Giving Control Back to Consumers with Open Federated Virtual Assistants

Monica Lam
Computer Science Department
Stanford University
lam@cs.stanford.edu

With Giovanni Campagna, Michael Fischer, Mehrad Moradshahi, Rakesh Ramesh, Richard Socher, Silei Xu, Richard Yang
Sponsors: AVG, Google, HTC, Hitachi, ING Direct, Nokia, Samsung, Sony Ericsson, UST Global
Consumer Privacy at Stake

• Facebook owns and sells 2-billion people’s personal data
• Cambridge Analytica incident
• EU GDPR (General Data Protection Regulation)
• There is no meaningful alternative
Unhealthy Commercial Ecosystem

- Platform Monopolies/Duopolies
- Google and Apple app stores: 30% revenues
- Google and Facebook: 60% digital marketing revenues

Monopoly — Open Competition — Innovation?
Virtual Assistants

50 millions in 2 years
Internet: 50 millions in 4 years

https://techcrunch.com/2018/03/07/47-3-million-u-s-adults-have-access-to-a-smart-speaker-report-says/
Virtual Assistants

Personalized: sees all personal info
Linguistic User Interface (LUI)
Intermediates all digital services
Controls choice of vendors

Amazon, Facebook, Google Combined!
Mobile & Ubiquitous: Graphical->Linguistic

<table>
<thead>
<tr>
<th>Graphical Web</th>
<th>Linguistic Web</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphical user interface (GUI)</td>
<td>Linguistic user interface (LUI)</td>
</tr>
<tr>
<td>Browser</td>
<td>Virtual Assistant</td>
</tr>
<tr>
<td>Web page addresses</td>
<td>LUIs as skills</td>
</tr>
<tr>
<td>Hosted by owners</td>
<td>In virtual assistants’ platform</td>
</tr>
<tr>
<td>Nonproprietary</td>
<td>Proprietary (like AOL)</td>
</tr>
</tbody>
</table>

We are witnessing the start of proprietary linguistic webs.
Almond Vision

- An open & nonproprietary linguistic web
  - Thingpedia: crowdsourced repository of skills open to all assistants
- Privacy: Open federated virtual assistants -> choice
- Adoption: Open-source best virtual assistant technology
  - Put the users back in the driver seat
  - Connect disparate resources
  - Share them with “who, what, when, where, how”

“Program” our virtual assistant in natural language!
Example: Asthma Patient

**devices**
“when I use my inhaler, record my GPS location in logfile on Box”

**location**
“Let my Dad know if I am at the hospital”

**environment**
“when the air quality index is above 500 and Bob is running, warn him”

**people**
Dr. Smith:
“if Bob’s peak flow-meter drops below 180L/min let me know”
Impact of Natural Language Programming

• Power of language
• Consumers: today’s software cannot satisfy the long tail of user needs
• Professionals: automate their own repetitive tasks
Natural Language Programming
Technology
Core Concepts

Natural Language → LUInet → ThingTalk → Open Interoperable Web

- Semantic Parser
- Neural Network
- Control Constructs
- Formal Virtual Assistant Language

+ Thingpedia

- Full API Signatures
Natural Language Programming

“When I use my inhaler, get my GPS location, if it is not home, write it to logfile in Box.”

- Event-driven program
- Multiple function calls
- Parameter passing
- Filters on values
“When I use my inhaler, get my GPS location, if it is not home, write it to logfile in Box.”

```
monitor @Inhaler-use(),
=> @GPS(), location <> "home"
=> @Box-write(file="logfile", data=location)
```
Thingpedia: Encyclopedia of Things

- Interoperability
  - API signatures + corresponding NL
  - Not just intents
- Open repository
  - Available to Alexa, Google Assistant, …

> 60 devices / 200 functions

<table>
<thead>
<tr>
<th></th>
<th>Natural Language</th>
<th>API Signatures</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHEN</td>
<td>@Stanford tweets</td>
<td>Monitor (@home_timeline(), …) author==“Stanford”)</td>
</tr>
<tr>
<td>GET</td>
<td>tweets matching “#Cardinal”</td>
<td>search(...), contains (hashtag, …)</td>
</tr>
<tr>
<td>DO</td>
<td>tweet “Stanford won!”</td>
<td>post (status)</td>
</tr>
</tbody>
</table>
THINGTALK Compound Statement

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

**FILTERS:** =, <, >, <=, >=, <>, contains, starts with, ends with

- When I use my inhaler, get my location, save them to Dropbox
- If I get taken to a hospital, let my dad know.
- When the air quality index is above 500, and I am running, send me an SMS.

- When the Bitcoin price reaches $10,000, search for a “bitcoin” picture, and tweet it with caption “I’m rich!”
Expressiveness of ThingTalk

- Inspired by IFTTT
- ThingTalk is a superset of IFTTT recipes (2 clauses)
- IFTTT has 250,000 unique recipes
- IFTTT provides a GUI: no formal or natural language
- IFTTT is proprietary: user must share credentials
Real Natural Language Input

When my car is at home, and it is not plugged in, send me a reminder email.

Remind me if my car is not plugged in at home. If I am not charging my car when it is home, let me know. Remind me to plug in my car whenever I’m home.
Technical Challenges

• Co-design of Thingpedia, ThingTalk and LUInet

• Natural language training-data acquisition
  —> methodology, tool: Genie
Genie: A Semantic Parser Generation Tool

- ThingTalk Grammar
- Thingpedia
- NL Templates
- Synthetic Sentence & Program Generator
- Crowdsourced Paraphrases
- Parameters & Data Augmentation
- Training Data
- Neural Network Model
- Trained Semantic Parser
LUInet Results

- Dataset (44 skills, 131 functions, 178 parameters)
  - 550K synthetic, 25K paraphrased, 15K augmented sentences
  - Test set: 1K paraphrased sentences
  - Training set after parameter expansion: 2.9 M sentences
- Model: Seq2seq bi-LSTM with attention, pointer network
- Accuracy: 87%
- Future work: real user data
Sharing with Privacy
Sharing is Broken Today

Let your virtual assistant help you share
General+Fine-Grain: ThingTalk Extension

Requester: 

GET-PREDICATE [FILTERS]

WHEN [FILTERS] → GET [FILTERS] → DO

FILTERS: =, <, >, <=, >=, contains, starts with, ends with

Let Dr. Smith monitor my peak-flow-meter, if it drops below 180L/min
Let my father monitor my security camera for motion,
Let my secretary, whenever I am out of town, read email messages whose subject is marked urgent.
Let my daughter, from 6-8pm, watch NetFlix
Let my boyfriend get pictures from my dropbox, taken on Feb 14, and post them on Facebook
Almond: 1st Federated Virtual Assistant

Expressiveness:
Any ThingTalk command

Privacy:
Remote execution model
Owner executes requests
Returns need-to-know

Giovanni, Xu, Ramesh, Fischer, Lam, Ubicomp 2018
Remote Execution

(a) Request

σ=＠dad, c=SELF: monitor @security_camera. event(), has_motion=true ⇒ return

(b) Check

σ=＠dad, c=SELF: monitor @security_camera. event(), has_motion=true && @phone.get_gps() {location≠home} ⇒ return

(c) Ask for permission

“@dad wants to get notified when any event is detected on your security camera and has motion is equal to true.”

(d) Respond

“Only if I’m not home.”

(e) Save

“Notification from monitor security camera: motion detected ...”

(f) Return detected events

“Ask @alice to notify me when her security camera detects motion.”
Conformance of Access Control

Natural Language

2nd-Party Program Access Control

Satisfiability Modulo Theories (SMT)
Conformance Algorithm

• SMT: Generalization of boolean satisfiability (SAT) with theories of strings, arrays, ...

• Provably correct algorithms for conformance, and synthesis of conforming code

• NP-hard, but fast enough in practice
Needs and Acceptance?
Do Consumers Need Access Control?

### Role-Based Permission
- **Teenage daughter** to use credit card: 0%
- **Amazon courier** to unlock door: 100%
- **Friends** to access cloud drive: 100%
- **Parent/kid** to see security cameras: 0%
- **10-year-old kid** to use Netflix: 0%

### Attribute-Based Permission
- With a $20 budget limit: 0%
- Restaurants only: 100%
- If the package is over $1000: 0%
- If your security camera is on: 100%
- Photos with their faces in them: 0%
- Photos in a specific folder: 0%
- If you are not at home: 0%
- Cameras facing the front yard/garage: 0%
- Between 7 PM to 9 PM: 0%
- Free G or PG rated movies: 0%

% People comfortable in giving permission (200 person survey)
More Examples

Willingness to share doubles with attribute-based access control
Expressiveness of ThingTalk?

Solicit use cases by showing AMT workers 3 examples, without describing ThingTalk

**Enforceable:**
Mom: “You need to follow this guy on Twitter, give me your Twitter account”.
Me: “OK, add him but don’t follow any other Twitter user”.

**Unenforceable:**
Friend: “Can I use your library card?”
Me: “OK, only if you return the book on time”.
ThingTalk is Expressive

60 people; 220 suggestions; 85 unique assets

Diverse use cases

- Business Accounts: 1%
- Social Media: 16%
- Services: 21%
- IoT: 27%
- Personal Data: 34%

85% in the scope of ThingTalk

- Existing API: 70%
- New API: 15%
- Out of Scope: 9%
- Unenforceable: 6%
User Study: “Sharing Without Passwords”

Like the concept?

Like the app?

Use the app?

# users

Ratings
Applying Genie

Accuracy on paraphrased sentences

- ThingTalk
- ThingTalk + Access Control
- ThingTalk + Aggregates

%
Summary

Programming systems + Distributed systems + HCI + AI

Open platform + Genie-like tools + Crowdsourcing + Real Data:

- NL Parsing Neural Network
- Constructs for work flow
- All Interoperable IoTs + Services

LUInet \( \rightarrow \) ThingTalk++ + Thingpedia
Artificial Collective Intelligence

Recorded workflow, behavior, outcomes
-> Predict human behavior, automate professional services

Protect privacy &
Make data available to scientists

Key: Open federated virtual assistants