NII Today

National Institute of Informatics News

Special Feature:
Get Exited by "NetCommons"

The Key to Creating an IT Infrastructure for Schools
Using Information Enjoyably and Efficiently
CommonsNet: Promoting NetCommons

Speech Corpora Promote Research and Development

Noriko Arai

Professor
Information and Society Research Division, NII

The Key to Creating an IT Infrastructure for Schools

An Open Platform System that Anyone Can Use

Itoya: The sites created with NetCommons incorporate a wide variety of functions: blogs, bulletin boards, databases and so on. What exactly is NetCommons?

Arai: NetCommons is an open source information sharing platform developed by NII. Currently it is widely used to build websites at schools and other public institutions. It allows you to freely combine different modules with one another. If you just want to create a simple website, even a beginner can do that in fifteen minutes.

Itoya: Is NetCommons used primarily by schools?

Arai: At present, yes. Although many schools have websites, from the outset a website was designed to publish research and the like on the Internet. In other words, it was a one-way architecture that wasn't designed for bidirectional exchanges with users. Also, website creation and system and server maintenance were a major burden.

Itoya: How did a professor whose specialty was mathematical logic become involved in NetCommons research?

Arai: Originally, I operated a BBS on a members-only Internet site whose users were primarily working mothers. I would post mathematical problems and the members would work out the answers to these problems. It was very meaningful, but I was managing it and answering the questions from the participants all by myself, and it was way too much. (laugh) I started to think that there must be a way to create a community by having people who want to learn and people who want to teach work together.

Itoya: Something more like a social networking service (SNS) than a BBS.

Arai: More than that. It could also combine the functions of a portal site and groupware. I wanted to integrate all of these aspects and create something that could be used as a one-stop resource.

Itoya: That's what led you to create e-classroom, the precursor to NetCommons?

Arai: That's right. e-classroom was a site devoted to learning. In each course, students studied by discussing problems submitted by teachers. But suppose it became necessary to provide a "faculty room" in which teachers could discuss the problems to be submitted? In NetCommons, you could do this by creating a faculty room on the "Group" level.

Itoya: Google and other search engines provide various features tailored to individual users, such as calendars, document management and so on. How is NetCommons different?

Arai: They may be similar in terms of user-friendliness. However, they're different in terms of who manages the data. In services from Google and so on, the services are stored on the databases of the company providing the service. In NetCommons, your own organization manages the data.

Itoya: When this system is introduced at a school, it means that the system is handling the personal information of many children, so I assume security measures are very important.

Arai: Enormous costs would be required to try to completely prevent risk from such

things as viruses, spam and server attacks. Moreover, only a few major corporations manage their own servers, and it would be difficult for a school or other organization to do so. At the same time, there would still be concerns if the management of the school registry, grades and other important personal information were left completely up to an outside organization. You want to take the lead in managing this kind of information yourself.

Itoya: There would be great advantages in being able to safely and easily manage and administer that kind of information yourself.

Arai: With NetCommons, you can do all of your writing and editing on the web, and this reduces the danger of files going astray. We also plan to keep working to improve our security structure.

Itoya: How about Internet-specific problems such as "Net-bullying?"

Arai: We have to be very careful about those. First, it's very important for principals to make sure that teachers, parents and students all understand the special properties of the Internet and the points to be careful about—namely, that once a file goes somewhere it's not supposed to go, there's no way to call it back. NetCommons is designed on the premise that it will be widely used at schools, and it's made so that, for the most part, people can't do whatever they like someplace out of sight of the administrators.

Itoya: In what other ways is NetCommons used in education?

Arai: There was one example in which Net-Commons was used as part of an elementary school class. The assignment was to write a one-page essay telling people about something that you do well and then post it on the Internet. It was designed to help students learn what they needed to do to effectively communicate their thoughts to others without saying either too much or too little. The student evaluations had comments like "I now have confidence in my ability to communi-

---

A Free, One-Stop System

---

cate information well" and "In order to communicate information, it's important to think of how it will by received by the other person."

Itoya: It sound like they really learned the essence of how to communicate.

Arai: In society from now on, more and more we will not be communicating our thoughts directly through speech but making judgments based on thoughts put down in written form. There is a lot of de-

bate about the plusses and minuses of the impact of the Internet on children. But I think learning how to master the ability to express yourself from the time that you're in elementary and junior high school will be a major help in enabling people to survive in the information age.

Itoya: You have also called for logical education through mathematics, and you clearly feel that NetCommons will become an important educational tool. What are your goals for the future?

Arai: People are the source of information. I'd like to increase the usefulness and independence of NetCommons and free people from their dependence on servers and terminals. Up to now, it's taken a great deal of knowledge and energy to transmit information and operate a website on the Internet. And it also required a lot of money. I'd like for NetCommons to remove these hurdles and allow anyone to safely communicate information on the Internet on an equal footing. I'd like to make it a type of social infrastructure—like the water supply.

✎ A Word from the Interviewer

Professor Arai has a clear research and development objective for NetCommons: to create a system that anyone can use as easily as they can use the water supply. From now on, educational approaches that help students master the ability to express themselves through the Internet and survive in the information age will be more and more necessary. I have high hopes for NetCommons which was designed to support this ambitious endeavor.

Get Excited by "NetCommons"

Using Information Enjoyably and Efficiently

On August 8, 2007, around 330 IT administrators from junior and senior high schools, educational centers, small and medium-sized companies and other entities throughout Japan gathered at the National Institute of Informatics (NII). They had come to participate in the CMS NetCommons user conference, which was provided free of charge by NII. Six years after development, how is NetCommons being used in these various settings? In this article, we report on the exchanges of information that occurred at the user conference.

In her congratulatory remarks on the holding of the third NetCommons User Conference, Professor Noriko Arai commented, "Even just counting the ones that I know about, more than 1000 educational institutions have introduced NetCommons."

NetCommons comprises a content management system (CMS) provided free of charge by NII. With this CMS, as long as you have mastered the IT skills sufficient to allow you to read and write e-mails, you can easily create a well-designed information portal site. In other words, NetCommons allows users to quickly create websites and the like.

School personnel were quick to recognize the advantages of such a system, and they currently account for the majority of users. Moreover, the development and use of the system has been designated a consignment project whose aim is the "introduction of IT technologies to conduct school affairs" by the Ministry of Education, Culture, Sports, Science and Technology.

However, Professor Arai also commented that what is important in modern-day society is to find ways to use information skillfully. "NetCommons was designed for the purpose of allowing users to edit information until it is in an easy-to-use form and enable it to be shared among the members of the group that needs that information." She hopes NetCommons will not be limited to school personnel but will be used by many people in various ways. For this reason, Professor Arai devotes her energies to searching for the ways in which NetCommons is being used. From time to time, she conducts Internet searches to see if NetCommons is being used in a new way; if she finds one, she talks to the persons involved. The purpose of the user conference is to introduce ways of using NetCommons gleaned through such field research in order to ensure that as many users as possible know about them, and to allow users to exchange information with one another. At the plenary session that began the conference, four intriguing examples discovered by Professor Arai were introduced.

The Process of Building Trust in Schools Begins with the Providing of Information

One of the main ways in which IT technologies are introduced to schools is through the use of a website to provide information. Takefunishi Elementary School in the city of Echizen began using NetCommons to build websites 15 months ago. The website now gets 300 or more hits each day. At a small school like Takefunishi Elementary School, which has 435 students, the reason that the school can maintain this number of hits is that the website is upda-

Teacher Yuki Mizushima
Takefunishi Elementary
School in Echizen City

Website translated into Portuguese for the benefit of Brazilian children.

Emcees Shigemichi Watanabe (Head Supervisor, Administration Office, Teacher Training Division, Educational Development Department, Chiba Prefecture Board of Education) and Professor Noriko Arai.

The arrival of CMS for the Web2.0 era

NetCommons 2.0 was unveiled at the end of the user conference.

ted without fail every school day. And it is NetCommons that makes this possible. NetCommons makes the process of updating the website extremely easy. "What keeps us going is the pressure of knowing that a lot of people are looking," laughs Yuki Mizushima, a teacher at the school. However, out impression was that they update the website not out of obligation but because they have things that they want to communicate.

Naturally, the posted content includes both text and photographs. When there are public readings or choral group recitals, the capability of the Internet to transmit audio and video is also utilized. Such services are very popular with families, since they allow the families to get a clear idea of their children's school experience. However, the school does not always hold major events. When there have been no such special events, the website features content that might even seem trivial—for example, the fact that the "furikake" (a seasoning powder that is sprinkled over rice) used in the school lunches was particularly tasty. Nevertheless, such content is a great way to promote dialogue between children and parents.

Schools also need to communicate important information—for example, whether or not there will be school the next day when a typhoon is approaching. Such important communications are sent to each family in the form of "Urgent E-mails." Other communications include notices of PTA meetings, changes to timetables, notifications of the time that children will get out of school and so on. All of these communications can also be viewed on the website. This facilitates communication with parents and the community and helps to create a safe and secure environment for children. It has also had a visible positive impact in terms of reducing the number of objects left behind by students and so on. Takefunishi Elementary School also has 50 non-Japanese students, and a link to a translation site has been provided to enable these children and their families to easily translate the website into other languages.

The result is a website that is both enjoyable and meaningful. It's no wonder that the website has more than 300 hits each day.

IT Technologies Bring Change to Class Instruction

The benefits of NetCommons are not always clearly evident to the general public, as in the case of a website. Makoto Kato, a teacher at Hokkaido Sapporo Tohoku High School, provides an example showing how NetCommons is being used in the classroom. In the previous school year, Tohoku High School began participating in the Academic Achievement Assessment Project sponsored by the National Institute of Educational Research (NIER). For this project, it was necessary to efficiently determine the academic ability of students.

At Tohoku High School, teachers prepare a teaching portfolio that collects and organizes data relating to class instruction and its effectiveness in terms of student learning. NetCommons is then used to digitize this teaching portfolio in order to centralize the tasks of lesson design, instruction, evaluation and validation to make the process more efficient.

This was not the only beneficial impact of NetCommons. Mr. Kato wanted to make his students able to think for themselves. The basic ability needed to ac-
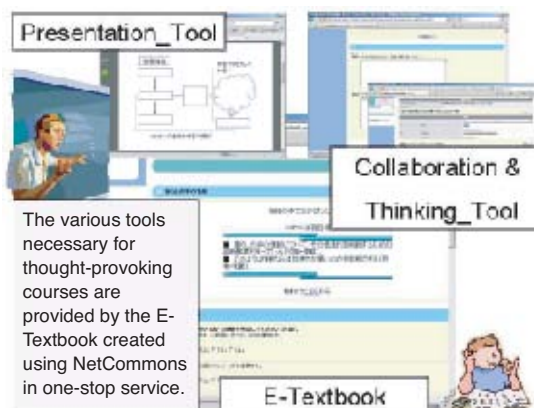
The various tools necessary for thought-provoking courses are provided by the E-Textbook created using NetCommons in one-stop service.

NetCommons is being used appropriately at various teaching levels.

Teacher Makoto Kato Hokkaido Sapporo Tohoku High School

Get Excited by "NetCommons"

Higashi-Katsushika High School in Chiba Prefecture

Times at which NetCommons is Used

8:00 a.m. The vice principal checks the blog
8:20 a.m. Teacher's meeting
8:35 a.m. Class starts
            Morning Individual teachers check the blog
            Teachers are only allowed to post comments
1:00 p.m. Blog writing begins
4:00 p.m. End-of-day short home rooms
            and notices to students are provided

The Higashi-Katsushika Blog is playing an important role in making school business more efficient.

complish this goal was reading comprehension, and he felt that using NetCommons in his classes would be an ideal way to foster this ability.

Mr. Kato divided his class into three stages: Presentation, in which the teacher presents a problem; Collaboration, in which students compare their opinions with those of other students; and Thinking, in which each student thinks about the problem in more depth. At the Presentation stage, Mr. Kato still writes on the blackboard, partly to heighten the motivation on the part of students to learn. In other words, he doesn't think that everything should be done on the computer. However, he says that the computer has made the learning process clearer by making it possible to display one's opinions next to those of other students and using the change history and so on to make it easy to see how student's opinions have changed.

"If the children are observed by others as they write, it increases their ability to think logically," says Professor Arai. "That's why we're currently at work developing a module that emphasizes writing by the students themselves." After quoting these words of Professor Arai, Mr. Kato sums up his view as follows. "I think NetCommons, which was developed from this approach, is an ideal tool for educational settings."

"Higashi-Katsushika Blog"Makes School Business More Efficient

Higashi-Katsushika High School is located in the northwest part of Chiba Prefecture. Although the school prepares students to go on to college, it is by no means advanced in terms of the introduction of information technologies. Shinya Ohashi, the teacher giving the report, says that when he joined the faculty in 2000, he was amazed to find that even the ac-

tivity plan for the school year that was handed out at the staff meeting at the beginning of the school year was a photocopy of a handwritten document. Yet last year, Higashi-Katsushika High School started up a NetCommons school blog called the Higashi-Katsushika Blog.

Since Saturday classes have been eliminated, teachers teach a seventh hour of classes two days a week. To enable students to use their after-school hours efficiently, it was necessary to eliminate the morning homeroom and to make the morning teacher's meeting as efficient as possible.

The most important considerations for the blog were threefold:
- to create an interface that would provide information at a glance,
- to enable all teachers to post blog entries easily, and
- to ensure security.

NetCommons was chosen because it allows anyone with the ability to use a word processor to post blog entries, and because of the peace of mind that came from using a product developed by NII. In addition, the "Group Room" function, with sections that can only be viewed by members of a specific group, was convenient for assigning school duties and creating sites to issue notifications for specific grade levels.

Nowadays, the order of priority for making announcements and discussing topics at the morning meeting is based on the entries on the Higashi-Katsushika Blog. Teachers also consult the blog to receive notifications that do not need to be formally announced. "Although I'm the blog administrator, I didn't formulate the use policy and the other rules," says Mr. Ohashi. "It's important for the entire faculty to recognize the need for a blog and consider how it should be used and so on.

Necessary Board of Education Support

In an increasing number of cases, the Board of Education and other organizations are taking the lead in introducing NetCommons at schools. The Chiba General Education Center is on the vanguard of this trend. "I want schools to compete with one another based on the content of the websites," says Takeshi Kokubo, a teacher who worked at the Center until March 2007. "For that reason, I think that the board of education has to provide the server and other components of the basic environment."

Mr. Kokubo noticed that the number of parents who had grown up in the IT age was increasing and

there was a need to provide information by website and email. So he began searching for a CMS that was suitable for use in an educational setting. At around the same time, he participated in the 2005 user conference and realized he'd found what he was looking for. Mr. Kokubo says the three-level hierarchy of NetCommons, which has a top page and below that "Group Rooms" designed to enable information sharing within specific groups, and below that "My Rooms" for individuals, is similar to the organization at schools, with a hierarchy of classes and club activities extending laterally. This appears to be one of the reasons that NetCommons has been accepted in so many educational settings.

In order for a system to be successfully introduced to many schools simultaneously, training must be conducted to enable teachers to learn the skills needed to use the system. NetCommons is designed to be easy for anyone to use, so only a short period of training is needed before teachers can begin to actually use the system.
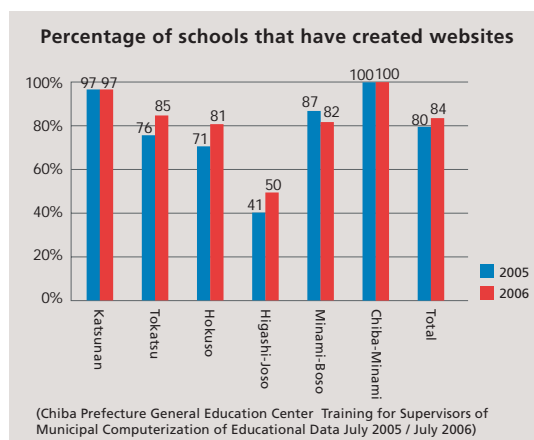
Future Development of NetCommons

After the plenary session, participants attended individual sessions, choosing the session that most interested them:
- Session 1 Efforts to computerize publicity, education and school business.
- Session 2 Case studies of the use of e-learning
- Session 3 Case studies of the increasing use of Net Commons by NPOs and small and medium-sized companies
- Session 4 Ways to use NetCommons to increase business opportunities
- Session 5 Hands-on training for beginners

These sessions were a valuable opportunity for participants to explore the possibilities of NetCommons. Finally, the participants once more returned to the plenary session at the Hitotsubashi Memorial Hall for the unveiling of NetCommons 2.0 by Ryuji Masuka-

Percentage of schools that have created websites

(Chiba Prefecture General Education Center Training for Supervisors of Municipal Computerization of Educational Data July 2005 / July 2006)

Educational centers must find ways to correct the regional disparity among schools in terms of infrastructure.

Vice principal Takeshi Kokubo (ex-supervisor of instruction, Chiba General Education Center) Takayanagi Junior High School in Kashiwa City

wa, the development supervisor. NetCommons 2.0 is scheduled for official release in 2008. One of the common complaints regarding websites created using a CMS is that they all look the same. In version 2.0, windows can be divided into blocks for editing, increasing the flexibility of design in order to resolve this problem.

"The most noteworthy feature of 2.0 is that teachers who are transferred to other schools can take the websites that they have developed with them, not only in terms of content but in that same form," emphasizes Professor Arai. In other words, the website belongs to the owners of the data. This is one step closer to the original NetCommons goal of making it possible for anyone to easily edit his or her data and making that data portable. "This change will radically alter the way we work. For this reason, it will become more and more important to learn what users want," says Professor Arai, who also called for user conferences to be held on an ongoing basis.

We also interviewed participants in the beginner's workshop who may become NetCommons users in the future. One young man, who said he was a children's nurse, commented that he wanted to start up a website because he felt the need to communicate information to the children's parents and guardians. Another man was holding Sunday School classes at an elementary school for a group of middle-aged and older people. "A free CMS is a big help for those of use who are trying an IT infrastructure at elementary schools," he said. Both of these participants are strongly leaning toward the use of Net-Commons.

We are already looking forward to seeing what new methods of use will be presented at next year's user conference.   (Written by Akiko Ikeda)

Session 5 Beginner's Workshop

CommonsNet: Promoting NetCommons

Kohei Teraguchi

CommonsNet is an NPO whose primary activity is to promote the use of NetCommons in addition to technology transfers and so on. NetCommons is growing to the point where the NII team alone will be unable to cope. From this point on, further growth in the public institution and company sectors as well as at schools is anticipated.

The NPO is the result of a chance encounter. Kohei Teraguchi, who is one of the members of the Net-Commons development team and currently serves concurrently as the director of the CommonsNet NPO, met NII Professor Noriko Arai in late 2000 at his workplace. At the time, the idea that would eventually become NetCommons was percolating in Professor Arai's head. This encounter began the process that eventually turned this idea became into reality.

The Launch of CommonsNet

Development of NetCommons as an information-sharing platform began in 2001. In August 2005, six years after the encounter of the previous year, free distribution of the open source NetCommons system began. Initially, the system was introduced gradually at schools and the like. Beginning this year, however, the number of downloads increased dramatically. Particularly remarkable has been the introduction of the system at the Board of Education level. So far, three prefectures—Tottori, Saitama and Yamagata—have formally decided to introduce the system, and Boards of Education in Ishikawa, Tochigi and other prefectures are currently studying the idea. "Introducing the system at all of the schools in the prefecture will make it possible for teachers to use the same NetCommons system no matter where they are transferred. This would make training in IT technologies much more efficient," says Mr. Teraguchi. "However, with the increase in the number of users, it's becoming impossible to cope using a bulletin board on an official website."

NetCommons is an NII research and development project. There are limits to the degree to which the organization can extend its coverage to include user support and other dissemination tasks. Accordingly, it was decided to set up a separate organization to serve as the liaison with the outside and handle tasks such as further dissemination, technology transfers

and so on. The organization that was established in June 2007 was the CommonsNet NPO.*1
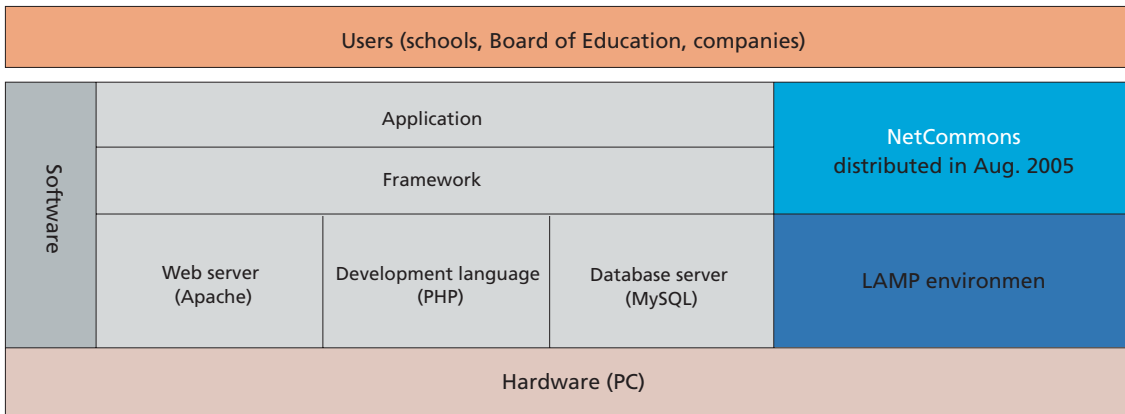
Takashi Matsumoto, vice director of the Commons-Net NPO, says there is a reason that CommonsNet was made an NPO. "NetCommons has been open source software from the outset. The whole purpose of open source software is to encourage distribution," he says. But relying on volunteer activities alone "would not provide a firm enough foundation," says Mr. Teraguchi. "We wanted to create a structure in which the public institutions that are the main users of NetCommons could introduce it with confidence." For NetCommons with its primarily public institution user base, the NPO mode was the best fit.

The Goal: Nationwide Deployment

Currently the primary activity of CommonsNet is the holding of various seminars: practical seminars for users, analysis seminars for software developers (photo), and management seminars on maintenance, troubleshooting and other issues for administrators. The purpose of these seminars is the same as the purpose for which CommonsNet was established: the dissemination of NetCommons as well as development and maintenance technologies (technology transfers). The user seminars enable people to start using NetCommons after only half a day of training, and they are free of charge to CommonsNet members.

On August 8 and October 29, NetCommons Source Analysis Seminars were held. Both seminars were extremely successful and there was a waiting list for seating. The seminars attracted many participants from both schools and IT companies. "We're planning to hold one more seminar before the fiscal year ends, but we'd like to hold it someplace other than Tokyo," says Mr. Teraguchi. Currently more than a thousand organizations nationwide have become

Software	Application			NetCommons distributed in Aug. 2005
	Framework			
	Web server (Apache)	Development language (PHP)	Database server (MySQL)	LAMP environmen
Users (schools, Board of Education, companies)				
Hardware (PC)				

NetCommons is a software application with its own framework (meaning the group of components used by that application). Up to now, the only easy-to-use frameworks were those found in commercial software, the use of which entails license renewal and other costs. However, the emergence of NetCommons has created a free, open source environment for the communication of information over the Internet and intranets.

NetCommons users, and further growth is anticipated.

Increased Business Opportunities

As NetCommons is provided free of charge, no costs are involved. More accurately, only the cost of server rental is required, but basically the software is free. That being the case, you'd think this wouldn't work as a business. In reality, however, the business opportunities are unexpectedly wide-ranging: maintenance to ensure continuing service, safeguard against the event of trouble, customization involving the change of programs, design of themes and so on.

Moreover, "it is a physical impossibility for NII and CommonsNet to provide operation and maintenance support for all of the users nationwide," points out Mr. Matsumoto. "We're thinking of searching for companies in each region that can handle these tasks and conducting technology transfer of support technologies nationwide to these companies as CommonsNet members." In reality, a company in Okinawa has already been contracted by local restaurants and a pharmacists association and so on to provide operational services. And several companies in Tottori, Osaka, Kyoto and the Kanto region are already become involved in the form of providing user support.

NetCommons 2.0 and Beyond

It was only recently that CommonsNet was established. Right now the network is chiefly engaged in holding seminars, but future plans call for the scope of CommonNet activities to be expanded to include survey research, information-gathering, the issue of publications and newsletters, and the development and sale of software. The original aim was to develop NetCommons as a general information infrastructure. However, as the majority of the initial users were educational institutions, many people probably mistakenly assume that it is an e-learning system. From this point on, consideration of ways to gain the participation of government agencies, private sector companies and so on as users will be needed.

"In 2008, a completely new version of NetCommons, NetCommons 2.0, will be put into general release," says Mr. Matsumoto.*2 "We think this will coincide with the beginning of full-fledged CommonsNet activity." Until that time, efforts will focus on using the current version, NetCommons 1.1.2, to build a foundation for full-fledged activity.

At the conclusion of the interview, we asked Mr. Matsumoto what he thought the next version of NetCommons, version 3.0, would be like. "For example, I envision a sort of network to link all of the services that will be constructed using NetCommons 2.0," he replied. The construction of the infrastructure for the "information society" has only just begun. NetCommons will continue to evolve, and the role of CommonsNet will continue to expand accordingly.   (Written by Tomoaki Yoshito)

*2 NetCommons 2.0 is scheduled for official release to CommonsNet members at the beginning of fiscal 2008.

NetCommons Source Analysis Seminar sponsored by CommonsNet and held on October 29. The display in the center of the three-person table shows the screen being operated by instructor Kohei Teraguchi. Attendees consult this display as they learn how to conduct the process.

Speech Corpora Promote Research and Development

The Speech Resources Consortium (NII-SRC*) was established a little over a year ago to collect and distribute speech data for research on speech recognition and speech synthesis. NII speech corpora have become a basis of speech research conducted at universities and in companies.

Shuichi Itahashi

Starting in the early 1990s, institutions in various countries around the world began to collect speech and language data from their respective regions and to create speech corpora.

There are now robots that react and respond to human speech, as well as computers that have ears and mouths so that they can exchange in conversations with human beings. Speech recognition and speech synthesis technologies are rapidly progressing, and the near future may have scenes that seem like they are from science fiction movies. Just look around you today and you will find technologies such as mobile phones, family game machines, and car navigation systems that use speech recognition or speech synthesis.

Establishment of NII-SRC

Research and development of speech recognition and speech synthesis technologies require enormous quantities of "voice" data. To enable the system to cope with the speech of various people in various situations, it is necessary to collect as many samples as possible.

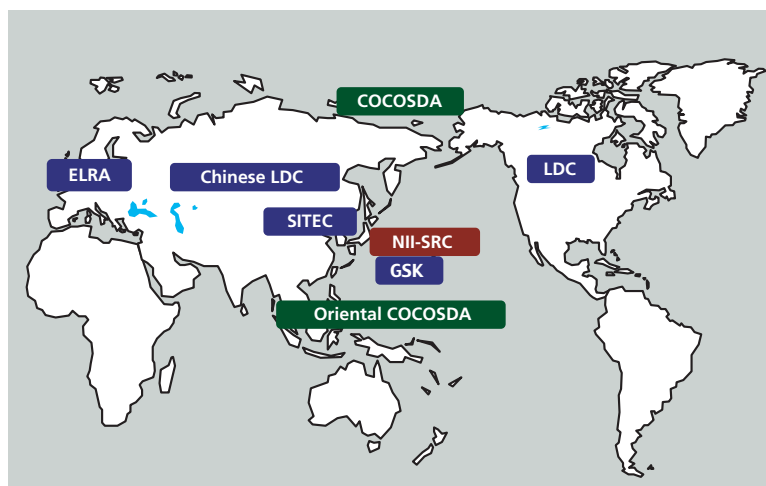So far, individuals, organizations, and companies have been independently collecting speech data. However, this independent approach requires a large amount of money and labor. Surely it would be easier if there were an organization that would collect data and make it publicly available to researchers. This is the objective of NII-SRC, which is operated by NII researchers, Professor Shuichi Itahashi and Ms. Kimiko Yamakawa. NII-SRC collects data and conducts research with the aim of creating speech corpora from a large collection of speech data.

The impetus for the creation of a common speech database began with a single research paper published in the United States in 1981. In this study, researchers collected speech recognition devices that had begun to appear on the market at the time and used common speech data to compare the performance of these units. This paper elucidated the concept of "sharing speech data" for the purposes of study and comparison.

Ten years passed. Following an international conference in 1991, organizations to create speech corpora were set up around the world. Japan set up the Linguistic Resources Sharing Initiative (LRSI) in 1994 and the *Gengo Shigen Kyokai* (GSK) or Language Resource Association in 1999. However, neither was able to adequately fulfill its functions.
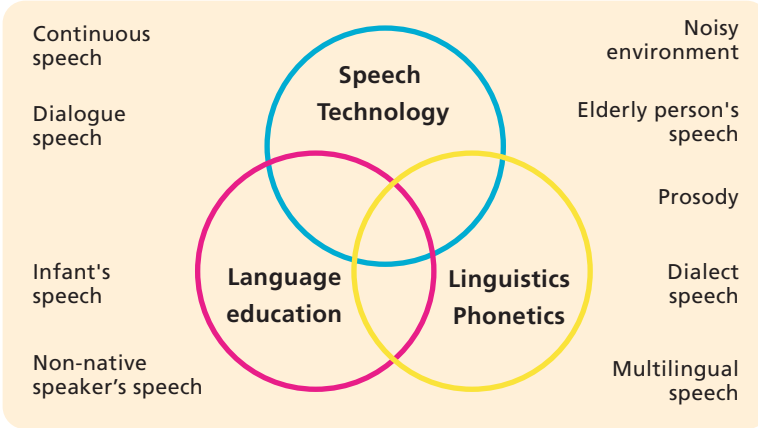
Considering the situation, Itahashi felt that NII — which manifests the e-Science concept saying that

Corpus Institutions Worldwide

Corpora available from NII-SRC

Continuous speech

Dialogue speech

Speech Technology

Noisy environment

Elderly person's speech

Prosody

Infant's speech

Language education

Linguistics Phonetics

Dialect speech

Non-native speaker's speech

Multilingual speech

Overview of data distributed by NII-SRC. The figure describes many different aspects of the data (environment, age, language, etc.). .

"essence can be found in a wide variety of data" — was the ideal organization to meet the needs. He became a professor at NII in 2005, and NII-SRC was subsequently established. His view was that NII, as Japan's organization for informatics, was the most appropriate entity for this project.

Collecting Speech Data

Currently, we mainly request representatives of various projects supported by research funds, universities, scientific societies and so on to supply speech data. NII-SRC, up to now, has created corpora from data collected independently by speech researchers. However, even if the data are collected at random, the issue remains as to how to supply suitable materials for research. As a result, NII-SRC is now studying how to classify the data.

Ms. Yamakawa has devised a classification system in which speech corpora can be expressed in terms of 58 attributes. The characteristics of each corpus can be visually represented by plotting the classification results on a two-dimensional coordinate plane. "This enables us to identify the content of existing corpora and to determine what types of corpora are lacking," she said. The visual representation makes it possible to determine the relevance and similarities of multiple corpora at a glance.

In this way, speech corpora can be steadily enriched.

NII-SRC's corpora have attracted attention from researchers and companies. In fact, the organization has received many more requests for use than it had expected. Collecting the data and making the corpora has required enormous amounts of time and money; hence, potential users have eagerly awaited their completion.

Liaison For Industry-Academia Cooperation

Speech corpora are used for an extremely wide array of purposes, from social welfare and medical care applications to car navigation systems. These days, speech corpora have uses not only in engineering but also in the humanities such as linguistics and education. Consider the cases of e-learning of English over the Internet and Japanese lessons for overseas students. NII-SRC has corpora containing English spoken by Japanese and Japanese spoken by overseas students that would be helpful for these purposes.

Other corpora such as one containing speech of elderly persons and one containing people speaking particular dialects can meet various needs. In fact, the range of use is extraordinarily broad for corpora. "I think that once you find out the demand for corpora, you'll be able to determine what technologies are needed." said Ms. Yamakawa.

Nevertheless, there are still many pending issues, such as corpora copyrights and how to protect personal information. It will be necessary to strengthen the system to protect the rights and safety of the people who have helped by allowing their voices to be recorded.

Professor Itahashi adds, "The corpora themselves are still far from perfect. The quantity of data at similar organizations in the United States and Europe is immensely greater. We have to keep working to make our data more complete."

NII-SRC will continue on its path toward the goal of becoming a center of speech resources in Japan and a clearinghouse for requests from countries around the world. (Written by Ryo Tanaka)

Kimiko Yamakawa

If the mathematician Gauss were alive today

Ken Hayami

Professor, Principles of Informatics Research Division, National Institute of Informatics

It was the German mathematician Carl Friedrich Gauss (1777 - 1855) who said "Number theory is the queen of mathematics." What he meant was that number theory, which studies the properties of numbers, is the most beautiful among all mathematical theories. The twentieth century English mathematician Hardy, who specialized in number theory, also expressed his view that the value of number theory lies in its beauty and not in its practical use. Gauss's great book on number theory, which he wrote in his youth, certainly had a huge impact on later mathematical research. But number theory also plays an essential role in today's internet society through encryption and security technologies. Thus, aesthetic value gave rise to practical value.

Gauss and applied mathematics

After writing the book on number theory, Gauss turned his interests to astronomy, geodetics and electromagnetism, involving himself actively in the research of applied mathematics for the majority of his life, although there may not have been a clear distinction between pure and applied mathematics in his times.

Through research in astronomy and geodetics, he discovered the least squares method and the normal distribution (Gaussian distribution). On the old German ten Mark note, Gauss's portrait is printed along with a surveying instrument which he invented, a triangulated map of the Hannover district, the Gaussian distribution curve, suggesting how proud the Germans are of Gauss. The study in geodetics also gave rise to Gaussian curvature and differential geometry, which is another major field of mathematics today. Differential geometry, in turn, was later used by Einstein to formulate the theory of gravity. In addition, his famous numerical methods for the solution of systems of linear equations may have evolved from practical problems that he was handling.

If Gauss were alive today

These are good examples of rich mathematical seeds hidden in practical problems. Today, there are active interactions between mathematics and its applications, and also within mathematics many fields are interrelated.

Turning our attention to today's information-oriented society, it must also contain all sorts of rich seeds of new mathematics. What would Gauss be doing if he were alive today? Would he be developing simulators on a supercomputer or the grid to analyze global environmental issues? Or would he be doing research on quantum computing? He may even be decoding DNA's to investigate the mystery of life, or constructing a mathematical model of the brain and developing human-friendly robots. There is no end to imaginations!