuously adjusting to interjecting answers even when no direct question was asked (with in reason). This will make interaction far closer to an interaction with a friend than any device we have ever used.

A New Ecosystem, Backend Cloud APIs

Once one really understands how people will use Siri it is not too hard to see that quite a number of very popular apps and sadly some business plans may become redundant or perhaps less useful. The new model may not be apps as much as structured Cloud APIs to deliver data to Siri. Over time it is perhaps easy to see the ecosystem that will develop around Siri and the APIs that are allowed to connect. I am not at all predicting the end of apps as we know it in any way of form. However I am predicting that we will see a Darwinian adaption to the new ecosystem Siri will create. It will be of very high

importance to see this trend developing and adjust business models accordingly. Perhaps the opportunities that will be available for Siri backend cloud APIs may be as large as the opportunity the iTunes app store has created.



Siri will be building on an ecosystem of Backend Cloud APIs. In its simplest form the API would declare the meaning of the data going in and out via ontologies that have been pre specified reachable by Siri on the Internet. Siri will then build a response on the Fly from the API data. The concept of ontologies-as-specification is the hallmark of Tom Gruber (<http://tomgruber.org/>) CTO and founder of Siri Inc and now at Apple working on Siri. Tom’s approach to the challenges of reaching out to the data on the Internet and getting back something useful is quite revolutionary in that it does not require a “Semantic Web” ecosystem. Through APIs and the land rush that I postulate will develop around this ecosystem getting at relevant data will become rapidly easier.

It is important to understand that the APIs that Tom speaks to are backend Cloud APIs only reachable by the Siri engine via a request that is deemed to be relevant. I am not speaking to APIs that are specific to iOS and its interaction on the platform from an App to operating system interaction. I have little doubt that Apple will open the endpoint APIs to 3rd parties. However I am certain they will retain the right to have the base APIs and data sources under direct relationship control.

If you are a developer, it would be very worth your time to understand ontologies and the semantic web. There is more detail to be found here: [Siri \(product\): What do application developers need to know about Siri to interface with it?](#)

Real Fears Of The Walled Garden

Apple always has been the example of the “Walled Garden”. This concept almost bankrupted the company prior to the return of Steve Jobs. However at the same time, it is the Walled Garden that has made Apple so successful. It is quite frustrating and perhaps almost Monopolistic at times and is feared for good reasons. We can reach as far back as the first Apple II all the way up to the iTunes store. Apple desires to own the garden, but

at the same time they have invited everyone in to play. Apple has created uncountable wealth within the Garden walls. Some may not like the way the Garden is run, but very few can argue with the success of 1000s of companies.

The Walled Garden will however be one of the last really large blows to lesser competitors and the Smart Phone offerings they may have. To compete other companies would have to find equal or greater technology that Apple owns with Siri. And Apple has about a 40 year head start with how well Siri uses the DARPA research. Google of course is in the position to compete but thus far they have a wonderful voice recognition system and really great example results for typical "Voice" searches. There does not seem to be the deep dimension we find in Siri. But I am certain this will change soon with a similar offering on Android paired with Google's approach to the semantic web problem.

It is also important to note that Apple has a Patent application that may limit how APIs connect to Siri and how competitors may be able to respond: [Siri \(product\): Does Apple have patents that may show the future of Siri?](#)

Just A Start

Clearly this new way of interacting with a device will continue to evolve. As I mentioned, there is little doubt that all the existing interfaces and modes of accessing information, via apps and the web will not disappear. However, there is also little doubt that Siri will have a very meaningful impact on how we interact with our devices and how they will interact with us.

This will all start with iPhone 4S but I see it moving rapidly to iPad 3 and in the home on Apple TV. Siri plus Bluetooth 4 and Bluetooth Low Energy (BLE) will have a profound interaction relationship. BLE devices like your front door lock could be controlled via Siri, "Siri, lock the front door" or "Unlock the front door when Sarah arrives". More details can be found here: [iPhone 4S: What impact will the addition of Bluetooth 4.0 have on the iPhone 4S?](#)

It Remains To Be Seen

It remains to be seen how well all of this research and technology really winds up working. I am certain there will be a number of rather large issues to start with. However clearly Apple is in many ways, "Betting the house" on this for at least the medium term. With 20/20 hindsight, perhaps 5 years from now, we will be able to judge the true impact of this technology. Will it be little more than a "Parlor Trick"? Or will it be the ultimate way for us to interact with our devices?