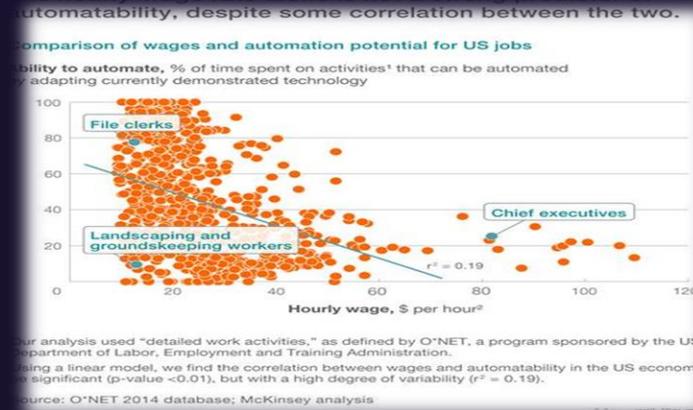

A.I. and Its Application in Language Assessment and Education

April 2018

A.I. Comes SOONER than You Could Imagine



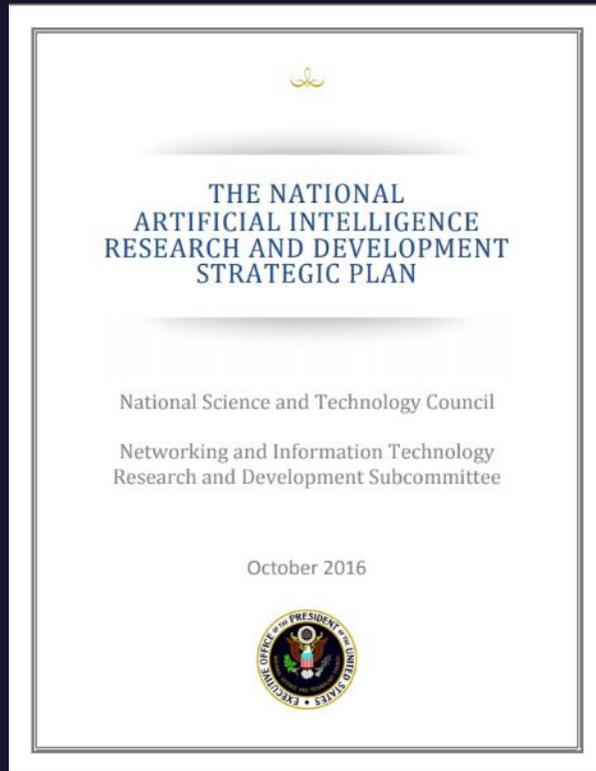
- 23rd. May 2017, AlphaGo Vs Kejie by 3-0
- 15th March, 2016, AlphaGo Vs Lee Sedol by 4-1.



- MIT: **45%** of US jobs, **20%** of CEOs' can be automated by currently techs. If natural language could be understood, another **13%** could be replaced.

—— McKinsey Global Institute : The redefinition of jobs

A.I. in National Strategy



In Oct 2016, White House introduced
National Artificial Intelligence Research and Development Strategic Plan



In July 2017, the State Council of China issued
Next Generation Artificial Intelligence Development Plan

A.I. in National Strategy

Nov. 2017, the Ministry of science and technology appointed 4 National AI innovation platforms



Dec. 2017, the Ministry issued "The Three Year Action Plan for Ai industry"

□ Key Supports for AI products of

- Smart Voice interaction
- Smart interpretation & translation
- Automated Vehicles
- Service robots
- Smart UAV(unmanned aerial vehicle)
- Medical Image Aided Diagnosis System
- Video image identity recognition
- Smart Home

A.I. in National Strategy



Aug. 2012

Speech Industry Alliance of China
iFLYTEK as president of the council



Nov. 2017

A.I. Industry-University-Research Innovation
Alliance of Chinese Academy of Sciences
iFLYTEK as president of the council



What is the current situation of A.I. industry?

Is it a concept? Is it a bubble?

Or is it actually changing the world?

3 Levels of Artificial Intelligence



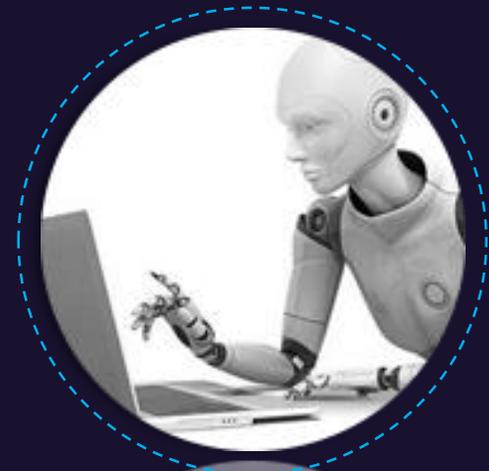
Computing Intelligence

Compute



Perceptive Intelligence

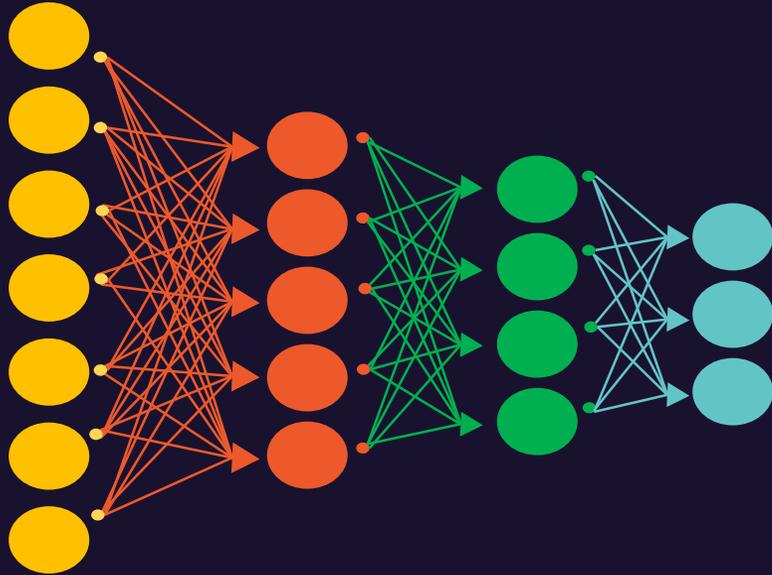
listen
Speak
Recognize



Cognitive Intelligence

Understand
Think

2 Main Approaches of AI

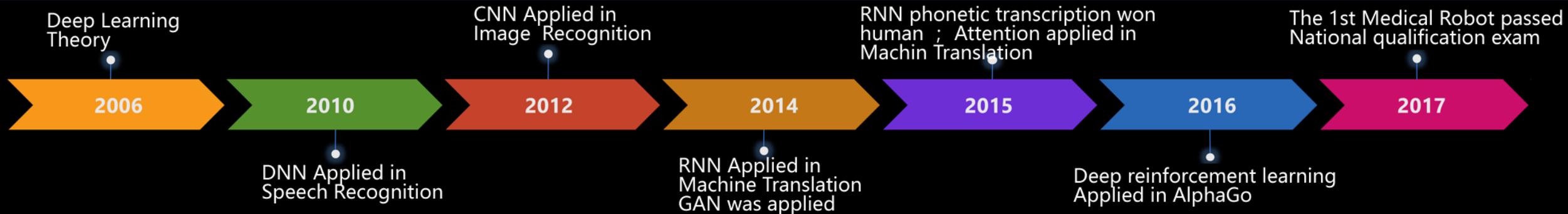


The progress of neural network algorithm
Such as DNN



The Study of Brain science

Deep Learning under the Third Wave of A.I.

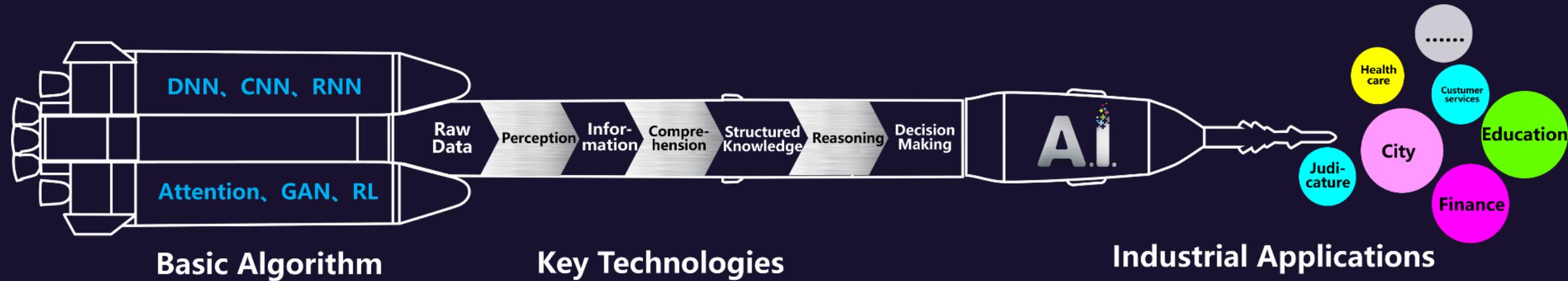


Overview of IFLYTEK

- Founded in 1999
- National Key IT Enterprise
- The largest public AI & Speech Technology company in Asia Pacific regions
- Over 9000 employees, over 10 Billion US dollars market value



The Components of IFLYTEK A.I. Research and Industry



Construction of the first National Key Laboratory of Cognitive intelligence in China

The first National Major AI Project of China --- Humanoid Answering Robot

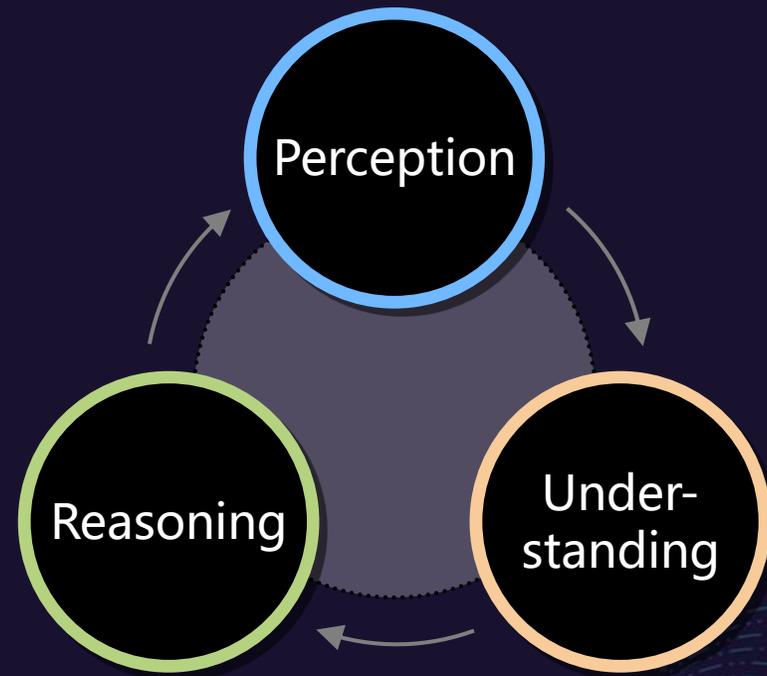
As Leading Unit, Iflytek launched the program "Hyper Brain of Iflytek", developing smart systems based on humanoid neural networks and cognitive intelligence



讯·飞·超·脑

National Engineering Laboratory for speech and language information processing

National Engineering Laboratory for Application of Brain-like Technologies



Speech Synthesis



Top1 of Blizzard Challenge, 2006~2017

ASR VS Human Dictation

In 2015, our machines outperformed human stenographers for **the 1st time**

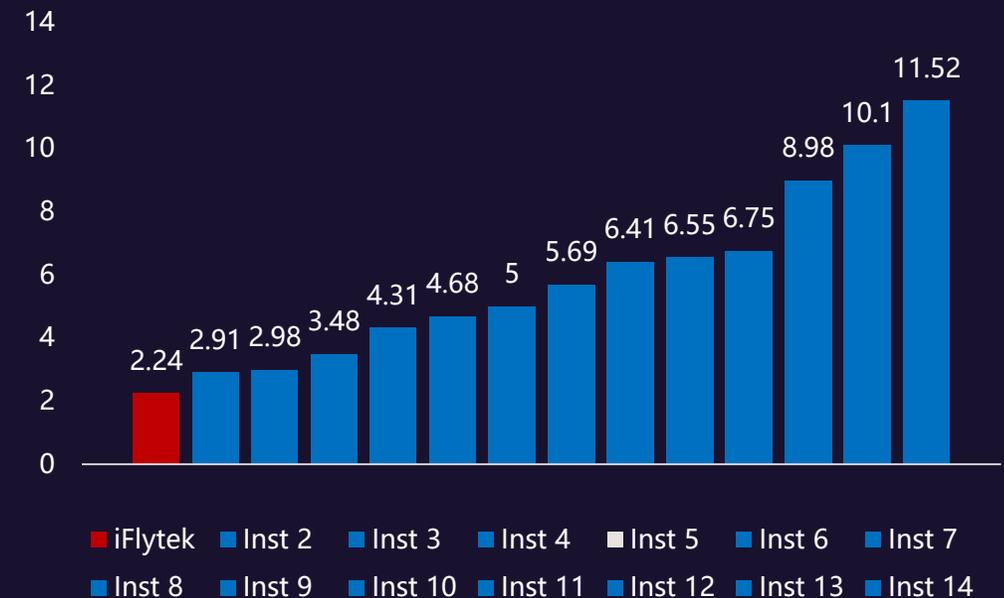
1st Place of 2016 CHiME Challenge



Contestants	iflytek	Steno-grapher A	Steno-grapher B	Steno-grapher C	Steno-grapher D	Steno-grapher E
Accuracy	98.70%	74.40%	69.60%	72.40%	60.10%	70.8%

iflytek Product Launch Event, Dec. 21, 2015
Under the supervision of notaries from state notarial organization

Error rates of participants (6-microphone scenario)



Participants including Stanford institute, Carnegie Mellon University, NTT, Hitachi, MITSUBISHI, Singapore Nanyang Technological University, France National Institute of information and automation etc.

Machine Translation VS Human Translation

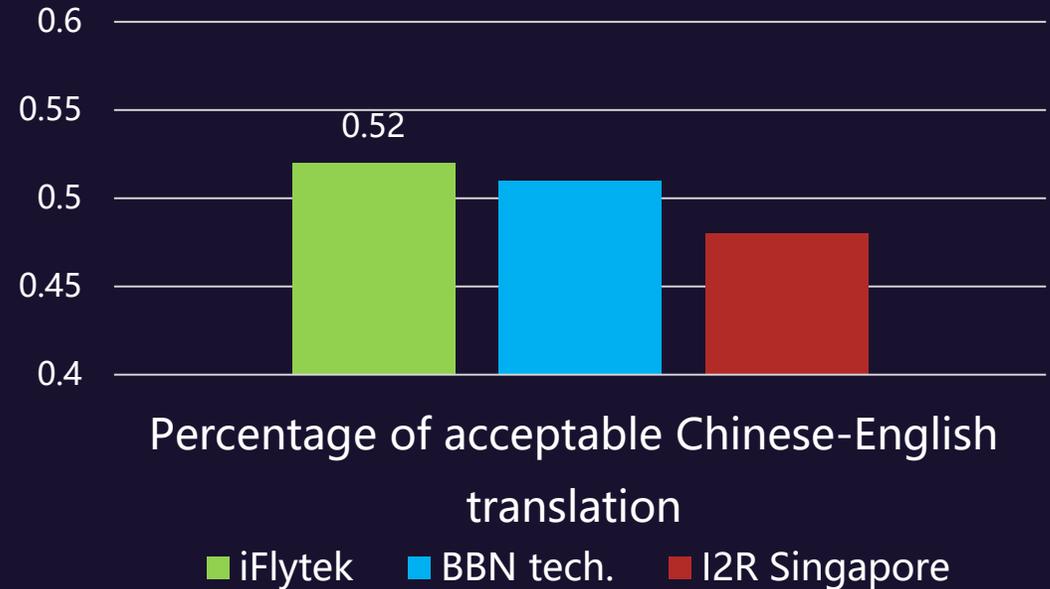
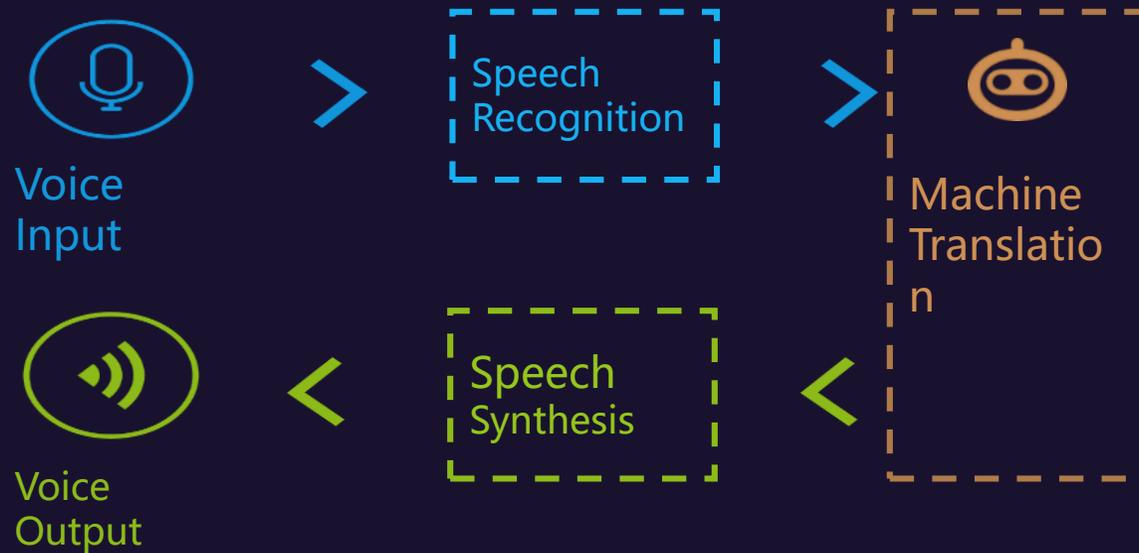
	SMT	GNMT	(Human) Full Marks
Chinese -> English	4.046	4.606	5.0
English -> Chinese	3.984	4.598	5.0

Source : Google' s Neural Machine Translation Systems, Yonhui Wu, etc., 2016.9, Test Data from Wikipedia and News Sites

	SMT	NMT	(Human) Full Marks
Chinese -> English	4.46	4.73	5.0
English -> Chinese	4.54	4.81	5.0

Source: iFLYTEK Verbal Test Collection On Tourism English , iFLYTEK On-line Engine Test Results on Jun,2016

Speech Translation VS Human Translation



Machine interpreting of iFlytek is sophisticated enough for daily conversations.

By Aug 2017, machines' language proficiency reached level of **CET-6**

1st Place in NIST Open Machine Translation 2015 Evaluation

Oral English Level Equals to Chinese CET-6 ,
Still a Great Distance from Simultaneous Interpretation

Translation Devices



Easy-trans (晓译) Translator machine

34 Main Languages

First Trans-machine with an off-line Engine

Translate whatever you see and whatever you hear



Yibei (译呗)

Translator machine with Language **practice function**

20000+ Dialogues for Daily lives

Machine Voice with a MOS of **4.2**

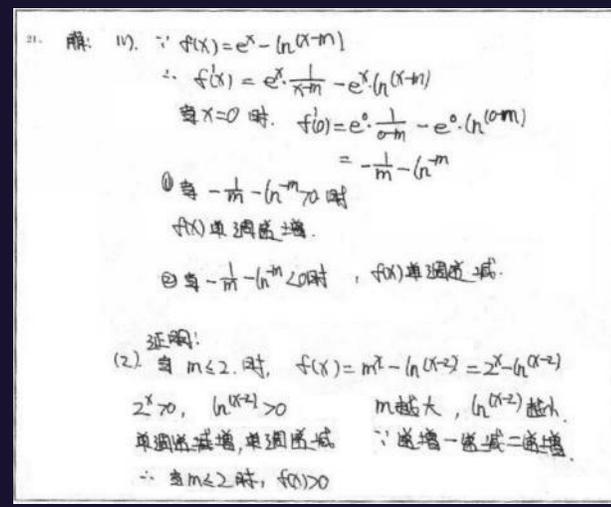
Breakthrough of Image Recognition and OCR/HWR

Accuracy of English
HWR

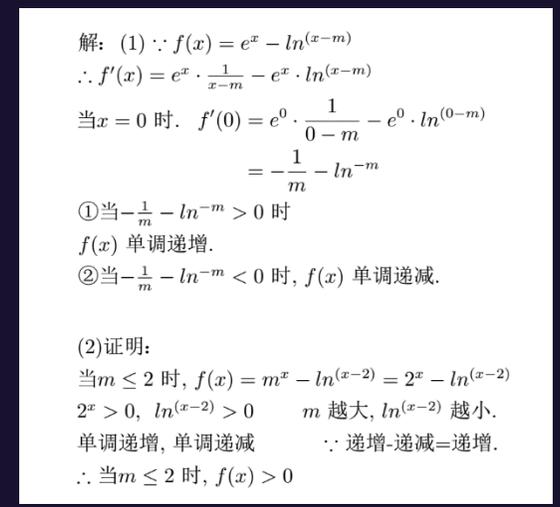
97%

92%

Mixed Image and Text Recognition:
Machines can Read Anything



Original
Paper

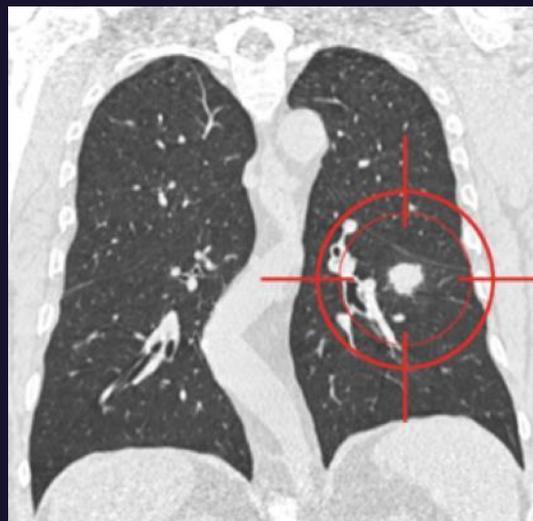


Recognition
Result

Medical Images

94.1%

LUng Nodule Analysis



Broke the LUNA world record in August 2017

The most authoritative international evaluation in medical imaging of lung nodule

Participants: Radboud University Nijmegen, Alibaba, the Chinese University of Hong Kong, Peking University, Zhejiang University, Mevis, and many start-up companies.

Machine diagnosis of CT medical images in Anhui AI Hospital
the accuracy reached 94%.

CCTV 13
新闻

CCTV.com



10月18日
周六

14:47 举行联合反导计算机演习。

CCTV新闻 中国

CCTV : Smart medical robot passed national Qualification exam

Auto-Drive

Cityscapes World Record



CITYSCAPES
DATASET

Initiated by Mercedes Benz An international authoritative dataset evaluation in autonomous driving

81.4%
urban scene
understanding

Broke Cityscapes world
record in Oct. 2017

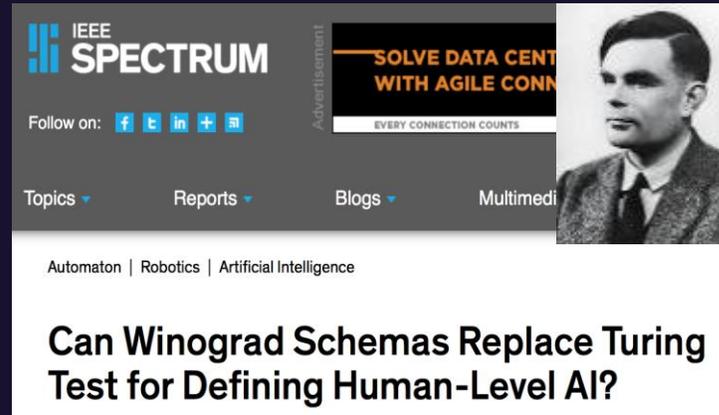
Participants: Google, the Chinese University of Hong Kong, and more than 40 innovative enterprises and top academic institutions at home and abroad.



Sample of Smart car system

Commonsense Reasoning

1st Place of Winograd Schema Challenge 2016



A father cannot lift his son, because he's heavy.

Q: Who is heavy? A: The son.

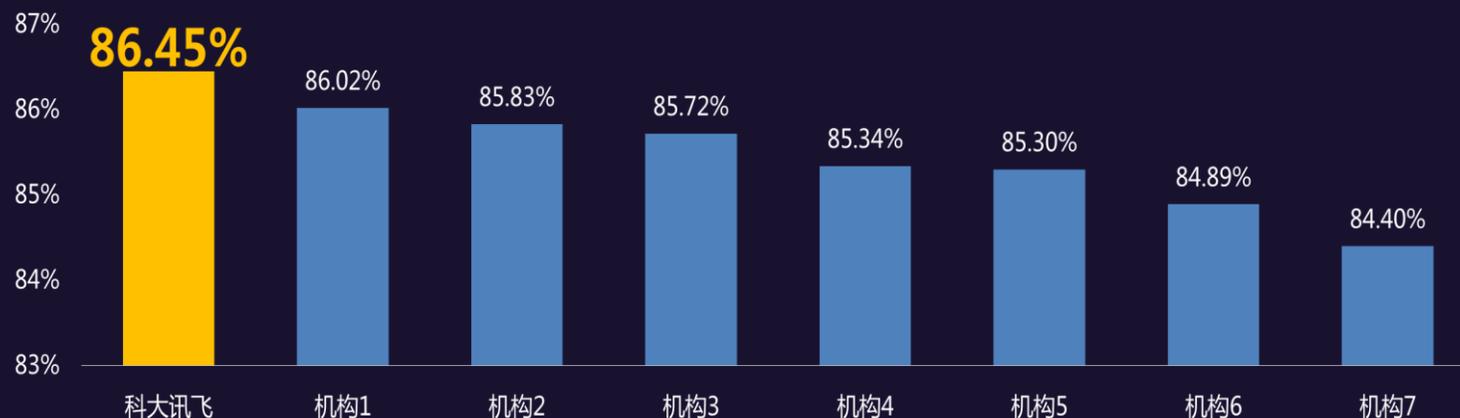
A father cannot lift his son, because he's weak.

Q: Who is weak? A: The father.



Reading Comprehension

Broke SQuAD (Stanford Question Answering Dataset) world record in October 2017 , by **86.45%**



Rank	Model	EM	F1
1 Jul 2017	Interactive AoA Reader (ensemble) Joint Laboratory of HIT and iFLYTEK Research	77.845	85.297
2 Jun 2017	r-net (ensemble) Microsoft Research Asia http://aka.ms/rnet	77.688	84.666

The most authoritative international evaluation in machine reading comprehension

Participants: Microsoft, Google, Facebook, IBM, Salesforce, Stanford University, Carnegie Mellon University, Tsinghua University, Peking University, Fudan University and other enterprises and research institutions at home and abroad.

IFLYTEK Open Platform



2017/12

Terminal Devices

1.6 b

Developer teams

500 k

Daily Use

4 b

2016/12



1 b



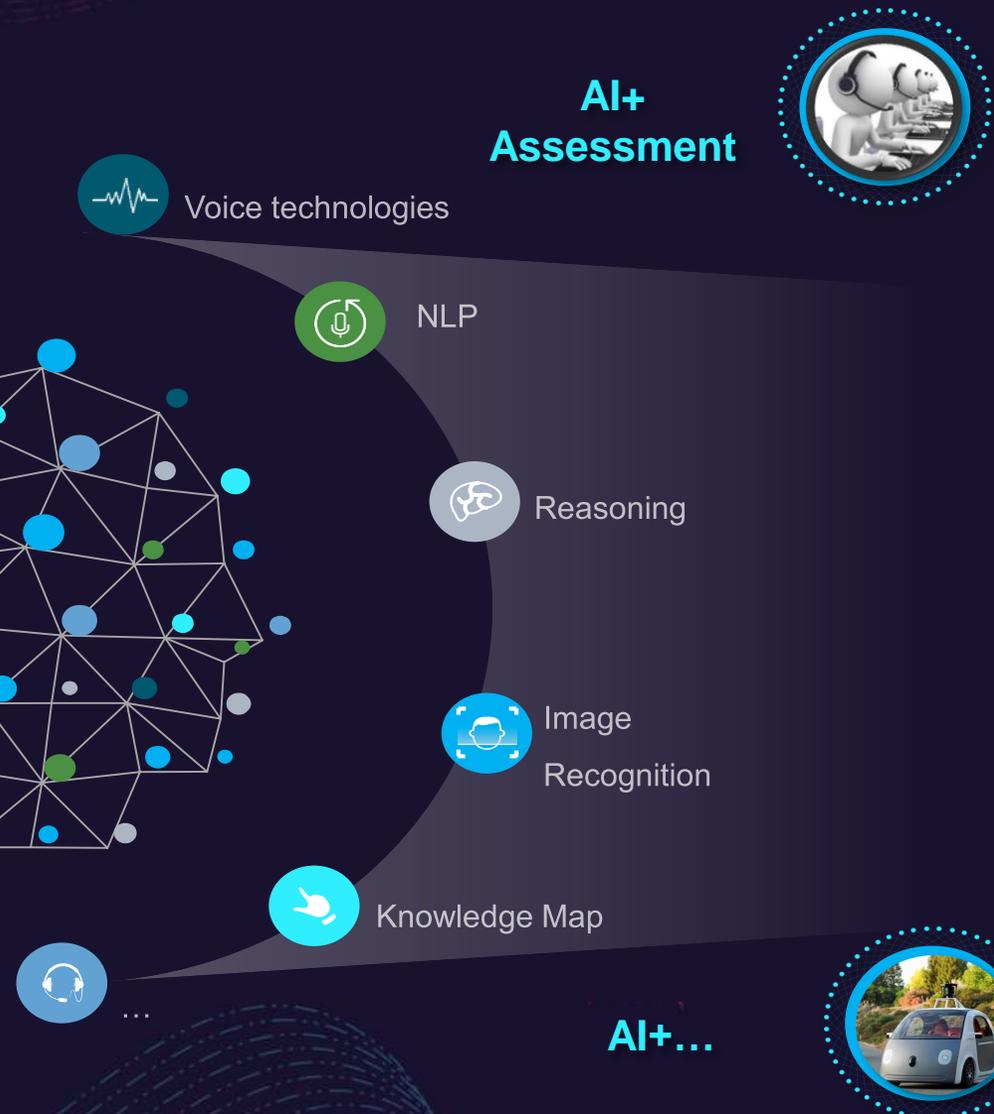
240 k



1.5 b

A.I. + Education (Language Assessment)

A.I. + Education



AI+ Targeted Teaching



AI+ Adaptive learning



AI + Management



AI+ Smart Learning Pal



AI + STEAM Education



AI+...

Evolutional Roadmap of Language Assessment

2003
Evaluation of
Pronunciation

work
[英][wɜ:k]
工作

峙 zhì

⚠️ 声母读得不正确

▶ 原声播放

▶ 我的录音

● 正确朗读

● 错误朗读及漏读

▶ 标准读音

▶ 我的读音

▶ 对比播放

2009
Oral Essay

成绩报告

本题满分100分

我的得分 65分

流畅度	64%
语速	55%
准确度	66%

我的答案

I like summer best because my favorite sport is swimming. I can go swimming with my friends often. And I can wear cool clothes because of the hot weather. In summer, we can see many beautiful flowers, green trees, blue sky and clear water in the lake. Ice-cream is the most popular in summer. But we can't eat too much. Eating too much ice-cream is bad for our health.

▶ 原声播放

▶ 我的录音

● 正确朗读

● 错误朗读及漏读

2014
Written Essay

In Chinese culture, red is often a symbol of good luck, longevity and happiness. Red can be seen everywhere. People put the cash as gifts to family members or close friends, usually in red envelopes. Red is popular in China as revolution and the communist party. However, red is not always represents good luck and happiness. Because once upon a time the dead's name written in red, ~~ink~~ to ~~write~~ commonly used in red ink to write the Chinese name is considered an offense.

OCR

AES

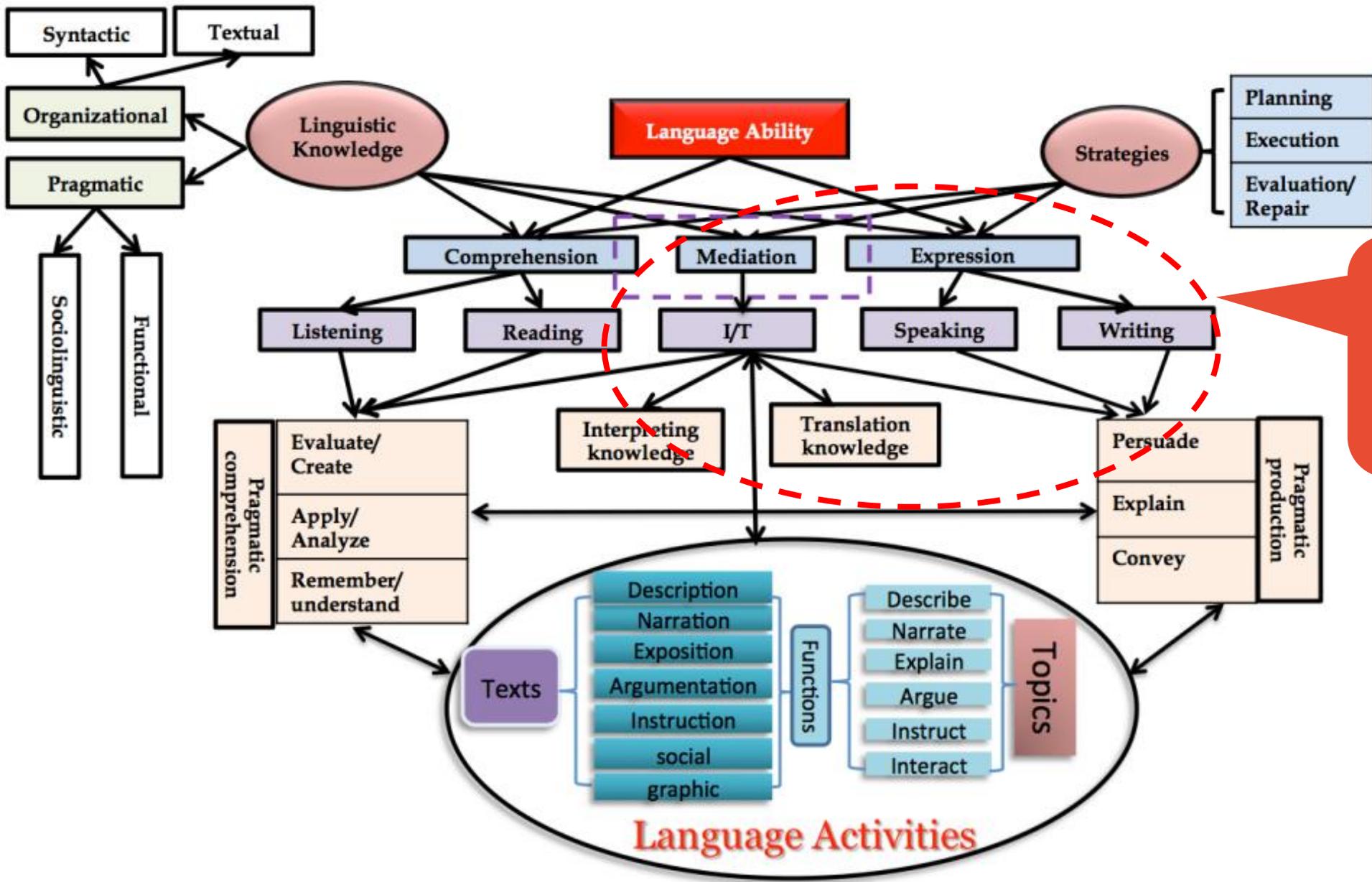
Correction: 97%+

2017
Semantic Parsing

SQuAD
The Stanford Question Answering Dataset

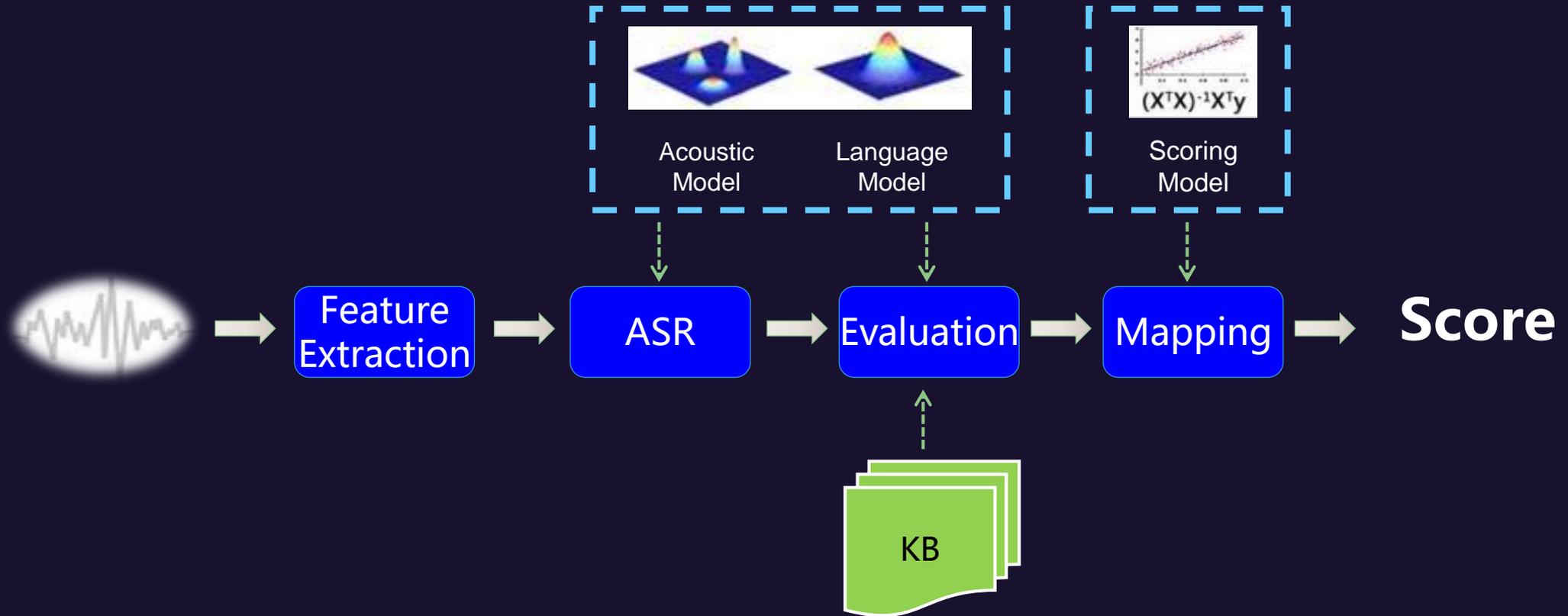
Rank	Model	EM	F1
1 Jul 2017	Interactive AoA Reader (ensemble) Joint Laboratory of HIT and iFLYTEK Research	77.845	85.297
2 Jun 2017	r-net (ensemble) Microsoft Research Asia http://aka.ms/rnet	77.688	84.666

- Won the top prize at the
1. Stanford Question Answering Dataset (SQuAD)
 2. Winograd Schema Challenge
 3. NIST TAC(Text Analysis Conference) Knowledge Base Population in Cognitive intelligence fields Competitions



Labor- consumed for testers to grade.

Core Procedure of Spoken Language Assessment



Speech Grading System

- Auto Grading System in Mandarin/English oral test. Agreement rate with experts is over 95%.
- Mandarin Chinese oral test : 31 provinces with 5.5 million students each year
- English oral test: 3 million students each year, the only official system applied in college entrance examination in China.

Total Score Correlation



Total average deviation



Comparison between experts and machine
Source: Scientific and technological identification Report

Hand-written Essay Evaluation



HWR :
Handwriting
Recognition



NLU : Natural
Language
Understanding



Assessing
With DL Learning
from human



Reducing
Workload

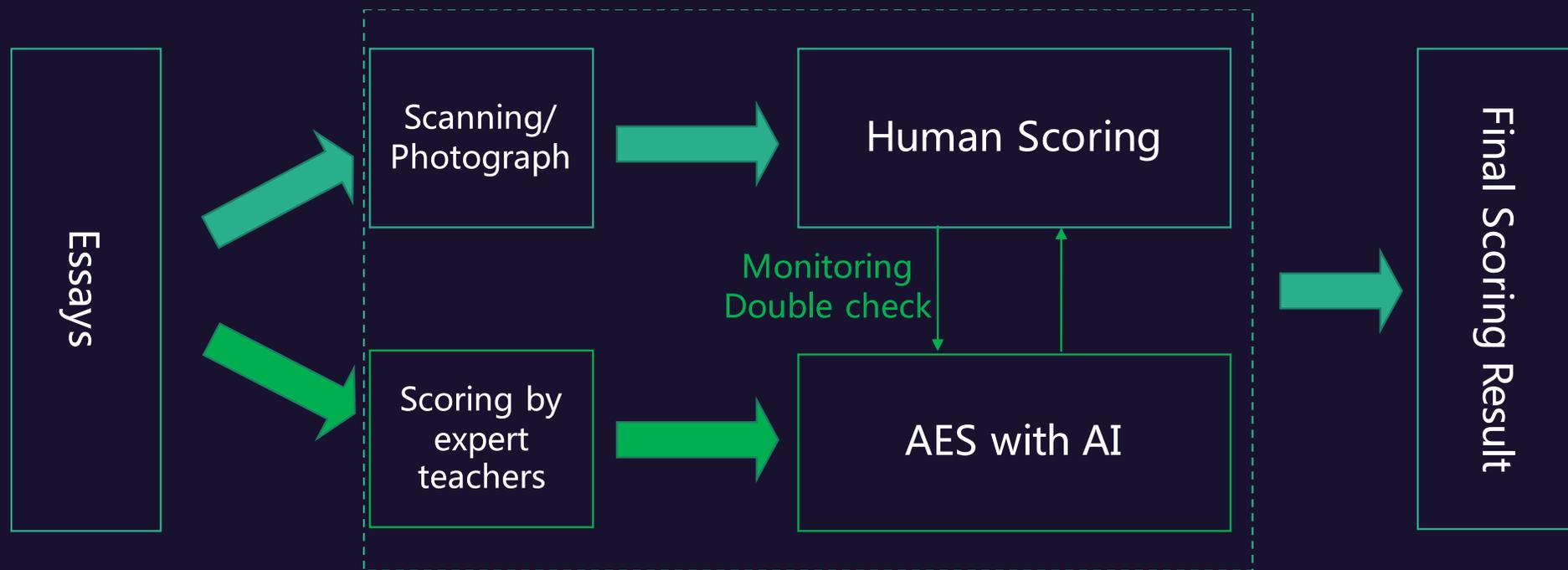


Unified
Measurement



Studying
Feedback

Human-Machine Collaboration

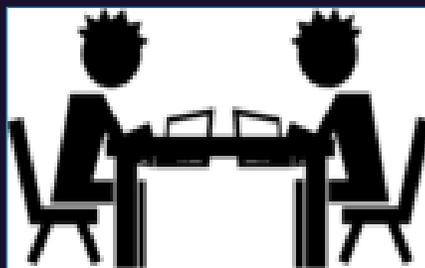


Human-machine collaboration in scoring

Language Assessment Applications



10,000+
schools



10,000,000+
students



100,000+
essays per month



Big and Bigger Data

Other Language Assessment Applications



MOE-iFLYTEK collaborative online platform for Chinese learning

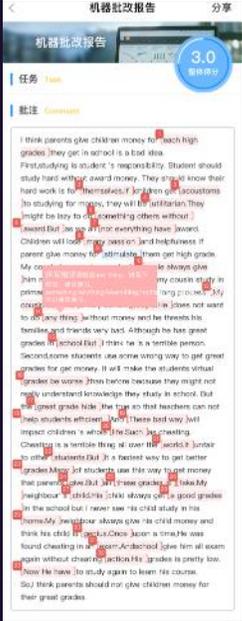


The most popular online system for learning and mock tests powered by iFLYTEK

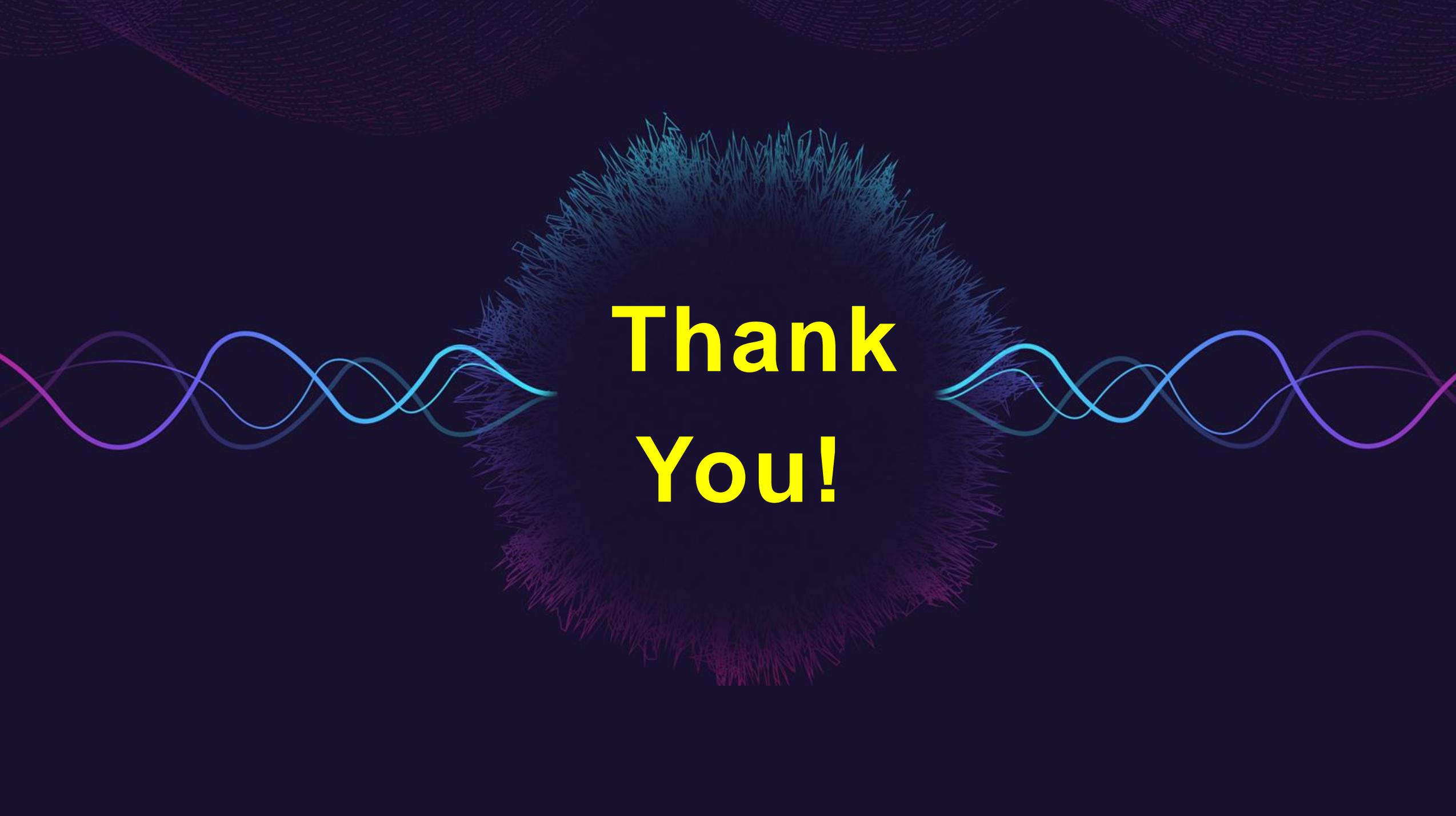


Mother Language learning Portal of Singapore MOE

RealSkill



RealSkill



**Thank
You!**