Anatomy of an Autonomous Virtual Assistant

Glenn Taylor

glenn@soartech.com

NDIA Human Systems Conference

June 2022

What is a Virtual Assistant?

- A software system that a user engages with to help perform one or more tasks

- Commercial examples: Amazon's Alexa, voice-based reservation systems, company chat bots, …

- Alexa lets you give voice commands to play music, get the weather, etc.

- "Alexa, play the Black Keys"

- Essentially a voice interface to some well-defined services – currently 100k such services ("skills") available for Amazon's Alexa


Image: Amazon.com

SOARTECH

This information is approved for release

Autonomous Virtual Assistant

- Broader view: not just a voice interface to some services, but something that helps you do a task, may even act as a teammate performing some tasks with/for you
  - Elements of delegation, authority

- Connected to range of computing systems / backend services – system diagnostics, planners, information systems, sensors
  - Doesn't have all the smarts, but knows where to get it

- Ideally also aware of the human performing the task, the task itself, …

Image: Seapower Magazine
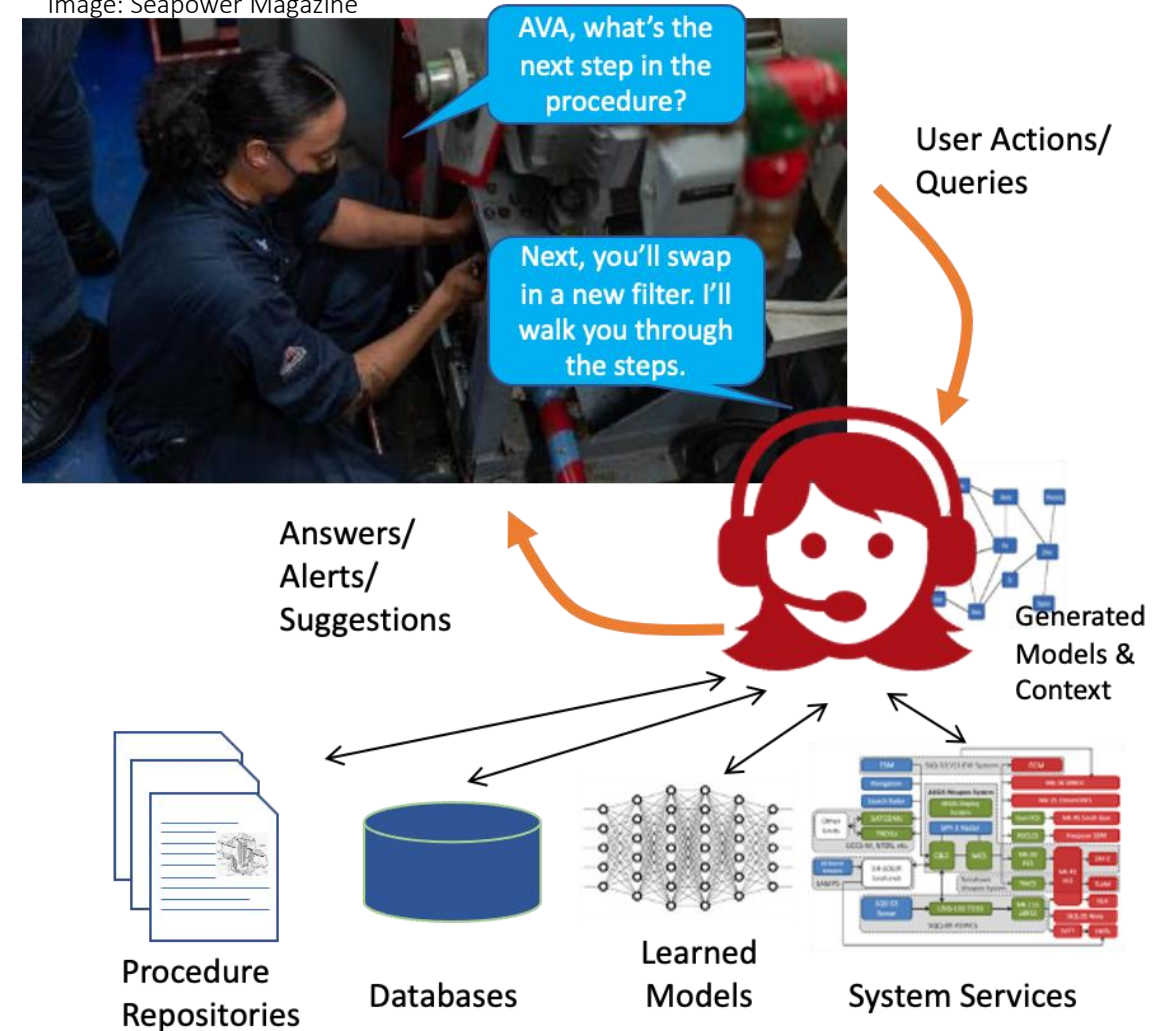
Potential Uses?

- Computing systems, autonomous platforms, vehicles/spacecraft are becoming more capable but more complicated and complex

  - Tasks overwhelming for a single person
  - Non-experts required to perform difficult tasks
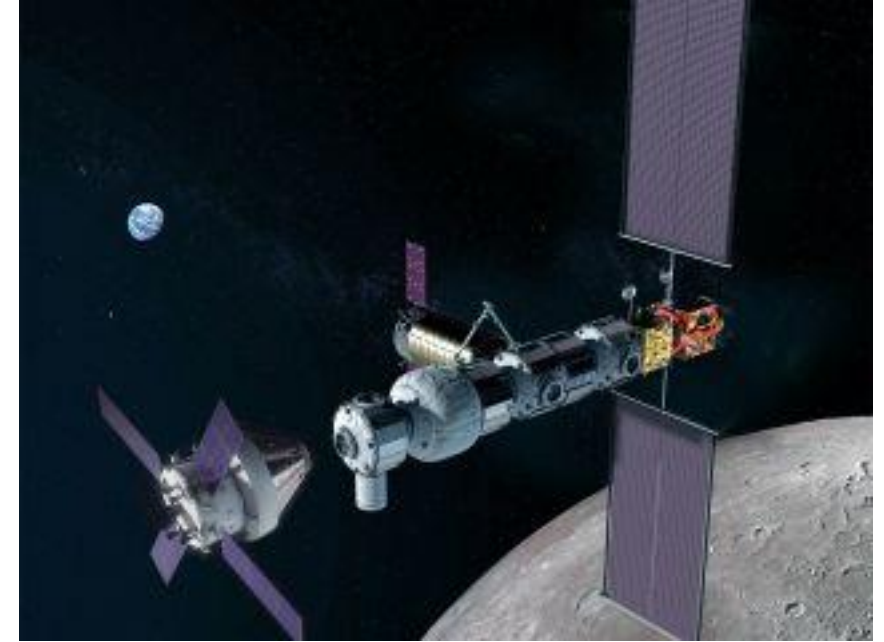 - Lack of 'reach back' to system/domain experts

Image: NASA

Image: Mitsubishi Electric

This information is approved for release

Autonomous Virtual Assistant Reference Architecture

- **Purpose of a reference architecture:** "…describes a family of similar systems… collects relevant solution patterns" (SEI 2017)

- Describe main functional elements of an autonomous virtual assistant

- For a particular implementation, different elements of this may be emphasized to different degrees
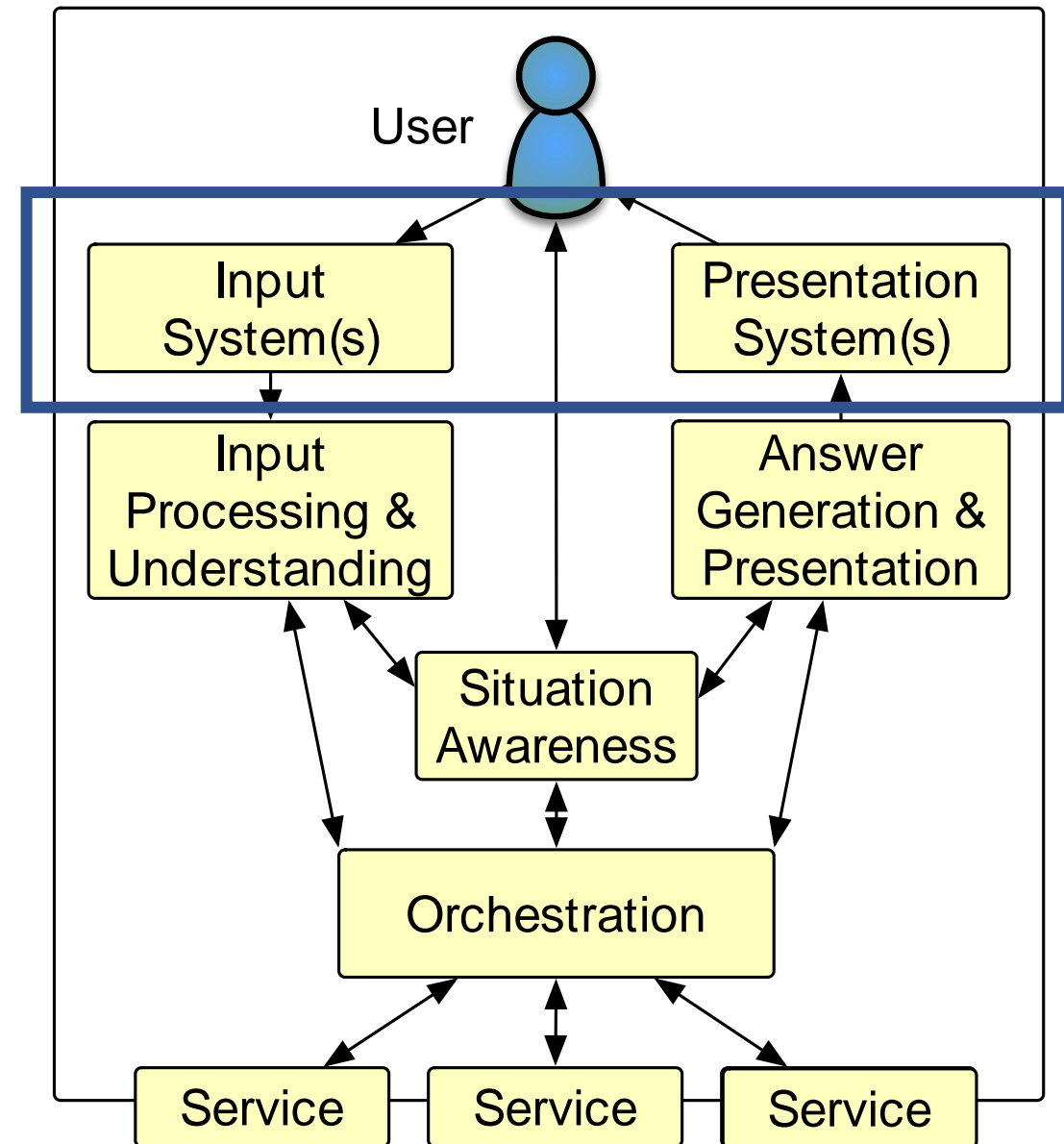
Autonomous Virtual Assistant Architecture

This information is approved for release
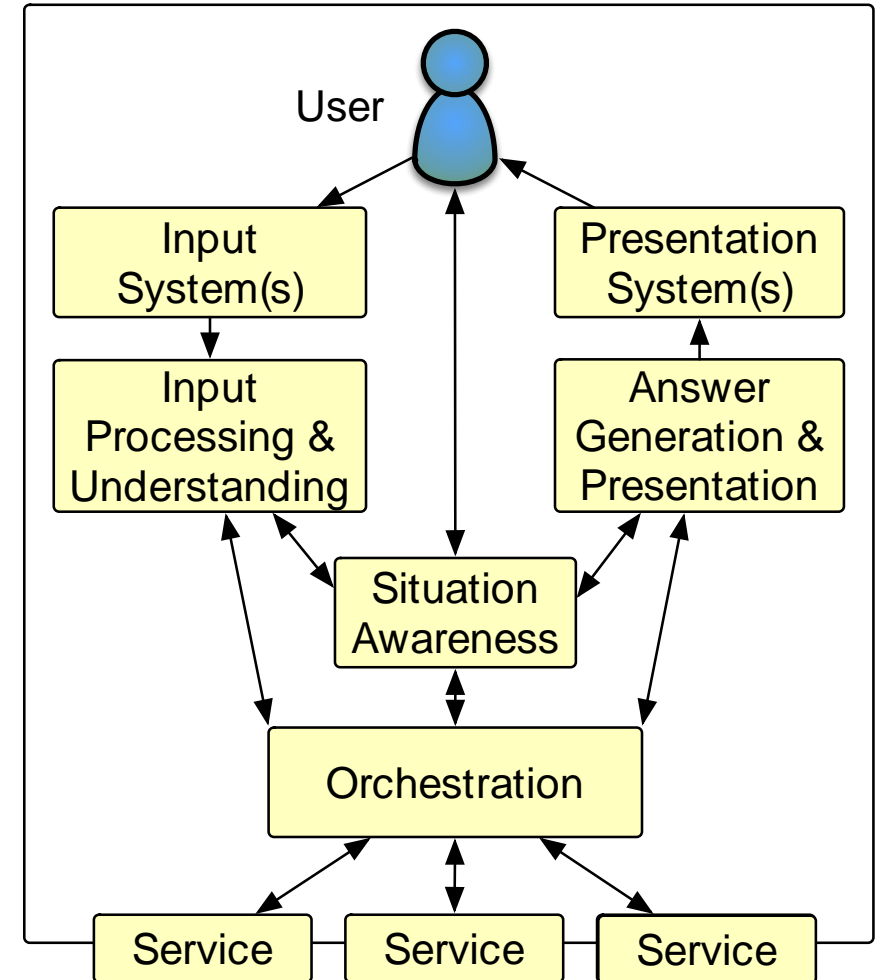
Input and Presentation Systems

- **Purpose: capture inputs from user and push outputs to user**

- Physical hardware (& software) layer that the user interacts with

- This could be a keyboard and a monitor, or could be the latest XR headset

- Will vary widely based on what makes sense for the task(s) that AVA is helping with

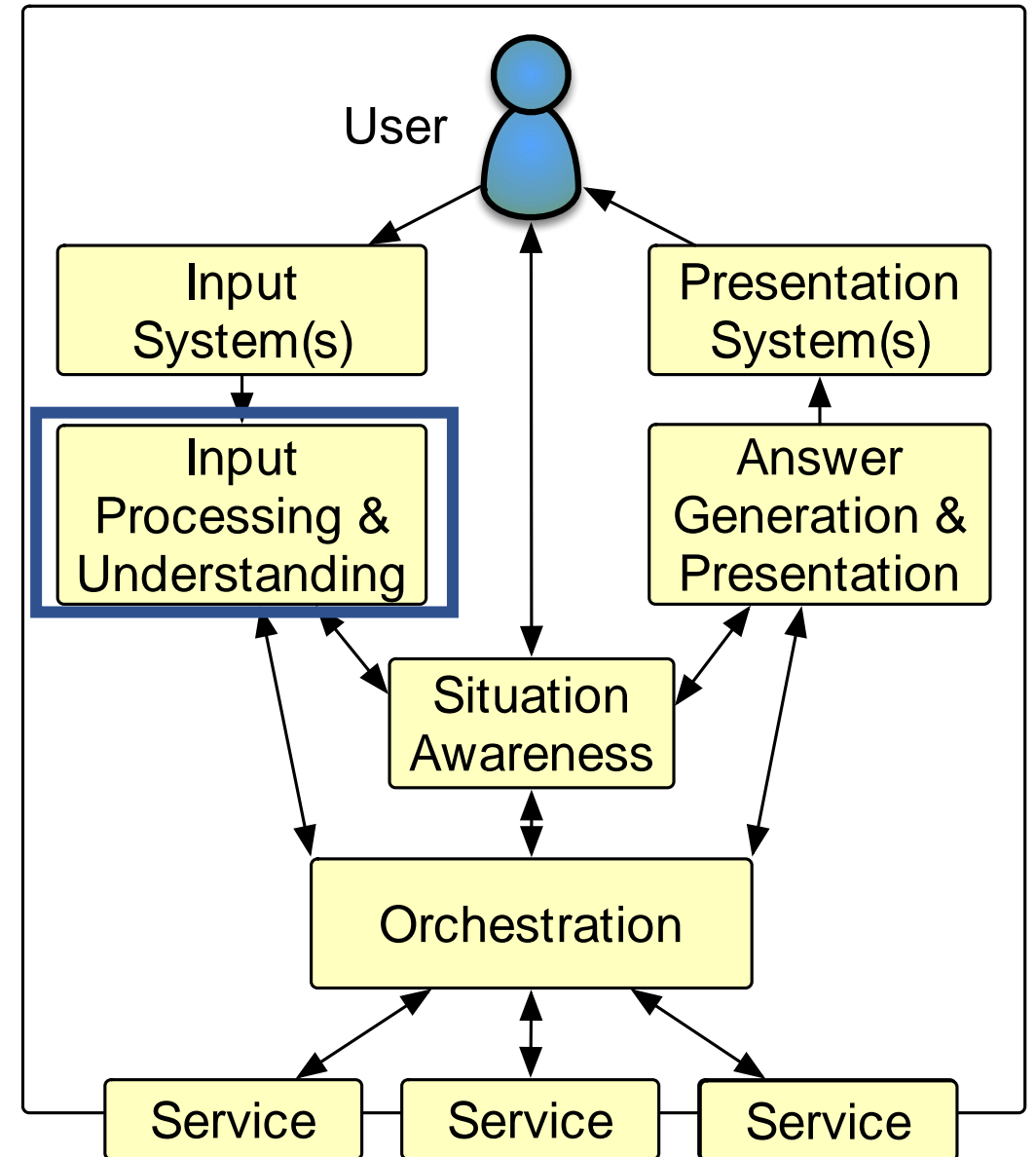**Autonomous Virtual Assistant Architecture**

User

Input System(s)

Presentation System(s)

Input Processing & Understanding

Answer Generation & Presentation

Situation Awareness

Orchestration

Service   Service   Service

This information is approved for release

Input Processing & Understanding

- **Purpose: make sense of user inputs**

- Sometimes called "intent recognition", but could be wider than this

- Given some raw stream of input, turn it into something meaningful
  - Intents – e.g., commands, requests
  - Gaze → attention
  - Heart rate variability → workload

**Autonomous Virtual Assistant Architecture**

User

Input System(s)

Presentation System(s)

Input Processing & Understanding

Answer Generation & Presentation

Situation Awareness

Orchestration

Service   Service   Service

This information is approved for release

Answer Generation and Presentation

- **Purpose: provide information(+) to user**
  - **Results of commands, answers to questions/requests**
  - **Alerts, notifications**

- This might be a pre-determined process:
  - Given data type X, show it in form Y

- Or could be deliberative: adapting presentation to content/user/situation

- Could be simply passing along results verbatim from a service query, could include consolidating, reformatting, etc., based on multiple services
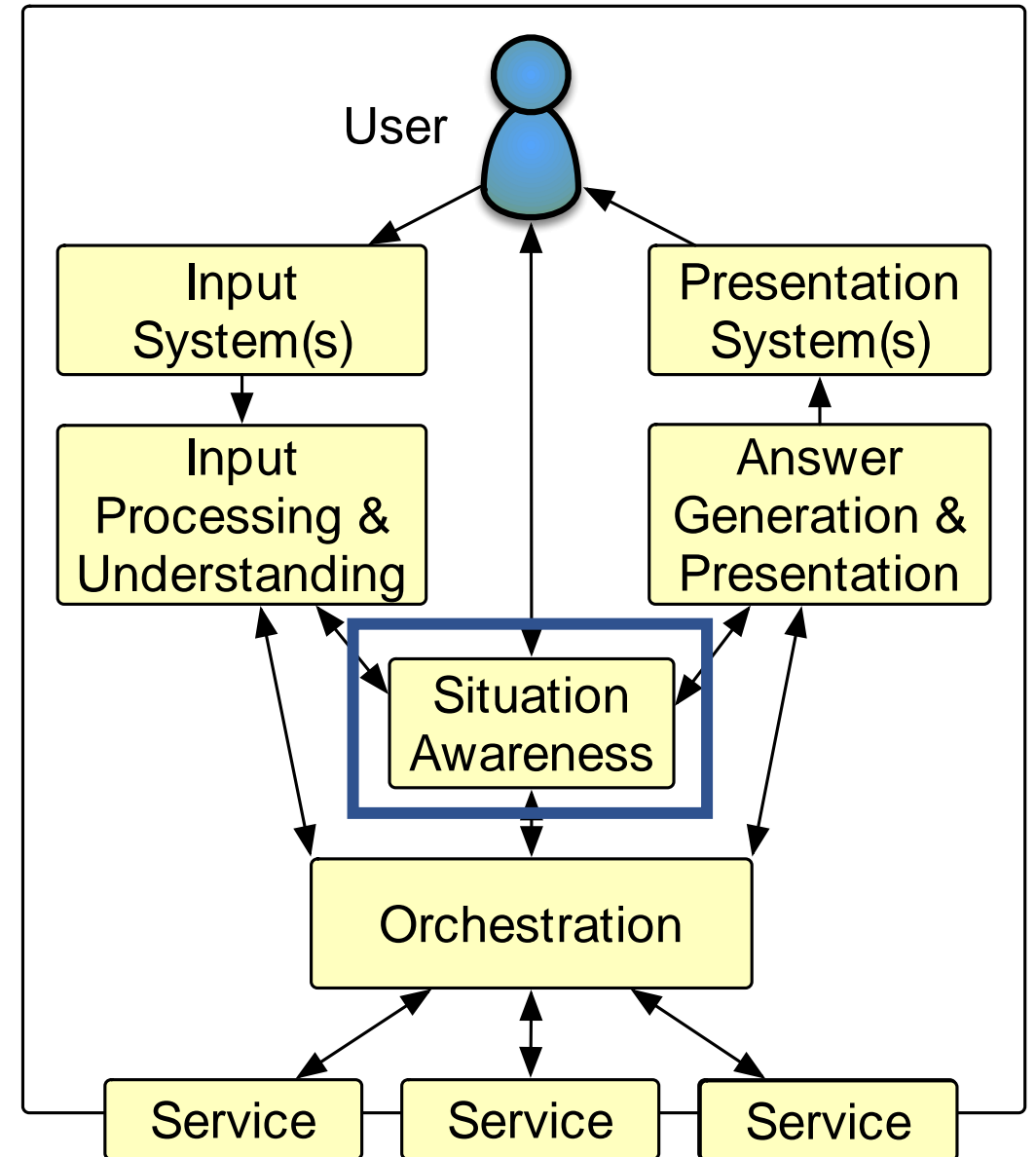
Autonomous Virtual Assistant Architecture

This information is approved for release
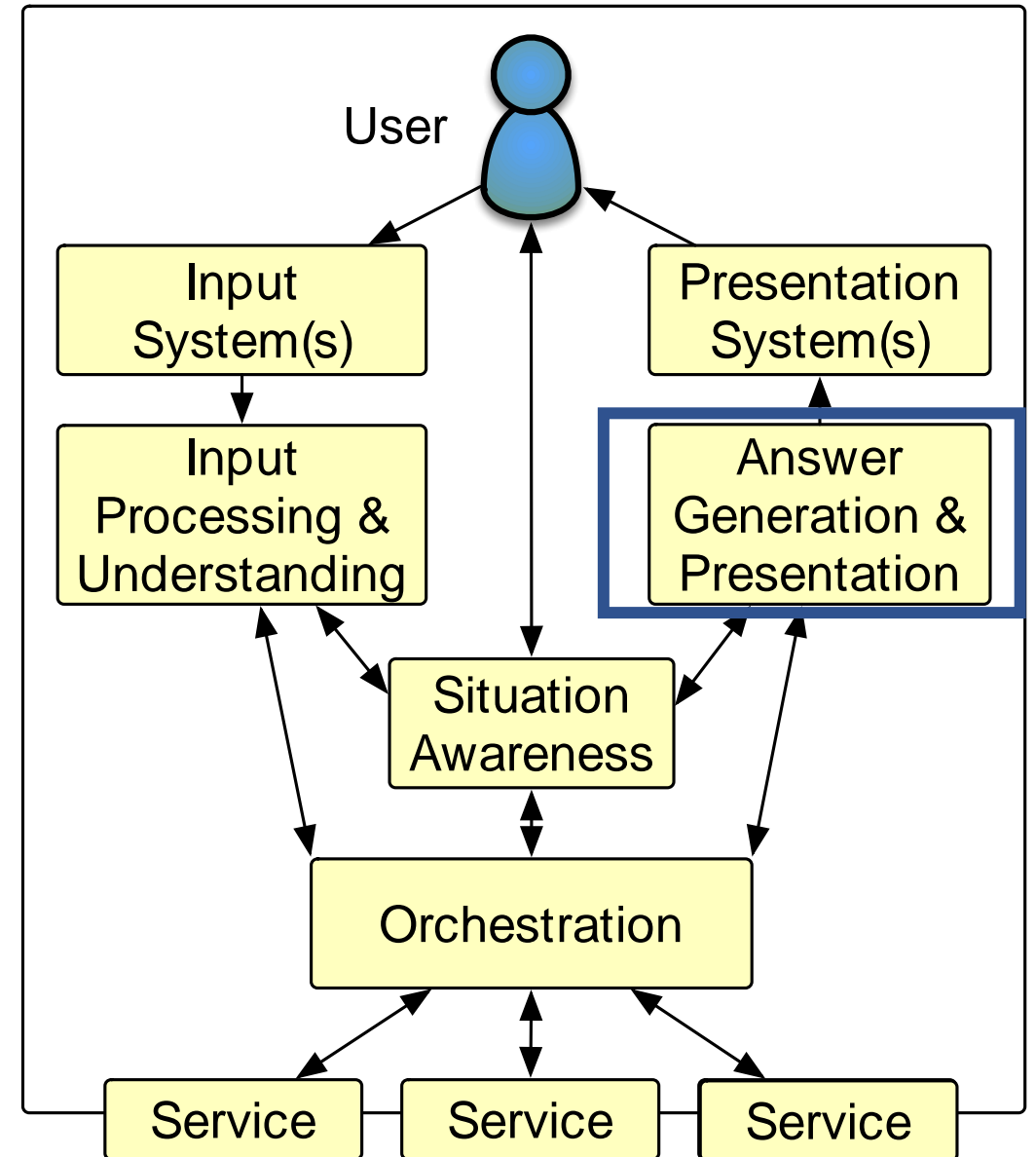
Situation Awareness

- **Purpose: build+maintain awareness about user, task, environment, … to support other processes**

- Provides context for:
  - understanding user input
  - deciding if/how/when to present info
  - …

- Not just a repository of information, but an **active process** to maintain awareness
  - Could be itself going out to find information, asking the user, etc.

Autonomous Virtual Assistant Architecture

User

Input System(s)

Presentation System(s)

Input Processing & Understanding

Answer Generation & Presentation

Situation Awareness

Orchestration

Service   Service   Service

This information is approved for release

Orchestration

- **Purpose: manage and coordinate AVA's connection to services, manage tasks in coordination with user**

- Could be simple connection – e.g., take a user request and push it out to the appropriate service, return result to user

- Could include planning and execution – e.g., responding to a complex user request that must be staged out over time and over multiple services

- Needs to know about APIs to services
- Needs to know about user-relevant tasks

Autonomous Virtual Assistant Architecture

This information is approved for release
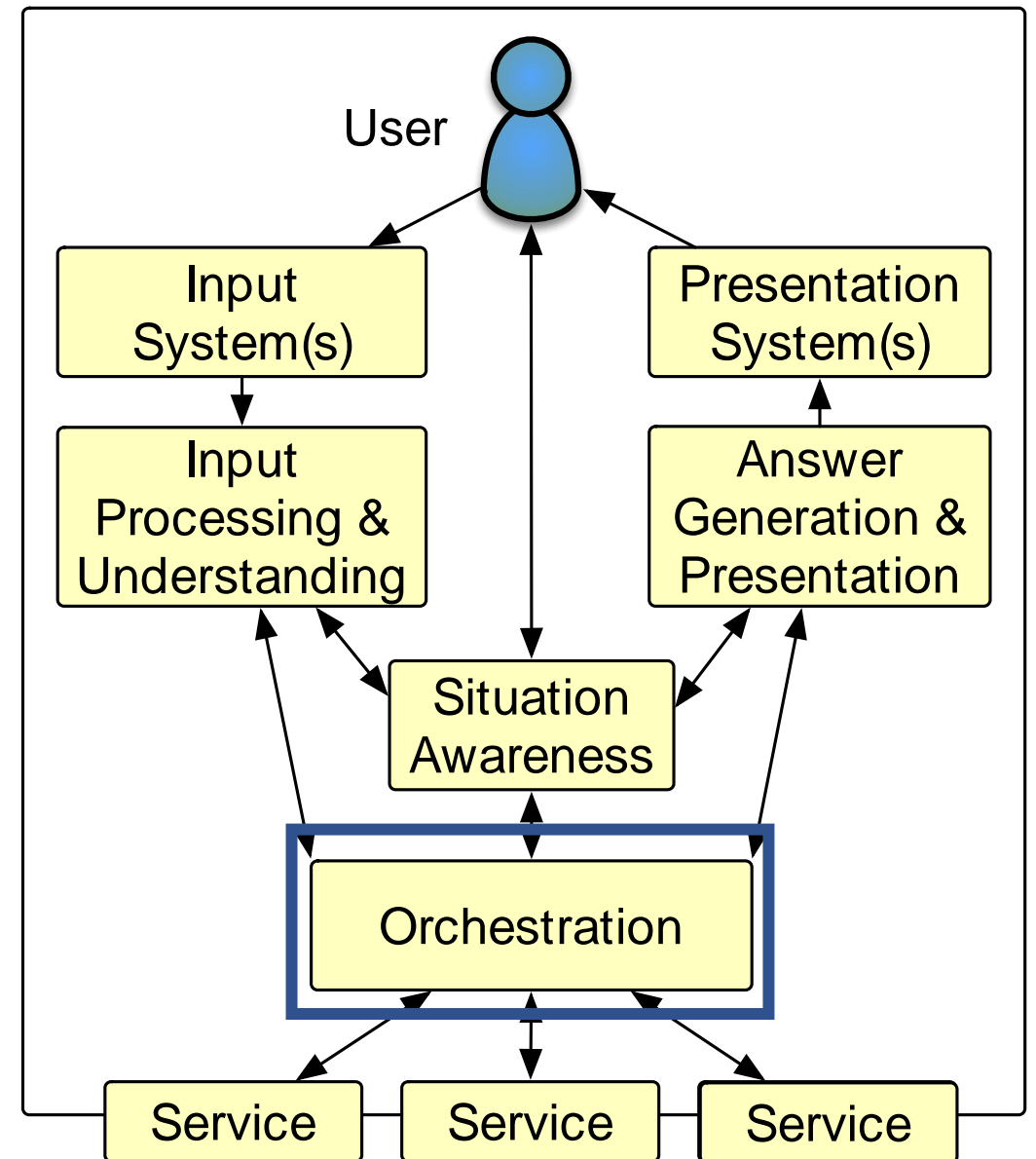
Services

- **Purpose: encapsulate functions, capabilities, etc., that are of interest to the user/task**

- Could be external: e.g., an existing database, vehicle capability/sub-system, …

- Could be "internal" to AVA if idiosyncratic to your project, but outside of core AVA functions
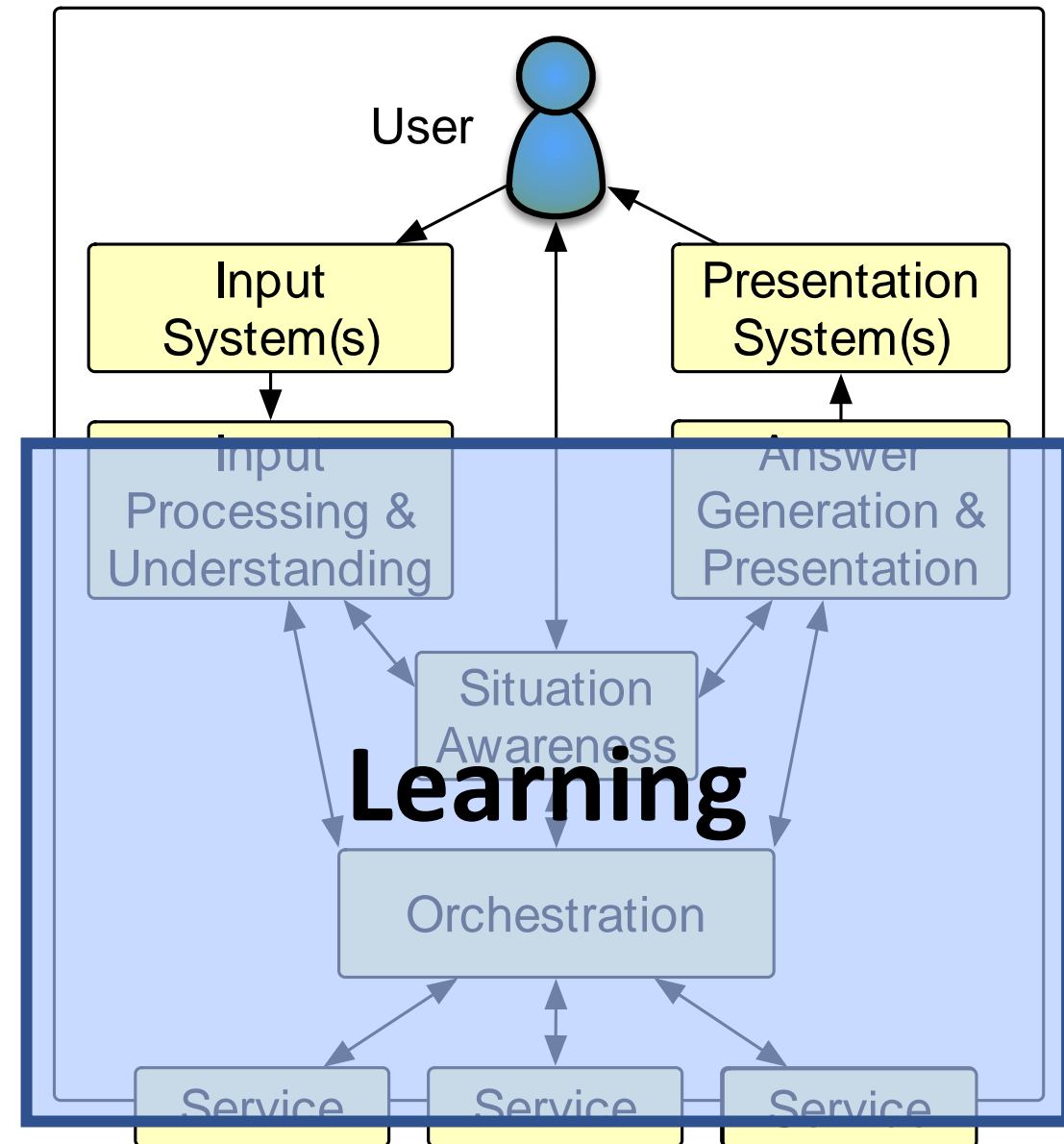
- Anything can be wrapped as a service

Autonomous Virtual Assistant Architecture

This information is approved for release
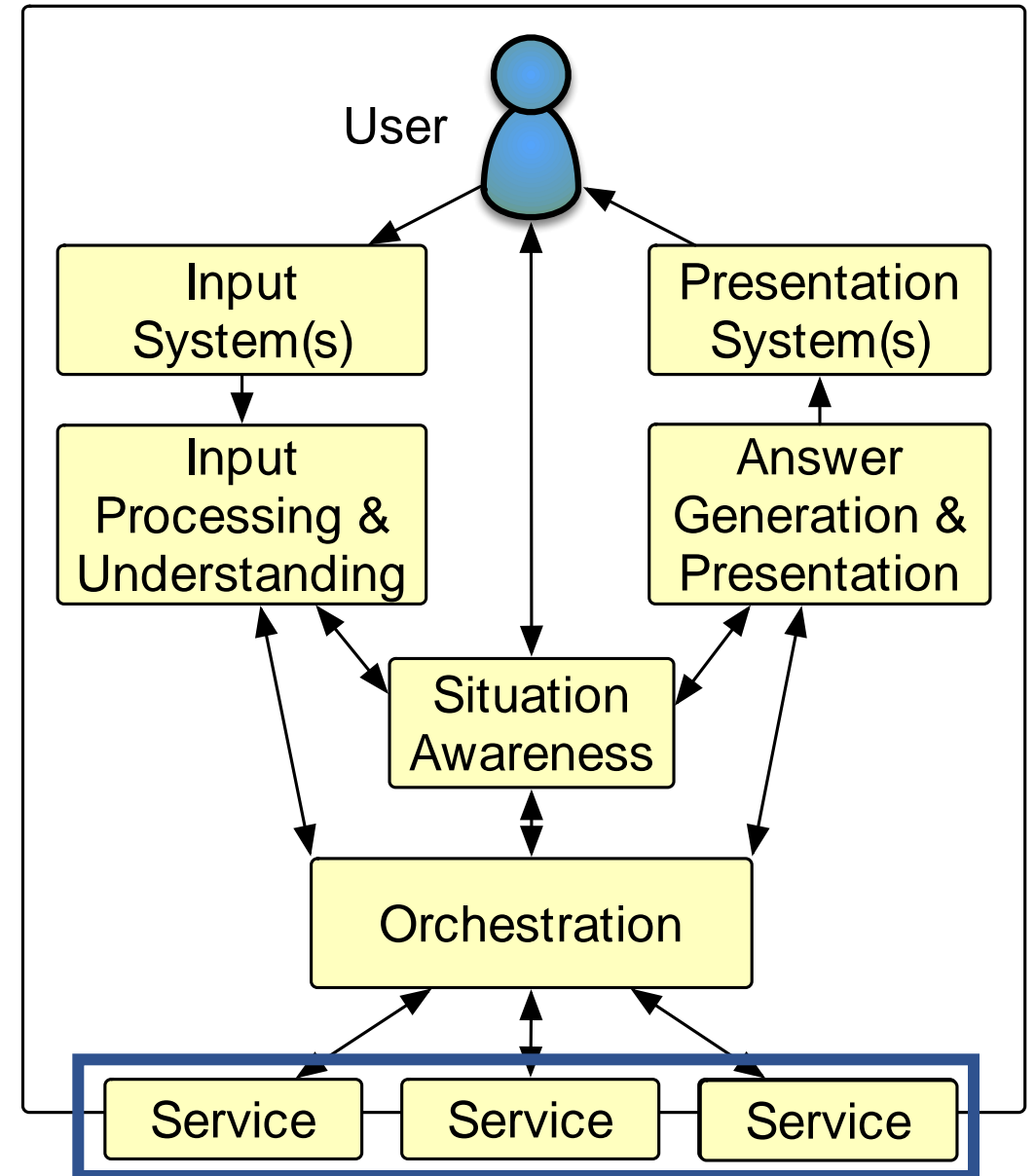
Learning

- **Purpose: adapt the behavior of the system to improve over time**

- Not a specific functional element, but cross-cutting : could be applied to much of the rest of the system, individual functions or across functions

- Examples:
  - Learning to understand new user inputs
  - Learning user preferences for information presentation
  - Learning how to do new tasks

Autonomous Virtual Assistant Architecture

User

Input System(s)

Presentation System(s)

Input Processing & Understanding

Answer Generation & Presentation

Situation Awareness

Learning

Orchestration

Service

Service

Service

This information is approved for release

AssistMATE™ example: Help managing multiple UxVs

- **Purpose: make UxVs easier to use**

- Multi-modal dialogue interface – specify plans, manage execution, get/give situation awareness

- Connects to underlying 'services' like route planning, video servers

- Provide multiple forms of feedback: visual (map, graphical), text/chat, verbal

- Provide alerts when certain events of interest occur (e.g., enemy crossing a line, units linked up)

This information is approved for release
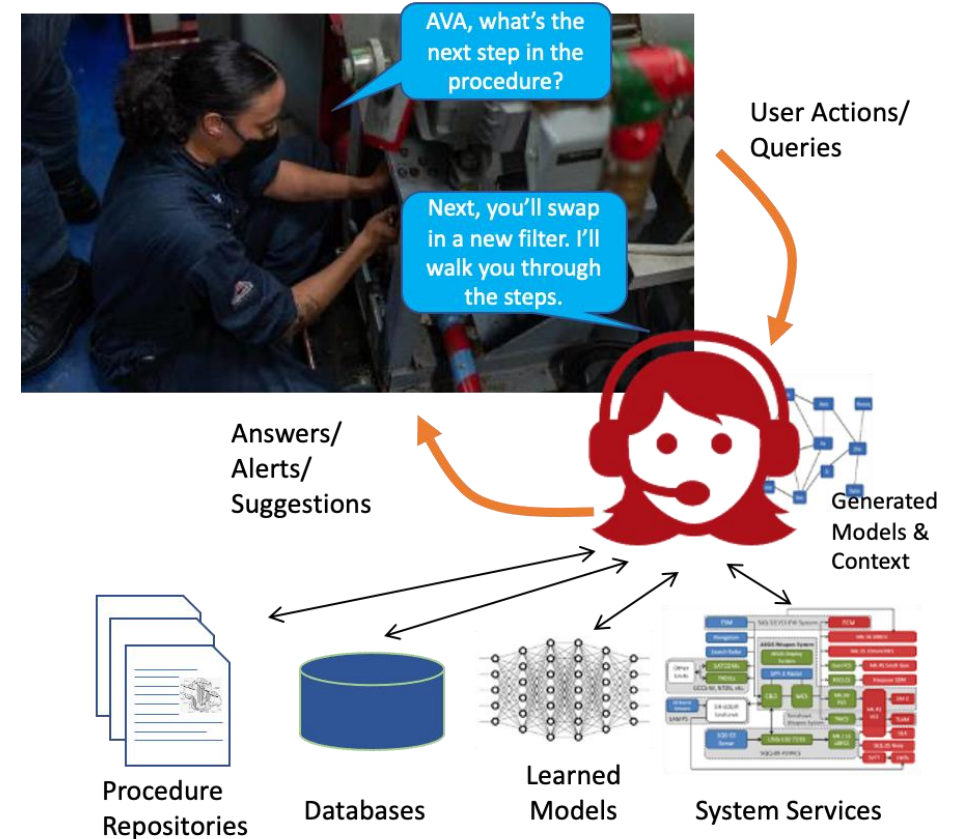
AssistMATE™ example: NASA procedure helper

- **Purpose: help users work through procedure-based task**

- Conversational interface: 'Where do I find the 3/8" torque wrench?' ... 'Can you read that step again?'

- Multi-modal presentation – e.g., present steps in different modalities depending on content, user requests; account for user situation, limitations

- Connected to tools/equipment database 'services' to know where to find tools, etc.

- Could be embodied on mobile platform like Astrobee or disembodied

This information is approved for release

Image: NASA

Conclusion

- An autonomous virtual assistant is more than Amazon's Alexa or Apple's Siri
  - Potential to be a teammate rather than a tool, perform tasks autonomously with delegated authority

- Many potential applications for autonomous virtual assistant in DoD and related domains
  - Help with complex tasks, non-experts, limited reach-back ability

- Reference architecture can serve as as starting point for developing virtual assistant for different use cases

This information is approved for release

Thanks!

This information is approved for release