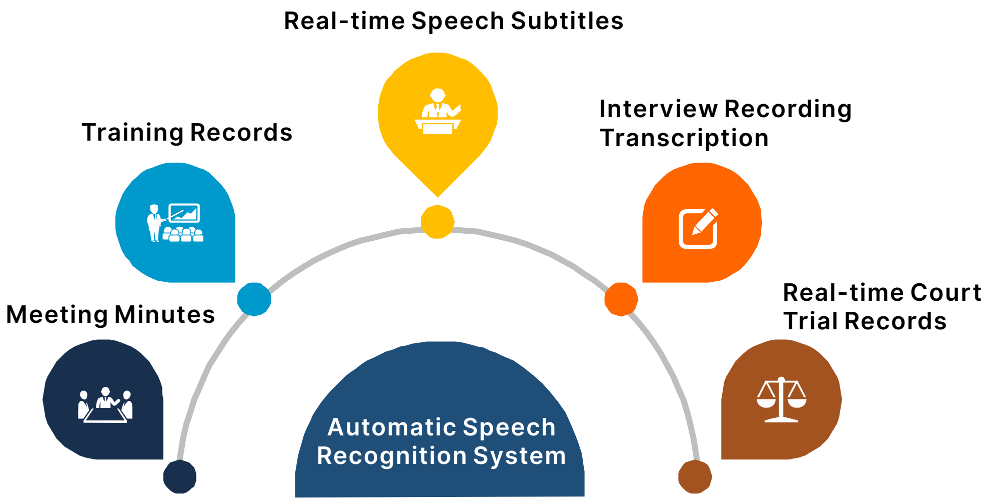
GONSIN®

Automatic Speech Recognition System

Automatic Speech Recognition System

As new development of modern conference solution, automatic speech recognition (ASR) system brings more intelligent human-computer interaction experience. For traditional conferences, the communication by sound and video cannot satisfy the modern conference needs any more. Besides, after meeting, the document processing, meeting minutes and legal procedures of specific users are also required to be presented in words format. Gonsin Automatic Speech Recognition System can achieve real-time, complete and orderly text transcription from sound, and ensures the text corresponding to each delegate's speech. The transcribed text can be displayed on the large screen, as well as Gonsin paperless conference system in real time.

ASR system suits various application scenarios, including meeting minutes, training records, real-time speech subtitles, interview records transcription, real-time court trial records, etc.



Advantages

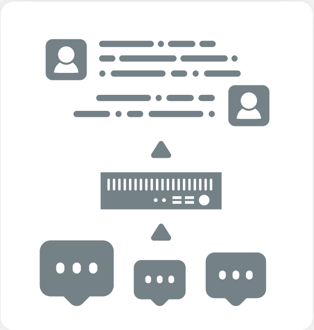
The noisy environment affects the transcription accuracy, proofreading is needed.	Equip with conference system, adapt to noisy environment, clear sound pickup.
Too many speaking delegates cause record confusion and information error.	Roles, delegate's voice matches the transcribed text.
If only with written documents, It is difficult to make omission verification and quality check.	Full voice recording and text transcription save the original content, and good for post-meeting verification.
Too much words and long-time manual input cannot guarantee transcription accuracy.	Real-time transcription from voice to text, and quickly documented, this makes accuracy rate over 97%.
Traditional technology makes high labor cost.	ASR supports cloud server rental deployment, and local server LAN deployment, artificial intelligence learning, and continuous system optimization.

System Solution

GONSIN ASR System offers three solutions: Online Speech Recognition and Translation Solution, Lightweight Private Deployment Solution, and Conference Room Cluster Private Deployment Solution.

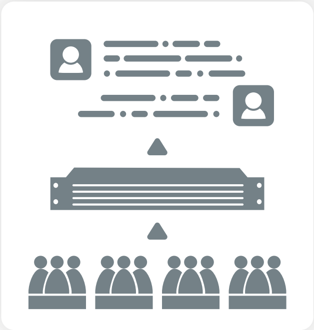
Online Speech Recognition and Translation Solution

- ♦ Adopt a cloud-based deployment with a package payment model, effectively reducing the investment cost and construction threshold for speech recognition
- ♦ Support translation in 10 major languages and provide on-site multilingual subtitle display
- ♦ Support role separation recognition, allowing simultaneous recognition and translation of multiple specified languages



Lightweight Private Deployment Solution

- ♦ The server is compact and easy to install, meeting the intelligent speech recognition needs of small and medium-sized meetings
- ♦ Local deployment with embedded ASR engine, ensuring low latency and accuracy rates exceeding 99%



Conference Room Cluster Private Deployment Solution

- ♦ Adopt local centralized deployment server, meeting the needs of speech recognition and transcription of multiple meeting rooms
- ♦ Adopt a flexible mechanism for authorizing language recognition capabilities, and authorize the number of intelligent languages in parallel according to the needs of the meeting room
- ♦ Support up to 50 channels of speech recognition on a single server

Equipment List

Online Speech Recognition and Translation Solution

Order	Product	Model	Qty.	Unit	Function
ASR Application Terminal					
A.1	ASR Engine	V3.1	1	SET	Provide online speech recognition and translation services Package available: V3.1: 500 hours; V3.2: 1000 hours
A.2	ASR Engine	V3.2			
A.3	ASR Software	V7.1.0	1	SET	Set up in computer for system settings and function display
A.4	ASR Subtitle Display Software	V7.1.0	1	SET	Set up in computer or Android terminal for real-time display of subtitles
A.5	Control Computer	Purchased by the customer	1	PCS	Set up GONSIN ASR Software and ASR Subtitle Display Software
Conference Discussion System					
Can co-work with GONSIN all series of discussion system, choose the product configuration according to the needs					

Lightweight Private Deployment Solution

Order	Product	Model	Qty.	Unit	Function
			Lightweight ASR Server		
A.1	Lightweight ASR Server	GX-AS201	1	SET	Built-in ASR Engine, with 1-way speech recognition
A.2	Lightweight ASR Server	GX-AS202	1	SET	Built-in ASR Engine, with 2-way speech recognition
A.3	Lightweight ASR Server	GX-AS205	1	SET	Built-in ASR Engine, with 5-way speech recognition
A.4	Lightweight ASR Server	GX-AS208	1	SET	Built-in ASR Engine, with 8-way speech recognition
			ASR Application Terminal		
B.1	ASR Software	V7.1.0	1	SET	Set up to Lightweight ASR Server for system settings and function display
B.2	ASR Subtitle Display Software	V7.1.0	1	SET	Set up in computer or Android terminal for real-time display of subtitles
B.3	Control Computer	Purchased by the customer	1	PCS	Set up GONSIN ASR Subtitle Display Software
			Conference Discussion System		
			Can co-work with GONSIN all series of discussion system, choose the product configuration according to the needs		

Conference Room Cluster Private Deployment Solution

Order	Product	Model	Qty.	Unit	Function
			ASR Server		
A.1	ASR Engine	V3.0	1	SET	ASR Engine
A.2	ASR Server	GX-AS301	1	PCS	Support up to 50 channels of speech recognition on a single server
A.3	Speech Recognition Module Authorization	V1.0	1	WAY	Based on the number of conference rooms for speech transcription requirements on the LAN
			ASR Application Terminal		
B.1	ASR Software	V7.1.0	1	SET	Set up in computer for system settings and function display
B.2	ASR Subtitle Display Software	V7.1.0	1	SET	Set up in computer or Android terminal for real-time display of subtitles
B.3	Control Computer	Purchased by the customer	1	PCS	Set up GONSIN ASR Software and ASR Subtitle Display Software
			Conference Discussion System		
			Can co-work with GONSIN all series of discussion system, choose the product configuration according to the needs		

Application Project

People's Government of Xinluo District, Longyan City, Fujian Province

Alibaba Group Guangmingding Conference Room

GONSIN Automatic Speech Recognition Software V7.1.0

Basic Functions

- Support public cloud and proprietary cloud voice server selection docking, which can meet different server deployment methods. Support installation to PC computer or speech recognition server, which can be flexibly applied to a variety of application scenarios.
- Support ASR server shutdown management function, ASR server, discussion system connection, search, and microphone role customization function, and support the public letter of each series of discussion system seamless docking, conference management, role separation, and automatic identification.
- Support personnel and equipment management, including equipment search, displaying unit number information, IP address information, and personnel name settings; support meeting information editing, including new meeting name, defining meeting time, location, and meeting content editing.
- Support simultaneous recognition of multiple microphone roles and anti-crosstalk function, which can effectively avoid mutual crosstalk when multiple microphones are recognized at the same time; support microphone status prompts, which can display the microphone on and off status in real time.
- Support language model learning function. It support importing common words such as names of people and places to learn the language model.
- Support automatic identification of participants' roles, automatic identification of participants' voice and transcription into text. The software support translation into other required speech (software functions vary according to engine capabilities)
- Support intelligent semantic understanding, which can automatically understand the semantics of the participants and automatically break sentences and segments according to the semantics. Support automatic conversion of consecutive numbers to Arabic format, and support automatic identification of cell phone numbers, ID cards and other consecutive numbers converted to Arabic format.
- Support meeting text editing and correction functions. Generate separate recording files for different roles, or merge the text records and recordings of each role. Voice and text records can be synchronized playback and display against the document correction.
- Support meeting record output function. Support text merge, generate meeting minutes, and export text.
- Support content search function, support text content search. Keywords can be searched, quickly locate the position of the corresponding content, greatly improving the efficiency of content retrieval.
- Support text split-screen output function. Installation to the PC computer, you can realize the transcription text real-time display in the main screen of the operating computer, support the expansion of split-screen output, real-time display of the text content of voice recognition. Support screen customization function, screen resolution adaptive, support text font, size settings, to provide high-quality split-screen text display service.
- Support recording file recognition, through the recording file import, automatically convert the recording file content into text content; support mp3, wav and other file formats.
- Support the selection of audio input devices, you can connect the computer's audio input devices, real-time audio input transcription text
- Support the computer to recognize the current playback sound content, and automatically convert it to text.
- Support more customized features: the software support Chinese and English switching, as well as other custom languages; support for secondary development, according to the project requirements of the open interface protocol or customized development.

Technical Parameters

System Requirements	WIN7/WIN8/WIN10 operating system 32/64 bit
CPU Requirement	i7 or above
Hard Drive Capacity Requirement	500GB or above
Memory Capacity Requirement	16GB or above
Graphics Card Requirement	Independent graphics card support VGA / HDMI / DVI interface, support for display split screen
PC Interface Requirement	1 RS-232 interface; 2 RJ45 interface
Resolution	Adaptive
PC Communication Method	Ethernet/RS-232

GONSIN Automatic Speech Recognition Subtitle Display Software V7.1.0

Basic Functions

- ♦ Good system compatibility, support subtitle display for Windows and Android devices.
- ♦ Support multiple subtitle display mode settings. Support full-screen mode and barrage mode
- ○ Full screen mode: display the transcription content in full screen in the form of a dialog box. Support background setting and font setting.
- ○ Barrage mode: display the transcription content in a floating barrage style. Support line setting and font setting
- ♦ Support video overlay subtitle function: support real-time subtitle function overlaying on the video screen, integrated with video conferencing and camera tracking applications.
- ♦ Support paperless overlay subtitle function: Enables real-time subtitle overlay on paperless screens, integrating with paperless systems, and displaying transcribed text in real-time on paperless terminals.

GONSIN Automatic Speech Recognition Engine V3.1 / V3.2

Basic Functions

- ♦ With industry-leading online speech recognition technology, deployed through the cloud to provide speech recognition services for local speech. Low latency, high recognition accuracy, accuracy rate can reach more than 99%
- ♦ The engine adopts a package payment model, effectively reducing the input cost and construction threshold of speech recognition. Users can purchase the package program of appropriate length according to the actual demand for the length of speech recognition (please purchase the package service in time to ensure the normal use of the engine)
- ♦ Support role-separated recognition: different original languages and translation languages can be selected according to different roles, so as to realize simultaneous recognition of multiple languages, transcription into corresponding text, and translation.
- ♦ Support multiple major languages, such as Chinese, English, French, Russian, Arabic and Spanish.
- ♦ With intelligent speech recognition subtitle display software, it can display the original text and translated text at the same time, or set to display the original/translated text separately, providing subtitle service for business negotiation and video conference in different languages.

GONSIN Automatic Speech Recognition Engine V3.0

Basic Functions

- Adopt intelligent language recognition model technology, based on AI technology to achieve speech recognition
- Support customized language recognition,such as Chinese, English, French, Russian, Arabic and Spanish.
- Support recognition in multiple application scenarios: education, judicial, medical, conference speech, news media, entertainment video, smart home, social, automotive and so on
- Adopt a flexible mechanism for authorizing language recognition capabilities, and authorize the number of intelligent languages in parallel according to the needs of the meeting room

Speech Recognition Module Authorization V1.0

Basic Functions

- Adopt a flexible mechanism for authorizing language recognition capabilities, and authorize the number of intelligent languages in parallel according to the needs of the meeting room

Lightweight Automatic Speech Recognition Server
GX-AS201/GX-AS202/GX-AS205/GX-AS208

Handiness and simplicity
Realize voice recognition, transcription and recording for small and medium-sized conference

Built in ASR Engine
Excellent performance, fast transcribing speed, high recognition rate, easy deployment, strong stability

Suit to fixed venue and temporary venue
System docking is simple for rapid device switching and system construction in different conference venues

Basic Functions

♦ With intelligent voice recognition software, it can realize Web access management

♦ Support automatic recognition of participant roles, automatic recognition of participant's voice, and transcription into text

♦ With built-in ASR Engine, adopt industry-leading online speech recognition technology, deployed through the cloud to provide speech recognition services for local speech. Low latency, high recognition accuracy, accuracy rate can reach more than 99%

♦ Speech recognition server can realize speech transcription of different channels:

○ GX-AS201: supports 1-way speech recognition capability

○ GX-AS202: Supports 2-way speech recognition capability

○ GX-AS205: supports 5-way speech recognition capability

○ GX-AS208: supports 8-way speech recognition capability

♦ Support customized language recognition, such as Chinese, English, French, Russian, Arabic and Spanish.

♦ Support recognition in multiple application scenarios: education, judicial, medical, conference speech, news media, entertainment video, smart home, social, automotive and so on

♦ Support multiple conference rooms to share the server. Support multiple conference rooms in the conference center to form a LAN and centrally deploy the server to meet the parallel speech recognition and transcription in multiple conference rooms.

♦ With intelligent speech recognition subtitle display software, provide subtitle display service for conferences.

Technical Parameters

Model	GX-AS201	GX-AS202	GX-AS205	GX-AS208
System Version	Centos7.4+			
CPU	i3	i7		
RAM	16G	32G		
Hard Disk	256G SSD	500G SSD		
Front Panel Interface	4×USB2.0 Type-A,1×3.5mm Line out,1×3.5mm Micin,1×Power button,1×Power LED			
Rear panel interface	4×USB3.0 Type-A,1×RJ4510/100/1000M,1×HDMI 1.4 out, 1×COM out,1×3.5mm Line out,1×3.5mm Mic in,1×WIFI/BT ANT			
Power Input	19V DC			
Operating Temperature	-5°C~45°C			
Storage Temperature	-20°C~60°C			
Volume	210(L)×210 (W)×56 (H) mm			

GONSIN Automatic Speech Recognition Server GX-AS301

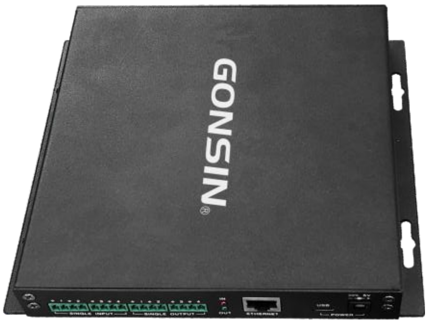
Basic Functions

- 2U standard rack-mounted server with stable and reliable performance, adopting SGCC galvanized steel plate, environmentally friendly exterior paint, fingerprint resistance, and resistance to contact 4kV strong magnetic interference
- Adopt high-performance configuration LINNUX server, install ASR Engine V3.0 software to realize automatic identification of participants' roles, automatic recognition of participants' voices and transcription into text.
- Support multiple conference rooms to share the server. Support multiple conference rooms in the conference center to form a LAN and centrally deploy the server to meet the needs of multiple conference rooms for parallel speech recognition and transcription.
- Co-work with intelligent speech recognition subtitle display software to provide subtitle display service for meetings
- High-efficiency CTC model, through the optional authorization, a single server supports a maximum of 50 concurrent recognition.
- The server adopts SSL encryption mechanism to effectively ensure the storage security and transmission security of sensitive information. RC4, MD5 and RSA encryption algorithms are used to ensure the security of platform data and avoid leakage of important information.
- Built-in power management embedded software. It can monitor the voltage status to avoid equipment failure caused by voltage fluctuation and realize all-weather protection.

Technical Parameters

System Version	Centos 7.4+
CPU Frequency	>2.30GHz
CPU Core	32 cores (64 threads)
Memory Type	DDR4 2667MHz
Memory Capacity	64G
Hard Disk Interface	SATA3.0/M.2
Hard Disk Capacity	1TGB
Network Interface	2 Gigabit LAN Ports
USB Interface	USB3.0 2pcs
VGA Output Interface	1pc
Power Supply Type	Hot-plug Power Supply
Number of Power Supplies	1pc
Power Supply Power	500W
Dimension(L×W×H)	550×430×88mm
Weight (L×W×H)	18kg
Operating Temperature	5~60℃
Storage Temperature	5~60℃

Audio Extender DCS-AE04



Basic Functions

- ♦ Black metal chassis design, rugged and durable, with mounting holes for easy installation
- ♦ Can realize the mutual conversion of audio digital and analog. The number of channels of a single audio expander is 4 channels, and can be cascaded to a maximum of 32 channels.
- ♦ Support two audio expanders to work together, can realize 4-channel audio long-distance transmission.
- ♦ Support two working modes of DA/AD for the channels, which can be set according to different application scenarios:
- ○ DA mode: digital audio can be converted to analog audio to achieve system expansion.
- ○ AD mode: analog audio can be converted into digital audio to achieve digital transmission. Can meet the analog audio, discussion system and intelligent voice recognition system cascade scenarios used to extend the voice transcription function
- ♦ Support device settings via WEB page. Including analog signal input interface selection, gain setting, noise gate setting, channel ID setting, channel mute time setting, volume detection and so on.

Technical Parameters

Color	Black
Installation	Hidden mounting, 4 hanging holes provided
Analog Input Interface	XLR input × 4 (Phoenix terminals) RCA input × 4 (Phoenix terminals)
Analog Output Interface	RCA output×4(Phoenix terminal)
Network Interface	ETHERNET×1(RJ45)
Power Supply Interface	USB(TypeC)×1 DC×1 (5V)
Dimension(L×W×H)	185×185×30mm
Weight (L×W×H)	0.9kg
Operating Temperature	0~45°C
Standard Accessories	TypeC USB cable×1

Audio Extender DCS-AE10

Basic Functions

- ♦ Black tempered glass panel design with navigational knobs, OLED screen displays mode status information for each channel, and the level of each channel can be adjusted independently or uniformly.
- ♦ Provides a wealth of interfaces. Including 11 groups of audio interfaces (Phoenix terminal), 1 dedicated simultaneous transmission interface (25P), 2 dedicated conference terminal interface (RJ45), 1 network communication interface (RJ45), 1 channel of the network communication interface (RJ45)
- ♦ Audio digital-analog conversion can be realized. The number of channels of a single audio expander is 10 channels, and the maximum cascade can be up to 32 channels.
- ♦ Support two sets of audio expanders to work together, can realize 10-channel audio long-distance transmission
- ♦ Support independent or unified setting of DA/AD two working modes for channels, which can be set according to different application scenarios:
- ○ DA mode: digital audio can be converted to analogue audio to achieve system expansion. It can meet the digital audio output of GONSIN all-digital conference system to FS-FHSS wireless simultaneous transmission and infrared simultaneous transmission system, so as to achieve system audio sharing.
- ○ AD mode: analogue audio can be converted into digital audio to achieve digital transmission. Can meet the FS-FHSS wireless simultaneous transmission, infrared simultaneous transmission system, other analog audio signal input to the GONSIN all-digital conference system, to achieve system audio sharing; AD mode supports the conversion of analog audio into digital audio, to meet the analog audio, the discussion system and the intelligent voice recognition system cascade scenarios to use, expand the voice transcription function
- ♦ Supports two working modes: master/slave:
- ○ The host mode has the management function of digital interpreter terminals and can access digital interpreter terminals to meet the requirements of FS-FHSS wireless simultaneous interpretation and infrared simultaneous interpretation system to realize the management of digital interpreter booths.
- ○ In the slave mode, it needs to be used with GONSIN10000N server, which is responsible for the management of digital interpreter booths, and is applied to GONSIN10000N server and other simultaneous interpretation systems. The GONSIN10000N server is responsible for digital interpreter room management, and it is used in occasions where the GONSIN10000N server is used simultaneously with other simultaneous interpretation systems.

Technical Parameters

Audio Frequency Response	20Hz~20kHz
Signal-to-Noise Ratio	>96 dBA
Channel Isolation	>85 dB
Total Harmonic Distortion	<0.05%
Audio Interface	11 groups of audio interfaces (Phoenix terminal); 1 dedicated simultaneous transmission interface (25P)
Terminal Interface	2-way dedicated conference terminal interface with private power supply protocol (RJ45)
Network Interface	1-way network communication interface (RJ45)
Display	OLED (61×37mm)
Operating Temperature	-0~45℃
Storage Temperature	-20~50℃
Dimension (L×W×H)	485×325×90mm
Installation	Desktop or 19" cabinet installation
Weight	8kg

GONSIN CONFERENCE EQUIPMENT CO.,LTD.
www.gonsin.com

English Wechat