How to Add Voice to Your Next Product

Robert Thompson

Director i.MX Ecosystem

June 2019 | Session #AMF-SMH-T3535

NXP
SECURE CONNECTIONS
FOR A SMARTER WORLD

NXP: Enabling Audio & Voice Processing at the Edge [1,2,3]

"Voice control front end to grow at 29% CAGR ('17-'22) in the Smart Home"

NXP Enabled
Smart Speakers

Google Voice Assist

Alexa Voice Services

Dev Kit for Amazon AVS

Voice Solutions for China

NXP i.MX 8M
Largest Voice Solution
provider for Mandarin

1. Source: Google Android Things developer website
2. Amazon AVS developer website
3. Source: ABI Research
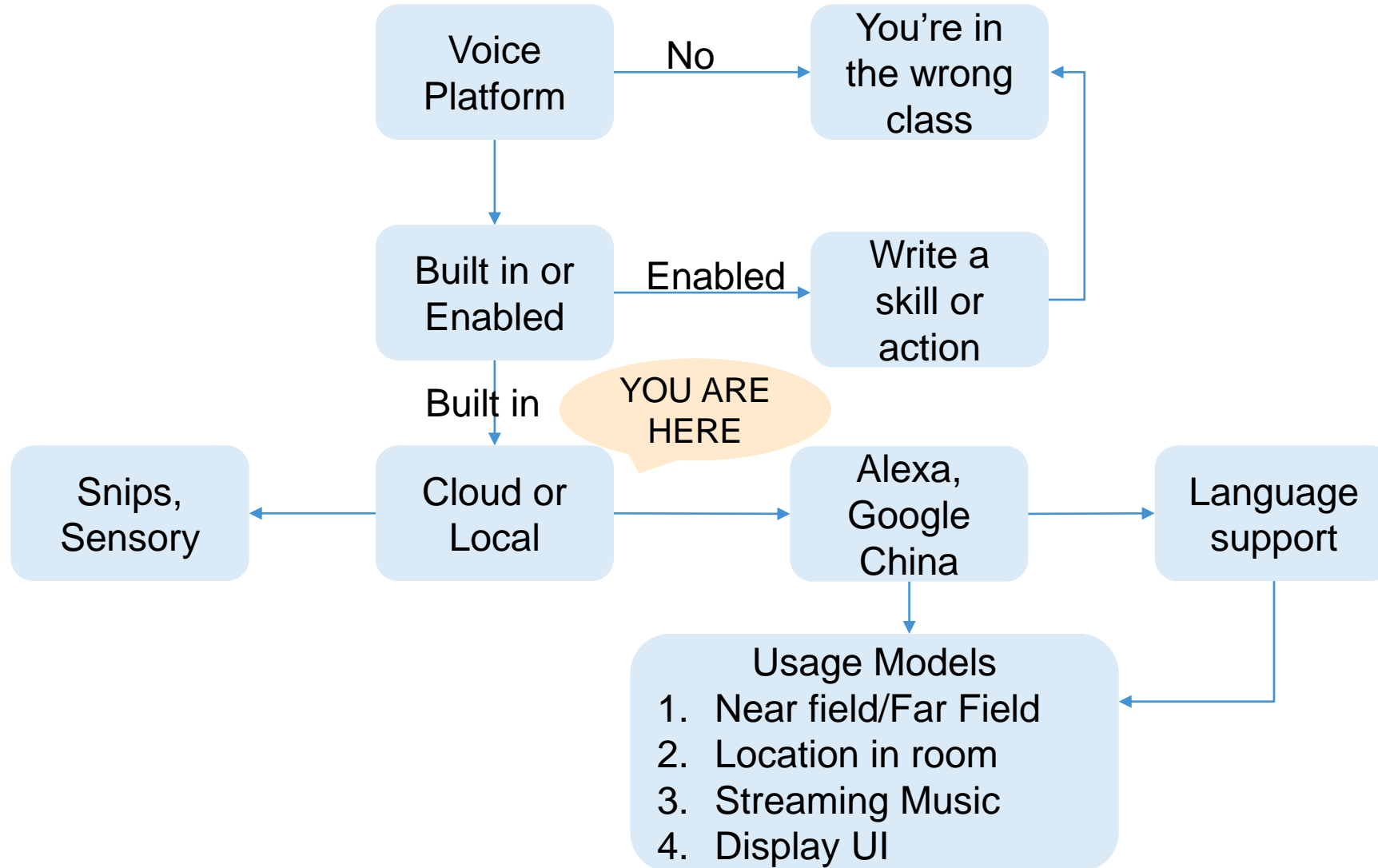
Agenda

- Introduction to Voice

- Selecting a Voice Front End

- Embedded Voice Solutions

  - Alexa Voice Services (AVS)

  - Google Voice Assistant (GVA)

  - Local Voice

- NXP Voice Ecosystem
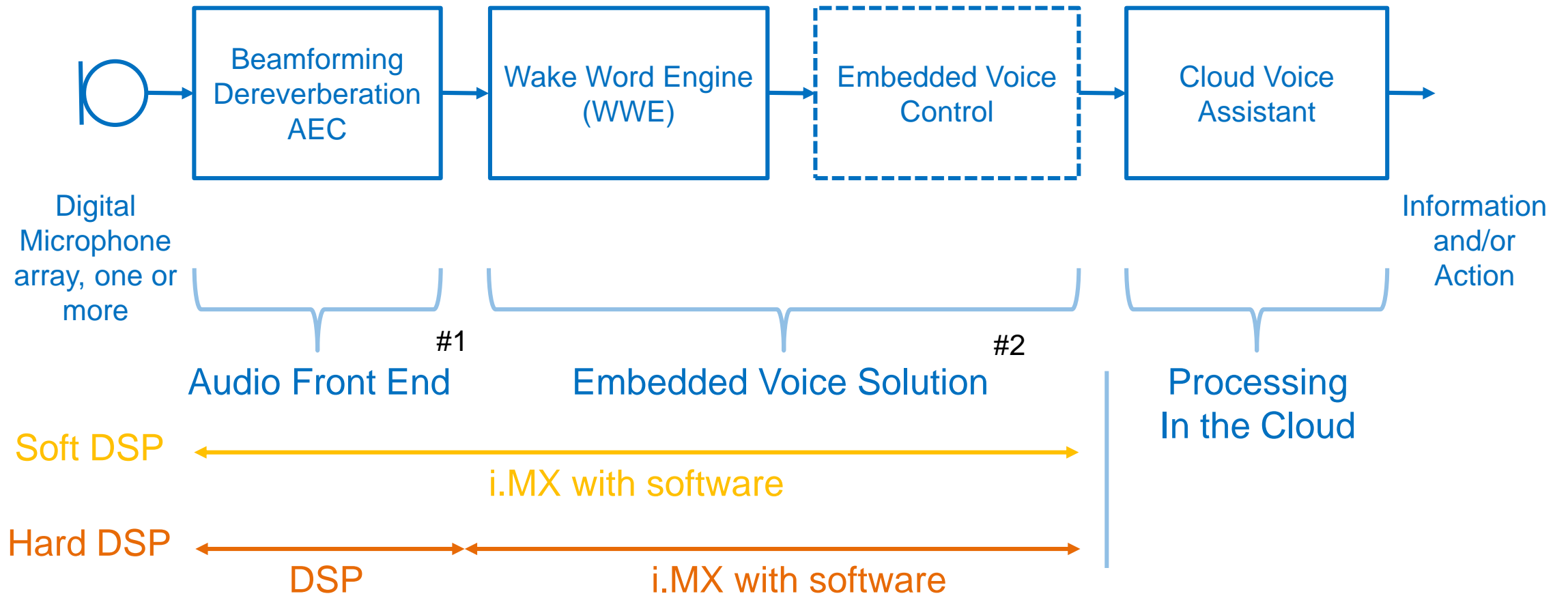
Key Takeaways from this Session:

By the end of this session you will know:

1. What is an Audio Front End & choices based on i.MX processors

2. What Embedded Voice Solutions options are available based on i.MX

3. What NXP partners can help develop and certify your voice enabled device

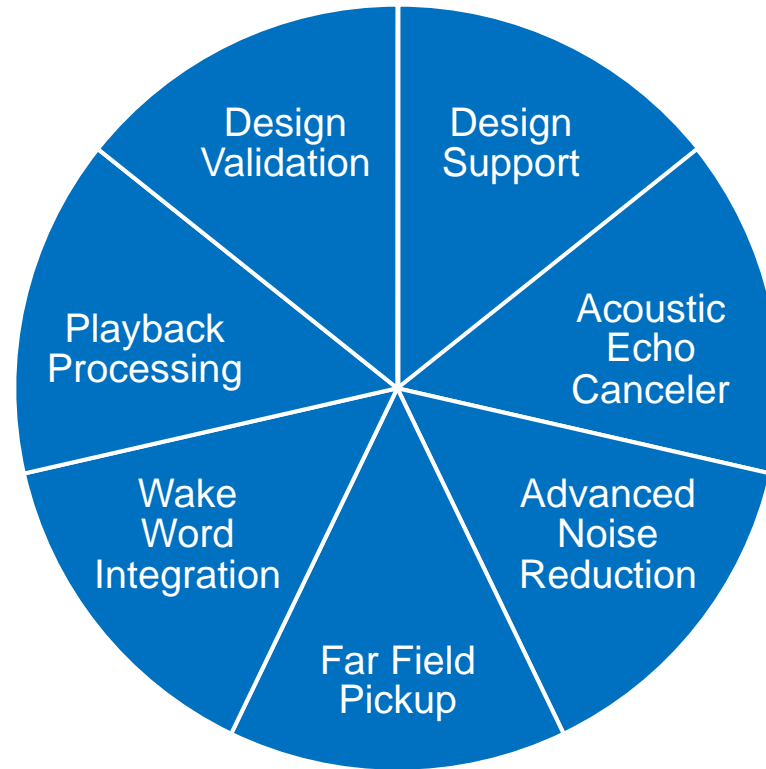# Adding Voice to Your Next Product…The Decision Flow

Voice Platform

No → You're in the wrong class

Built in or Enabled

Enabled → Write a skill or action

Built in ↓

YOU ARE HERE

Snips, Sensory ← Cloud or Local → Alexa, Google China → Language support

Alexa, Google China ↓

Usage Models
1. Near field/Far Field
2. Location in room
3. Streaming Music
4. Display UI

Typical Far Field Voice Systems

#1: Selecting a Voice Front End

Full-Featured Voice UI Solution

Acoustic Echo canceler – allows microphones to hear voice commands even while music is being played back
Beamforming – combines multiple microphone signals to improve the quality of voice recognition
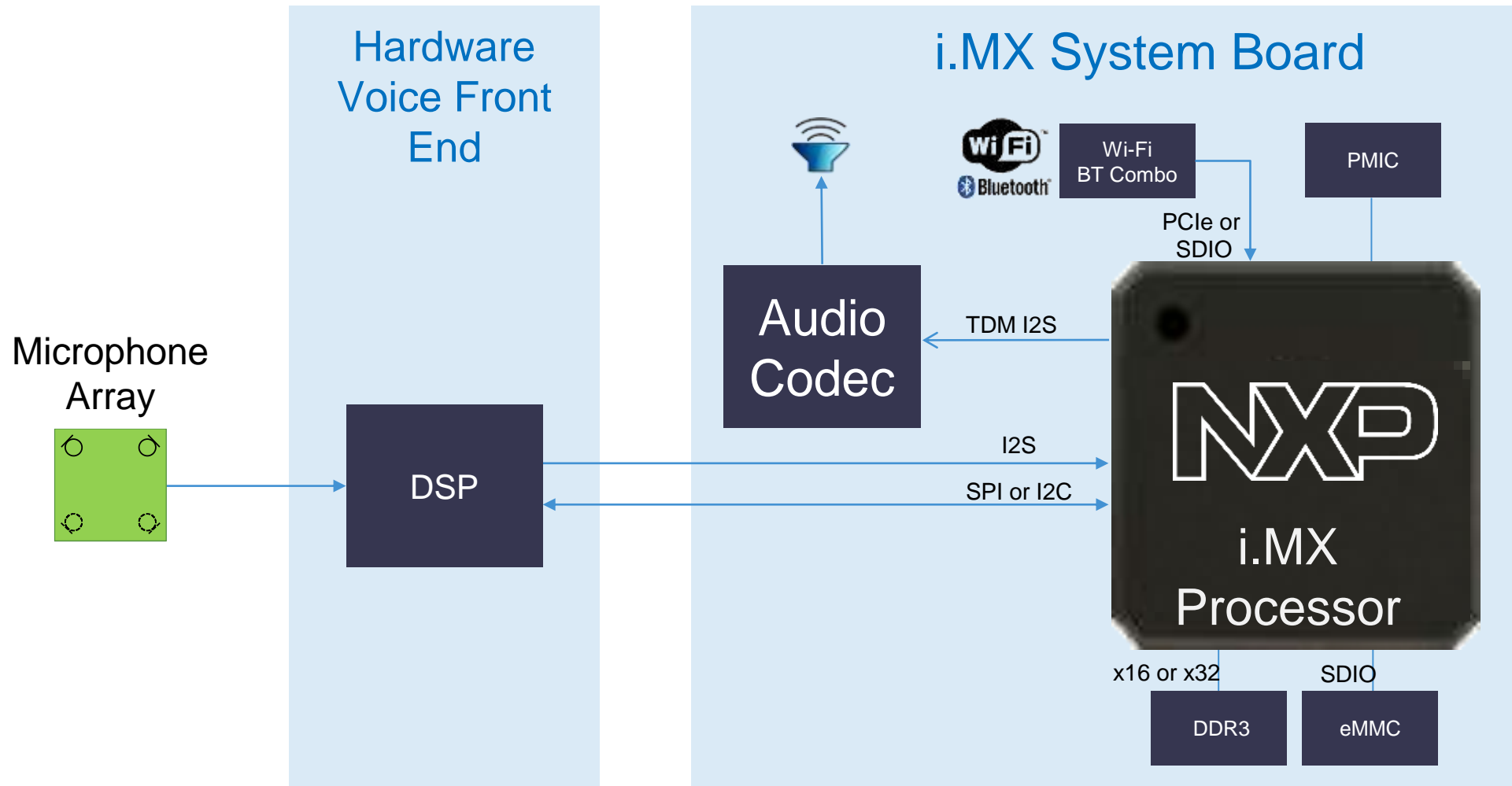Direction of Arrival – determines direction that the voice is coming from. Used in conjunction with beamforming
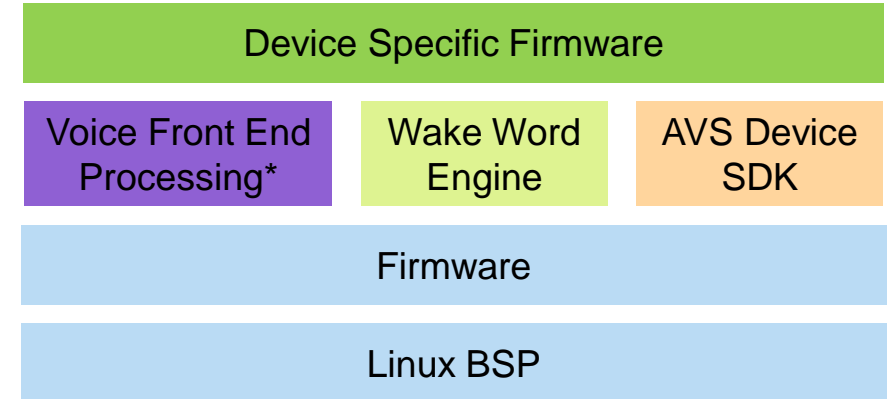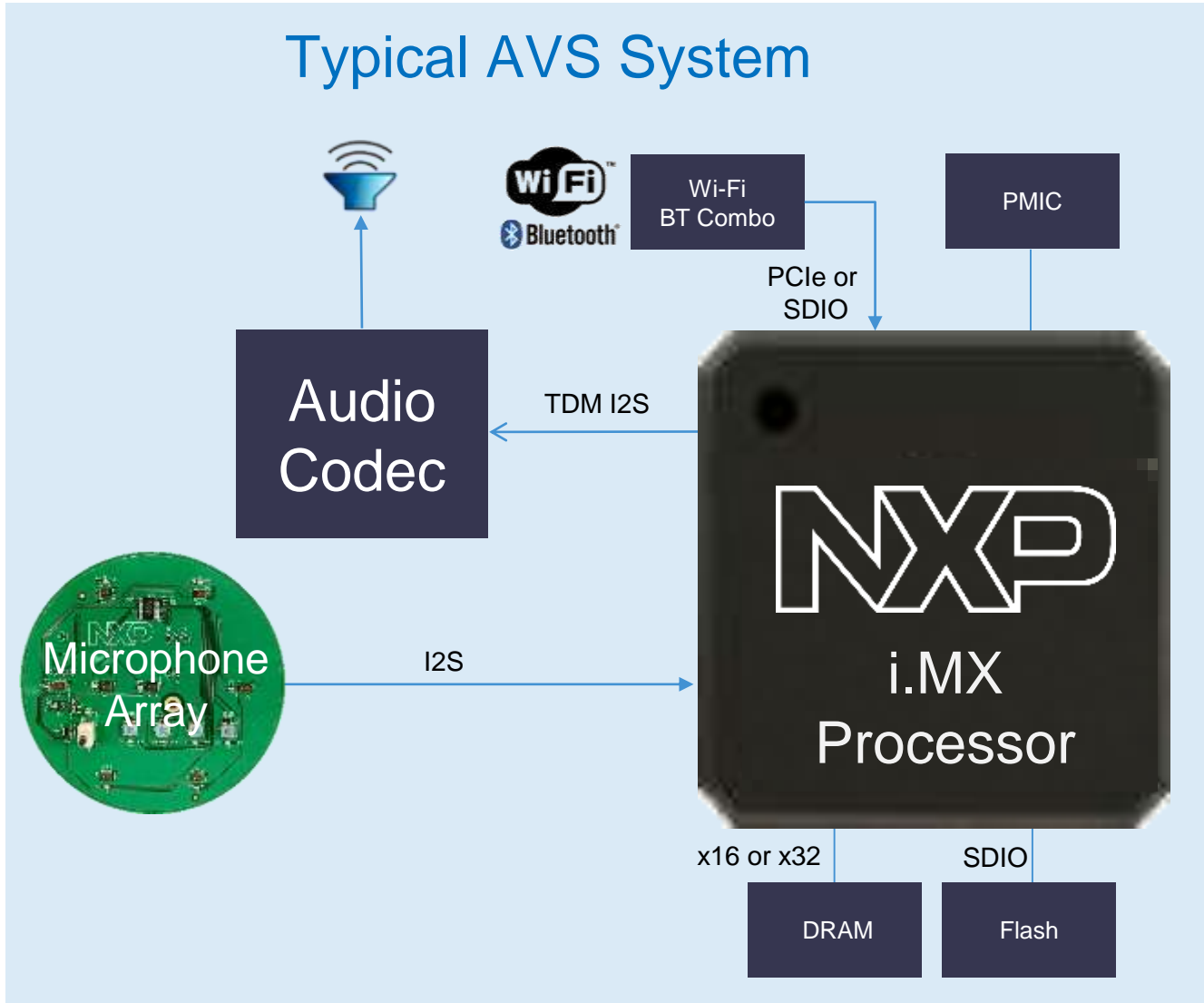Noise Reduction – removes background noise to further improve voice recognition
Wake Word – recognizes the trigger word "Alexa"
Playback Processing – makes the loudspeakers sound good

Hardware Voice Front End Example

Software Voice Front End Example

Typical AVS System

Audio Codec

Microphone Array

Wi-Fi BT Combo

PMIC

NXP i.MX Processor

Wi-Fi Bluetooth

TDM I2S

PCIe or SDIO

I2S

x16 or x32

SDIO

DRAM

Flash

Device Specific Firmware

Voice Front End Processing*

Wake Word Engine

AVS Device SDK

Firmware

Linux BSP

- Device developer
- AVS
- AVS
- Software Algorithms
- NXP

Optional, software or hardware implementation

i.MX - Audio Front End (AFE) Solutions

- The market is currently seeing a wide range of usage models for voice integration
- Customers can choose the audio front end best aligned to their needs
- The starting point is to share benchmarks with the different AFE's

Partner	Microphones	Front End Processing	Amazon Approved	Interfaces	Demo Availability
Synaptics	2	Hardware	Yes	USB I2S SPI	Now
	4				
Microsemi	2	Hardware	Yes	I2S	Now
DSP Concepts	2 to 7	Software	Yes	API	Now
Sugr	1-4	Software	Yes	API	Now

NXP i.MX8M Development Kit for Amazon Alexa Voice Service

- **First** AVS development kit approved to run the Alexa wake word & **Multi-Room–Music**

- **Modular** design shortens time-to-market- 2 Mic board & production ready SOM

- Complete software stack based on Yocto Linux with Audio front end from DSPConcepts

- Targets from **high end audio devices** to **smart home** for Alexa built-in

- Availability now for $219

DSP Concepts Voice Algorithms – Barge in Performance

The most difficult function to achieve is "barge-in" and requires multiple algorithms to be working together seamlessly

Barge-In Benchmarking

- Reflects performance of the overall system
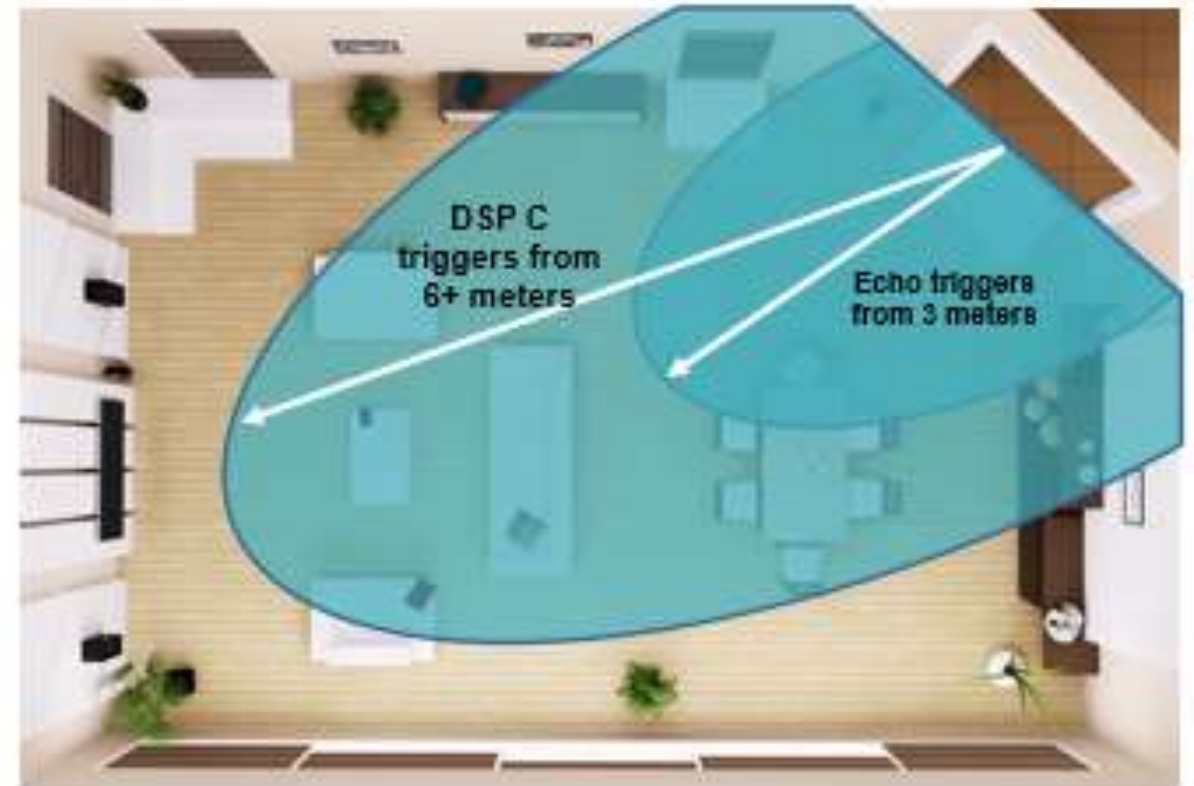- *DSP Concepts beats Amazon Echo by 6 dB*

Example Use Case

- Music playing at 90 dB SPL
- Normal voice at 61 dB SPL

Google Home triggers 48%

Amazon Echo triggers 72%

DSPC triggers 92%

DSP C triggers from 6+ meters

Echo triggers from 3 meters

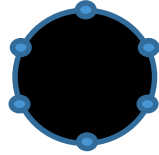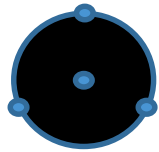# Appliance / Tablet Designs

Good

Better

- 180 degree operation
- 2 or 4 microphone linear array
- 25 to 75 mm design
- Physically separate microphones and loudspeakers for best performance
- Mono or stereo playback

Set-Top Box Designs

High-End

Standard

- **Top of Device**
  - 180 degree operation
  - Microphones on top of product
- **Tethered "puck"**
  - 360 degree operation
  - Microphones on top of product

- **Support for optional internal speaker for voice playback**
- **Audio playback through HDMI**

Comparison of Hardware and Software Front End CPU Loading

Average CPU % Remaining Voice Front End Comparison (Higher is Better)
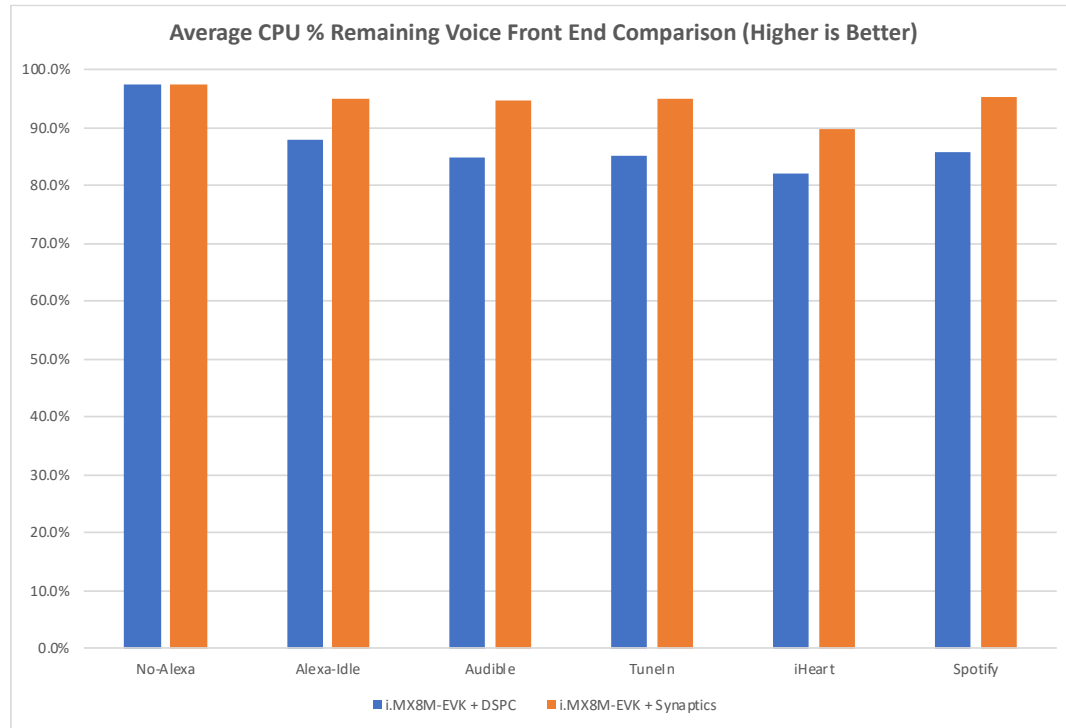
| System Configuration	No-Alexa	Alexa-Idle	Audible	TuneIn	iHeart	Spotify
i.MX8M-EVK + DSPC	97.5%	87.8%	84.9%	85.1%	81.9%	85.8%
i.MX8M-EVK + Synaptics	97.5%	94.9%	94.8%	94.9%	89.6%	95.2%

System Configurations:
- AVS v1.9 (Amazon Benchmarks)
- i.MX8M EVK
- i.MX8M system clock = 1.5 GHz
- 2 Microphones
- Hardware Front End;
  - Synaptics CX20921 evaluation board
- Software Front End;
  - Voice HAT
  - DSP Concepts (DSPC) software

Note:
Graph shows average CPU throughput percentage available for application per AVS function (higher is better)

i.MX8M EVK with DSP Concepts 2 Mic FE AVS v1.9 Performance
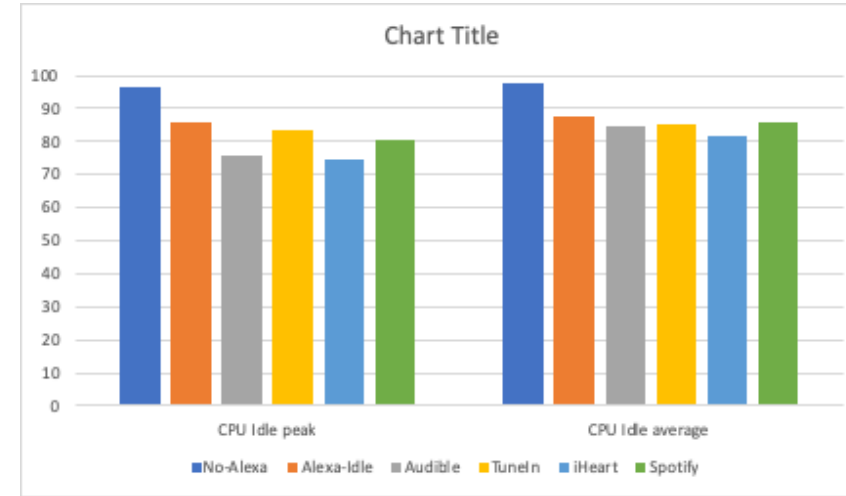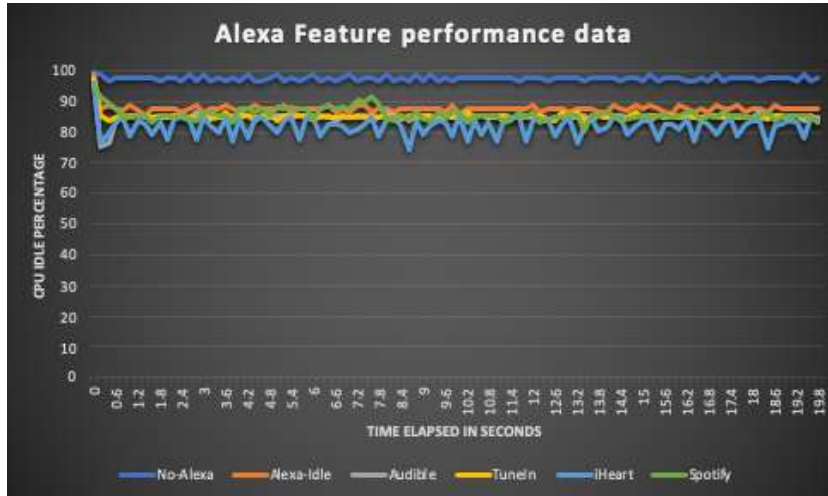
Graph and table show processor throughput available for other functions (*higher is better*)

Performance of i.MX8M-EVK with 2 mic Voice HAT and DSP Concepts software front end running Amazon benchmarks with AVS SDK 1.9 and 1.5 GHz system clock

		No Alexa	Alexa Idle	Audible	TuneIn	iHeart	Spotify
Peak CPU Available	CPU%	96.6%	85.7%	75.6%	83.5%	74.4%	80.4%
	DMIPS	13330	11826	10432	11522	10267	11095
Average CPU Available	CPU%	97.5%	87.8%	84.9%	85.1%	81.9%	85.8%
	DMIPS	13461	12119	11724	11752	11312	11844

Total DMIPs available = 13800 (1500 * 2.3) * 4

8_Mic Rev D Connected to i.MX 8M Mini EVK

Mic arrays and UI are suitable for AVS, and are GVA compatible
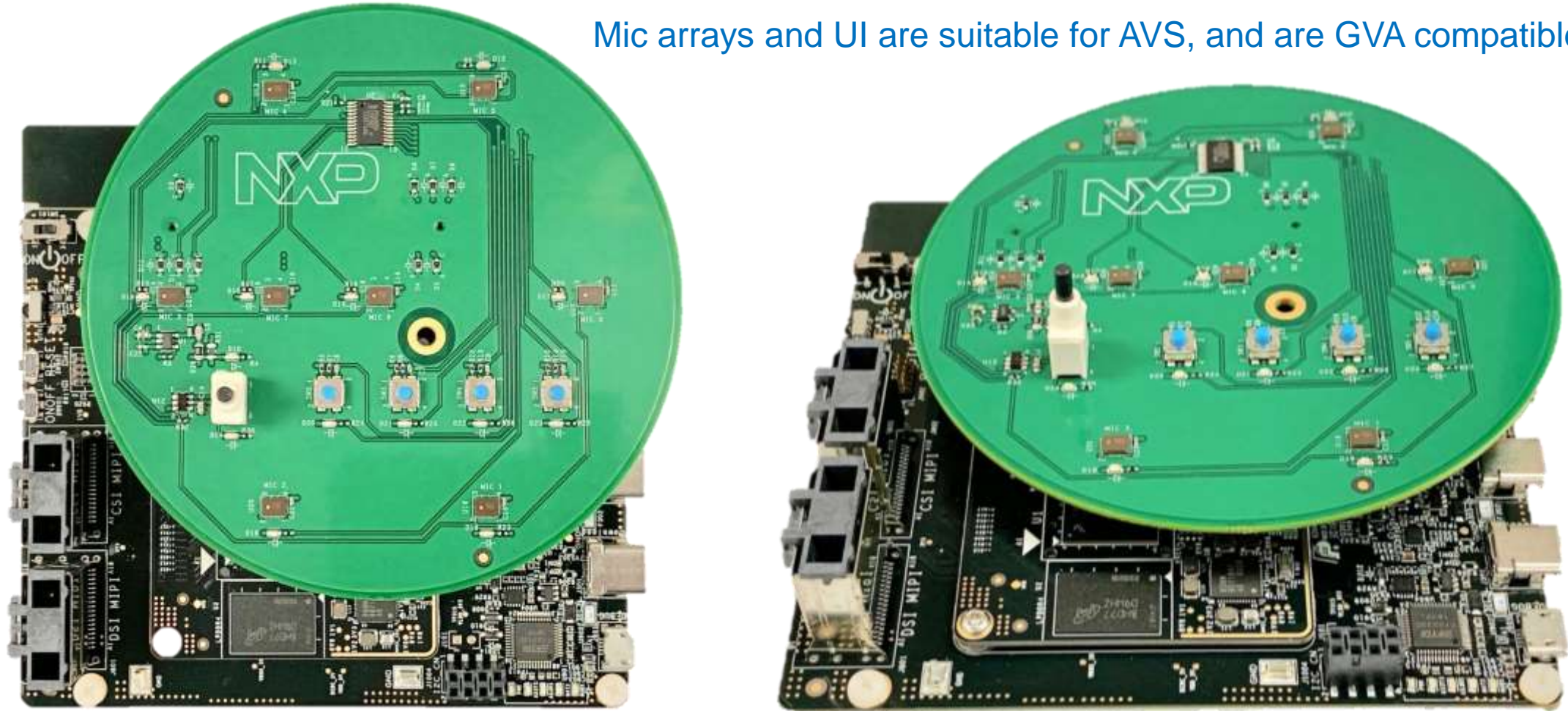
8_Mic Rev D plugs directly into the i.MX8M Mini EVK 40-pin RPi connector

#2: Embedded Voice Solutions: Alexa Voice Services (AVS)

Alexa Voice Service Features Comparison

Feature	AVS (MCU)	AVS (MPU)
Dialogs (inc. multiturn)	✓	✓
Alerts	✓	✓
Skills & 3rd Party Skills	✓	✓
Kindle	✓	✓
Audible	✓	✓
Notifications	✓	✓
Routines	✓	✓
Shopping	✓	✓
Alexa announce	1Q19	✓
Alexa Messaging	1Q19	✓
Alexa Calling	2Q19	✓
Video calling	X	X
Amazon Music	✓	✓
iHeart Radio, Pandora, TuneIn	1Q19	✓
Spotify	X	✓
Multiroom music & Bluetooth sync	X	✓
Basic GUI	X	X
Rich Interactive GUI	X	X
Amazon Video	X	X

Current AVS Approved Development Kits

System Dev Kits

Complete, end-to-end system reference designs for creating products with Alexa built-in.

Amazon Alexa Smart Screen Dev Kit
Features a production-ready system reference to create beautiful screen-based products with Alexa built-in.
Limited Availability
Learn More »

NXP i.MX RT MCU Alexa Voice Service Solution
End-to-end Alexa built-in reference solution for low-cost MCU-based AWS IoT Core connected devices running Amazon FreeRTOS.
Limited Availability
Learn More »

Qualcomm Mesh Networking Dev Kit
For mesh routers based on Qualcomm's single-SoC Wi-Fi router platform and Mecomi's ClearVoice audio algorithms.
$1,395.00
Learn More »

Qualcomm Smart Headset Reference Design
The first solution qualified by Amazon for Bluetooth headsets and with support for the Alexa Mobile Accessory (AMA) protocol.
$299.00
Learn More »

HiSilicon STB Push-to-Talk Dev Kit
Enables set-top-box (STB) device makers and TV operators to build Alexa into new designs, and retrofit deployed STBs to support Alexa.
Limited Availability
Learn More »

NXP i.MX 8M Dev Kit
Pre-integrated Multi-Room Music capabilities, Far-Field voice capture, and with the Amazon Wake Word Engine.
$219.00
Learn More »

An AVS approved Dev Kit has passed the AVS test suite. A final product has to pass the same test suite.

Amazon uses approved kits to roll out latest features (MRM) and Alexa for MCUs

NXP i.MX RT MCU AVS Solution

- First MCU-based AVS development kit

- Targets smart home for Alexa built-in

- Enables significantly lower cost

- Turnkey design shortens time-to-market

- Availability

  - Now (limited access)

  - May (distribution $49)

NXP i.MX 8M Development Kit for Amazon Alexa Voice Service


Kit focus
1) Easy out of box hardware set up
2) Detailed User Guide
3) Software set up simplified by scripts to automatically install 3rd party libraries at runtime

Kit Hardware	Software
SOM + Baseboard: Pico-Pi i.MX8M	Yocto 4.14 Linux
Mic Array: 2 Mic Voice HAT	Alexa Device SDK & Wake Word
Antennae, Speakers + cables	Audio Front End: DSP Concepts

Development Kit Software Image
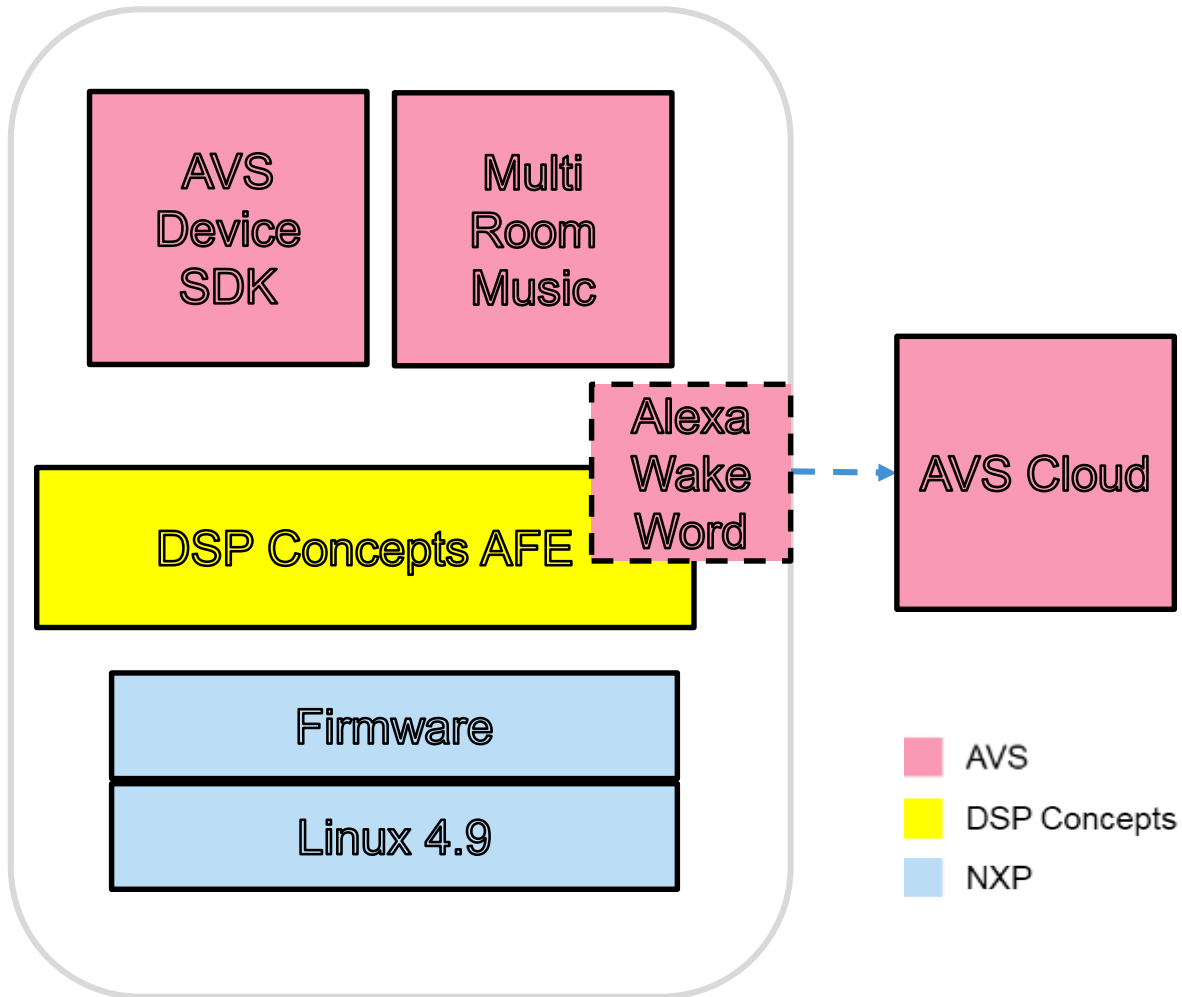
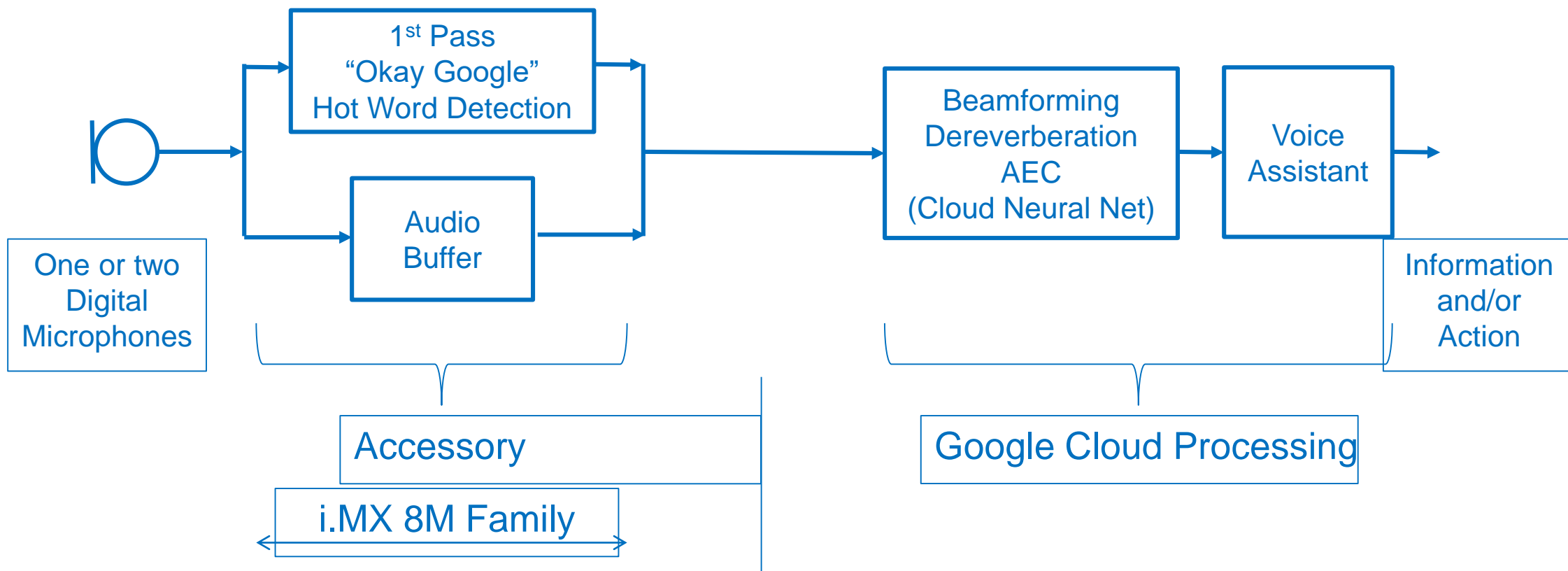- Complementary to the NXP BSP for i.MX

- Allows to build the AVS SDK or install it at runtime

- Includes scripts to install runtime libraries for easy build

- Easily customizable for use with other i.MX boards

AVS Device SDK

Multi Room Music

Alexa Wake Word

AVS Cloud

DSP Concepts AFE

Firmware

Linux 4.9

- AVS
- DSP Concepts
- NXP

#2: Embedded Voice Solutions: Google Voice Assistant (GVA)

Google Assistant Built In Voice System

One or two Digital Microphones

1st Pass "Okay Google" Hot Word Detection

Audio Buffer

Beamforming Dereverberation AEC (Cloud Neural Net)

Voice Assistant

Information and/or Action

Accessory

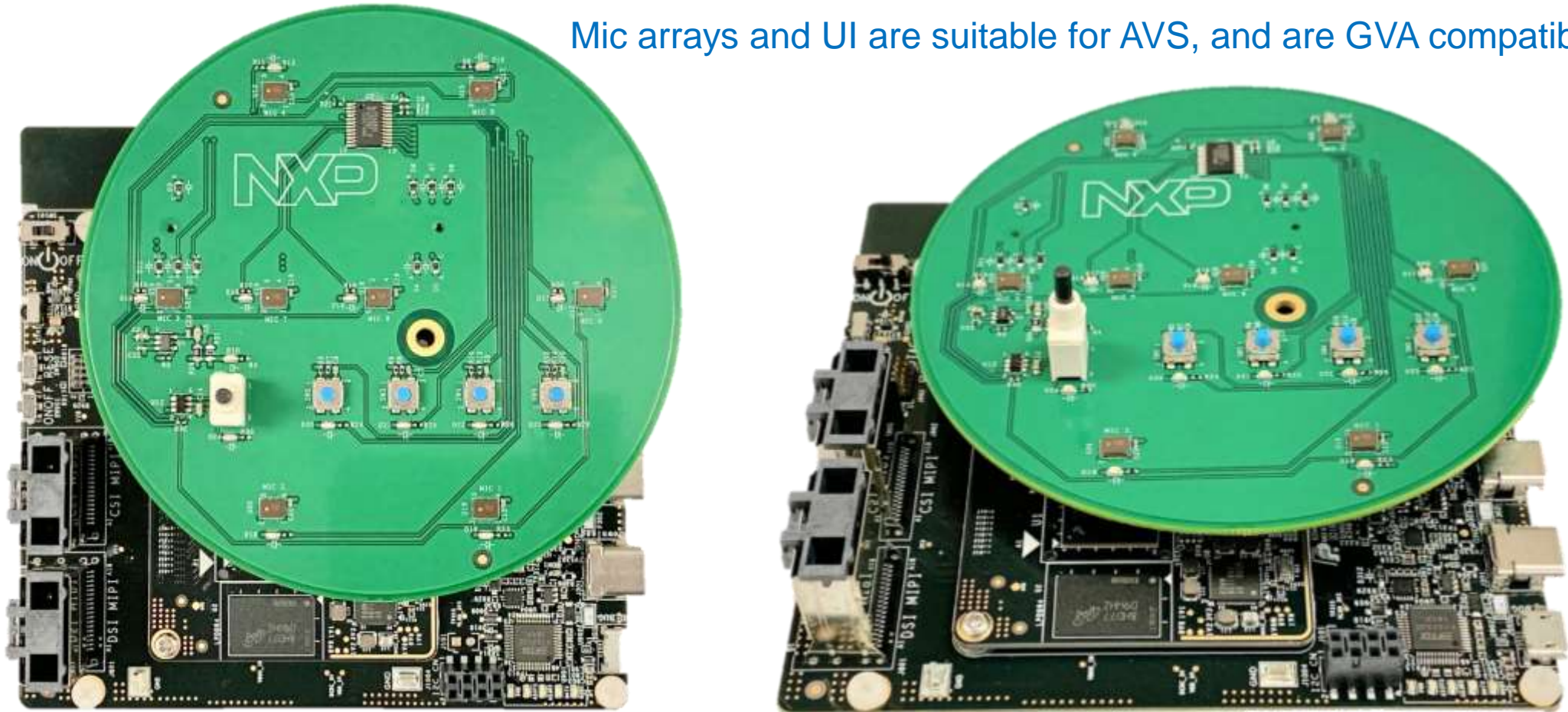i.MX 8M Family

Google Cloud Processing

AEC = Acoustic Echo Cancellation;   Important distinction: Far field (across the room) vs Near field (ie. on a headset)

i.MX 8M Mini GVA Dev Kit

Mic arrays and UI are suitable for AVS, and are GVA compatible

8_Mic Rev D plugs directly into the i.MX8M Mini EVK 40-pin RPi connector

Google Voice Assistant (GVA) System Requirements

Google Voice Assistant (GVA) includes (Google) Cast audio streaming

- Minimum processor requirements:
  - Quad core A7 at 1.2 GHz
    - 1 core for voice assistant
    - 1 core for Cast audio streaming
    - 2 cores for product software (UI, audio processing, etc)

- Minimum memory requirements
  - 1 GB DRAM
  - 4 GB flash

- 1 microphone, Recommended 2 microphone 66 to 71 mm apart

We recommend customers should leave some headroom over that…

# GVA Certified System Integrators – NXP Engagement

Company	Location	Current SOM Partners	DSP	Key Value	NXP engagement
Frontier Silicon	UK	AzureWave	DSPC	Own full featured streaming audio platform, offers integration and engineering services	Support for i.MX8M Mini
StreamUnlimited	Austria	InnoComm	DSPC	Own full featured streaming audio platform, offers integration and engineering services	Support for i.MX8M Mini

NXP is in the "Onboarding" process to become a GVA Certified System Integrator, expecting completion 2Q2019

#2: Embedded Voice Solutions: Local "Embedded" Voice

The Only Solution that Covers Voice Commands to Natural Language
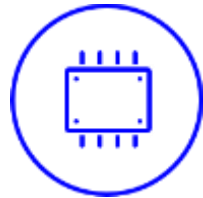
MCU

Simple voice commands
20-30 commands
Cortex M7
512kb

Light On
Dim lights

MPU

Natural Language
Complete Sentences
A9 / 1.2GHz
80MB + 15 MB

Hey Snips, please turn the lights in
Dani's bedroom to warm white and
dim to fifty percent .

snips

Available in Multiple Languages

We Invented a Way to Generate Synthetic Data to Train our Models Without Relying on Real Users

Languages available today

English   French   German

Q2 and Q3 2018

Japanese   Korean   Spanish

Next

Chinese   Portuguese   Arabic

Italian   Russian

With more added on a case-by-case basis

snips

NXP

#3: NXP Voice Ecosystem Summary

Voice Ecosystem Components
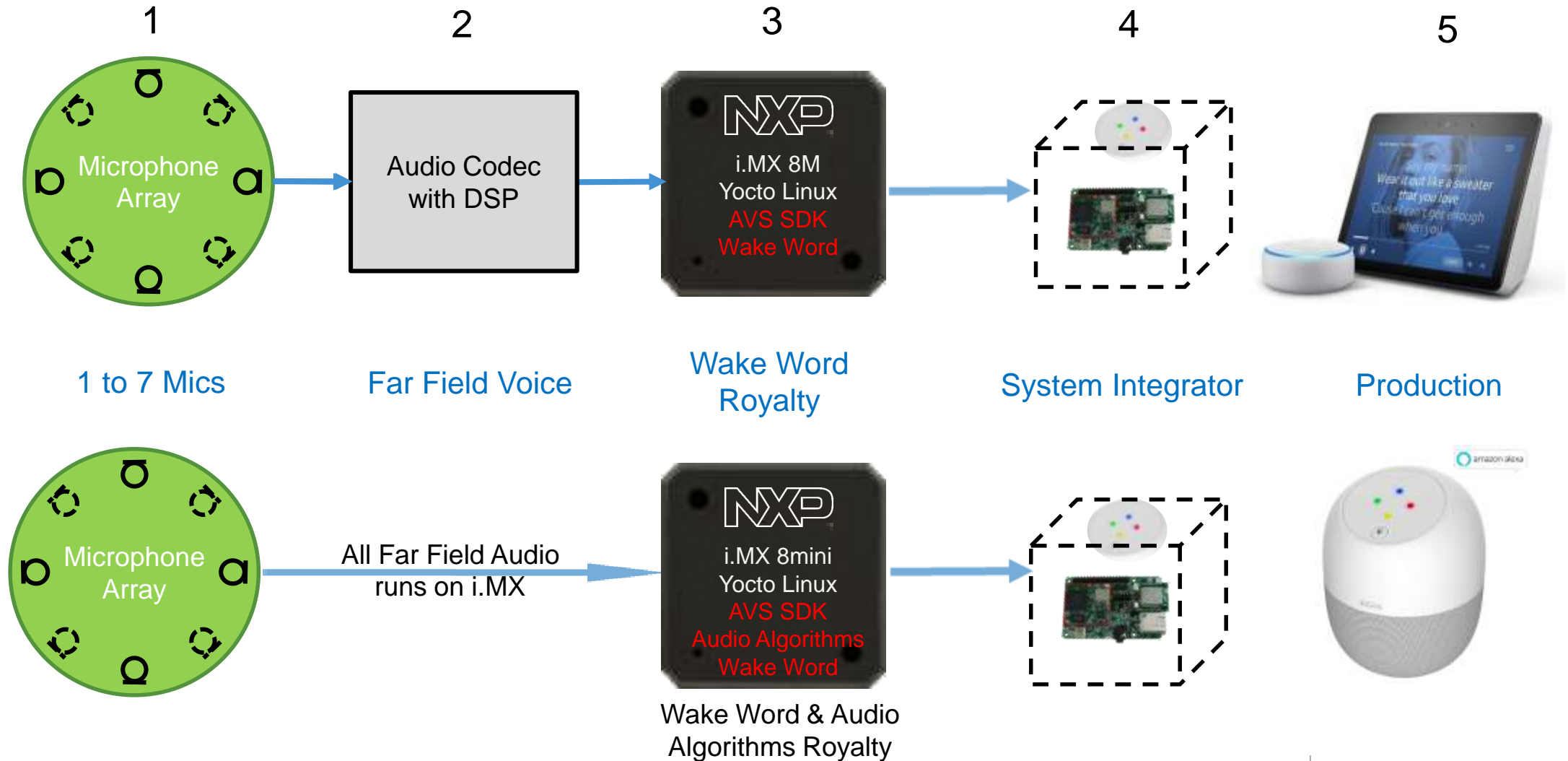
1 — Microphone Array — 1 to 7 Mics

2 — Audio Codec with DSP — Far Field Voice

3 — i.MX 8M Yocto Linux, AVS SDK, Wake Word — Wake Word Royalty

4 — System Integrator

5 — Production

Microphone Array — All Far Field Audio runs on i.MX

i.MX 8mini Yocto Linux, AVS SDK, Audio Algorithms, Wake Word — Wake Word & Audio Algorithms Royalty

AVS Approved System Integrators – NXP Engagement

sugr

i.MX 7D & i.MX 8M Mini

CREOIR
BRANDS GO MOBILE

i.MX 7D & i.MX 8M

CARDINALPEAK

i.MX 8M – scaling partner
with DSP Concepts

FRONTIER SILICON

i.MX 8M Mini

AVS Approved System Integrators – NXP Engagement

Company	Location	DSP	Key Value
SuGR	Shenzhen	Software	Sugr Sense full SW stack solution w soft DSP > 10 products certified by AVS, low power, low cost, Linux & Android supported
Creoir	Finland	Synaptics	Audio & acoustics design house offer complete solution, voice tuning/ certification
Cardinal Peak	USA	DSP Concepts	Audio & acoustics design house offer complete solution, voice tuning/ certification
Frontier Silicon	UK	Various	Own full featured streaming audio platform, offer integration & engineering services
StreamUnlimited	Austria	Various	Own full featured streaming audio platform, offers integration & engineering services

Note: NXP is working to enable several additional system integrators that are not yet AVS approved

Leading i.MX 8M SOM Partners

Board Partner	Solution
Compulab	CL-SOM-iMX8 System-on-Module
Emcraft	i.MX 8M SOM (SOM-IMX8M)
Innocomm	iMX8M System-on- Module: WB10
Seco	SECO SM-C12
Solidrun	SOM iMX8 M
TechNexion	8M – PICO-PI
Variscite	DART-MX8M System-on-Module

Embedded Board Partners & ODMs

Key Takeaways

- ## Selecting an Audio Front End (AFE):
  - NXP offers customers solutions with both hard and soft AFE solutions
  - Buy an NXP Dev kit, start prototyping/benchmarking to understand your usage model and performance requirements

- ## Embedded Voice Solution
  - NXP offers Dev kits/demos based on AVS, GVA and local solutions
  - Select your solution early based on ability key features & ability to get to market

- ## NXP Voice Ecosystem
  - Most companies do not have voice expertise to design and certify a voice product
  - NXP voice partners can help get your voice enabled product to market

SECURE CONNECTIONS
FOR A SMARTER WORLD