



“ The biggest difference between human intelligence and animal or machine intelligence is cognitive intelligence. It comes from our mastery of language and how we express knowledge.

”

— Hu Yu, Executive President and Consumer BG President of iFlytek

# iFlytek: The voice of AI

AI is on a clear upward trajectory and is reshaping all aspects of life. According to Hu Yu, Executive President and Consumer BG President of iFlytek, AI is starting to approach human intelligence. Serving hundreds of millions of users with its world-leading technologies, iFlytek started off as a pioneer in China's voice recognition industry and has now evolved into a global leader in AI. But it all started with a little twist of fate.

By Xu Shenglan, Xue Hua

## From intelligent voice to Super Brain

**F**ounded in 1999, iFlytek's primary goal was to make machines talk, something that even today is reflected in the company's mission: "We want the world to hear our voice." And that's starting to happen – the company is now at the forefront of the AI phenomenon.

Hu smiles as he recalls, "We had no idea at the

time that we were working on AI. At least we weren't sure what AI really was. We also weren't aware that 1999 was a bad year for AI, as the second wave of AI innovation had just peaked." Slightly tongue-in-cheek, he says, "If we'd known that AI was going to be such a tough business, we might never have started the company. I guess it was just fate."

Around 2004, AI wasn't the hot tech it is today, says Hu, but his team had come to realize that they were holding a key piece of AI. "The biggest difference



“

The cognitive revolution around voice and language is the peak of human intelligence and the biggest challenge for AI today.

”

between human intelligence and animal or machine intelligence is cognitive intelligence. It comes from our mastery of language and how we express knowledge, which allows us to do logical reasoning and complex decision-making,” he says. The cognitive revolution around voice and language, Hu believes, is the peak of human intelligence and the biggest challenge for AI today.

Hu is the leader of the iFlytek Super Brain Project, which was launched in 2014, “It’s much more than just a fancy name. We announced our definition of AI as computational intelligence.” He asserts that machines were much more powerful than humans since the day they were invented, citing AIs that play the board game Go as an example of computational intelligence. “Humanoid machines possess both perceptual intelligence and motion intelligence. That means they can see, hear, and feel the surrounding world. Today there are some impressive humanoid and animal-like machines,” he says. “However, the reason we’re at the top of the planet’s food chain is language, or ‘cognitive intelligence’.” According to Hu, one of the goals of the Super Brain Project is to evolve machines from the level of perceptual intelligence, where they can hear, talk, see, and recognize, to the level of cognitive intelligence, where they can understand and think.

Currently, Super Brain is using big data to train and

optimize its algorithms. They’re not trained by simply cramming all kinds of data into the system; instead, the system actively processes data from interactions in real-world scenarios, and uses that data to update itself. Hu believes this style of self-enhancement is like the ripple effect, where the volume of data grows exponentially as the product reaches more people, enabling his team to more rapidly iterate and optimize the product experience.

## No shortage of awards

iFlytek boasts leading tech in areas like speech synthesis, voice recognition, voice assessment, and translation. From 2005 to now, the company has racked up 13 consecutive wins at the Blizzard Challenge, the world’s leading speech synthesis contest. It’s also won various machine translation championships, including the IWSLT 2014 and NIST 2015. Over the past six years, iFlytek’s voice recognition accuracy has improved from 60.2 percent to over 98 percent. The company’s strengths in voice tech became a natural bridge into the world of AI and its industrial application.

iFlytek is also researching the dynamic of AI and neurology. Through computing based on the human brain, iFlytek is trying to unlock the mystery of our intelligence. If they succeed, it may pave the way towards Artificial General Intelligence, meaning human levels of intelligence, one of AI’s holy grails.

“

Offering real-time interpretation between Mandarin and 33 other languages and Chinese dialects, iFlytek Translator also translates text in photographs and can be used on 4G or Wi-Fi networks or offline.

”

## Translation on the fly

iFlytek started applying AI to the real world in the shape of natural language processing (NLP) back in 2010, when it developed China's first voice input product and the second of its kind in the world, after Google. iFlytek's system has an accuracy of more than 98 percent and supports 22 different Chinese dialects.

In 2016, iFlytek released its first smart device, the iFlytek Translator, which it followed up in April 2018 with the 2.0 incarnation. Offering real-time interpretation between Mandarin and 33 other languages and Chinese dialects, it also translates text in photographs and can be used on 4G or Wi-Fi networks or offline. Most of its users – 86 percent – use it on vacation. Translator 2.0 has also mastered the accents of four major dialects in China's complex and voluminous linguistic web: Cantonese, Sichuanese, Northeastern Mandarin, and Henan dialect, with support for more expected in the future. In an advance for NLP, the product can recognize different situations and adapt to its users' language tics.

“There are some who say that there's no need to build a translator device because the translation function can be integrated into a smartphone. But we made a deliberate decision to sell our translator as a hardware device,” says Hu. First, he explains, we tend to hold our phones close to our faces, which might

not always be possible depending on the scenario. Second, phones are affected by ambient noise. Third, Hu believes that intelligent hardware must be easy to use. The best experience is something that works with a single click, but using an app on a smartphone isn't always easy or intuitive. Fourth, the translation process should allow for natural and intuitive interaction, and sticking your smartphone in someone's face isn't always socially acceptable.

In 2012, iFlytek launched its voice cloud platform as part of its efforts to build an ecosystem for the AI industry. Since then, more than 860,000 developers have worked on the platform, which connects 1.9 billion devices and provides nearly 4.6 billion interactive services each day.

In 2015, iFlytek launched the human-machine interaction interface AIUI, hitting a milestone in the AI industry. AIUI redefined the standards for human-machine interaction in the connected era. Hu adds, “In 2017, iFlytek was announced as one of China's first open innovation platforms for next-generation AI and our platform will focus on intelligent voice technology. The government clearly recognizes the importance of the ecosystem built on our company's AI.”

## AI: An industry enabler

iFlytek is also applying intelligent voice and AI technology

“

In the intelligence era, all AI applications will run on the cloud. As cloud computing consumes a lot of resources, device computing and edge computing will better support AI.

”

to different sectors, including the judiciary and education.

In the justice system, iFlytek is working with China's Supreme People's Court and Supreme People's Procuratorate (public prosecutors). In 2016, a test in Anhui Province showed that an AI system could identify phone scams with a very high level of accuracy. Moreover, a pilot study found that trials were 30 percent shorter when intelligent voice recognition was used instead of a human reporter.

In education, AI has outperformed all expectations in scoring test papers. In a test in Jiangsu Province, two different AIs scored a series of college entrance test papers. For Chinese essay questions, the two AIs differed by an average of less than seven points per paper. They were 92.82 percent consistent – more than 5 percent higher than the average consistency of two human teachers. A trial in Hunan showed similar scores.

iFlytek is currently working with China's National Education Examinations Authority to build an AI lab to jointly develop more advanced technologies for education.

## A partnership covering multiple markets

iFlytek and Huawei have formed a strategic partnership to develop practical applications for voice and AI

technology in the areas of telecoms and smart devices, building on nearly a decade of collaboration. In 2010, the two companies deployed the world's first open cloud platform for Chinese voice recognition.

In May 2018, Huawei and iFlytek signed a strategic agreement covering four areas: public cloud services, ICT infrastructure, smart devices, and office IT systems. Huawei also integrated iFlytek's AI technology into its smartphones to gain an edge over its competitors. Huawei and iFlytek are working on smart devices and device cloud services based on iFlytek's voice AI technologies and capabilities, including voice recognition, speech synthesis, iFlyrec, and iFlytek translation.

In the enterprise space, Huawei uses iFlytek's technology and products in its infrastructure and its own office applications. The iFlytek speech engine will form a key component of Huawei's Enterprise Intelligence cloud platform. Hu believes that in the intelligence era, all AI applications will run on the cloud. As cloud computing consumes a lot of resources, device computing and edge computing will better support AI.

Each with its own strengths and ecosystems, we're certain that Huawei and iFlytek will help build a strong AI ecosystem and make AI a valuable asset to life, business, and society. 