BRIDGE REPORT　　　　　　　　　　　　　　　　Bridge Salon

Representative Director Daisuke Yoshida	AI, Inc. (4388)
	Qiエーアイ

Corporate Information

Exchange	TSE Mothers
Industry	Information and communications
Representative Director	Daisuke Yoshida
Address	KDX Kasuga Building 10F, 1-15-15 Nishikata, Bunkyo Ward, Tokyo
Year-end	March
URL	https://www.ai-j.jp/english

Stock Information

Share Price	Shares Outstanding		Total Market Cap	ROE (Actual)	Trading Unit
¥1,817	5,138,000 shares		¥9,335 million	16.0%	100 shares
DPS (Estimate)	**Dividend Yield (Estimate)**	**EPS (Estimate)**	**PER (Estimate)**	**BPS (Actual)**	**PBR (Actual)**
8.00	0.4%	¥40.53	44.8x	¥208.84	8.7x

*The share price is the closing price on June 11. Each number is taken from the brief financial report for the FY March 2020.

Earnings Trends

Fiscal Year	Net Sales	Operating Income	Ordinary Income	Net Income	EPS	DPS
March 2017 (Actual)	451	115	116	76	19.57	0.00
March 2018 (Actual)	591	146	147	109	24.73	0.00
March 2019 (Actual)	737	211	202	150	30.84	8.00
March 2020 (Actual)	819	273	273	172	34.12	7.00
March 2021 (Estimate)	840	280	280	205	40.53	8.00

*Unit: Million yen, yen. The estimated values are provided by the company.

*DPS ¥8.00 for FY 2019 includes ¥3.00 commemorative dividend.

This report presents AI, Inc.'s earning results for the fiscal year ended March 2020, etc.

BRIDGE REPORT   Bridge Salon

Table of Contents

Key Points

- **AI, Inc. offers a speech synthesis engine and solutions regarding speech synthesis. The company provides corporations and consumers with products and services based on "AITalk®," a speech synthesis engine developed by the company, for automatic answering systems, car navigation systems, anti-disaster wireless systems, smartphones, communication robots, in-vehicle devices, and games. The company has unrivaled characteristics and strengths; for example, it can synthesize high-quality speeches from a few voice samples and offer many speakers.**

- **The sales for the term ended March 2020 were 819 million yen, up 11.1% year on year. The sales of corporate products and services grew, while the sales of consumer products shrank. Gross profit margin improved 3.0 points, and gross profit increased 15.4% year on year due to covering outsourcing expenses with the in-house system, etc. Operating income rose 29.6% year on year to 273 million yen. SG&A expenses also augmented due to personnel increase, etc., but the increase in sales and gross profit offset this. Ordinary income increased 35.1% year on year to 273 million yen because there were no longer temporary costs associated with the listing. Both sales and profit exceeded the initial estimates. Considering the strong performance, the company increased the dividend forecast by 1 yen/share from 6 yen/share to 7 yen/share.**

- **The sales for the term ending March 2021 are expected to increase 2.5% year on year to 840 million yen, and operating income to rise 2.3% year on year to 280 million yen. Although there are concerns about the impact of the spread of the novel coronavirus, sales and profits are expected to increase due to the expansion of the speech synthesis market. The dividend is to be 8 yen/share, which is 1 yen/share higher than the previous year's 7 yen /share. The expected payout ratio is 19.7%.**

- **The company started providing the next-generation speech synthesis engine AITalk 5.0. AITalk 5.0 actualizes more human-like, natural, high-quality speech synthesis. Also, by using deep learning, it can offer a speech synthesis engine at a low cost by greatly reducing the recording time and voice dictionary creation time. Thus, AITalk 5.0 is expected to expand the usage scenes of voice synthesis rapidly. We would like to expect the release of concrete application cases in the future.**

- **Moreover, although sales and profits are on the rise, we would like to pay attention to how the profitability of the company, whose top line needs to grow further as a growing company, will change with AITalk 5.0.**

BRIDGE REPORT   Bridge Salon

1. Company Overview

AI, Inc. offers a speech synthesis engine and solutions regarding speech synthesis. "AITalk®," which is a speech synthesis engine developed by the company, is offered to corporations for producing voices for automatic answering, car navigation, and anti-disaster wireless systems, and also as an audio communication system for smartphones, communication robots, in-vehicle devices, and automated call center operation. It also sells products targeted at consumers, including VOICEROID.

【1-1 Corporate history】

When the founder Daisuke Yoshida (representative director of AI, Inc.) was working for Advanced Telecommunications Research Institute International*, he encountered a speech synthesis technology, and had an intuition that it is a promising technology that would contribute to society. The technology was still immature, but he established AI, Inc. in April 2003, for the purpose of substantiating, diffusing, and commercializing that technology.

In 2007, the company started granting the license of the series of "AITalk®," which is a speech synthesis engine developed by the company. Later, it developed a variety of products and services based on "AITalk®." Its unique features, including "a wide array of speakers and languages" and "reduction of time and expenses with a small amount of voice samples," were highly evaluated. Since it was adopted by the government for anti-disaster wireless communication, it has been adopted by many institutions and applied in a wider variety of cases.

In June 2018, the company was listed in Mothers of Tokyo Stock Exchange.

*Advanced Telecommunications Research Institute International (ATR)

It was established in 1986, under the concept of the preparatory meeting held by the then Posts and Telecommunications Ministry, NTT, Japan Business Federation, Kansai Economic Federation, universities, etc., with the mission to promote pioneering, unique research in the field of information and communications based on the international collaboration among government, industry and academia. 111 companies hold a stake in the company such as NTT and KDDI.

【1-2 Corporate Mission, Vision】

On November 11, 2019, the company renewed its logo, corporate philosophy, and vision; it newly added a mission, value, and action guidelines.

Corporate Mission	Enriching our society with sound technology To create a new culture of sound information and contribute to the improvement of daily life culture through application development and service provision of sound technology.
MISSION	Providing "convenience" and "joy" through creating voices
VISION	To continue providing sound technology to thrive our society
VALUE	To keep being the pioneer and NO.1 company for sound technology 1.To provide Joy and happiness through our service and technology 2.To grow and create a prosperous future with our customers and employees 3.To thrive each day with each step we take

ACTION GUIDELINES	・To always achieve new skills and technology ・To be a considerate employee and progress with our customers and friends ・To thrive with ambition and achieve a prosperous growth

Amid the changes in internal or external business environments, the company has set forth the way they should be as its new corporate mission, corporate logo, and action guidelines, through which it aims to get recognized and become a company that provides value to society.

【1-3 Market environment】
**(1) Market environment**
The development of the speech synthesis technology has a long history. However, the expansion of the scope of application was slow because the mainstream method has been to produce audio data mechanically although it has been adopted for automatic answering machines, anti-disaster announcement, voice interaction via smartphones, etc.

As the technology for producing sounds pronounced by human beings has advanced and artificial intelligence (AI) has evolved in recent years, we have seen the improvements in functions, including the shift from voice-over recording to "the utilization of speech synthesis," the shift from unilateral provision of information to "the actualization of interactive communication," and the shift from the Japanese language only to "multiple languages." Going forward, the scope of application is expected to expand rapidly, and it will be used for e-learning, mobility, robots, AI speakers, etc.
A private research firm predicted that the scale of the global market of voice recognition and speech synthesis technologies will grow from about 47 billion dollars in 2011 to 200 billion dollars in 2025 (compound annual growth rate [CAGR]: about 10%).

Following "Phase 1: One-way information provision" and "Phase 2: Interactive dialogue and dissemination to consumers," AI, Inc. believes that the speech synthesis market has entered "Phase 3," a period of rapid growth, with alternatives to narration by speech synthesis, development of multiple languages, creation of new markets, etc.
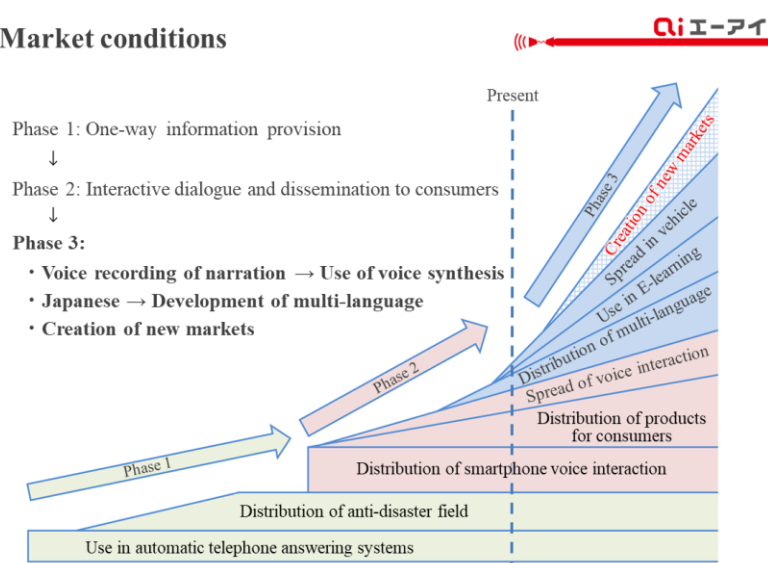
(Taken from the reference material of the company)

(2) Competitors
Major competitors of "AITalk®," a speech synthesis engine of AI, Inc., include HOYA Corporation (1st section of TSE, 7741, product name: Voice Text) and Toshiba Digital Solutions Corporation (unlisted, product name: To Speak).
Specializing in speech synthesis, AI, Inc. meets the requests from users swiftly and flexibly and secures its market share, by offering services of R&D, product development, sale, and support in an integrated manner.

【1-4 Business contents】
(1) What is the speech synthesis technology?
The voice technology can be roughly classified into the "voice recognition technology" for recognizing voices and translating them into characters, etc., and the "speech synthesis technology" for converting text information into audio data. AI, Inc. has been conducting the "speech synthesis" business since it was established.

BRIDGE REPORT   Bridge Salon

R&D in the speech synthesis field has a long history and dates back to around the 1850s. "Speech synthesis" reminds us of "mechanical sounds and robot voices" developed in around 1940, but AI, Inc. adopted the "corpus-based text-to-speech method."
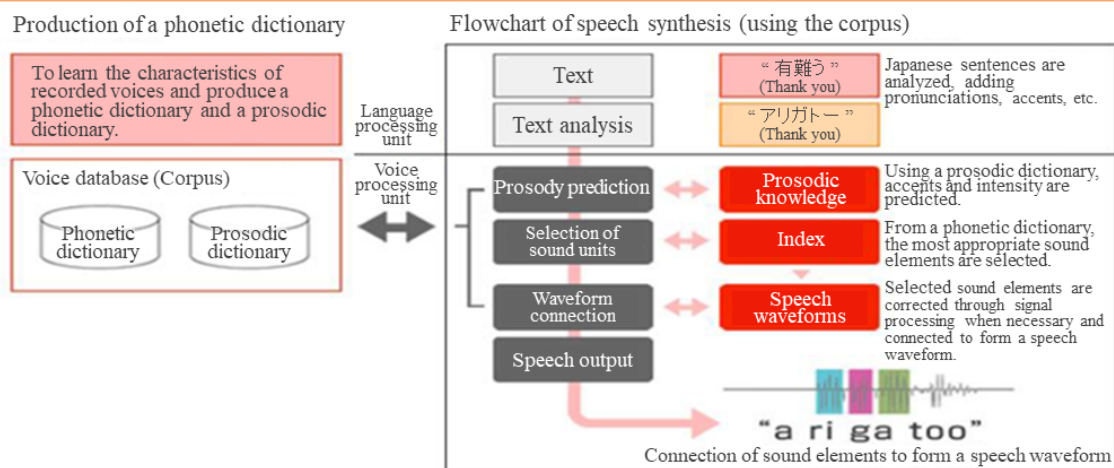
(Outline of the corpus-based text-to-speech method)
While the conventional "speech synthesis by rule" produces audio data mechanically, the "corpus-based text-to-speech method" produces a waveform by combining recorded human voices in units of vowels and consonants. Accordingly, sounds are derived from human voices rather than mechanical sounds.

The technology for "corpus-based text-to-speech synthesis" is constituted by the two technologies: "a technology for producing a phonetic dictionary" and "a speech synthesis technology for producing audio data from text information."

Technology for producing a phonetic dictionary	This technology records the voices of a specific person, breaks down recorded voices into sound elements, that is, vowels and consonants, and produces a phonetic dictionary (a collection of sound elements) and a prosodic dictionary (prosodic information of recorded voices). The precision of the task of producing a phonetic dictionary is essential for enhancing the reproducibility of recorded human voices.
Speech synthesis technology	This technology is composed of "a language processing unit," which analyzes Japanese text and adds information on pronunciations and accents, and "a voice processing unit," which predicts prosodic information with reference to the prosodic dictionary, selects the most appropriate sound elements from the phonetic dictionary, connects them to the sound waveform again, and outputs a speech.
	Both units require the precisions in the analysis of the Japanese language, prosody prediction, and the connection to sound waveforms.
	When these precisions are improved, it is possible to produce synthetic sounds that are like recorded human voices, as the sound elements of recorded voices are recombined to output a speech.

BRIDGE REPORT   Bridge Salon

	Conventional speech synthesis method		The company's speech synthesis method
Synthesis type	Recording and editing	Synthesis by rule	Corpus-based text-to-speech synthesis
Synthesis method	Replay of recorded voices as they are	Mechanical voice production	Production of a speech waveform by extracting and combining vowels and consonants
Voice quality	○	✕(Robot voice)	○
Degree of freedom	✕	○	○

(Taken from the reference material of the company)

(2) "AITalk®," a high-quality Japanese speech synthesis engine

"AITalk®" is a high-quality speech synthesis engine researched and developed by the company based on the "corpus-based text-to-speech synthesis technology," which produces sounds based on human voices.

The following section will describe the features of "AITalk®," which can synthesize speeches freely with more human-like and natural voices, major application cases, and outlines of products based on "AITalk®."

①Characteristics of "AITalk®"

*A diverse lineup of speakers and languages

Currently, Japanese speakers of this system range from adults to kids and speak 17 kinds of male or female languages (15 kinds of standard languages and 2 kinds of Kansai dialects). From this diverse lineup of voices, customers can choose appropriate ones for various scenes.

*Please try the "demonstration of speech synthesis" in the company's website at https://www.ai-j.jp/demonstration/.

BRIDGE REPORT

Bridge Salon

Introducing the Japanese Speakers

Nozomi　☺ 通常 ☺ 喜 ☺ 怒 ☺ 悲
Corresponding Expression of Emotion: Normal, joy, anger, sadness
Her voice is pleasant and youthful. Her voice can be used for various situations such as for narrations, automatic telephone answering system, wireless-activated disaster warning system, entertainment, etc.

Kaho
Her voice is extremely clear and easy to understand. Available for a wide range of use including automatic telephone answering (CTI, IVR) and narration for the making of animation.

Akari
Her voice gives a cheerful and bright impression. Most suitable for the use of product guidance and promotions.

Reina　☺ 通常 ☺ 喜
Corresponding Expression of Emotion: Normal, joy
Features a very sweet and gentle voice. This voice goes great with anime characters, apps and toys, and games.

Osamu
His voice features high applicability. Applicable to various scenes.

Taichi　☺ 通常 ☺ 喜
Corresponding Expression of Emotion: Normal, joy
His voice gives a youthful and unique impression. Most suitable for using in the field of entertainment.

Chihiro
A charming nasal voice.

Yuuto
A brisk and intelligent sounding boy's voice.

Sumire
Her voice gives a mature and sophisticated impression. Applicable to various scenes, it is also the same voice * used for VOICEROID and Yuzuki Yukari.

Maki　☺ 通常 ☺ 喜 ☺ 怒 ☺ 悲
Corresponding Expression of Emotion: Normal, joy, anger, sadness
Her voice is best suited for the use of characters in animes and games or other areas of entertainment.

Nanako
Features a very calming voice. Her voice is best suited for reading news and audio guidance.

Seiji
His voice has a very sincere tone. Suitable for persuasion and calling attention.

Hiroshi
His voice gives a homely and honest impression. Very suitable for literary recitation.

Anzu
Features a very loving and earnest voice.

Koutarou
Features a slow-paced and cute voice.

Japanese Speakers (Kansai)
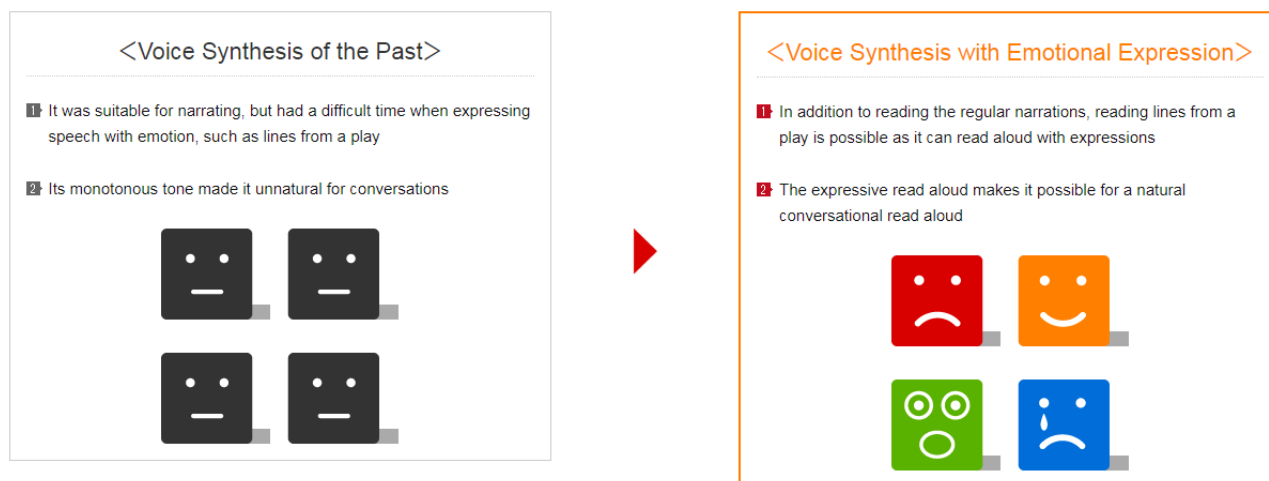
Miyabi
She speaks slowly and gently.

Yamato
He speaks clearly, briskly, and youthfully.

(Taken from the website of the company)

*It is also possible to express emotions.
It is possible to express emotions, including delight, anger, sorrow, and pleasure, according to situations and purposes of use.

(Taken from the website of the company)

*Anyone's voice can be converted into synthetic data.
The voices of entertainers, voice actors, and users recorded for a short period of time can be converted into data for speech synthesis.
Since it is possible to easily produce speeches of real people just by inputting text, it is possible to offer a variety of contents, including online campaigns, smartphone applications, and games.

②**Customer segments and major application cases**
As the "corpus-based text-to-speech synthesis technology" has advanced, the speech synthesis engine has been adopted in various scenes where recorded voices of voice actors and narrators had been used.
AI, Inc. has a broad range of client enterprises in the fields of communications, disaster prevention, finance, railways, transportation, in-vehicle devices, games, sightseeing, municipalities, and libraries. Over 500 companies adopted the system, and we heard that the number of clients is increasing by 20-30% every term.

As IoT and robots have been popularized and the number of sightseers visiting Japan has increased over the past several years, there are an increasing number of cases in which the system is used as a dialogue solution combining voice recognition and the interpretation of intentions or a speech translation solution combining translation and multilingual speech synthesis. The company expects that the speech synthesis technology will be used for interactive dialogue as part of artificial intelligence, indicating the evolution from the conventional unilateral information provision.

Application case	Outline
(1) Anti-disaster wireless communication	Many municipalities use the system for producing audio announcements to citizens in anti-disaster wireless communication and the national early warning system (J-ALERT).
(2) Smartphone voice interaction	The voice interaction apps for smartphones, such as "Shabette Chara®," which is provided by NTT Docomo, Inc., and "Yahoo! Audio Assist," which is provided by Yahoo Japan Corporation, are increasingly used.

(3) Road traffic information and car navigation	The system is utilized for road traffic information, which offers real-time road traffic information, such as "road traffic information" of Japan Road Traffic Information Center and car navigation, which guides an enormous number of place-names throughout Japan, such as "Docomo Drive Net Info" of NTT Docomo.
(4) E-learning	Lightworks (CAREERSHIP®), Tokyo Customs, Chugai Pharmaceutical, Taiho Pharmaceutical, etc. use the system.
(5) Broadcasting	The system is used by TBS (IRASUTO Virtual Caster), TV Tokyo (Morning Satellite), BS JAPAN (Nikkei Morning Plus), etc.
(6) Communication Robot	The system is used in many robots such as "Pepper" by SoftBank Robotics Corp and Matsukoroid by Matsukoroid Production Committee.
(7) Public-address in buildings and stations	The system is utilized for announcing information at stations, airports, commercial facilities, such as JR Kyoto Station and Memanbetsu Airport Bldg.
(8) Automatic answering system	The system is used for notifying library users of the dates when a library is closed by telephone, answering customers' calls at banks, and attending to customers at call centers. It is applied broadly to automatic answering systems, including telephone banking.
(9) Reading of websites	The system is utilized as a tool for giving information of websites of municipalities and enterprises throughout Japan with synthesized voices.
(10) Production of audio files	The system is utilized as a tool for producing audio files used for narrations of e-learning content, guidance about equipment, such as ticket dispensers, and so on.
(11) Video games	The system is utilized for voice-overs of video games, such as the series of "StarHorse," an arcade horse racing game provided by SEGA Interactive Co., Ltd., and "Kuma-Tomo (Teddy Together)" of BANDAI NAMCO Entertainment Inc.
(12) Packaged products for consumers (Package for reading contents aloud)	The system is utilized for producing audio files for packaged products for consumers, including the "VOICEROID®" series offered by AHS Co., Ltd.

Matsukoroid

This is an android entertainer developed by making a cast of the entire body, including the head and toes, accurately mimicking facial expressions, behavior, habits, etc., and applying the cutting-edge android technology, with the aim of producing an android that is like two peas in a pod with Matsuko Deluxe.

It was born under the supervision of Professor Hiroshi Ishiguro of Osaka University, who is a pioneer in android research.

AITalk®, a speech synthesis engine of AI, Inc., was adopted for producing some voices of "Matsukoroid." AI, Inc. recorded the actual voices of Matsuko Deluxe in a short period of time, and produced "AITalk® CustomVoice®," an original phonetic dictionary for speech synthesis. This enabled Matsukoroid to read a variety of texts aloud with the voices of Mastuko Deluxe.

Going forward, Matsukoroid will speak with AITalk®, which synthesizes speeches with the voices of Matsuko Deluxe, at events, etc.

(Taken from the website of the company)

BRIDGE REPORT　　　　　　　　　　　　　　　　　Bridge Salon

③**Major products**

Based on AITalk®, AI, Inc. develops and sells products and services suited for various scenes of corporations and individuals.

Product name	Outline	Application cases
AITalk® Koe-no-shokunin (Voice Craftsman)	Software for producing narrations, with which you can produce audio files easily just by inputting text into your PC. Anyone can produce high-quality narrations with easy, intuitive procedures. The latest version "AITalk® 4" can adjust emotions.	Narrated video manuals for e-learning, sightseeing guides, public-address announcements, etc.
AITalk® Koe Plus (Voice Plus)	Add-in software for PowerPoint®, which can add voices to the slides of PowerPoint® easily. You can easily produce high-quality voices in PowerPoint® files.	Production of narrated e-learning content with PowerPoint® only, addition of voices to presentation material for use inside and outside your company, etc.
AITalk® SDK	This software development kit (SDK) can synthesize speeches freely from human-like, natural voices and offer them via libraries. The latest version "AITalk® 4 SDK" can adjust emotions.	To integrate into package software / voice of automatic telephone answering system / integration into devices/ WEB campaign and WEB service
AITalk® Server	This engine is suited for cases where a network is used and synthesis is conducted with multitasking, such as automatic answering and online services.	Voice for automatic telephone response / WEB campaign, WEB service
AITalk® Custom Voice®	This is a service of recording the voices, etc. of entertainers, voice actors, and customers and producing an original Japanese phonetic dictionary for speech synthesis. Just by inputting text, it is possible to produce speeches with real voices.	It can be applied to a variety of content, including online campaigns, smartphone apps, and video games.
Kantan (Easy)! AITalk®	Packaged software for individual users, with which you can produce high-quality narrations just by inputting text.	Inputting your own voices for narrations of videos, production of original audio teaching material which can be used in trains and vehicles for listening.
AITalk® Anata-no-koe (Your Voice)	Your voice, etc. can be reproduced with the speech synthesis technology. With your PC and this packaged software, including Custom Voice®, you can produce speeches in various words anywhere, anytime.	It is possible to read a closing address of a funeral with the voice of the deceased. You can give lectures and presentations without speaking, by synthesizing speeches with your voice.

(3) The next-generation speech synthesis engine, AITalk®5 .

In May 2020, AI, Inc. released the next-generation speech synthesis engine, AITalk®5 (provisional name), which utilizes the Deep Neural Network (DNN) to express emotions with the speech synthesis engine.

(Background for development)

For the company's current AITalk®4, which is a corpus-based speech synthesis engine, it is necessary to create a separate emotion sound dictionary for emotions, such as happiness, sadness, and anger, needed to create interactive sound synthesis.

BRIDGE REPORT   Bridge Salon

This has problems like the large cost and that the change in the emotion of the synthesized speech was random and not smooth. Therefore, the company has been working on "subsidies for developing new products and technologies project" for 18 months from July 2017 to December 2018 in order to successfully transition from a calm state to an emotional state smoothly, by predicting emotion change filters from DNN and producing emotion elements from normal elements; they succeeded in commercialization. The company is currently applying for a parent on that system.

(Overview of the next-generation speech synthesis engine, AITalk®5)

(1) Characteristics

According to the usage scene, it offers the option of a conventional corpus-based speech synthesis system or the DNN speech synthesis system and has the following characteristics:

1 The system can synthesize a more natural and human-like high-quality sound thanks to the sound improvement achieved by using deep learning. Moreover, it eliminates the problem of jarring transition between emotions associated with AITalk®4 and now can synthesize an emotion-rich voice that transitions smoothly between happiness, sadness, and anger.

2 The conventional AITalk®4 required separate emotion sound dictionaries for each emotion and entailed recording separate sounds assigned to each emotion happiness, anger, and sadness. Comparatively, the next-generation speech synthesis engine, AITalk®5 utilizes Deep Learning to create sound dictionaries from a much shorter recording than usual. Therefore, shortening the time of recording and creating sound dictionaries lead to reducing the costs of creating sound dictionaries and allowed the company to offer the speech synthesis engine at a much lower price.
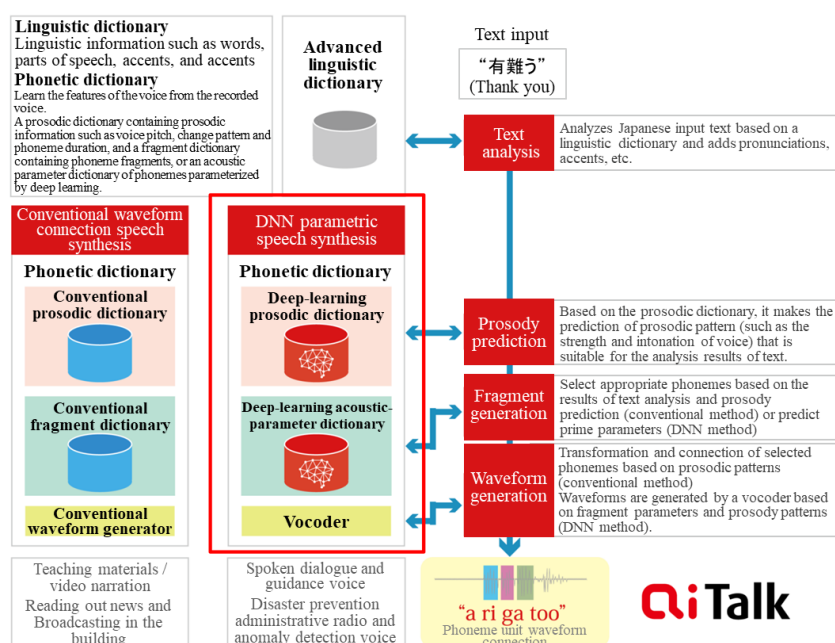
(2) Products lineup

AITalk®5 SDK: development kit/library

AITalk®5 Custom Voice®: Original sound dictionary creation service

AITalk®5 Editor: A narration/guidance sound creation software

AITalk®5 Server: server-based speech synthesis

In May 2020, the company started providing "AITalk®5 Koe-no-shokunin (Voice Craftsman) ® Package Edition" and "AITalk®5 SDK."

(Taken from the reference material of the company)

**(4) Business model and commercial distribution**

The company's products and services are classified into "products for corporations," "services for corporations," and "products for consumers."

To corporations, AI, Inc. offers the most appropriate products or cloud services according to the characteristics of each client.

As for marketing targeted at corporations, the company owns "Inside sales" staff, who deal with inquiries through sales promotion (SEO, email newsletters, news releases, etc.), and "Field sales" staff, who strive to increase new customers and orders from existing customers, and sales partners sell packaged software.

As for marketing targeted at consumers, the company does not sell its products directly to customers, but entrusts distributors with sale, and receives royalties from them on a quarterly basis.

Sales composition by each segment

■ Products for corporations ■ Services for corporations ■ Products for consumers
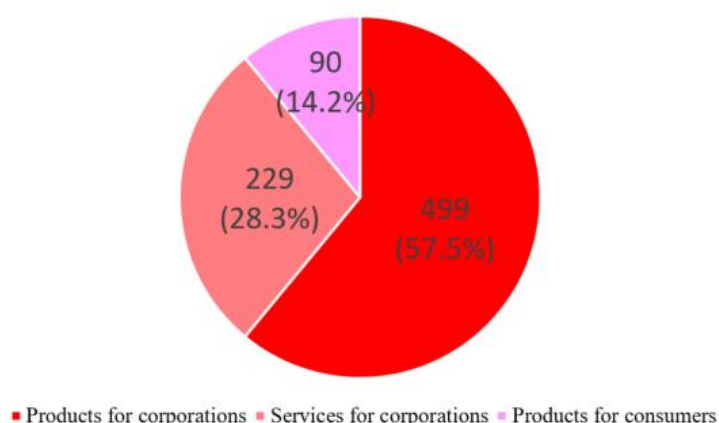
＊ Unit: million yen, FY3/20

①**Products for corporations**

AI, Inc. sells packaged software, grants licenses, and carries out entrusted development.

◎Sale of packaged software

The company sells packaged software with which you can easily produce audio files just by inputting text into your PC.

Through easy, intuitive operation, it is possible to produce high-quality voice-overs.

Major products and services	Business model	Fee example
AITalk® Koe-no-shokunin (Voice Craftsman)® AITalk® Koe Plus (Voice Plus)	One-shot revenue type	900,000 yen for eternal use

◎Licensing

This is a major business model of AI, Inc. The company concludes a licensing contract for use with each client and receives some fees for the use of the speech synthesis engine.

The company individually set the basic license fee, monthly fees for use, royalties, which depend on sales results, and so on. The company offers the most appropriate speech synthesis engine according to the purposes of use.

Major products and services	Business model	Fee example
AITalk® SDK AITalk® Server micro AITalk®	Recurring-revenue type	Basic license fee + Royalties (set individually)

BRIDGE REPORT   Bridge Salon

◎Entrusted development

AI, Inc. is entrusted by clients with the development of original phonetic dictionaries for respective clients.

Major products and services	Business model	Fee example
AITalk® Custom Voice®	One-shot revenue type	400,000 to 5,000,000 yen according to plans

②**Services for corporations**

◎Cloud service

The company offers speech synthesis services utilizing the cloud environment. Users can use services utilizing speech synthesis via the Internet.

Major products and services	Business model	Fee example
AITalk® WebAPI AITalk® Web-yomi Shokunin (Website Reading Expert)® AITalk® Koe-no-shokunin (Voice Craftsman)® Cloud Version	Recurring-revenue type	5,000 yen/month

◎**Support services**

The company provides clients of products for corporations with continuous technical support.

Major products and services	Business model	Fee example
Technical support	Recurring-revenue type	Annual contract

③**Products for consumers**

The company sells packaged software, with which you can easily produce audio files.

Major products and services	Business model	Fee example
Kantan (Easy)! AITalk® AITalk® Anata-no-koe (Your Voice)® VOICEROID® Series－Kotoha, Akane® and Aoi®	One-shot revenue type	The company outsources sales and sets royalties according to sales performance.

(5) R&D structure

As of March 31, 2020, the number of R&D staff members was 11. The total R&D cost for the term ended March 2020 was 115 million yen.

The three groups, "language processing", "voice processing" and "engine development", are working on improving Japanese language processing technology for speech synthesis, developing a new high-quality speech synthesis engine, and putting new algorithms for language and speech developed in the early stages of development to practical use, respectively.

【1-5 Characteristics, strengths, and competitive advantage】

AI, Inc., which developed AITalk®, a high-quality speech synthesis engine, and offers products and services, has the following characteristics, strengths, and competitive advantage.

(1) The required number of voice samples is small.

The general approach for improving speech synthesis quality in the "corpus-based text-to-speech synthesis" is to increase voice samples. However, it has a disadvantage; if voice samples increase, then recording time is prolonged and the size of a phonetic dictionary increases, augmenting the cost for producing the phonetic dictionary.

BRIDGE REPORT　　　　　　　　　　　Bridge Salon

AI, Inc. is proceeding with R&D, with the aim of synthesizing high-quality speeches with a small number of voice samples. In general, it is necessary to record voices for several tens of hours (several to ten thousand sentences), but the company can produce a phonetic dictionary with 2 to 6 hours of recording (200 to 600 sentences).

(2) Provision of a variety of speakers
Since a phonetic dictionary can be produced with a small number of voice samples, it is possible to offer a wide array of phonetic dictionaries. At present, the company offers a total of 15 speakers, including 7 female speakers, 4 male speakers, 2 boyish speakers, and 2 girlish speakers.

(3) Multiple introductions and sales results
The production of a phonetic dictionary used to cost tens of millions of yen, but the company developed a technology for producing it with a small number of voice samples at a cost of 0.5 to 5 million yen. As a result, it is now possible to inexpensively produce a phonetic dictionary desired by each user, including the voices of specific voice actors, narrators, and characters, and the scope of application of the speech synthesis engine has expanded.
Up until now, the company has produced over 300 custom voices.

Also, the company's technology has been exceptionally highly evaluated, with 1,200 companies using AI's products, 648 local governments using AI's products for anti-disaster wireless systems, 1,300 licenses for corporate package software sales, and more than 60,000 licenses for consumer package software sales.

(4) System for offering services of R&D, product development, sale, and support in an integrated manner
Most competitors that offer speech synthesis engines are large makers, in which R&D and product development/sale sections are separated.
Meanwhile, AI, Inc. deals with almost all processes including R&D, product development, sale, and support, by itself, so that it can operate business flexibly and swiftly. For the speech synthesis engines for foreign languages, it collaborates with overseas makers.

【1-6 ROE analysis】

	FY 3/16	FY 3/17	FY 3/18	FY 3/19	FY 3/20
ROE (%)	**15.0**	**15.4**	**17.8**	**16.7**	**16.0**
Net income margin [%]	15.09	17.03	18.51	20.38	21.07
Total asset turnover [times]	0.84	0.77	0.83	0.73	0.68
Leverage [times]	1.19	1.17	1.16	1.12	1.12

The net income margin for the term ending March 2021 is estimated to be 24%, and it is expected that ROE will remain high this term.

【1-7 ESG activities】
In the term ended March 2020, AI, Inc. carried out the following activities.

ESG	Theme	Outline
S: society	(1) Empowerment of women	・Among 40 employees, 21 (52.3%) are female ones. ・Among 11 managers, 3 (27.3%) are female ones.
	(2) Promotion of child-care support	・A child-care leave was taken by 3 employees.
	(3) Promotion of the reform of ways of working	・Working environment where the overtime work amount is small. Average overtime hours: 10.96 hours/month

		・Working environment where employees feel free to take a day off: Ratio of paid leaves taken: 85.7%
	(4) Promotion of social contribution activities	・Accepted the tour of students of　school for social studies 6 schools (82 students)
G: governance	(1)　Dialogue　with shareholders and investors	・A briefing session for individual investors held once (August). ・A briefing session for institutional investors held twice (May, November). ・A 1-on-1 meeting with an institutional investor held 43 times. ・Interviewed by magazine reporters and others 5 times. ・Appeared in a TV or radio program 2 times.

2. Fiscal Year March 2020 Earning Results

(1) Earnings Results

	FY 3/19	Ratio to net sales	FY 3/20	Ratio to net sales	YOY	Ratio to the estimates
Net sales	737	100.0%	819	100.0%	+11.1%	+2.4%
Gross profit	576	78.3%	665	81.3%	+15.4%	-
SG&A expenses	365	49.6%	392	47.9%	+7.2%	-
Operating income	211	28.6%	273	33.4%	+29.6%	+24.1%
Ordinary income	202	27.4%	273	33.4%	+35.1%	+24.1%
Net income	150	20.4%	172	21.1%	+14.9%	+7.5%

*Unit: Million yen.

Increase in sales and profits. The results exceeded the initial estimates.

The sales for the term ended March 2020 were 819 million yen, up 11.1% year on year. The sales of corporate products and services grew, while the sales of consumer products shrank.

Gross profit margin improved 3.0 points, and gross profit increased 15.4% year on year due to covering outsourcing expenses with the in-house system, etc.

Operating income rose 29.6% year on year to 273 million yen. SG&A expenses also augmented due to personnel increase, etc., but the increase in sales and gross profit offset this.

Ordinary income increased 35.1% year on year to 273 million yen because there were no longer temporary costs associated with the listing.

Net income was up 14.9% year on year to 172 million yen. The company thought that investment securities owned are losing excess profitability, and it posted a 45 million yen loss on the revaluation of investment securities due to impairment.

Both sales and profit exceeded the initial estimates. Considering the strong performance, the company increased the dividend forecast by 1 yen/share from 6 yen/share to 7 yen/share.

(2) Sales in each segment

	FY 3/19	Composition ratio	FY 3/20	Composition ratio	YOY	Ratio to the initial estimates
Products for corporations	423	57.5%	499	61.0%	+17.8%	+10.6%
Services for corporations	208	28.3%	229	28.0%	+9.9%	0.0%
Products for consumers	104	14.2%	90	11.0%	-13.5%	-25.0%
Total	737	100.0%	819	100.0%	+11.1%	+2.4%

*Unit: Million yen.

BRIDGE REPORT   Bridge Salon

Products for corporations

In the anti-disaster field, multilingual projects with translation functions increased, and orders from manufacturers rose. In addition, the sales of packages for producing narration, such as e-learning and video, were strong.

The use of speech synthesis in the broadcasting industry has expanded.

Services for corporations

"my daiz®" service of NTT Docomo contributed significantly. (my daiz® gives the most appropriate proposal to each user through the dialogue with the characters of my daiz and agents of each service.)

Products for consumers

The sales of the VOICEROID series were sluggish due to the delay in releasing new products.

(3) Financial Conditions and Cash Flow

◎Major BS

	End of March 2019	End of March 2020		End of March 2019	End of March 2020
Current assets	1,115	1,137	Current liabilities	105	138
Cash and deposits	970	964	Trade payables	3	13
Trade receivables	130	159	Other payables	35	55
Noncurrent assets	96	51	Noncurrent liabilities	2	2
Property, plant, and equipment	13	16	Total liabilities	108	141
Intangible assets	15	8	Net assets	1,103	1,047
Investments and other assets	67	26	Retained earnings	761	894
Total assets	1,211	1,189	Total liabilities and net assets	1,211	1,189
			Equity ratio	91.1%	94.6%

*Unit: Million yen

Equity ratio decrease by 3.0% from the end of previous term to 88.1%.

◎Cash flow

	FY3/19	FY3/20	Increase/decrease
Operating CF	135	217	+81
Investing CF	-59	6	+65
Free CF	76	223	+147
Financing CF	257	-229	-486
Cash and cash equivalents	970	964	-5

*Unit: Million yen.

In the term ended March 2020, there was no purchase of investment securities, and investing CF turned positive, and free CF increased.

The cash position was almost unchanged.

BRIDGE REPORT   Bridge Salon

3. Fiscal Year March 2021 Earnings Estimates

(1) Earnings Estimates

	FY 3/20	Ratio to net sales	FY 3/21 (Est.)	Ratio to net sales	YOY
Net sales	819	100.0%	840	100.0%	+2.5%
Operating income	273	33.4%	280	33.3%	+2.3%
Ordinary income	273	33.4%	280	33.3%	+2.5%
Net income	172	21.1%	205	24.4%	+18.8%

*Unit: Million yen. The estimates were announced by the company.

Sales and profit estimated to grow.

The sales for the term ending March 2021 are expected to increase 2.5% year on year to 840 million yen, and operating income to rise 2.3% year on year to 280 million yen. Although there are concerns about the impact of the spread of the novel coronavirus, sales and profits are expected to increase due to the expansion of the speech synthesis market. The dividend is to be 8 yen/share, which is 1 yen/share higher than the previous year's 7 yen /share. The expected payout ratio is 19.7%.
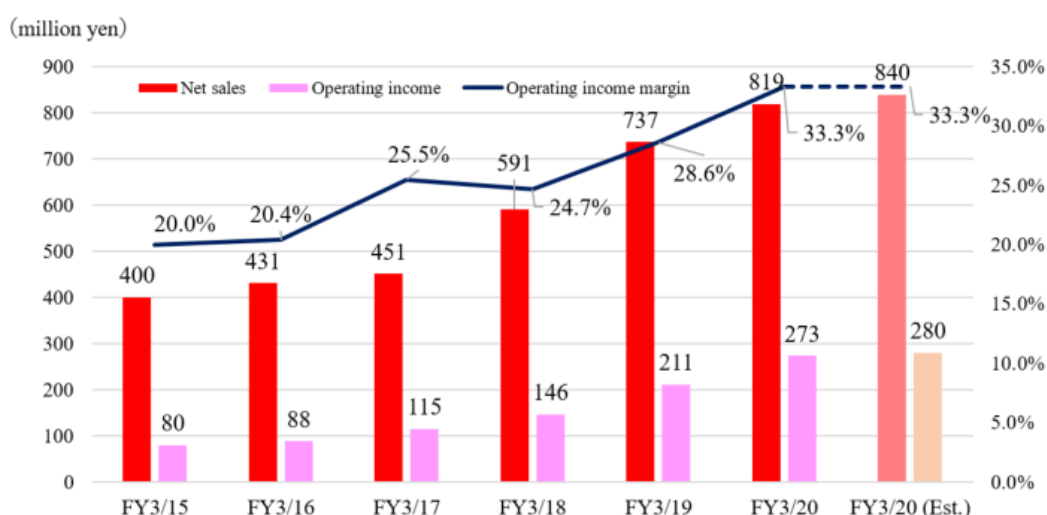
(2) Sales in each segment

	FY 3/20	Composition ratio	FY 3/21 (Est.)	Composition ratio	YOY
Products for corporations	499	61.0%	500	59.5%	+0.1%
Services for corporations	229	28.0%	240	28.6%	+4.6%
Products for consumers	90	11.0%	100	11.9%	+10.7%
Total	819	100.0%	840	100.0%	+2.5%

*Unit: Million yen.

(Products for corporations)

In addition to the market growth, inquiries regarding the use of e-learning teaching materials are increasing due to the expansion of telecommuting and online learning. On the other hand, sales are expected to remain at the same level as the previous term as there are concerns over the impact of the self-restraint request made by the Japanese government on sales activities due to the spread of the novel coronavirus.

(Services for corporations)
In addition to the contribution of NTT Docomo, Inc.'s "my daiz" service, the company expects that demand for low-priced "AITalk® Web API" and "AITalk® Koe-no-shokunin (Voice Craftsman) Cloud Version" will increase.

(Products for consumers)
The sales of VOICEROID and other products are expected to rise because of stay-at-home requests caused by the spread of the novel coronavirus.

(3) Major Initiatives

(1) Commercialization of the next-generation speech synthesis engine AITalk 5.0
AI, Inc. has commercialized the speech synthesis engine that utilizes deep learning (DNN: Deep Neural Network) and started providing "AITalk®5 Koe-no-shokunin (Voice Craftsman) ® Package Edition" and "AITalk®5 SDK" on May 7, 2020.

Using deep learning improves sound quality and actualizes more human-like and natural high-quality speech synthesis. It also reduces the voice dictionary creation costs by shortening recording time and voice dictionary creation time. Hence, it is possible to provide a speech synthesis engine at a lower cost.

(2) Promotion of work style reform
The company has been working to create a comfortable working environment for employees. In this term, it will further promote work style reforms through "adoption of a flextime system," "adoption of telecommuting," and "reviewing the personnel evaluation system."

(3) To build a foundation for further growth
In the speech synthesis market that has entered the "Phase 3," AI, Inc. will take the following measures to create a new market and further increase the growth rate.

(I) To accelerate collaboration (in the automotive field) with Cerence, Inc. (U.S.)
Cerence, Inc. has expertise in AI, natural language understanding, voiceprint recognition, gesture, gaze detection, augmented reality (AR), etc. It collaborates with major automobile manufacturers around the world as an innovation partner to provide unique solutions and promotes business development in connected cars, autonomous driving, electric vehicles, etc.

AI, Inc. has signed a license agreement in November 2019 to provide the basic technology of the high-quality Japanese speech synthesis engine "AITalk®" to Cerence, Inc. and will further accelerate this collaboration toward concrete commercialization.

(II) Development of the new market utilizing voice synthesis
AI, Inc. will work on concrete servitization and commercialization through alliances with various companies in fields such as the e-learning field.

(III) Promotion of joint research with Professor Toda of Nagoya University
The company will promote research and development of state-of-the-art deep learning-based speech synthesis technologies such as WaveNet (one of the deep neural networks for generating speech waveforms).

(IV) Utilization of M&A
The company will proactively conduct M&A to speed up commercialization and creation of new markets.

BRIDGE REPORT　　　　　　　　　　　　　　Bridge Salon

4. Conclusions

The company started providing the next-generation speech synthesis engine AITalk 5.0. AITalk 5.0 actualizes more human-like, natural, high-quality speech synthesis. Also, by using deep learning, it can offer a speech synthesis engine at a low cost by greatly reducing the recording time and voice dictionary creation time. Thus, AITalk 5.0 is expected to expand the usage scenes of voice synthesis rapidly. We would like to expect the release of concrete application cases in the future.

Moreover, although sales and profits are on the rise, we would like to pay attention to how the profitability of the company, whose top line needs to grow further as a growing company, will change with AITalk 5.0.

<Reference: Regarding Corporate Governance>

◎**Organization type and the composition of directors**

Organization type	Company with audit and supervisory committee
Directors	5 directors, including 3 outside ones

◎**Corporate Governance Report**

Last update date: June 27, 2019

<Basic policy>

Recognizing that for an enterprise to grow and develop stably, it is indispensable to enhance the efficiency and soundness of business administration and establish a fair, transparent management system, the company considers thoroughgoing corporate governance as the most important mission.

<Reasons for Non-compliance with the Principles of the Corporate Governance Code (Excerpts)>

Our company follows all the basic principles of the Corporate Governance Code.