Why an Open Virtual Assistant and How We’re Building It

Monica Lam
Computer Science Department
Stanford University

With Giovanni Campagna, Rakesh Ramesh, Silei Xu, Michael Fischer
Funded by an NSF Expedition, 2008
Lab Sponsors: AVG, Google, HTC, ING Direct, Nokia, Samsung, Sony Ericsson
Software Eats the World

Marc Andreesen
# The Five Kingdoms

<table>
<thead>
<tr>
<th>Type</th>
<th>Platform</th>
<th>Source of Revenue</th>
<th>Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>Microsoft Windows</td>
<td>Licensing</td>
<td>$9 B</td>
</tr>
<tr>
<td>Open $</td>
<td>Amazon store</td>
<td>Net Revenues</td>
<td>$154 B</td>
</tr>
<tr>
<td>Open $</td>
<td>Google search</td>
<td>Ads</td>
<td>$80 B</td>
</tr>
<tr>
<td>Closed $</td>
<td>Apple app store</td>
<td>Revenue Share</td>
<td>$28 B</td>
</tr>
<tr>
<td>Closed $</td>
<td>Facebook social network</td>
<td>Ads</td>
<td>$26 B</td>
</tr>
</tbody>
</table>

M. Lam  
Stanford University
# Who Will Eat These Eaters?

<table>
<thead>
<tr>
<th>Open</th>
<th>Microsoft Windows</th>
<th>Licensing</th>
<th>$9 B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open $</td>
<td>Amazon store</td>
<td>Net Revenues</td>
<td>$154 B</td>
</tr>
<tr>
<td>Open $</td>
<td>Google search</td>
<td>Ads</td>
<td>$80 B</td>
</tr>
<tr>
<td>Closed $</td>
<td>Apple app store</td>
<td>Revenue Share</td>
<td>$28 B</td>
</tr>
<tr>
<td>Closed $ Privacy</td>
<td>Facebook social network</td>
<td>Ads</td>
<td>$26 B</td>
</tr>
</tbody>
</table>
Today’s Virtual Assistants

Amazon Alexa & Google Home — Closed Platform

- Always listening: “Alexa, …”, “OK Google”
- Skill / Action partners
  - Turn on the lights
  - Ask Uber to request a ride
  - Ask Pizza Hut to place an order
- Alexa: purchases —> Amazon
- Home: questions —> Google search

Google and Siri — Proprietary

- Commands on the phone
- Contextual suggestions based on personal info
Virtual Assistants Eat the Internet

Uniform language based interface for the post-PC era
Access to all private data
Personalized & customized (Unlike chat bots)
Concerns

Privacy
Monopoly
Choice
Open competition
Innovation
Solution: Open System

Remember AoL?
Solution: Open System

Nonproprietary Open Platform ➔ Researchers & Contributors ➔ Early Adopters ➔ Main Stream

Unix OS: Bell Labs, Berkeley
Mach: CMU
Linux: Torvalds
Sun OS: Sun
NextStep: Next
Linux desktops
Server OS
MacOS
iOS, Android

SUIF Compiler: Stanford
Compiler Researchers
LLVM: Illinois
SGI, Intel, Tensilica

Mosaic browser: Illinois

Compiler Researchers
Apple
Google
Mozilla Firefox

M. Lam
Stanford University
Almond: Building the World’s Best Virtual Assistant
Can we predict / invent the future?
A Day in the Life with Almond

<table>
<thead>
<tr>
<th>Time</th>
<th>Activity</th>
<th>Service</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>7am</td>
<td>food and coffee</td>
<td>Starbucks</td>
<td>“order an Americano from Starbucks if I’m up before 8am”</td>
</tr>
<tr>
<td>9am</td>
<td>social media</td>
<td>Twitter</td>
<td>“if a tweet in my area gets more than 100 likes notify me”</td>
</tr>
<tr>
<td>11am</td>
<td>retail</td>
<td>Amazon</td>
<td>“when my printer runs out of paper order more paper”</td>
</tr>
<tr>
<td>2pm</td>
<td>search</td>
<td>Google</td>
<td>“alert me when the stanford women basketball team plays”</td>
</tr>
<tr>
<td>5pm</td>
<td>transportation</td>
<td>Uber</td>
<td>“pick me up 20 minutes after my last meeting of the day”</td>
</tr>
<tr>
<td>9pm</td>
<td>entertainment</td>
<td>Netflix</td>
<td>“if I turn off the lights in the bedroom by 10 start playing a movie”</td>
</tr>
</tbody>
</table>
In collaboration with Prof. Mark Gaynor, Univ. St. Louis
If This Then That

- Proprietary cloud-based service
  - Over 100K recipes (trigger - action)
  - Access to all personal account credentials and data
  - Web interface only
- Natural language research (quirk et al)
  - Convert IFTTT descriptions to recipes
  - “Ig to fb”:
    “When I post a new Instagram picture, post the same picture on Facebook with the same caption”
  - Data set is too noisy to infer parameters: 0% accuracy on parameters
The World’s 1st Open Virtual Assistant

- Interoperability across things
- Interoperability across assistants
- Customization:
  - Natural language programming
  - With privacy and generality
- Almond prototype available in Play Store

In 26th International World Wide Web Conference, 2017
Why?  What?  How we’re building it?
The World’s 1st Open Virtual Assistant

Interoperability across things
Interoperability across assistants
Customization:
  Natural language programming
  With privacy and generality
Almond prototype available in Play Store

In 26th International World Wide Web Conference, 2017

M. Lam

Stanford University
1. Privacy

- Almond is open-source software
- Can be scrutinized for security
- Can be run on phone, home servers, in the cloud
Open Architecture: e.g. Email

Inter-operability
Choice
Privacy if desired
Open Competition

Company Mail
University Mail
Personal Mail
2. Generality

ALMOND Virtual Assistant

DATA Personal Accounts

THINGPEDIA Interface Repository

CODE Application Programming Interfaces (API)

Natural Language
**Thingpedia: Encyclopedia of Things**

Currently:
- 45 devices
- 187 functions

<table>
<thead>
<tr>
<th>Natural Language</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHEN @Trump tweets</td>
<td>source(text, ..., from, ...)</td>
</tr>
<tr>
<td>GET tweets matching “Trump”</td>
<td>search(..., hashtag, ...)</td>
</tr>
<tr>
<td>DO tweet “I am in Hong Kong”</td>
<td>sink (status)</td>
</tr>
</tbody>
</table>
WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN

WHEN
3. Interoperability

**WHEN** [FILTERS] → **GET** [FILTERS] → **DO**

When my alarm goes off, open my blinds
Get my tweets and save them to Dropbox
Every day at 9 am, get my tweets and send them to Dropbox
When I receive an email from my advisor then send the message on SMS
When I tweet, share the text on LinkedIn

@twitter.mytweet (text)
⇒ @linkedin.share (status), status = text
Show me tweets with hashtag #Stanford

TwitterHashtagFilter(HashTag : String) {
    @twitter.source(text, hashtags, _, from, _, false),
    $contains(hashtags, HashTag)
⇒ @$notify("Tweet from @" + from + ": " + text);
}

M. Lam

Stanford University
4. Real Natural Language Input

**Natural Language**

Post all my tweets on LinkedIn
Share my tweets on my LinkedIn

**Thingpedia Snippets**

When I tweet, share the text on LinkedIn

**ThingTalk**

@twitter.mytweet (text)
⇒ @linkedin.share (status), status = text
Real Natural Language Input

- if receive an email on gmail, send an email to from address

Send back an email whenever I receive an email on gmail
Reply to all email I received on gmail
Auto reply on gmail
Programming by natural language is hard

many combinations:
3 functions, parameter pairings, filters

Machine Learning needs lots of training data

Wish to handle new devices
Bootstrapping the Machine Learner

- **Thingpedia**
- **Crowdsourcing natural training data**
- **Machine Learning Model**
- **Incremental Learning**
  - training with GUI
High-Level Parsing Algorithm (Sempre)

- **input**: play 'presidential debate' from Youtube on my TV
- **generate many candidates using canonical grammar**
  1. Canonical: search video on youtube with query “presidential debate” then play url on tv with video url video url
     ThingTalk: @youtube.search_video, query = “presidential debate”, v0 = video_url ⇒ @tv.play_url, video_url = v0
  2. 
  3. ...
- **use machine learning to best match candidate**
  Canonical: search video on youtube with query “presidential debate” then play url on tv with video url video url
  ThingTalk: @youtube.search_video, query = “presidential debate”, v0 = video_url ⇒ @tv.play_url, video_url = v0
Generating Synthetic Sentences with a Grammar (Sempre)

If receive tweet on twitter then send message on slack → @twitter.new_tweet ⇒ @slack.send

monitor if WHEN [w] → RULE [ w ⇒ notify ]

GET [g] → RULE [ now ⇒ g ⇒ notify ]

DO [a] → RULE [ now ⇒ a ]

if WHEN [w] then DO [a] → RULE [ w ⇒ a ]

if WHEN [w] then GET [g] → RULE [ w ⇒ g ⇒ notify ]

GET [g] then DO [a] → RULE [ now ⇒ g ⇒ a ]
Parameters!

\[
\begin{align*}
c(x) &: \text{canonical form of Thingtalk code } x \\
t(x) &: \text{type of Thingtalk code } x \\
\text{NT}_t &: \text{NonTerminal of type } t
\end{align*}
\]

\[
\begin{align*}
c(f) \text{ with } c(p) \text{ NT}_{t(p)}[v] & \rightarrow \text{ NT}_{t(f)}[f, p = v] \\
\text{send message on slack} \text{ with message STRING } [v] & \rightarrow \text{ DO } [@\text{slack.send}, \text{ message } = v]
\end{align*}
\]

\[
\begin{align*}
c(r) \text{ with } c(p_1) c(p_2) & \rightarrow \text{ RULE } [r, p_1 = p_2]
\end{align*}
\]

If receive tweet on twitter then send message on slack with message text
\[ \rightarrow @\text{twitter.new_tweet} \ (v = \text{ text}) \Rightarrow @\text{slack.send} \ (\text{message } = v) \]
Crowdsourced Training Challenges

- People don’t know what rules are
- People can’t think of rules themselves — coverage?
- If we tell them what we want, they won’t come up with new sentences
Paraphrasing Examples

When I tweet share the text on LinkedIn

Share my tweets on my LinkedIn
Whenever I tweet, post the same message on LinkedIn
Post all my tweets on LinkedIn
Testing on Crowdsourced Paraphrases

**Training set**
- Thingpedia contributors: 2394
- Paper authors: 628
- Paraphrase: 4466

**Test set**
- Paraphrase: 1874

![Graph showing paraphrase comparison]
Alice lives in California, where you know, it doesn’t rain very often. So she rarely checks the weather forecast and always forgets to bring her umbrella when it rains. Could you make a rule for her to solve this without her checking the weather every day?

If the chance of rain is greater than 50%, send an email
Alert me to weather forecast only on rainy days
Testing on Scenarios (71 sentences)
Typical Users

• Show the cheat sheet
• Remove the cheat sheet
• Ask for examples of rules

When an asteroid passes close to the earth, tweet "God's throwing rocks at us".
DO notify 911 WHEN my pulse drops below 40
Extensibility of Almond

The diagram shows the extensibility of Almond across different devices and domains. It compares the performance of existing devices (Gmail) and new devices (Slack and communication domain) in terms of primary and compound compression. The chart highlights the percentage of top 5, top 3, and top 1 performances.
Incremental Learning
Here's what I can do for you on twitter.

- If I post a picture on Instagram then post it on Twitter
- Tweet
- Send a direct message on Twitter to
Almond

make a rule

Pick a command category:

When
Get
Do
User Tests

- Present Almond description, cheatsheet, scenarios
  - 15 users
  - Success: 40%
    - 24% out of scope; 9% quoting issues
  - Typing is easier than menu scrolling
  - Natural language is preferred for “low risk”
5. Interoperability Across Assistants
Sharing Economy: Uber

I'm ready to give rides

IoT
GPS

Natural Language

ThingTalk

Database

Thingpedia

I need a ride to the airport

IoT
GPS

Natural Language

ThingTalk

M. Lam
Stanford University
I want to sell a bike for $25

I'll pay $30 for a red bike, and $20 for a blue bike.
You should not run if cottonwood pollen is high.

Let me know if your peakflow meter reading is below 400L/m.
Overnight Development Platform

- Instant audience — no need to create a network for each interest
  - Users can tap into different markets with natural language
  - Markets connect interested users instantaneously
- Instant rich interface with efficiency

Long tail of automated Interactions
Summary

- The world needs an open virtual assistant
- Platform:
  Thingpedia: all digital dialogs
  Almond: Privacy-preserving virtual assistant
- Natural language for interoperating things & interoperating virtual assistants
- Can we attract researchers, contributors, early adopters, mainstream developers?

By HenkvD - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=5702715