

## Truly Innovative Portal Services

*Masahiro Oku<sup>†</sup> and Katsuyoshi Tanabe*

### Abstract

We have been conducting research and development on core technologies to achieve innovative portal services that will provide interesting and valuable information to a user in the ubiquitous broadband era.

### 1. Tasks for portal services

What do you use the Internet for? People have various different purposes. For example, some people use it to get up-to-date international news or the latest rumors about celebrities. Others want to find the location of an interesting-looking hot-spring hotel they saw on TV, want to know what others say about a product they are thinking of buying, or want to read articles in blogs (Weblogs) written by people with similar interests. As **Fig. 1** shows, portal services are the gateways to such information. They are designed to offer various search services to meet various requests. However, the current search services are not useful enough to lead users to the information they want if definite keywords are unavailable. You may have had the experience of failing to get desired information simply because the keywords you came up with were not detailed enough.

Another common problem is that too many search results are presented, so it is hard to extract the desired information from them. If the user has a clear purpose and some idea of the desired information,

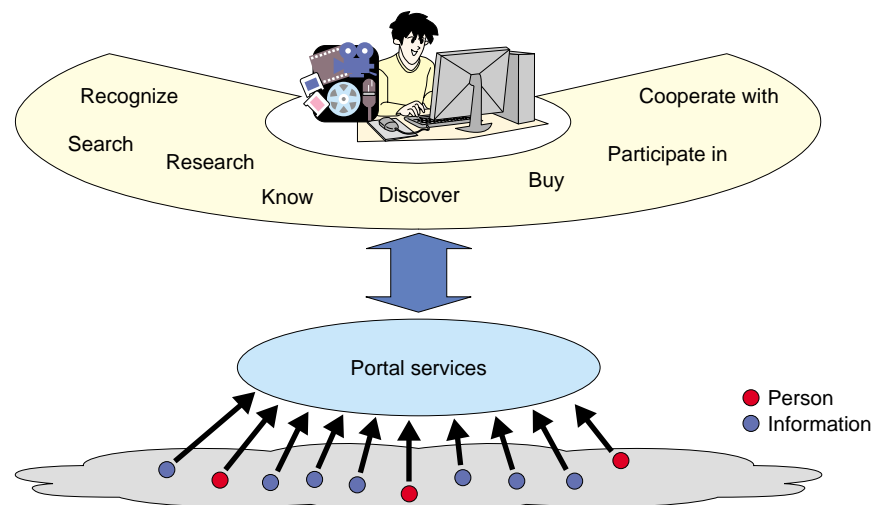


Fig. 1. Various user demands for portal services.

he/she will try to look for the desired result by checking the huge number of returned search results one by one, which is a very laborious task. Most users just glance at the search results on the first one or two pages and if they cannot find the desired information, they give up and do not bother to check the remaining results or they try a different approach.

**Figure 2** shows how users get information over the Internet. Users who are very experienced at accessing information (information experts) can continuously expand their own “areas that are easy to reach” by adding appropriate keywords based on the feedback provided by the first search results (i.e., by refining information in an easy-to-reach area) or by following links skillfully to get information from “areas that have useful information but are difficult to reach”

<sup>†</sup> NTT Cyber Solutions Laboratories  
Yokosuka-shi, 239-0847 Japan  
E-mail: oku.masahiro@lab.ntt.co.jp

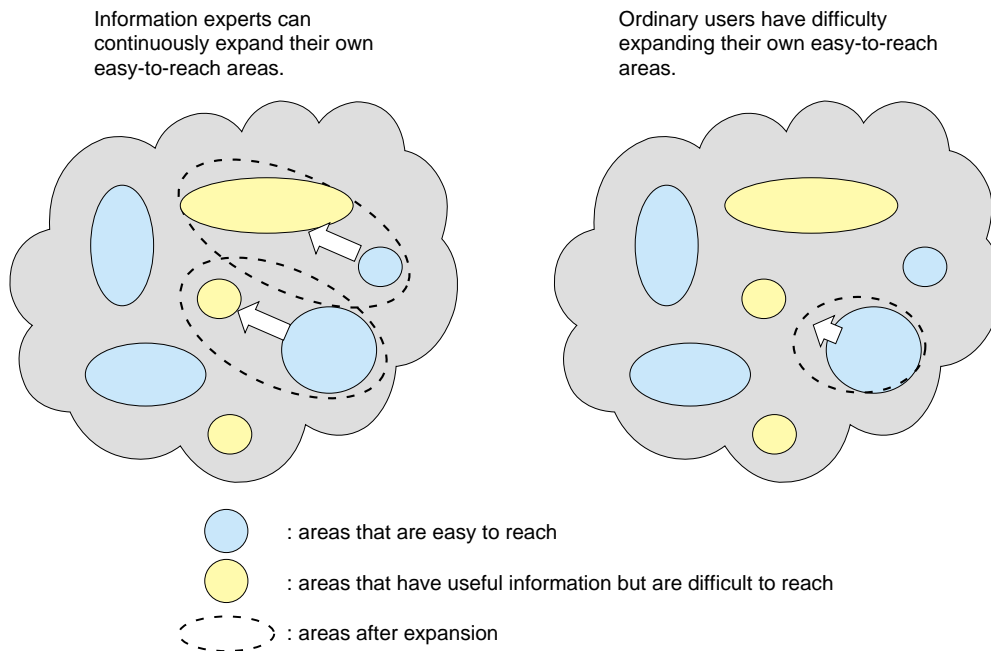


Fig. 2. How users get information over the Internet (current usage).

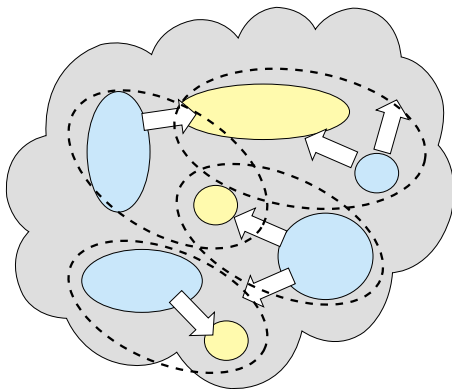


Fig. 3. How users could get information over the Internet (improvements provided by future portal services).

(see the area expansion arrows in Fig. 2). Unlike information experts, ordinary users often fail to think of appropriate keywords or get lost while following the links, even if they see the same search results as the experts. In other words, ordinary users have difficulty in expanding their own easy-to-reach area. Next-generation portal services should enable as many users as possible to easily reach the information they want regardless of their level of information literacy. It is important to expand easy-to-reach areas while reducing areas that contain useful information but are difficult to reach by developing navigation technology to guide users to the information they want, as shown in **Fig. 3**.

## 2. Technologies for the next-generation portal services

Some examples of information in areas that contain useful information but are difficult to reach are given below.

- Information that users cannot get by the simple search method:  
 Information that is difficult to get by simple keyword search; for example, information that includes more than one answer (e.g., information regarding methods or reasons), information that is discovered while searching with no definite purpose, information stored in various databases, etc.
- Information other than text:  
 Multimedia information (e.g., images/videos/sounds/music) that is difficult to locate by using keywords.
- Information buried in an enormous amount of information:  
 Information that users cannot get by a search-based approach; for example, information that is important to only a limited group of users or information that is being discussed in a limited field.

We believe that the new portal technologies shown in **Fig. 4** can deal with such information ((a)-(c)) by connecting information sources and users. Will our

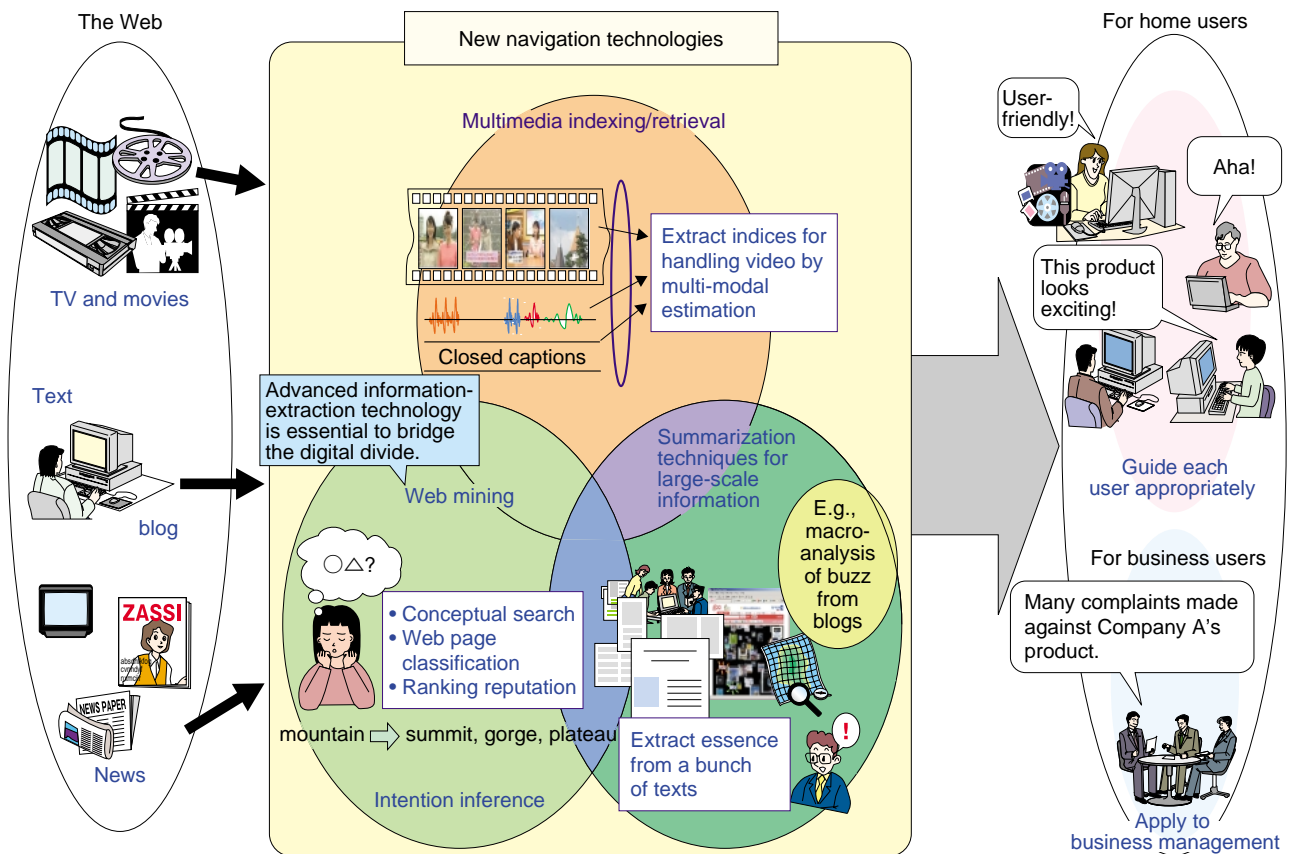


Fig. 4. Navigation technologies for next-generation portals.

new portal technologies be accepted by users when they are put into actual use or will they need to be improved? It may be that we can answer such questions only after the technologies have been tested by users in practice. NTT Laboratories offers users an online site, called “goo Labs” [1], where they can experience our newest technologies including the next-generation search technologies. Our portal technologies are being tested there too. The site is used to identify our future study tasks based on user reactions; their responses allow us to refine the technologies before putting them into commercial use. For example, the goo multimedia search service [2] for images, moving pictures, and music is a graduate of “goo Labs”. This service uses a unique search technique developed by NTT. Before putting it into commercial use, we tested it at “goo Labs” about for eight months to verify its technical and operational usefulness.

NTT Laboratories is engaged in R&D of technologies that can handle all types of information ((a)-(c)). Our R&D activities on portal technologies have four main themes, which are discussed in more detail in

the other articles in this Special Feature:

(1) Turning online web information into “pearls of wisdom”

For information that users cannot get by a simple search method, a navigation-type search system assists users in finding the information they need in a number of ways. We are researching and developing technologies that will enable users to easily reach the information they want by arranging and categorizing a huge volume of information from a different viewpoint from that used in traditional ranking methods. We are also studying search technologies that enable users to reach information even though it does not include the input keyword but a word with a similar meaning.

(2) Making the fullest use of broadband

There is no doubt that multimedia information will become more popular with the proliferation of broadband. Unlike text information, however, multimedia information is harder to handle. Because multimedia information requires the system to first indicate the information it wants to handle. Based on an analysis of the content of multimedia information, we have

been developing techniques that will enable users to handle the media depending on the content so that users can easily arrange, search, understand, and send the multimedia information, such as videos, at will.

(3) Giving users an overview of an enormous amount of information at a glance

Information that is buried in an enormous amount of information cannot be found simply based on the frequency of appearance. Therefore, we are studying the latest technique for selecting topics based on how frequently they appear on websites by category and how fresh they are, by analyzing blogs and news websites, etc. We are also studying interface techniques for displaying a large number of keywords and contents on small screens, such as those used in cellular phones.

(4) Providing services customized for individuals

We need to know what kind of information users want in order to lead them to it. However, users cannot always specify what information they want. Therefore, we are now developing technologies that will help them achieve their goals by showing them useful information, such as their behavior history or a list of websites browsed by many users.

### 3. Future work

Since the Internet market contains many rival companies and we learn by trial and error, it is becoming essential to shorten the cycle of testing and feedback provided by “goo Labs” in order to advance our R&D on a timely basis. The top page of goo Labs on February 28, 2006 is shown in Fig. 5, but by the time you read this article and try and access the URL [1], it will probably look rather different. We will proactively test our new technologies at goo Labs to prove their usefulness and utilize user feedback to refine our R&D activities.

From the viewpoint of technology development, we think that we must be really careful about how we handle information sent by individuals. Due to the explosive proliferation of blogs and social networking services (SNSs), not only the traditional type of information or information released by companies but also information released by individuals is becoming increasingly common on the Internet. Information sent by individuals is often of interest to other users, for example, word-of-mouth reports or opinions. Such information can represent useful marketing data for companies and essential information

The screenshot shows the 'goo Labs' website interface. At the top left is the 'goo ラボ' logo with the tagline '次世代の検索機能を体感'. At the top right is the 'NTT' logo. Below the header is a navigation bar with 'goo ラボ' and 'goo Labs'. The main content area is divided into several sections:

- 次世代型ブログ検索実験 (Next-Generation Blog Search Experiment):** Announced on Dec 19, 2005. Features 'BlogRanger Multifaceted blog search system' and 'BLOGRANGER' search results.
- プッシュ型トピック通知実験 (Push-based Topic Notification Experiment):** Announced on Nov 1, 2005. Features 'Push-based topic notification system'.
- エリア情報検索実験 (Area Information Search Experiment):** Announced on Oct 20, 2005. Features 'Local area information search system'.
- 日本語概念フィルタリング実験「関心事アンテナ」 (Japanese Language Concept Filtering Experiment 'Kanshinji Antenna'):** Announced on Aug 4, 2005. Features 'Japanese-language concept search system'.
- 店舗情報検索エージェント実験 ホテル版 (Store Information Search Agent Experiment Hotel Edition):** Announced on Aug 16, 2005.
- 店舗情報検索エージェント実験 グルメ版 (Store Information Search Agent Experiment Gourmet Edition):** Announced on Feb 28, 2005.
- 新着情報 (New Information):** A list of recent updates, including 'Newly arrived information' and 'Multimedia search experiment'.
- goo βサービス (goo beta services):** A section for beta services, including 'Web search support'.

Fig. 5. Top page of “goo Labs”.

for individuals when buying or making decisions. We think it is necessary to develop technologies that can organize, process, and categorize the information sent by individuals and pass it on. However, it is important to keep in mind that information provided by individuals is sometimes inferior to that from companies in terms of reliability and credibility. Therefore, we think we need to research and develop technologies taking into consideration safety and reliability; for example, we need technologies that can evaluate the reliability of information. We will continue development of the next-generation portal services, which will, we believe, satisfy your expectations.

---

### References

- [1] <http://labs.goo.ne.jp/>
- [2] <http://bsearch.goo.ne.jp/>



#### **Masahiro Oku**

Senior Research Engineer, Supervisor, Media Computing Project, NTT Cyber Solutions Laboratories.

He received the B.E. and M.E. degrees in electronic engineering from Osaka Prefecture University, Osaka, in 1982 and 1984, respectively, and the Ph.D. degree in engineering from Niigata University, Niigata, in 2001. In 1984, he joined the Electrical Communication Laboratories of Nippon Telegraph and Telephone Public Corporation (now NTT). Since then, he has been engaged in R&D of natural language processing, especially Japanese language processing, such as analyzing Japanese idiomatic expressions, detecting and correcting Japanese homophone errors, and creating automated directory assistance systems. He is now engaged in R&D of broadband Internet services, such as multimedia information retrieval. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan, the Information Processing Society of Japan (IPSJ), and the Association for Natural Language Processing.



#### **Katsuyoshi Tanabe**

Senior Research Engineer, Media Computing Project, NTT Cyber Solutions Laboratories.

He received the B.E. and M.E. degrees in informational engineering from Yokohama National University, Kanagawa, in 1983 and 1985, respectively. He joined NTT Electrical Communication Laboratories in 1985 and mainly engaged in R&D of image processing systems, such as image understanding, similarity measures between images, similarity image retrieval, and automatic synthesis of color ocular fundus images. He is now engaged in R&D of broadband Internet services, such as multimedia information retrieval. He is a member of IEICE, IPSJ, and the Institute of Image Information and Television Engineers.

---