Infineon Smart Speaker Solutions

Essential Parts and Innovative Features in Smart Speakers' Design

www.infineon.com/smart-speaker

Contents

Introduction

The first, most effective way for a human to communication with another was through speech. The first, most effective way for a human to communicate with a computer was through a keyboard and a display. This is rather cumbersome and slow which is why the trend is to replace it with the long established, intuitive technique of speaking as evidenced by the growing proliferation of Smart Speakers with personal voice assistants.

Developing reliable speech recognition software was the first hurdle which was solved by devices being able to connect to powerful computers over the internet and now there are solutions being developed to solve this locally or "on the edge" to save energy and improve security as no uploading is required to the cloud.

This article looks at the design considerations of Smart Speakers and the features that will be incorporated in the future especially as Smart Speakers increasingly act as the central control hub for a Smart Home. The latter is going to be the ultimate target as it controls the whole user experience from media access to lighting and from climate control systems to door locks. In addition, the article will cover some new application areas that are opening up for Voice User Interfaces (VUIs) based on Smart Speaker technology.
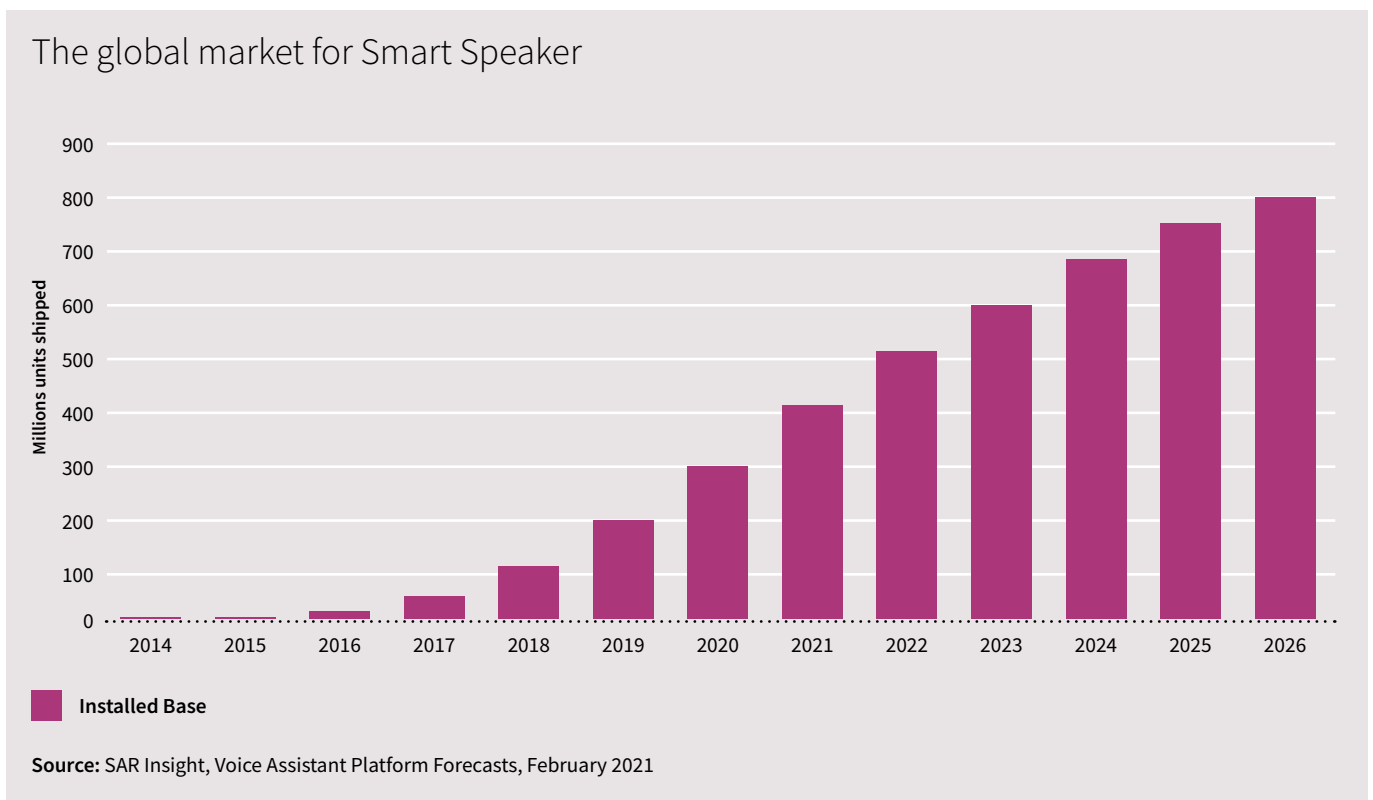
1. The world of Smart Speaker & smartness

What is a Smart Speaker?

It is categorised as a type of loudspeaker and a VUI with hands-free activation usually by a key or wake word of an integrated virtual assistant that provides intuitive interaction to access almost anything on the Web. The next stage that is being worked on currently is recognition of an individual's voice so that key word activation is not required.

Why do people like Smart Speakers?

People lead busy lives so love having their own personal assistant to help organise it, play the music they want, tell them the latest news and the weather forecast, and search the internet for answers to questions to name but a few of the ever-growing list of helpful features. Most of these could be done via typing on a computer or smartphone, but the convenience of using a VUI is compelling as it is hands-free and can be done from anywhere in the room when doing something else.

This is reflected in the Voice Assistant Platform Forecasts from SAR Insight & Consulting, which predicts that global shipments of Smart Speakers will grow to almost 200 million units per year by 2026. The installed base is forecast to almost double from 415 million units in 2021 to 798 million units by 2026, representing a CAGR (Compound Annual Growth Rate) of 14%. Total cumulated revenues are expected to top US$80 Billion from 2014 to 2026. To put the explosive growth in perspective, in 2015, global sales revenue was a mere US$0.2 Billion. Not bad for a product category that did not even exist before November 2014 and indicates that these forecasts could turn out to be much lower than reality as this is a very hot growth area.

The global market for Smart Speaker

Installed Base

Source: SAR Insight, Voice Assistant Platform Forecasts, February 2021

This demonstrably positive acceptance of Smart Speakers means that the adoption of its VUI technology is likely to rapidly be taken up into other devices as the favoured form of HMI (Human Machine Interface) from cars to TVs and from washing machines to robot vacuum cleaners. As a result, the growth rate for Smart Speaker-based VUI technology as a whole area could be significantly greater.

Ever Smarter Speakers

As it often the case with electronic devices, Smart Speakers are continually being improved with additional features in each year's new model. To achieve this cost effectively, designers should create a design that can have features added to it without having to redesign the whole device. The best way to do this is by partnering with Infineon who brings to the table not only a comprehensive range of key parts required for a Smart Speaker but is also driving the addition of new features to differentiate tomorrow's smarter Smart Speakers to improve the user experience.

2. Key parts of a Smart Speaker design today

What are the key parts of a Smart Speaker design today?

Microphone. Nowadays, with a perfect audio input, algorithms have few problems understanding the spoken word — apart from when confronted with a strong dialect or accent! The big remaining problem is that, in the real world, audio is rarely perfect. The Speaker could be in a room where there is ambient noise such as a TV, outside car noise, a dog barking and so forth. There are several techniques to improve the quality of audio capture such as using a high-quality microphone and locating the source of the voice so that extraneous noises can be ignored.

In almost all applications, big Electret Condenser Microphones have been replaced with tiny, Micro-Electro-Mechanical Systems (MEMS) microphones as they offer many advantages:
1) Size at only a few cubic millimetres that can be easily incorporated into Smart Speaker designs and even mounted directly onto PCBs as they are packaged just like other electronic chips and are actually made in a similar way in foundries.
2) This means that the power consumption and costs are also less.
3) Their design has a consistent performance over their lifetimes so there is no drift to confuse the algorithms.
4) They can handle a very wide range of sound input levels from whispers right up to very loud noises without distortion that would confuse the algorithms.

Locating the person. To make the capture of the sound even better, several microphones can be used and the difference between the signals can be used to pinpoint the location of the person speaking. This enables unwanted sounds to be filtered out and is only possible because of the virtually identical performance characteristics of MEMS microphones with the same part number.

Another technique for achieving a similar result is to use a radar IC to locate someone who is speaking. This has the added benefit of presence detection for the situation where someone is in the room but not speaking. The lack of a human means that the device can power down to save energy with just enough left to monitor if someone has returned within range.

Audio output. Gone are the days of large speakers as Smart Speakers can produce a surprising loud output from a tiny form factor, which is increasingly important as the use cases are now including on-the-go such as outdoors. This requires a quality audio chip that can be fine-tuned to provide a balanced audio output that compensates to the challenges of base output from a small speaker. With the latest audio processing techniques, a 360-degree audio experience can even be created.

Connectivity. Smart Speakers must have reliable wireless connectivity. The first is Bluetooth® for short range connection to a smart phone for setting up the Smart Speaker via an app and also perhaps an additional sound system. The second is Wi-Fi to provide connection to a router and then to the Web so the quality of this connection directly impacts the user experience so quality components are paramount. Third, Bluetooth® mesh enables the Smart Speaker to serve as a hub of the smart and connected home, controlling other IoT edge devices such as light and air conditioner.

The next generation of Smart Speakers will inevitably have more processing power which opens up the possibility of handling VUIs on the edge, i.e., a self-contained device that does not need to connect to the Cloud. This would enable designs to be created for on-the-go use anywhere such as outdoors where Wi-Fi reception may be patchy or non-existent.

Security. Smart Speakers are cloud connected as the voice recognition software is run on servers, so it is vital that security is built in and this can be done by incorporating an IoT security chip that provides the Speaker with its unique identity and integrity to protect personal information. Security is becoming even more important as the Speaker takes a central role as the control hub for the Smart Home with secondary speakers in the other rooms of the house.

Touch control. There is still a place for the occasional use for buttons or similar touch interfaces such as on/off and muting the microphone for privacy. Mechanical switches are a potential point of failure so capacitive touch control offers a better, more elegant option and has become intuitive thanks to gesture control of smartphones. In fact, the HMI is a key way to differentiate designs to stand out from rivals by using a combination of lights, screen, touch, VUI, and gesture as these are the primary interactions that influence the user experience.

Power. Naturally a Switched Mode Power Supply (SMPS) is required to convert mains power to the DC required for the electronics but, in addition, it is wise to include ESD (Electrostatic Discharge) protection to protect the electronics. A useful feature is to include a USB output port for charging devices such as smart phones, smart watches, etc. or a contactless charging facility which is an increasingly popular way to recharge smartphones. As mentioned earlier, portable Smart Speakers are a new emerging trend as, being battery-powered they can be used anywhere such as outdoors, in the garden or when travelling.

3. Future features for Smart Speakers

Sophisticated gesture control. Touch control could be complemented with gesture control using the radar IC as an easy, alternative way to control the device at a distance. For example, moving a hand up to increase volume and down to decrease it without using VUI that would interrupt the enjoyment of the music. Or simply waving a hand over the top of a Smart Speaker to turn it on or off.... A flat hand to stop an alarm... There is a world of possibilities opening up to create HMIs that combine voice and gesture control for intuitive, touchless interactions giving better user experiences and greater convenience.

Environmental sensing. The growing awareness of air quality and, in particular, the importance of CO_2 levels means that this is a feature that users will welcome. Elevated levels of CO_2 can quickly affect concentration and productivity. CO_2 levels in a closed room can rise surprisingly quickly especially with several people in it but just how detrimental this can be is only now becoming widely recognised. That tired feeling in a 'stuffy' room is not due to lack of oxygen but CO_2 build up. So an alarm that alerts people that they should open a window becomes an important health feature plus studies are showing that a poorly ventilated room can be linked to the transmission of infections such as COVID.
A CO_2 sensor chip is a simple addition to provide an alert should that happen. Some countries are already recognising this with legal requirements for CO_2 sensors in homes and school rooms with more expected to follow along with legislation for CO_2 levels in the workplace.

Presence sensing. Presence sensing using radar could be used to monitor people in the house. If an elderly or ill person deviates from their usual routine, the Smart Speaker could send an alert for them to be checked on. A fallen person can call via a Smart Speaker for urgent assistance to be sent giving them greater peace of mind that help can be with them quickly if needed. One could also imagine the scenario where someone collapses unconscious and the Smart Speaker works out that something is not right and requests help. Using radar for this avoids any issues of having video cameras as there are no images of people that would compromise privacy.

Health monitoring Radar is so sensitive that it can actually monitor the health of people as it can detect the micro-movements of vital signs such as breathing and pulse rates. A Smart Speaker by a bed could monitor someone while they slumber using radar to give very useful data on the quality and amount of their sleep.

Smart Home. Smart Speakers provide a logical, intuitive interface hub to a Smart Home to provide voice control of lighting, heating, curtains, and other domestic appliances. And even instruct the oven to turn on by voice control or the bath to run. Your personal assistant is then more like your personal butler.

Screens. As Smart Speakers provide more and more functionality, screens provide the ability to provide information in a visual way that is better that vocal response such as the chance of rain throughout the day. Naturally, designers need to take into account that the user might well be reading the screen from some distance way across the room. The screen could even show a computer-generated talking face to provide a more natural human interaction.
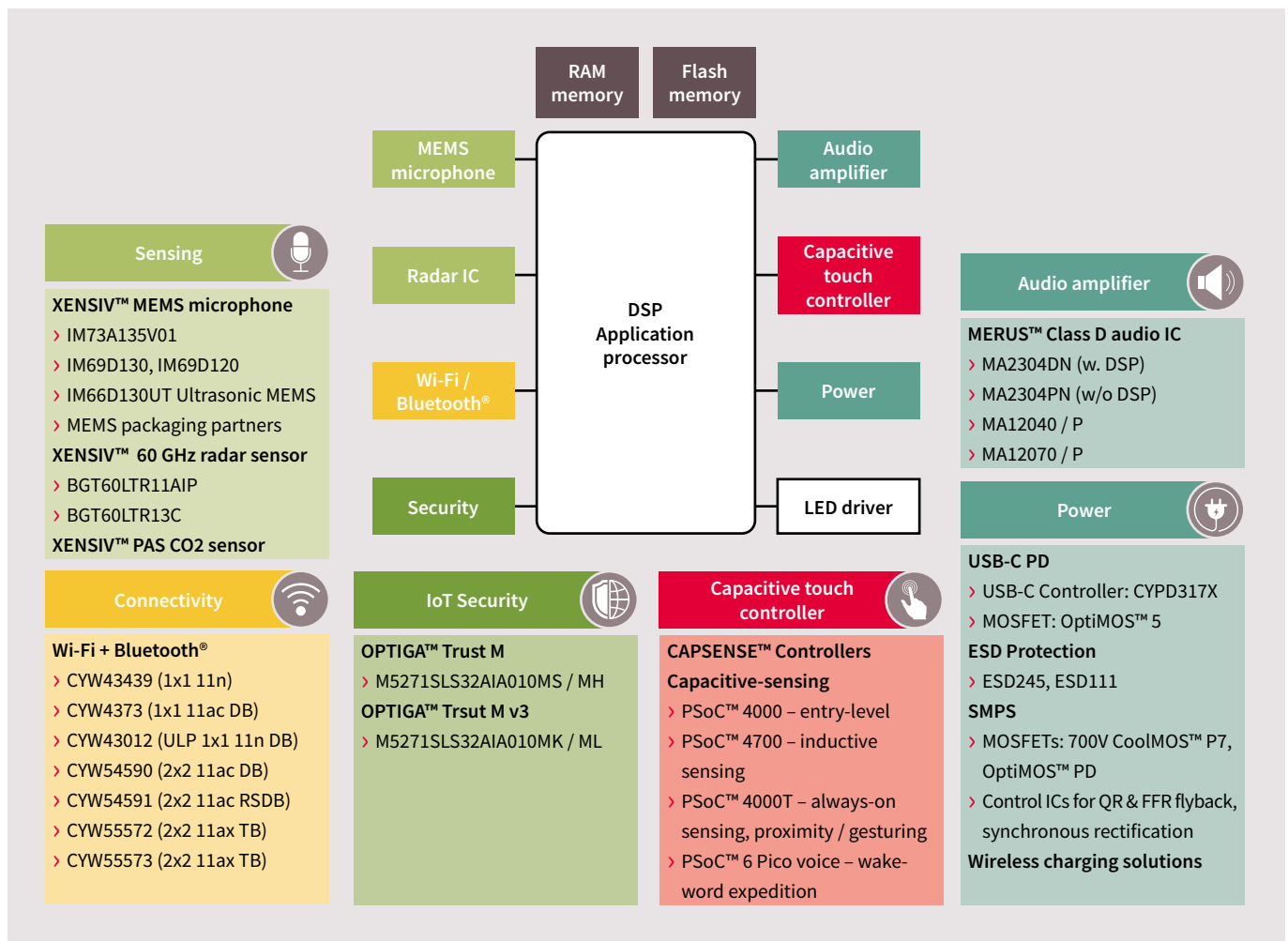
Next generation personal assistants. The software will constantly evolve to provide more features. There is a whole new opportunity for the kinds of apps that are on smartphones that use biometric face or fingerprint recognition to be done using biometrically verified voice control for financial transactions such as paying bills and ordering on line.

Automotive. Many of the features and benefits of Smart Speakers can be deployed in cars as a VUI providing an intuitive way to control a car that reduces risk as the driver does not have to take a hand off the wheel to press buttons or switches. Some cars have already proved how useful this is with the ability to make a phone call simply by using voice instructions. VUIs have already been built into some new models and these computationally demanding features will become more popular as 5G rolls out to provide fast mobile internet connection to voice service providers and the power of the Web.

4. Components and ecosystem for Smart Speaker designs in one place

Smart Speakers are evolving rapidly as new, smarter features are being developed to enhance the user experience and to enable them to play an increasingly important function as a central, co-ordinator in a Smart Home. It is therefore vital to ensure that Smart Speakers can be relied to work perfectly all the time. If a standalone radio fails it is a minor inconvenience but if the device that controls the heating, lighting, front door access, etc. fails then that is a major issue. The design must therefore embody quality, accuracy and reliability throughout from the components through the software to robust security.

The design of a Smart Speaker is a complex mix of sensors with digital, analog and mixed signal circuits along with power management plus software that could be challenging to source from many suppliers. However, Infineon has a comprehensive offering of all the key components needed to make Smart Speakers along with an ecosystem of software partners to create a one-stop-shop. As a leading supplier in this market, having specialised in it for a number of years, Infineon's offerings represent best-in-class solutions in terms of accuracy and quality.

RAM memory	Flash memory		
MEMS microphone			Audio amplifier
Radar IC	DSP Application processor		Capacitive touch controller
Wi-Fi / Bluetooth®			Power
Security			LED driver

Sensing 🎤

XENSIV™ MEMS microphone
› IM73A135V01
› IM69D130, IM69D120
› IM66D130UT Ultrasonic MEMS
› MEMS packaging partners

XENSIV™ 60 GHz radar sensor
› BGT60LTR11AIP
› BGT60LTR13C

XENSIV™ PAS CO2 sensor

Audio amplifier 🔊

MERUS™ Class D audio IC
› MA2304DN (w. DSP)
› MA2304PN (w/o DSP)
› MA12040 / P
› MA12070 / P

Connectivity 📶

Wi-Fi + Bluetooth®
› CYW43439 (1x1 11n)
› CYW4373 (1x1 11ac DB)
› CYW43012 (ULP 1x1 11n DB)
› CYW54590 (2x2 11ac DB)
› CYW54591 (2x2 11ac RSDB)
› CYW55572 (2x2 11ax TB)
› CYW55573 (2x2 11ax TB)

IoT Security 🛡

OPTIGA™ Trust M
› M5271SLS32AIA010MS / MH

OPTIGA™ Trsut M v3
› M5271SLS32AIA010MK / ML

Capacitive touch controller 👆

CAPSENSE™ Controllers
Capacitive-sensing
› PSoC™ 4000 – entry-level
› PSoC™ 4700 – inductive sensing
› PSoC™ 4000T – always-on sensing, proximity / gesturing
› PSoC™ 6 Pico voice – wake-word expedition

Power 🔌

USB-C PD
› USB-C Controller: CYPD317X
› MOSFET: OptiMOS™ 5

ESD Protection
› ESD245, ESD111

SMPS
› MOSFETs: 700V CoolMOS™ P7, OptiMOS™ PD
› Control ICs for QR & FFR flyback, synchronous rectification

Wireless charging solutions

Microphones. Infineon XENSIV™ MEMS microphones are designed for VUI applications requiring low self-noise (high SNR), wide dynamic range, and a high acoustic overload point. The XENSIV™ MEMS microphones offer crystal clear audio signals, an extended pick-up distance and sensitivity to both soft and loud signals in applications. The best-in-class, mic-to-mic matching results in identical audio signals from multiple microphones that can be used for noise cancellation or ultra-precise beam forming to hone in on a sound source to identify and recognise a particular speaker amongst several speakers. Details on the full range of innovative applications solutions that have been developed by Infineon and its ecosystem of partners can be found at online. These play a key role in differentiating and advancing next generation Smart Speaker designs.

Radar. Infineon's highly sensitive XENSIV™ 60 GHz Radar sensors can provide precise presence detection within a configured distance along with macro movements on the whole body scale through to micro movements to the sub-millimetre level for gesture control and even breathing and heartbeat rates for heath monitoring. One IC can even identify several people and their locations within a space and track their movements to help identify who is speaking, giving context information to the Smart Speaker to enable it to better understand what is happening. Infineon offers the smallest motion sensor on the market with integrated antennas as well as integrated detectors for motion and direction of motion, providing a complete, easy-to-implement feature that does not require specialist knowledge. Further details

Connectivity. Infineon's Wi-Fi and Bluetooth® technologies are the industry's most widely deployed solutions ensuring ultra-reliable interoperability with state-of-the-art security built in – the same as used in over 50% of credit cards and 40% of digital passports. Their ultra-low power design ensures long battery life in portable designs. Their long ranges of more than double that of some competitive solutions enable operation throughout the many rooms of the home with multi-channel audio and RSDB (Real Simultaneous Dual Band) for speaker sub-networks. Infineon's AIROC™ Wi-Fi & Combos portfolio integrates IEEE 802.11a/b/g/n/ac/ax Wi-Fi and Bluetooth® 5.2 in a single-chip solution to enable small-form-factor designs. Further details

Audio amplifier. Outstanding audio from a choice of best in class MERUS™ class D amplifier solutions to provide the perfect solution for the specification of the Smart Speaker to create rich, immersive sound. The advanced design reduces power consumption to provide up to 58% longer battery playback compared to competing solutions in portable designs. This also benefits wired designs as less waste heat is generated. Their multi-level compact designs decrease the number of filtering and external components to reduce the BOM costs. Further details

Capacitive touch controller. There is always going to be a need for a form of touch-based HMI such as in a suddenly noisy room. Rather than old fashioned switches that can easily fail due to dirt, liquids or mechanical failure, capacitive sensors are an elegant, more reliable solution that enable futuristic, novel control interfaces to be designed to enhance the overall design and can even recognise gesture input. Infineon's CAPSENSE™ Controllers has established itself as the industry leader solution due to its state-of-the-art noise immunity with a Signal to Noise Ratio (SNR) of greater than 100 to 1 and excellent resistance to water and dirt. Further details

Security. As Smart Speakers evolve to handle more and more sensitive personal data, they must be designed with high levels of security built-in to protect this as any device connected to the cloud is a target for hackers. Infineon's OPTIGA™ Trust M embedded security solutions provide an anchor of trust for connecting IoT devices to the cloud, giving every IoT device its own unique identity and integrity to ensure secure cloud connection for the lifetime of the device. This pre-personalized turnkey solution offers secured, zero-touch onboarding and the high performance needed for quick cloud access. Further details

Power. Infineon is a world leader in power with an extensive portfolio of energy efficient, space saving solutions for the various power requirements from power conversion to charging other devices including ESD protection. Power transfer is becoming increasingly popular as it enables phones and other devices to be charged without wires by simply putting them on top of a Smart Speaker, ending the annoying hunt for the appropriate charging cable. For details on Infineon's range of safe and efficient charging solutions, please see here.

USB-PD. EZ-PD™ USB-C controller supports all USB PD profiles commonly used in USB-C power adapters and requires no firmware development, ensuring specification compliance and interoperability.

5. Infineon – The smart choice of partner for Smart Speakers

When designing a Smart Speaker, it is important to have a trusted supplier that understands this application area with products that have been optimised for use in it along with system knowledge to assist with design to ensure faster time to market. This is vital not only for the first model but strategic for subsequent models where Infineon's expert knowledge of new use cases and features is crucially important to differentiate products in the market place with innovation and thus gain market share.

Further information on Infineon' comprehensive range of solutions for Smart Speaker design can be found at:
www.infineon.com/smart-speaker

Where to buy

Infineon distribution partners and sales offices:
www.infineon.com/WhereToBuy

Service hotline

Infineon offers its toll-free 0800/4001 service hotline as one central number,
available 24/7 in English, Mandarin and German.

> › Germany .... 0800 951 951 951 (German/English)
> › China, mainland .... 4001 200 951 (Mandarin/English)
> › India .... 000 800 4402 951 (English)
> › USA .... 1-866 951 9519 (English/German)
> › Other countries .... 00* 800 951 951 951 (English/German)
> › Direct access .... +49 89 234-0 (interconnection fee, German/English)

* Please note: Some countries may require you to dial a code other than "00" to access this international number.
Please visit www.infineon.com/service for your country!

Mobile product catalog
Mobile app for iOS and Android.

Please note!
This Document is for information purposes only and any information given herein shall in no event be regarded as a warranty, guarantee or description of any functionality, conditions and/or quality of our products or any suitability for a particular purpose. With regard to the technical specifications of our products, we kindly ask you to refer to the relevant product data sheets provided by us. Our customers and their technical departments are required to evaluate the suitability of our products for the intended application.

We reserve the right to change this document and/or the information given herein at any time.

Additional information
For further information on technologies, our products, the application of our products, delivery terms and conditions and/or prices, please contact your nearest Infineon Technologies office (www.infineon.com).

Warnings
Due to technical requirements, our products may contain dangerous substances. For information on the types in question, please contact your nearest Infineon Technologies office.

Except as otherwise explicitly approved by us in a written document signed by authorized representatives of Infineon Technologies, our products may not be used in any life-endangering applications, including but not limited to medical, nuclear, military, life-critical or any other applications where a failure of the product or any consequences of the use thereof can result in personal injury.