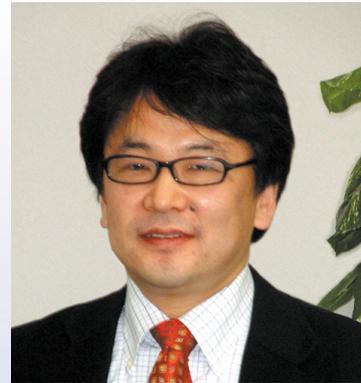


R&D Spirits

Toward Flexible Media Distribution in a Ubiquitous Environment

Masahito Kawamori
 Senior Research Engineer
 Promotion Project 1
 NTT Cyber Solutions Laboratories

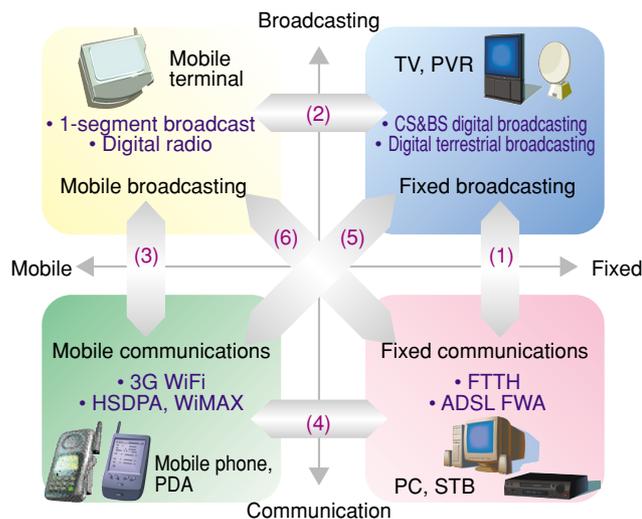


Researchers at NTT Cyber Solutions Laboratories are working on the development of an advanced media distribution system based on interfaces between broadcasting and communications and between mobile and fixed terminals. In this research, metadata is attracting much attention as a key technology for achieving a practical media distribution system. To find out how metadata might change the content business and what problems must be solved to incorporate it, we talked to senior research engineer Masahito Kawamori, one of the key members of the metadata development project and a contributor to international standardization activities in this field.

Development of metadata technology as the key to convergence of broadcasting and communications

—Dr. Kawamori, could you outline for us the R&D work that you are currently engaged in?

We are researching interfaces between different media environments such as mobile and fixed broadcasting and mobile and fixed communications (Fig. 1) with the aim of developing a metadata distribution system (Fig. 2). In the past, the goal was to get broad-



- (1) Server-based broadcast service
 - Viewing of missed programs
 - Online shopping
 - Internet-based electronic program guide
 - Program retrieval
 - TV portal
- (2) Expansion of broadcast viewing area
 - Continuous viewing (indoors to outdoors)
 - Continuous viewing (outdoors to indoors)
- (3)–(5) Mobile server type broadcasting
 - Digest viewing
 - TV remote control i-appli
 - TV-linked i-appli
 - Recommendations for individuals
 - Music downloading
- (6) Internet radio
 - Area expansion
 - Music downloading

PVR: personal video recorder
 CS: communication satellite
 BS: broadcast satellite
 3G: third-generation
 HSDPA: high-speed downlink packet access
 WiMAX: worldwide interoperability for microwave access
 PDA: personal digital assistant
 FTTH: fiber to the home
 ADSL: asymmetric digital subscriber line
 FWA: fixed wireless access
 PC: personal computer
 STB: set-top box

Fig. 1. Interfaces between media.

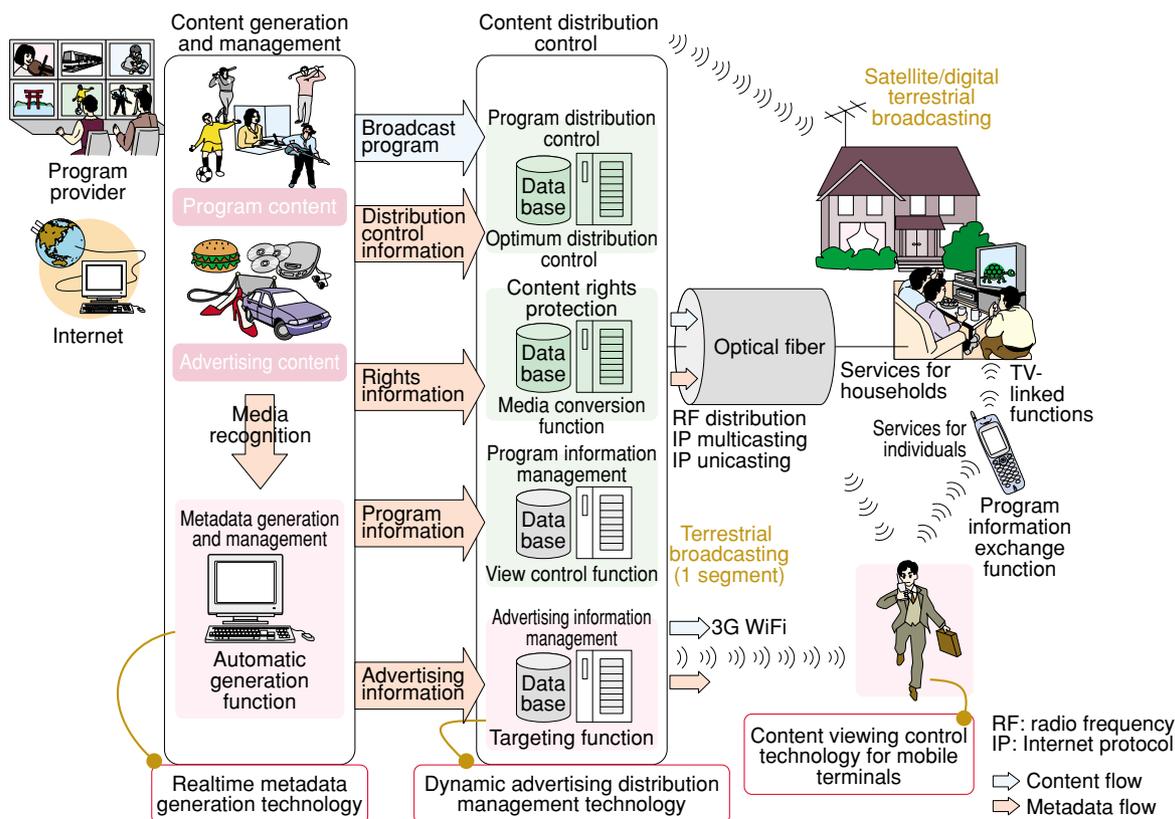


Fig. 2. Media distribution system.

casting and communications to coexist on specific service platforms, but now the idea is to achieve seamless media distribution in a ubiquitous environment based on the convergence of mobile and fixed terminals. To this end, we are constructing end-to-end systems by collecting elemental technologies possessed by NTT Laboratories and improving ones that we think are insufficient for our current needs.

The theme that I am spending the most time on is metadata, which can be regarded as an index to content. For example, metadata for a broadcast program might include delivery control information such as the delivery format and broadcast schedule, content rights information including conditions of use, program information such as program genre and a detailed program summary, and advertising information related to sales promotions (Fig. 3). As simply an index, metadata might not sound like an interesting research theme of value, but in reality, it is an extremely important technology that can provide the key to smooth and flexible media distribution. Without metadata, a user would not be able to find the part of a very large program that he or she wants to watch. Metadata also plays the role of a “common language” that is essential for the interfacing of different media

environments. In short, metadata not only serves as a guide to services for users, but also allows smooth interfacing between different media formats. In addition to working on the generation, retrieval, storage, and display of metadata, I am also involved in issues related to its standardization and systemization.

—What are some key technical issues associated with metadata?

One is how to go about generating metadata because the process of obtaining metadata requires that human ideas and matters be recognized and assimilated by mechanical means. Although this might be a simple task for the human mind, it turns out to be unexpectedly difficult to automate. This process must also be made efficient to be practical. To this end, we have exploited various elemental technologies developed by NTT, such as speech and image recognition, text analysis, natural language processing, and the use of context, and have developed easy-to-apply and efficient metadata generation technologies.

Another issue concerns the storage and transmission of metadata. For example, should metadata,

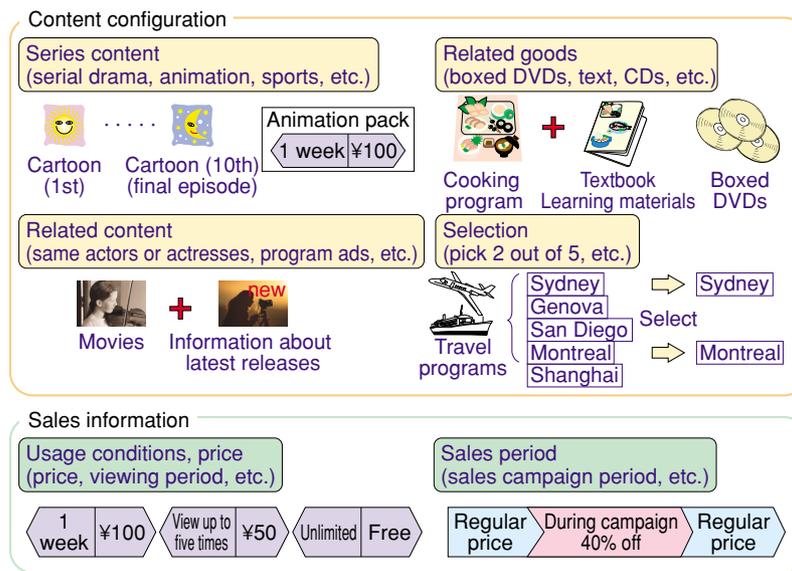


Fig. 3. Examples of representing metadata.

which can accumulate quickly into a huge amount of data, be compressed? And if so, in what way? And for that matter, how should metadata be transmitted to various kinds of environments? If we consider the sending of metadata to cellular phones, there's no doubt that metadata must be transmitted in compressed form given the relatively small data capacity of handsets. However, when we compress metadata, we lose the ability to search it, just as with ordinary compressed data files. For this reason, we have independently developed a technique for searching compressed metadata.

—How is your R&D progressing?

We have already finished making a practical, working system and its use for services provided by NTT Group companies is close to being approved. I cannot give you detailed information about its deployment here, but there should be an announcement soon. We have great expectations for this system.

—What problems, if any, are you dealing with at this time?

From a technical perspective, one problem is finding an appropriate method of protecting information. It is extremely important to protect copyrights and personal information related to content in either broadcasting or communications. At the same time, a higher level of security means less freedom in providing services, so one issue that must be addressed is how best to achieve a balance here. Another problem

concerns language. We have already received questions and inquiries regarding the generation of metadata from various countries, but our system currently supports only Japanese. This is a big problem, but we should be able to deal with it given a little more time.

There is yet another problem that relates to the social system. To get people to actually use the system, we must do more than just solve the technical issues. We must also take into account social concerns. The most difficult of these relates to business models. For example, the explosive spread of cellular phones has been accompanied by a parallel decrease in the number of public telephones as well as the migration of various types of content business to cellular-phone-compatible sites. In short, advances in technology have often meant drastic changes in business models. In such an environment, we must strive to raise awareness of this issue throughout society and to help establish new infrastructures including those related to laws and regulations. Although this is not directly related to R&D, it might be our biggest problem.

—What is the future outlook for this system?

When researching and developing something that has never been used before, you cannot say that you are done once the results of your research have made it to the outside world. Many things that you never considered may surface once a system that you have developed enters actual use. For example, problems associated with scalability will no doubt occur as the amount of metadata in circulation increases. Support

for dealing with problems unforeseen at the research stage will therefore become increasingly important in the future. The demand for more efficient systems will no doubt increase as well, and the shortcomings of current standards may also become an issue. In the face of these problems, it is vital that we continue with our R&D efforts with an eye toward supporting new needs and environments.

Promoting metadata technology to the world: contributing to Japanese and European standardization efforts

—Please tell us about R&D and standardization trends in the field of metadata.

Well, the field of metadata is quite broad, but if we limit ourselves to broadcasting and communications, I would say that not much research is going on at this time. One exception is the TV-*Anytime* Forum [1], which is an international standardization organization involved with broadcasting and communications. NTT is a member along with BBC R&D, France Telecom, Nielsen Media, Microsoft, and others. The standards that have been drafted there have been adopted directly as European standards, which means that the Forum is especially influential in Europe though it was originally a global organization. That being said, I should point out that I am currently chairing the metadata working group at the TV-*Anytime* Forum. Also, in Japan, the Association of Radio Industries and Businesses (ARIB) is charged with specifying broadcasting standards in general including those of digital terrestrial broadcasting and BS and CS (broadcasting satellite and communication satellite) broadcasts. At ARIB, metadata issues are handled in the Metadata Task Group (TG) of the Home-server-based Broadcasting Systems Industry Group. I am also the leader of that TG. Against this background, Japan and Europe have come to adopt similar standards. The United States, on the other hand, tends to attach greater importance to *de facto* standards and appears to have little concern about standardization at present. But this trend may change soon.

I would therefore say that the research that our group is working on is in no way inferior to broadcasting- and communications-related metadata research in Europe and the United States and elsewhere in Japan. Indeed, given that we can quickly create tools using the overall power of NTT Laboratories and that we are actively participating in standardization activities, I would say that we are in a superior position at this time.

—How have your research achievements in metadata been received in Japan and overseas?

Well, research results that I took to international conventions, meetings, and conferences over the last few years attracted much interest. I especially remember the positive response I received when presenting a metadata demonstration and being asked the price of the system from people that wished to purchase it! This was not an isolated event—I actually received similar queries from three companies headquartered in Europe and South America. One German company, which had a client with pressing needs, wanted to buy the whole system. The system they wanted was not even for broadcast use, but for public use such as for retrieving content from a public library. But since the demo system that I took along had everything including terminals, it was easy to see how it could be adapted for product-related applications. Since researchers do not become directly involved with the business aspects of research results, we are sometimes uncomfortable with such questions, but such inquiries indicate that our research is right on track, which is a great encouragement for researchers.

—Is your group collaborating with any companies or research institutions overseas?

To begin with, we are participating in standardization activities in step with other carriers and broadcasters such as France Telecom and the BBC. We have also worked with a French venture-capital company in the joint development of a media distribution system having common specifications for Japan and Europe. At present, that company's products are being used as metadata solutions at leading European communication companies. We are also in talks with various entities about future collaboration.

From a talking computer to a system that can handle meaning

—What is the foundation of your technical knowledge?

My main interest in school was language processing, and I was involved in research on artificial intelligence with the aim of creating an intelligent computer that could converse with people, just like HAL in the movie *2001: A Space Odyssey*. Specifically, I took up the problem of how people express meaning when carrying on a conversation and how language processing could be applied to teach that to computers. It might seem that this has nothing to do with

metadata research, but actually, that is not the case. Metadata is not restricted to broadcasting, and it is gradually being applied to the Internet. In fact, there is considerable research going on right now concerning the “Semantic Web,” that is, a Web that can handle meaning. Of course, the target of my research while at school was not the Internet—it was just the computer in front of me. But it was conceptually similar to this work on the intelligent Web.

—*Why did you choose NTT to pursue your research?*

Because NTT Laboratories has long been known for its work in language processing. The group that I first entered after joining the company had been researching language processing from early on, and at that time, there was much activity in research on interactive systems. This kind of research site was quite rare even in Japan, and I was eager to get in.

—*What research themes have you been involved with to date?*

In 1989, on entering the Information Science Laboratory of NTT Basic Research Laboratories, I worked on the development of agent systems based on language understanding, knowledge representation, speech recognition, and image processing and on the conversion of interactive speech data into a digital corpus. At that time, however, the advanced nature and significance of our research was not fully appreciated. For example, we noticed that indicating agreement or understanding during a conversation by grunts or words—which we call *aizuchi* in Japanese—plays an important role in interaction, and we constructed a system that included such a function. We thought this was a landmark discovery in the research of language understanding systems, and for this reason, the members of the group worked very hard to complete the system. Unfortunately, the response that we received from those around us was a bit lukewarm. Perhaps this is because *aizuchi* is something that people do unconsciously when conversing, and seeing a machine do it in a system designed to talk with people is consequently not that surprising and may even be expected. Furthermore, considering that bipedal walking robots have only now appeared after all these years, perhaps it was simply too early for a computer that could converse with people to be accepted by society. As a result, though the research was quite enjoyable in itself, this experience taught me that shifting to a more realistic research theme like metadata might be more worthwhile.

—*What gave you the opportunity to research metadata in earnest?*

I think that would be the fact that metadata research was initially given a low priority. Until recently, there was almost no R&D work on metadata. People thought that metadata did not represent any problems from a technical viewpoint. Specifically, they believed that the process of extracting meaning from data and creating an index—a simple task for human beings—could be automated without difficulty. In reality, however, it turned out to be more difficult than expected once metadata became a true need and an R&D theme. In fact, the difficulty of collecting and combining various elemental technologies to generate metadata turned out to be far greater than any difficulties associated with individual technologies. Knowing this, we decided to make good use of NTT’s wealth of diverse elemental technologies to give us a head start in dealing with this problem. Metadata can be considered to be a very difficult research theme in terms of achieving desired results.

—*What has been your goal through R&D?*

In short, my goal has been to change the way we live for the better. Wishing that our lives could be more enjoyable, happier, and more fulfilling, I would consider a variety of possibilities and then set out to create a system toward that end. I therefore expected that my ideas would always be up for scrutiny by the outside world through my R&D efforts. It was consequently a bit disappointing when research results that I achieved by this approach were met by no response at all—neither good nor bad. But in the last few years, I have been receiving more and more positive reactions to my work, and I feel that I am pursuing R&D that is useful to society.

The need for researchers and developers to think “outside the box”

—*Dr. Kawamori, how would you like to see your research develop in the coming years?*

Right now, my main objective is to see our system become used as a metadata infrastructure in society. I would also be glad if our system could help change the world for the better as I just touched upon. And, as an extension of my research on language processing and as a way to continue my current research from a slightly different point of view, I would like to study media and communication in a social network.

—*What do you see as your future direction?*

Well, I can't say that I have given it much thought. But I like the thinking process involved in R&D, so as a researcher, I'm quite satisfied with my current situation. At the same time, you cannot directly submit your results to the world from a research laboratory, so giving the business side of R&D a try might be one way for me to apply my ideas to society, and finding good partners for that purpose would probably be the way to go.

—*How do you like working at NTT Laboratories?*

NTT Laboratories is conducting much research of global value. Unfortunately, there are few research institutions of this type left in the world, and it is exactly for this reason that NTT Laboratories and its researchers should be all the more dedicated to its work. In particular, we must focus on fields that have yet to be fully developed and invest more resources in research dealing with ways of combining individual technologies. We should study the past as a way of making new discoveries. This is because NTT research is always ahead of the times. As I mentioned earlier, language processing at NTT was first researched more than 20 years ago, but it is only in recent years that the fruits of this research have seen the light of day. In other words, it is vital that we as researchers refrain from throwing out a research theme simply because it is old. Instead, we should continue with that research in some way while always checking whether it can now be useful from a new perspective. This, I believe, is the power of tradition and one reason why NTT Laboratories stands out.

—*What would you say to young researchers?*

I would like them to look beyond the 'paradigm' that they are presently working in. I say this, because, while I think that people tend to strive for excellence within the current paradigm, excellence in the old paradigm is not necessarily excellent in the new paradigm. It is more important to create something that will change our frame of mind than just work within the same old paradigm. We have to remember we are living in a truly amazing era as the Internet creates a paradigm shift, and the printing-press society of the last 300 years is coming to an end, which means that the concept of research itself may change, just as Gutenberg's invention changed the way people perceived the world. I would like to see young researchers take up their work with the spirit that they themselves can create a new paradigm for research.

That is, I would like to see them think "outside the box."

Reference

- [1] M. Kawamori, "Recent Activities of the TV-Anytime Forum," NTT Technical Review, Vol. 4, No. 3, pp. 77-80, 2006 (this issue).

Interviewee profile

■ Career highlights

After joining the Information Science Laboratory, NTT Basic Research Laboratories in 1989, he did research mainly on artificial intelligence, semantic databases, and natural language understanding. He was reassigned to the Communications Science Laboratories in 1994 and there researched dialogue understanding and interactive agents, utilizing such technologies as speech recognition, natural language processing, and digital graphic processing. He developed a conversation system that conversed with humans with "natural" features, such as turn-taking, nodding, hesitation, and politeness. During this research he got involved with the taxonomy and ontology of dialogue tasks and also with the international effort for standardizing the tagging system of digitally collected conversation data as a member of the Discourse Resource Initiative Group and Special Interest Group in Dialogue of the Association of Computational Linguistics.

His current main interests are in the convergence between telecommunications and broadcasting within the context of broadband and mobile environments.

He has been involved in the development of a PVR-based broadcasting system based on the TV-Anytime/DVB specifications. He is currently working on IPTV and mobile service systems.

He is also active in standardization. He heads the Metadata Task Group of ARIB and is a committee member of the Ubiquitous Information Sharing Working Group of Information Technology Research and Standardization Center, Japan Standardization Agency, where he is doing research on digital rights management. He has also been active in the TV-Anytime Forum as a convener for the Phase 2 metadata specification.

He has co-authored books, including "Digital Broadcasting Handbook" (2003), Ohmsha and "Digital Broadcasting Textbook" ("2004"), Impress.