



Shanghai, China
18:00 - 22:45



Mumbai, India
15:30 - 20:15



Paris, France
11:00 - 15:45



London, UK
10:00 - 14:45



New York, USA
06:00 - 10:45

SUNDAY, October 25th - Program at a Glance

TIME (Shanghai)	Room 1	Room 2	Room 3	Room 4
16:30 - 18:00	Tutorial 1 Efficient and flexible implementation of machine learning for ASR and MT	Tutorial 2 Spoken dialogue for social robots	Tutorial 3 Meta learning and its applications to human language processing	Tutorial 4 Intelligibility evaluation and speech enhancement based on deep learning
Coffee break 18:00 - 18:15				
18:15 - 19:45	Tutorial 1 contd.	Tutorial 2 contd.	Tutorial 3 contd.	Tutorial 4 contd.
Coffee break 19:45 - 20:00				
20:00 - 21:30	Tutorial 5 'Speech 101' - What everyone working on spoken language processing needs to know about spoken language	Tutorial 6 Neural approaches to conversational information retrieval	Tutorial 7 Neural models for speaker diarization in the context of speech recognition	Tutorial 8 Spoken language processing for language learning and assessment
Coffee break 21:30 - 21:45				
21:45 - 23:15	Tutorial 5 contd.	Tutorial 6 contd.	Tutorial 7 contd.	Tutorial 8 contd.

MONDAY, October 26th - Program at a Glance

TIME (Shanghai)	Room 1	Room 2	Room 3	Room 4	Room 5	Room 6	Room 7	Room 8	Room 9	Room 10	Room 11	Room 12
17:00	Opening session Keynote 1:	Janet B. Pierrehumbert, The cognitive status of simple and complex models										
18:00 - 19:00												
Coffee break 19:00 - 19:15												
19:15 - 20:15	Mon-1-1 ASR neural network architectures I	Mon-1-2 Multi-channel speech enhancement	Mon-1-3 Speech processing in the brain	Mon-1-4 Speech Signal Representation	Mon-1-5 Speech Synthesis: Neural Waveform Generation I	Mon-SS-1-6 Automatic Speech Recognition for Non-Native Children's Speech	Mon-1-7 Speaker Diarization	Mon-1-8 Noise robust and distant speech recognition	Mon-1-9 Speech in Multimodality (MULTIMODAL)	Mon-1-10 Speech, Language, and Multimodal Resources	Mon-1-11 Language Recognition	Mon-S&T 1 Speech processing and analysis
Coffee break 20:15 - 20:30												
20:30 - 21:30	Mon-2-1 Speech Emotion Recognition I (SER I)	Mon-2-2 ASR neural network architectures and training I	Mon-2-3 Evaluation of Speech Technology Systems and Methods for Resource Construction and Annotation	Mon-2-4 Phonetics and Phonology	Mon-2-5 Topics in ASR I	Mon-SS-2-6 Large-Scale Evaluation of Short-Duration Speaker Verification (SdSV)	Mon-2-7 Voice Conversion and Adaptation I	Mon-2-8 Acoustic Event Detection	Mon-2-9 Spoken Language Understanding I	Mon-2-10 DNN architectures for Speaker Recognition	Mon-2-11 ASR model training and strategies	Mon-S&T 2 Speech annotation and speech assessment
Coffee break 21:30 - 21:45												
21:45 - 22:45	Mon-3-1 Cross/multi-lingual and code-switched speech recognition	Mon-3-2 Anti-spoofing and Liveness Detection	Mon-3-3 Noise reduction and intelligibility	Mon-3-4 Acoustic Scene Classification	Mon-3-5 Singing Voice Computing and Processing in Music	ISCA-SAC "2nd Mentoring"	Mon-3-7 Acoustic model adaptation for ASR	Mon-3-8 Singing and Multimodal Synthesis	Mon-3-9 Intelligibility-enhancing Speech Modification	Mon-3-10 Human speech production I	Mon-3-11 Targeted Source Separation	Diversity Meeting 21:00-22:00

TECHNICAL AREAS	1 Speech Perception and Production	3 Paralinguistic Analysis	5 Analysis of Speech and Audio Signals	7 Speech Synthesis	9 Speech Recognition II: Architecture	11 Spoken Dialog Systems	Special Sessions
	2 Phonetics, Phonology, and Prosody	4 Speaker and Language Identification	6 Speech Coding Enhancement	8 Speech Recognition I: Signal Processing	10 Speech Recognition III: New Applications	12 Spoken Language Processing	Show & Tell

TUESDAY, October 27th - Program at a Glance

TIME (Shanghai)	Room 1	Room 2	Room 3	Room 4	Room 5	Room 6	Room 7	Room 8	Room 9	Room 10
18:00	Keynote 2: Barbara Shinn-Cunningham, Brain networks enabling speech perception in everyday settings									
Coffee break 19:00 - 19:15										
19:15 - 20:15	Tue-1-1 Speech Translation and multilingual/multimodal learning	Tue-1-2 Speaker Recognition I	Tue-1-3 Spoken Language Understanding II	Tue-1-4 Human speech processing	Tue-1-5 Feature extraction and distant ASR	Tue-SS-1-6 Voice Privacy Challenge	Tue-1-7 Speech Synthesis: Text Processing, Data and Evaluation	Tue-1-8 Search for speech recognition	Tue-1-9 Computational Paralinguistics I (CP I)	Tue-1-10 Acoustic Phonetics and Prosody
Coffee break 20:15 - 20:30										
20:30 - 22:00	ISCA General Assembly									

TECHNICAL AREAS	1 Speech Perception and Production	3 Paralinguistic Analysis	5 Analysis of Speech and Audio Signals	7 Speech Synthesis	9 Speech Recognition II: Architecture	11 Spoken Dialog Systems	Special Sessions
	2 Phonetics, Phonology, and Prosody	4 Speaker and Language Identification	6 Speech Coding Enhancement	8 Speech Recognition I: Signal Processing	10 Speech Recognition III: New Applications	12 Spoken Language Processing	Show & Tell

WEDNESDAY, October 28th - Program at a Glance

TIME (Shanghai)	Room 1	Room 2	Room 3	Room 4	Room 5	Room 6	Room 7	Room 8	Room 9	Room 10	Room 11	Room 12	Room 13
18:00	Keynote 3: Lin-shan Lee, <i>Doing Something we Never could with Spoken Language Technologies from early days to the era of deep learning</i>												
Coffee break 19:00 - 19:15													
19:15 - 20:15	Wed-1-1 Tonal Aspects of Acoustic Phonetics and Prosody	Wed-1-2 Speech Classification	Wed-1-3 Speech Synthesis: Paradigms and Methods I	Wed-SS-1-4 The INTERSPEECH 2020 Computational Paralinguistics Challenge	Wed-1-5 Streaming ASR	Wed-SS-1-6 Alzheimer's Dementia Recognition through Spontaneous Speech (ADReSS)	Wed-1-7 Speaker Recognition Challenges and Applications	Wed-1-8 Applications of ASR	Wed-1-9 Speech Emotion Recognition II (SER II)	Wed-1-10 Bi- and multilinguality	Wed-1-11 Single-channel speech enhancement I	Wed-SS-1-12 Deep Noise Suppression Challenge	Industry Forum
Coffee break 20:15 - 20:30													
20:30 - 21:30	Wed-2-1 Voice and hearing disorders	Wed-2-2 Spoken Term Detection	Wed-SS-2-3 The Fearless Steps Challenge Phase-02	Wed-2-4 Monaural Source Separation	Wed-2-5 Single-channel speech enhancement II	Wed-2-6 Topics in ASR II	Wed-SS-2-7 Neural Signals for Spoken Communication	Wed-2-8 Training strategies for ASR	Wed-2-9 Speech transmission & coding	Wed-2-10 Bioacoustics and Articulation	Wed-2-11 Speech Synthesis: Multilingual and Cross-lingual approaches	Wed-2-12 Learning Techniques for Speaker Recognition I	Industry Forum
Coffee break 21:30 - 21:45													
21:45 - 22:45	Wed-3-1 Pronunciation	Wed-3-2 Diariation	Wed-3-3 Computational Paralinguistics II (CP II)	Wed-3-4 Speech Synthesis: Paradigms and Methods II	Wed-3-5 Speaker Embedding	ISCA-SAC "Students Meet Experts"	Wed-3-7 Single-channel speech enhancement III	Wed-3-8 Multi-Channel Audio and Emotion Recognition	Wed-3-9 Computational resource constrained speech recognition	Wed-3-10 Speech Synthesis: Prosody and Emotion	Wed-SS-3-11 The Interspeech 2020 Far Field Speaker Verification Challenge	Wed-3-12 Multimodal Speech Processing	Industry Forum

TECHNICAL AREAS	1 Speech Perception and Production	3 Paralinguistic Analysis	5 Analysis of Speech and Audio Signals	7 Speech Synthesis	9 Speech Recognition II: Architecture	11 Spoken Dialog Systems	Special Sessions	Industry Forum
	2 Phonetics, Phonology, and Prosody	4 Speaker and Language Identification	6 Speech Coding Enhancement	8 Speech Recognition I: Signal Processing	10 Speech Recognition III: New Applications	12 Spoken Language Processing	Show & Tell	

THURSDAY, October 29th - Program at a Glance

TIME (Shanghai)	Room 1	Room 2	Room 3	Room 4	Room 5	Room 6	Room 7	Room 8	Room 9	Room 10	Room 11	Room 12
18:00	Keynote 4: Shehzad Mevawalla, Successes, Challenges and Opportunities for Speech Technology in Conversational Agents											
Coffee break 19:00 - 19:15												
19:15 - 20:15	Thu-1-1 Speech Synthesis: Neural Waveform Generation II	Thu-1-2 ASR neural network architectures and training II	Thu-1-3 Neural networks for language modeling	Thu-1-4 Phonetic Event Detection and Segmentation	Thu-1-5 Human speech production II	Thu-SS-1-6 New Trends in self-supervised speech processing	Thu-1-7 Learning Techniques for Speaker Recognition II	Thu-1-8 Spoken language evaluation	Thu-1-9 Spoken Dialogue System	Thu-1-10 Dereverberation and echo cancellation	Thu-1-11 Speech Synthesis: Toward End-to-End Synthesis	Industry Forum
Coffee break 20:15 - 20:30												
20:30 - 21:30	Thu-2-1 Speech enhancement, bandwidth extension and hearing aids	Thu-2-2 Speech Emotion Recognition III (SER III)	Thu-2-3 Acoustic Phonetics of L1-L2 and Other Interactions	Thu-2-4 Conversational systems	Thu-SS-2-5 The Attacker's Perspective on Automatic Speaker Verification	Thu-2-6 Summarization, Semantic Analysis and Classification	Thu-2-7 Speaker Recognition II	Thu-2-8 General topics in speech recognition	Thu-2-9 Speech Synthesis: Prosody Modeling	Thu-2-10 Language learning	Thu-2-11 Speech enhancement	
Coffee break 21:30 - 21:45												
21:45 - 22:45	Thu-3-1 Speech in Health II (HEALTH II)	Thu-3-2 Speech and Audio Quality Assessment	Thu-3-3 Privacy and Security in Speech Communication	Thu-3-4 Voice Conversion and Adaptation II	Thu-3-5 Multilingual and code-switched ASR	Thu-3-16 Speech and voice disorders	Thu-3-7 The Zero Resource Speech Challenge 2020	Thu-3-8 LM adaptation, Lexical Units and Punctuation	Thu-3-9 Speech in Health I (HEALTH I)	Thu-3-10 ASR neural network architectures II - Transformers	Thu-3-11 Spatial Audio	
23:00 - 24:00	Closing											

TECHNICAL AREAS	1 Speech Perception and Production	3 Paralinguistic Analysis	5 Analysis of Speech and Audio Signals	7 Speech Synthesis	9 Speech Recognition II: Architecture	11 Spoken Dialog Systems	Special Sessions	Industry Forum
	2 Phonetics, Phonology, and Prosody	4 Speaker and Language Identification	6 Speech Coding Enhancement	8 Speech Recognition I: Signal Processing	10 Speech Recognition III: New Applications	12 Spoken Language Processing	Show & Tell	

iFLY Star

iFLYTEK Research Institute Campus Recruitment 2021 - iFLY Star Project

Who are we?

iFLYTEK Research Institute

iFLYTEK A.I. Research Institute was established in 2005, featuring 'aim high with technologies, stand firm with applications; conduct useful research via correct method'. We are committed to the research of core A.I. technologies. We focus on research in following areas: intelligent speech, computer vision, natural language processing, etc. iFLYTEK boasts multiple world's leading research achievements in the field of artificial intelligence and has won champions in global competitions for many times.

Here in iFLYTEK. Appreciate the world's leading A.I. technology. Enjoy a simple and sincere culture. Explore the profound mystery of A.I. with scientists.

iFLY Star Project

'iFLY Star Project' is a global recruitment looking for highly-motivated graduates of leading universities to join iFLYTEK Research. Through the tailored growth plans, we will help you pave your way to rising stars in the A.I. field.

You will get

1. Luxury offer package: Competitive salary + long-term incentives.
2. Tailored growth plan: An integrated system of production, education and research; broad participation in industrialization and application; 'Ace Pilot' Elite training program; Adequate guidance provided by top tutors of executives/scientists; Systematic training for infinite possibilities; Interdisciplinary research opportunities.
3. Opportunities and abilities to change the world together: Work side by side with the world's top research teams; Create a better world with artificial intelligence

Recruitment Positions

Researcher

Responsible for the research of core A.I. algorithms, including intelligent speech / computer vision / natural language processing, etc. Solve technical problems in application and industrialization of core algorithms.

Algorithm Engineer

Responsible for the R&D engineering, industrialization, application and promotion of core A.I. algorithms, including intelligent speech / computer vision / natural language processing, etc.

Locations

Hefei / Shanghai / Suzhou / Beijing / Guangzhou / Xi'an / Changchun

We need you

2021 undergraduates, postgraduates and doctoral graduates
Apply if you meet one of the following criteria:

- Rising academic stars with publications in top international academic conferences or journals.
- Legendary champions of college entrance examinations, gifted youths in special classes, and TOP 20% students in well-known universities at home and abroad.
- Medal winners of ACM-ICPC regional or global competitions.
- Candidates with rich internship experience in top companies or extensive project practices on campus are preferred, and those with unique understanding or breakthroughs in a certain technical field are preferred.

How to submit your resume:

Scan the QR code below via WeChat



>>>>
Enable machines to listen, speak, understand and think.
Let us join hands
Create a better world with artificial intelligence.

For more information, please visit the official website of campus recruitment (campus.iflytek.com)

AI.



Magic Data

WHO ARE WE?

Magic Data Technology, an AI data service provider.

We are committed to providing a wide range of customized data services in the fields of Automatic Speech Recognition, Text To Speech, Computer Vision Recognition and Natural Language Processing. With human-in-the-loop data processing, we significantly improved the efficiency and quality of AI data labeling. Established in 2016, we have collected more than 100,000 hours of standard multilingual speech corpora under various scenarios. We help our clients gain easy and timely access to data with accuracy up to 99% and also provide them customized solutions. Magic Data employs a vast team of skilled data specialists and has a wide network of consultants around the world to assist with specific data needs.

MAGIC DATA PROVIDES

- **Quality:** Pre-packaged datasets available for immediate application in AI system development.
- **Scalability:** Capable to cover 50+ languages and dialects for audio recording and collecting services.
- **Compliance:** Fully legitimate and traceable data, prepared under strict data encryption, in supervised system, promises its reliability and security.
- **Expertise:** Strict quality management system, ensuring a continuous output of high-quality data products.

DATASETS

We provide valuable and reliable training data to empower your AI models. You can find datasets in different languages, styles, and solutions. Our datasets can improve your AI models' performance, thus accelerating the commercialization of AI initiatives.

Selected Language Datasets

ASR	Read Speech Corpus
	American English Speech Corpus
	French Speech Corpus
	Peninsular Spanish Speech Corpus
	Korean Speech Corpus
	Japanese Speech Corpus
	Bahasa Indonesian Speech Corpus
	Thai Speech Corpus
	Mandarin Chinese Speech Corpus
	Shanghai Dialect Speech Corpus
	Guangzhou Cantonese Speech Corpus
	Minnan Dialect Speech Corpus
	Wuhan Dialect Speech Corpus
	Shanxi Dialect Speech Corpus
	Guangzhou Cantonese In-Vehicle Speech Corpus
	Chinese-English Code-Mixing Speech Corpus

Conversational Speech Corpus
English Conversational Telephone Speech Corpus
Japanese English Conversational Speech Corpus
Korean English Conversational Speech Corpus
Japanese Conversational Speech Corpus
Korean Conversational Speech Corpus
Bahasa Indonesian Conversational Speech Corpus
Turkish Conversational Speech Corpus
Malay Conversational Speech Corpus
Mandarin Chinese Conversational Speech Corpus
Mandarin Chinese Conversational Telephone Speech Corpus
Guangzhou Cantonese Conversational Speech Corpus
Shanghai Dialect Conversational Speech Corpus
Sichuan Dialect Conversational Speech Corpus
Uyghur Conversational Speech Corpus
Hangzhou Dialect Conversational Speech Corpus

TTS
Chinese female voice emotion TTS dataset
Chinese female customer service TTS dataset

American English Speech Corpus for TTS
Mandarin Chinese Speech Corpus for TTS

Popular Launched

ASR
German Conversational Speech Corpus
French Conversational Speech Corpus
Peninsular Spanish Conversational Speech Corpus

Italian Conversational Speech Corpus
Brazilian Portuguese Conversational Speech Corpus
Peninsular Arabic Conversational Speech Corpus



Scan QR code
for more information

For more dataset products, please consult customer service: 400-900-5251
 Email: business@magicdatatech.com
 Website: www.magicdatatech.com



英特尔® 至强® 可扩展平台 高效实践人工智能

深度学习加速，显著提升性能；创新存储技术，优化批量训练；
广泛硬件组合，满足多样推理；强大软件工具，加速开发部署。

打破理论与现实的壁垒



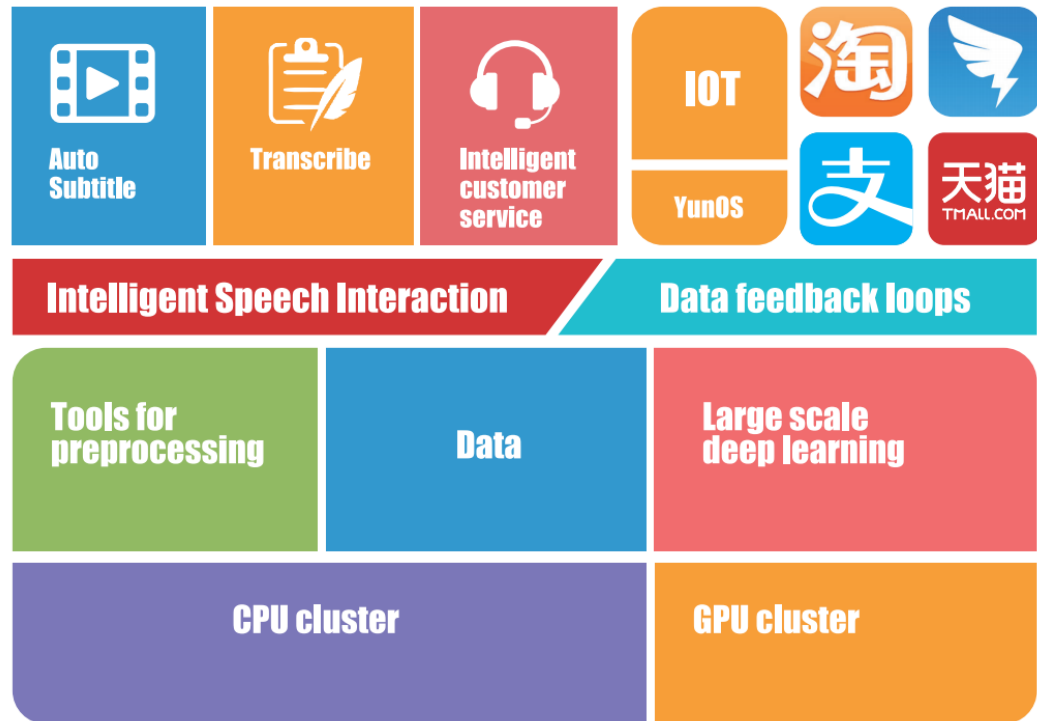
加速AI实践，请访问www.intel.cn/ai

性能测试中使用的软件和工作负荷可能仅在英特尔微处理器上进行了性能优化。诸如SYSmark和MobileMark等测试均系基于特定计算机系统、硬件、软件、操作系统及功能。上述任何要素的变动都有可能测试结果的变化。请参考其他信息及性能测试（包括结合其他产品使用时的运行性能）以对目标产品进行全面评估。更多信息，详见www.intel.com/benchmarks。英特尔技术特性和优势取决于系统配置，并可能需要支持的硬件、软件或服务得以激活。产品性能会基于系统配置有所变化。没有任何产品或组件是绝对安全的，更多信息请从原始设备制造商或零售商处获得，或请见intel.com。英特尔、INTEL、英特尔标识、至强、Xeon标志等是英特尔公司或其子公司在美国和/或其他国家的商标。©2019英特尔公司版权所有。

Intelligent Speech Interaction Team @ Alibaba Group

Our mission is to enable speech & NLP to become one of the technical infrastructures supporting Alibaba's business platforms and the small businesses in our ecosystem.

We are building state-of-the-art technologies in speech recognition, speech synthesis, speaker identification and verification, emotion detection, natural language understanding, dialogue system, QA system. We are also building large-scale deep learning infrastructure, and creating scalable speech & NLP services on top of Alibaba Cloud computing infrastructure. With all these technologies and platforms, we are the hub of natural user interface between human and Alibaba services.



The mission of Meituan is “We help people eat better, live better” . As China’s leading e-commerce platform for services, Meituan operates well-known mobile apps in China, including Meituan, Dianping, Meituan Waimai, and others. Meituan offers over 200 service categories, including catering, on-demand delivery, car-hailing, bike-sharing, hotel, and travel booking, movie ticketing, and other entertainment and lifestyle services, which covers 2800 cities and counties across China.



Meituan Speech Interaction

Meituan Speech Interaction Department is established in 2017, our mission is to develop advanced speech interaction technology to meet the needs in Meituan's commercial ecosystem. We focus on speech interaction technologies such as speech signal processing, speech recognition, speech synthesis, speaker recognition, natural language understanding, dialog and knowledge graph. We are looking for talents (both full time and internship) who are interested in speech interaction technology.

We Need You – Speech Engineer

Responsibilities

- Architect, design and develop ASR/TTS/speech processing algorithms.
- Improve the system performance in production with cutting edge technologies.

Basic Qualifications

- Advanced knowledge of software engineer, familiar with C++, python.
- Knowledge of speech signal processing, machine learning or deep learning.
- Experience in developing online speech service.

Preferred Qualifications

- Expert knowledge of ASR, TTS or other related technologies.
- Authors of papers in Interspeech, ICASSP or other related top tier journals/conferences.
- Experience with distributed training system, GPU optimization.

Contact us: meituan.oi@meituan.com (Titled with “Interspeech2020”)

与AI同行——2021思必驰校招进行中



如果你对人工智能极具热情

如果你具备优秀的学习能力

如果你善于沟通与协

思必驰以诚挚之心邀请你的加入

我们一起 可以做更酷的事情!

招聘岗位

苏州 / 上海 / 北京 / 深圳

NLP算法工程师	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
自然语言处理研究员	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
语言技术研究员	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
语音算法工程师 (识别/唤醒/合成)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
语音信号处理算法工程师	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
JAVA研发工程师	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
安全工程师	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
硬件工程师	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C/C++研发工程师	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
嵌入式开发工程师	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
助理测试工程师	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
安卓开发工程师	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
产品经理	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
产品运营	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
后端开发工程师	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

岗位详情, 请扫描下方二维码

简历投递

简历投递邮箱 (邮件正文请备注, 消息来自INTERSPEECH)

—— xiaozhao@aispeech.com



Sogou Vocational Avatars(搜狗分身): Digital human-based Multimodal Human AI Interaction system



Multimodal recognition (ASR+lip reading)

- More than 1.4 billion speech recognition requests per day, including multilingual & dialect recognition
- Initiated personalized speech recognition, which greatly improved personalized speech recognition experience
- First issued Chinese lip recognition system, command word recognition accuracy can reach 90% in vehicle, smart home and other scenes
Based on modality attention, initiated multimodal recognition technology, which can significantly improve the recognition effect in noisy environment

Speech synthesis

- Leading neural network modeling technologies , Self-developed StyleTTS system
- Supporting multilingual, multi-tone and multi-style diversified broadcast
- First issued personalized speech synthesis and emotion transfer in the industry
- Won the first place in the 2 subtasks of Bizzard Challenge 2018, the most authoritative speech synthesis competition in the world

Voice conversion

- Research of innovative technologies based on representation learning and transfer learning
- Industry's first voice conversion product which can support real-time any-to-one voice conversion.
- With high similarity and sound quality, it has supported dozens of speakers including stars, anime, games and more types.

Digital human

- Integrating Sogou Multimodal HCI technology including core technologies of speech, computer vision and natural language processing.
- The industry's first commercial application level of multi-mode audio and video generation technology with broadcasting effect similar to human.
- Has been applied in the media, customer service, justice, education, entertainment and other fields, and has gone abroad.

Multimodal recognition



The greatest voice input method in China

- More than 1.4 billion daily speech recognition requests



Language and Dialect



Speech enhancement




Personalized interaction

Sogou AI Hardware



Multimodal representation



Speech synthesis
Personalized speech synthesis

- Upload 10 sentences about 3 to 5 minutes of audio
- The timbre is close to human, customizing the proprietary sound archive.
- Featuring audio companionship, you can tell stories, read private book lists, etc

Story tellers
Let baby love your voice



Voice conversion
Pioneering in the industry, double breakthrough of representational learning and transfer learning.

- High similarity
- Free to convert your voice
- Multi-representation transfer





AI news anchor

As the world's first synthetic news anchor, Sogou Vocational Avatars have been used in media.

- Vivid image
- Extremely low customization costs
- Mature solutions
- Quick customization

Machine translation

Sogou simultaneous interpretation



Ranked first in IWSLT 2018.
Machine translation won first prize in Chinese-English translation competition at premier international academic competition WMT2017.



Conference



Live Captioning

Sogou Translator



 Please send your resume to
Join us zhaopin@sogou-inc.com

 Please visit Sogou AI Open Platform:
Learn More ai.sogou.com