

The W3C Voice Interaction Community Group

TPAC, 2022



About the Voice Interaction Community Group

- <https://www.w3.org/community/voiceinteraction/>
- Started in 2017, following on from the Voice Browser Working Group and the Multimodal Interaction Working Group
- 42 current members
- 4 Co-chairs
- Goals:
 - develop new use cases for voice interaction
 - develop requirements for applications such as virtual assistants
 - explore areas for possible standardization
 - possibly producing specifications if appropriate
- Topics:
 - **interoperability for intelligent personal assistants**
 - discovery of virtual assistants with specific expertise, for example a way to find a virtual assistant that can supply weather information
 - standard formats for statistical language models for speech recognizers
 - standard representations for references to common concepts such as time
 - work on dialogue management or 'workflow' languages

What is an intelligent personal assistant?

- Users interact with IPA's primarily by voice
- General personal assistants are often included with smart speakers
 - Ambient assistants can be activated by “wake words”
 - User doesn't have to touch the device
 - User doesn't have to be able to see a screen
 - Answer general questions about current news, sports, history, science facts and play music
- Enterprise personal assistants have deep information about an enterprise or other organization
 - Can be invoked through a general personal assistant
 - Can be invoked directly through an app

Some intelligent personal assistant platforms

- Includes smart speaker
 - Mycroft (open source)
 - Amazon Alexa
 - Google Assistant
 - Apple Siri
- Runs on other devices (mobile phone, web page)
 - Microsoft LUIS
 - Nuance Mix
 - Samsung Sam
 - IBM Watson AI

Potential for standardization

- Current platforms are all walled gardens
- All have their own communications protocols and messages
- Developers can't mix and match components



[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)

The Web and intelligent personal assistants

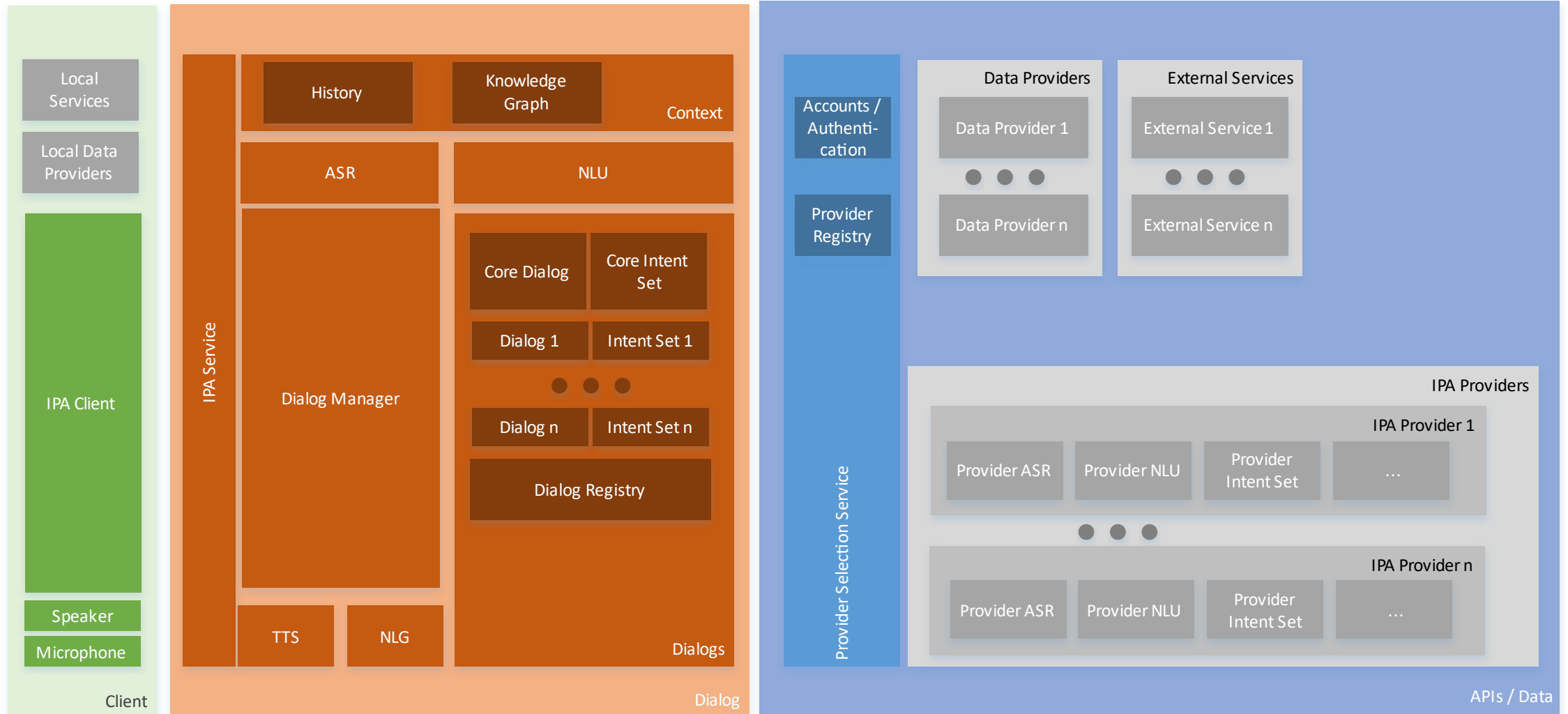
The Web

- Content can run on any browser (HTML, CSS, Javascript)
- A web page can transfer to any other web page (HTTP(S), DNS)
- Opportunity for innovation in components like browsers and search engines

Intelligent personal assistants

- Content can only run on the platform it was developed for
- Applications can only transfer to other approved applications on the platform it was developed for
- No opportunity for innovation in search or platform

Intelligent Personal Assistant Architecture



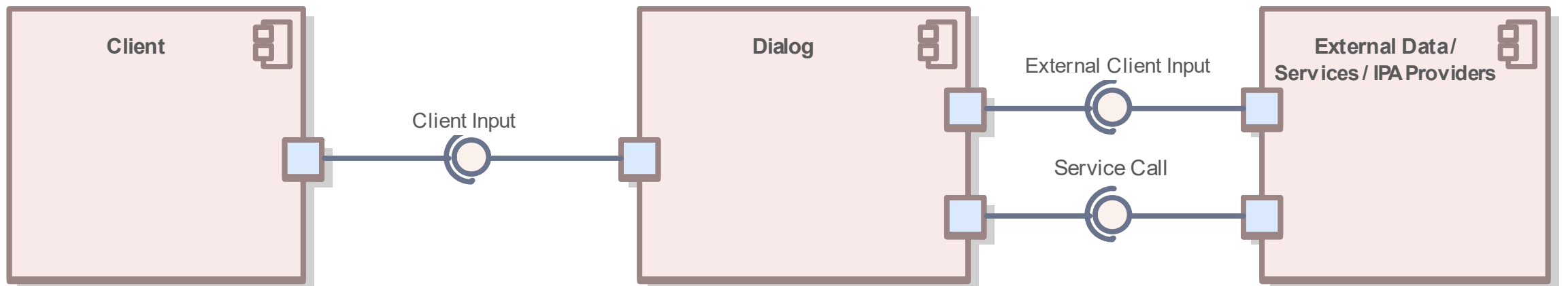
Device?

Browser?

Search Engine/DNS?

Websites?

High level interfaces



Current Specifications

- JSON Representation of Semantic Information;
A common format for natural language results
<https://w3c.github.io/voiceinteraction/voice%20interaction%20drafts/emmaJSON.htm>
- Architecture (Version 1.2)
Components of conversational systems and their functions
<https://w3c.github.io/voiceinteraction/voice%20interaction%20drafts/paArchitecture-1-2.htm>
- Interfaces between components:
How components communicate
<https://w3c.github.io/voiceinteraction/voice%20interaction%20drafts/paInterfaces/paInterfaces.htm>

How to get involved

- Join the Community Group
 - <https://www.w3.org/community/voiceinteraction/>
- Read the draft specs
- Comment on the public list
 - <https://lists.w3.org/Archives/Public/public-voiceinteraction/>
- Help edit the drafts