

## The Evolution of Basic Research

*Eisaku Maeda*

### Abstract

Basic research gives rise to innovation through new discoveries and inventions, and this can lead to changes in the structure of industry and our lifestyles. However, it is also true that such success stories are rare, and basic research carries a high degree of risk. This article analyzes the historical evolution of technologies born from the activities of NTT Communication Science Laboratories and clarifies the future strategy for the promotion of basic research and some issues that should be addressed based on previous cases.

*Keywords: basic research, innovation, research and development*

### 1. Introduction

It is 23 years since the establishment of NTT Communication Science Laboratories (NTT CS Labs), and 14 years since the start of basic research under a new system after the reorganization of NTT. During this period, a definite framework was formed at NTT CS Labs, and technologies born out of our basic research have gradually made their way out into the world. I believe that we have now reached a major turning point on our way towards the next phase. This series of Feature Articles introduces the technical details and future prospects of seven research achievements presented at the NTT CS Labs Open House 2014 exhibition. I will start by discussing the nature and role of basic research in research and development (R&D), the way in which basic research leads to innovation, and our strategy for promoting innovation.

### 2. Invisible innovation in information technology

Innovations in information technology occupy quite a special position compared with innovations in mechanical technologies such as the steam engine or letterpress printing. Information specialists complain that even though people talk about innovations for certain products and services, and even though information technology is used in these products and services, the importance of information technology is

not being conveyed to people in general. That is, information technology cannot be *seen* in the innovation. Even if we use Google's search engine or an Apple iPhone every day, we cannot necessarily tell precisely what sort of information technology lies at the core of products and services such as these. So what sort of role does information technology play behind innovations such as the Nobel prize-winning work on iPS (induced pluripotent stem) cells or the *Hatsune Miku* singing software that is creating a stir in the music industry? Most people probably have no idea. This is because by the time new scientific discoveries have been incorporated into products and services that affect our everyday lives, they are no longer visible from the outside.

Furthermore, even though major technological innovations occur every few years in each field of information technology, these innovations are not immediately reflected in products or services. Even among people that work in information technology, these technical innovations are only noticed by specialists in related fields. These are the invisible innovations that I am referring to.

For example, in recent years, speech recognition technology has been put to use in a wide range of practical services, such as a system that records transcripts of debates in the House of Representatives. One of the things that made this system possible was the introduction of speech recognition using a weighted finite state transducer (WFST) in 2003, which enabled the implementation of speech

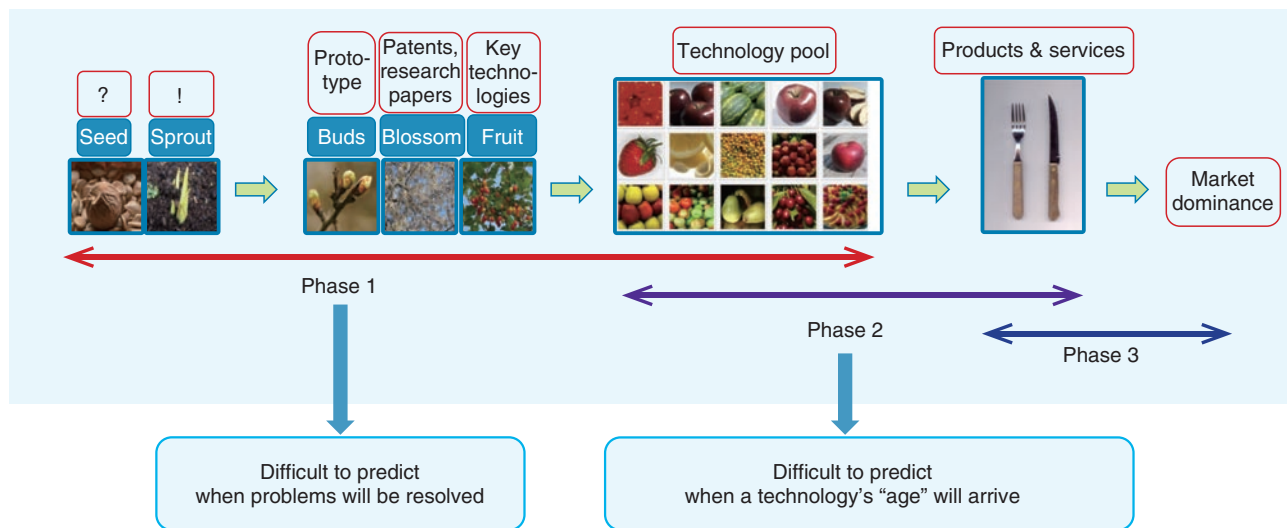


Fig. 1. The road from “basic research” to services.

recognition with a huge vocabulary of 2 million words [1]. It goes without saying that grasping this sort of hidden innovation as quickly as possible is the key to success in R&D, the development of new services, and the creation of new markets.

### 3. The cultivation and evolution of fruitful technologies

So when are hidden innovations produced in basic research, and how do we go about creating more of them? I have designed an illustration to show the workings of basic research (**Fig. 1**). The seeds of new research arrive as small flashes of inspiration, for example, the discovery of a new problem, or a new way of looking at a simple concept. If they are kept watered, some of these seeds will germinate and grow after a while, and given fertilizer and sunlight, they will blossom into research papers and patents. With a bit of luck, they will eventually bear fruit that is ripe for harvesting. Phase 1 of this process—from sowing the seeds to cultivating the plants and harvesting the fruit—forms the core of basic research and is the most important phase. Even plants that produce fine blossoms are sometimes of no value. There is also no guarantee that fine fruit will be produced by plants given plenty of water and fertilizer. While storing a range of different fruits, we obtain new seeds from these fruits, and once again, these are sown to promote the evolution of even more technologies.

The mission and value of basic research lies in tack-

ling difficult problems without worrying about the risks. It is therefore usually impossible even for experts to predict when these problems will be solved. This is the biggest issue in Phase 1. One of the recent cases at NTT CS Labs is the invention of *Buru-Navi3*, which is a device that uses human sensory characteristics to create a tugging sensation. After we invented *Buru-Navi1* in 2004, we thought it would be difficult to miniaturize this device while maintaining the same tugging effect [2]. The solution to this issue suddenly came in 2014, when we discovered that it is possible to make a device 20 times smaller than the *Buru-Navi1* without impairing its tugging effect. This new device is introduced in the article “*Buru-Navi3* Gives You a Feeling of Being Pulled” [3].

### 4. Fruit provides people with nourishment

In the same way that fruit provides nourishment only when eaten, the fruits of research have value only when they are used as technologies. The basic research achievements of NTT CS Labs are finding their way into technologies that are put to practical use in the real world. Some representative examples are listed in **Table 1**. By analyzing these examples, we can see that it can often take 10 years or more for seeds to grow to fruition, and that a lot of time is also needed to make this fruit available for consumption (Phase 2). Once a technology has been perfected, it will not be used unless it meets the needs of the current age. In most cases, it is difficult to predict when

Table 1. Examples of NTT CS Labs' achievements.

		Start of research	Representative results of basic research	Contributions to actual services
(1)	Robust media search [RMS]	1993	1995: Fast image search 1998: Fast time-series (audio/video) search 2004: Extremely robust media search with respect to severe noise and distortions 2013: Instance search	2004: Name-that-tune service via mobile phones 2007: Music copyright clearance for broadcasters 2008: Identification of known audio/video content embedded in media files in the Internet 2010: Second screen using smartphones
(2)	WFST-based speech recognition	2000	2004: Extremely large vocabulary (2 million words) recognition decoder 2010: Unified acoustic-language model learning	2011: Minute-taking system for House of Representatives
(3)	Reverberation control [REVTRINA]	2003	2005: Development of basic principles 2009: Surround technology based on reverberation control	2009: Pro reverberation control software 2012: Consumer surround speakers 2013: Pro surround equipment
(4)	Question answering	2001	2003: Japanese Q&A system SAIQA 2009: NAZEQA, capable of answering "Why?"	2012: Shabette Concier (talking concierge)
(5)	Statistical machine translation	2003	2009: Dependency parsing with semi-supervised learning 2010: Pre-ordering by head finalization	2011: Outperformed rule-based translation 2013: JAPIO starts Chinese patent translation service
(6)	Material perception information science	1998	2007: Proposal of a new theory of material perception (Published in <i>Nature</i> ) 2012: Finding of a novel object after-effect	2010: Inauguration of Scientific Research on Innovative Areas "Shitsukan (material perception)"
(7)	Pseudo-Attraction Force [Buru-Navi]	2003	2004: Development of pseudo-attraction force haptic compass (Buru-Navi1) 2014: Miniaturization (Buru-Navi3)	—

JAPIO: Japan Patent Information Organization

SAIQA: System for Advanced Question Answering

the *age* of a technology will arrive, which is the main issue in Phase 2. Until such time, a technology must be maintained and protected in a technology pool, and a system must be established for bringing the technology out into the world as soon as its time has come.

The following paragraphs describe some specific examples of technologies listed in the table.

(1) Robust media search (RMS) began in 1993 with research into image search techniques that can quickly find an image fragment in a larger image (as in the "Where's Wally?" picture books). This technology was developed to include music search and video search functions, and in around 2008 it came to play a major role in the identification of uploaded video content and the protection of music copyright in broadcasting [4]. Today, it is still evolving to allow searching for specific instances of a specific object in video images, as discussed in the article "Instance Search Technology for Finding Specific Objects in Movies" [5].

(2) WFST-based speech recognition is a technology that was given a major boost by the introduction of

WFST as mentioned above, but the introduction of a deep-learning technique has led to a new wave of development [1]. Also, the basic principles of (3) reverberation control technology (REVTRINA) were figured out about five years ago, and this technology is now being introduced into a wide variety of professional and consumer devices. This technology is introduced in the article "Enhancing Speech Quality and Music Experience with Reverberation Control Technology" [6].

(4) Question answering technology is a basic element of NTT DOCOMO's *Shabette Concier* service, and it grew out of SAIQA (System for Advanced Question Answering) that NTT CS Labs started researching back in 2001 [7, 8]. At the 2003 NTT CS Labs Open House exhibition, we connected it to a speech recognition system with a vocabulary of 2 million words to produce a speech-based question answering demonstration. However, it took 10 years for this technology to find its way into the real world. The fruits of this research in question answering technology led to further advances with the introduction of statistical machine learning methods into natural

language processing, which has progressed rapidly since the turn of the century. This introduction of statistical machine learning technology has also played a large role in the paradigm shift whereby (5) statistical machine translation is replacing conventional rule-based machine translation [9].

(6) Material perception information science came to the fore in 2010 as a new field of technical research promoted by the Ministry of Education, Culture, Sports, Science and Technology, and academic activity in this field is expanding rapidly. This came about through a joint study by NTT CS Labs and MIT (Massachusetts Institute of Technology) that started in 2000. Since the results of this study were published in *Nature* in 2007 [10], the research has been extended to include the other senses as well as sight [11]. More information can be found in the article “Recognizing Liquid from Image Motion and Image Deformation” [12]. This material perception information science and (7) *Buru-Navi* can be described as the fruits of research that are still waiting for their age to arrive.

A lot of work is also being done on technologies that are not yet complete. These technologies are introduced in the articles “Reading the Implicit Mind from the Body” [13], “Quantum Computing Beyond Integer Factorization—Exploring the Potential of Quantum Search” [14], and “Capturing Sound by Light: Towards Massive Channel Audio Sensing via LEDs and Video Cameras” [15].

## 5. Evolution of basic research

The pursuit of basic research will never be a special endeavor that takes place far away from the ordinary world. If we are to take on the challenges of the real world, we must be a part of it. John Pierce, the former executive director of Bell Laboratories, once said that, based on his own experience, “Ideas and plans are essential for innovation, but the time has to be right [16].” After leaving Bell, Pierce also became well known as one of the academic pioneers in computer music, but he is also known among speech recognition researchers as the person that pulled the plug on research on speech recognition at Bell Laboratories in the 1970s. In an article for the Acoustical Society of America, he wrote, “General-purpose speech recognition seems far away. Special-purpose speech recognition is severely limited. It would seem appropriate for people to ask themselves why they are working in the field and what they can expect to accomplish [17].” This had a large influence on

research in the area of speech recognition in the US. This example shows that even a highly experienced research manager can sometimes make bad judgments, and it highlights the difficulties of basic research administration. In fact, at that time, the speech research at Bell Laboratories had been assigned to visiting researchers from overseas, including Dr. Fumitada Itakura, whose work at the time led to results in the field of line spectrum pairs (LSP). In 2014, LSP was confirmed as a milestone by IEEE (Institute of Electrical and Electronics Engineers), and you can read more about this pioneering research in the article “LSP (Line Spectrum Pair): Essential Technology for High-compression Speech Coding” [18].

In the 21st century, the information environment that surrounds us in our daily lives is changing rapidly, and speed has become an essential requirement for entry into the market, corresponding to Phases 2 and 3 in Fig. 1 [19]. Even in basic research, the choice of issues to study changes over time, and researchers need to contribute to the commercialization of their results with a sense of speed that matches the speed of the age. While NTT aims to create new markets through co-innovation with other industries, the value of basic research and the expectations of the fruit of this research are likely to grow in the future. Every one of the results from basic research is a valuable seed of innovation and is just waiting for its age to arrive. It is therefore necessary to always be on the lookout for hidden technologies with the potential to lead to new innovation, wherever it may be, and to put them to use as fast as we can. They will serve us well in the competition in R&D and in the creation of new services.

## References

- [1] Y. Kubo, A. Ogawa, T. Hori, and A. Nakamura, “Speech Recognition Based on Unified Model of Acoustic and Language Aspects of Speech,” NTT Technical Review, Vol. 11, No. 12, 2013. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201312fa4.html>
- [2] T. Amemiya, H. Ando, and H. Ho, “Nonverbal Communication via a Five-senses Interface.” NTT Technical Journal, Vol. 19, No. 6, pp. 35–37, 2007 (in Japanese).
- [3] T. Amemiya, S. Takamuku, S. Ito, and H. Gomi, “*Buru-Navi3* Gives You a Feeling of Being Pulled,” NTT Technical Review, Vol. 12, No. 11, 2014. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa4.html>
- [4] K. Kashino, R. Mukai, K. Otsuka, H. Nagano, T. Izumitani, A. Kimura, T. Kurozumi, and J. Yamato, “Fast Media Search,” NTT Technical Journal, Vol. 19, No. 6, pp. 29–32, 2007 (in Japanese).
- [5] M. Murata, H. Nagano, R. Mukai, K. Hiramatsu, and K. Kashino, “Instance Search Technology for Finding Specific Objects in Movies,” NTT Technical Review, Vol. 12, No. 11, 2014.

- <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa2.html>
- [6] K. Kinoshita, “Enhancing Speech Quality and Music Experience with Reverberation Control Technology,” NTT Technical Review, Vol. 12, No. 11, 2014.  
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa3.html>
- [7] E. Maeda, H. Isozaki, Y. Sasaki, H. Kazawa, T. Hirao, and J. Suzuki, “Question Answering System: SAIQA—A “Learned Computer” that answers any questions,” NTT R&D, Vol. 52, No. 2, pp. 122–133, 2003 (in Japanese).
- [8] R. Higashinaka, K. Sadamitsu, K. Saito, and N. Kobayashi, “Question Answering Technology for Pinpointing Answers to a Wide Range of Questions,” NTT Technical Review, Vol. 11, No. 7, 2013.  
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201307fa4.html>
- [9] M. Nagata, K. Sudoh, J. Suzuki, Y. Akiba, T. Hirao, and H. Tsukada, “Recent Innovations in NTT’s Statistical Machine Translation,” NTT Technical Review, Vol. 11, No. 12, 2013.  
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201312fa2.html>
- [10] I. Motoyoshi, S. Nishida, L. Sharan, and E. Adelson, “Image Statistics and the Perception of Surface Qualities,” Nature, Vol. 447, No. 7141, pp. 206–209, 2007.
- [11] J. Watanabe, “Communication Research Focused on Tactile Quality and Reality,” NTT Technical Review, Vol. 9, No. 11, 2011.  
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201111fa6.html>
- [12] T. Kawabe, M. Sawayama, K. Maruya, and S. Nishida, “Recognizing Liquid from Image Motion and Image Deformation,” NTT Technical Review, Vol. 12, No. 11, 2014.  
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa5.html>
- [13] M. Kashino, M. Yoneya, H.-I. Liao, and S. Furukawa, “Reading the Implicit Mind from the Body” NTT Technical Review, Vol. 12, No. 11, 2014.  
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa6.html>
- [14] S. Tani, “Quantum Computing Beyond Integer Factorization—Exploring the Potential of Quantum Search,” NTT Technical Review, Vol. 12, No. 11, 2014.  
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa7.html>
- [15] G. P. Nava, Y. Shiraki, H. D. Nguyen, Y. Kamamoto, T. G. Sato, N. Harada, and T. Moriya, “Capturing Sound by Light: Towards Massive Channel Audio Sensing via LEDs and Video Cameras,” NTT Technical Review, Vol. 12, No. 11, 2014.  
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411fa8.html>
- [16] J. Gertner, “The Idea Factory: Bell Labs and the Great Age of American Innovation,” Penguin Press, March 2012.
- [17] J. R. Pierce, “Whither Speech Recognition?,” The Journal of the Acoustical Society of America, Vol. 46, No. 4B, pp. 1049–1051, 1969.
- [18] T. Moriya, “LSP (Line Spectrum Pair): Essential Technology for High-compression Speech Coding,” NTT Technical Review, Vol. 12, No. 11, 2014.  
<https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201411in1.html>
- [19] F. Vogelstein, “Dogfight: How Apple and Google Went to War and Started a Revolution,” Sarah Crichton Books, November 2013.



#### Eisaku Maeda

Director, NTT Communication Science Laboratories.

He received the B.E. and M.E. in biological science and the Ph.D. in mathematical engineering from the University of Tokyo in 1984, 1986, and 1993, respectively. He joined NTT in 1986. He was a guest researcher at the University of Cambridge, UK, in 1996–1997. His research interests are statistical machine learning, intelligence integration, and bioinformatics. He is a senior member of IEEE and a fellow of IEICE (Institute of Electronics, Information and Communication Engineers of Japan), and a member of IPSJ (Information Processing Society of Japan).