16-DPC-06

Databases in Japan
2004
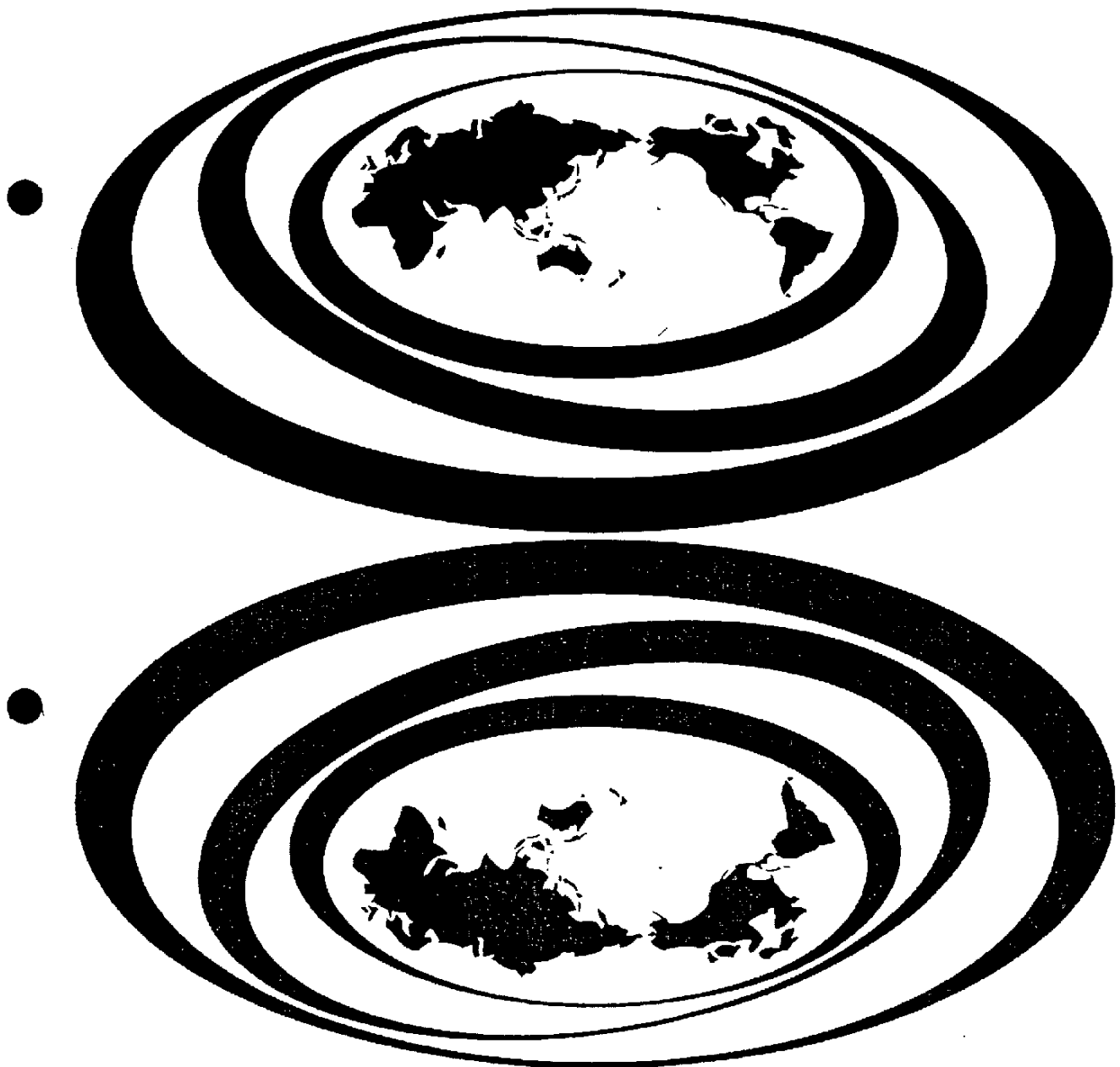
Database Promotion Center, Japan

Databases in Japan 2004

Contents

List of Figures

List of Tables

Venturing into a New Area—Database Evaluation

Prof. Tamiko Matsumura
Sugiyama Jogakuen University

1. Introduction

The dramatic progress of computer networking technologies in recent years has brought about dramatic changes in the production, distribution, and use of information. Concerning new media, Marshall McLuhan once said that "the medium is the message". According to McLuhan, media create a new paradigm of communication because they not only define content parameters but also alter the way messages are presented. Moreover, they change what people perceive as messages. Although many people still argue about this hypothesis, the fact is that computers and the Internet are exerting great influence on today's information communication.

Conventionally, databases were built and offered mostly by specific information service providers. In addition to such databases, government and public institutions as well as corporations and individuals have come to create and supply on the Web information which is greatly diversified in content and quality. Users of these information resources are also going through major changes, and Web information is now becoming part of the daily lives of the general public. Besides databases, the variety of Web information created and accumulated on computers and provided via the Internet is developing at a rapid pace. Under such circumstances, quality of information poses significant issues in terms of both its use and creation.

2. Databases and Web Information

The term "database" is said to have been used for the first time in the 1950s to refer to computer-accumulated data about US military bases. Generally, however, this term came into widespread use when PubMed (former MEDLARS) and other major scientific indexing and abstracting journals introduced computers and started providing "database search" services along with publication by the conventional print medium. Also at this time, databases were constructed for various publications such as newspapers and encyclopedias to form a new market. Thus, the term "database" initially referred to a resource of computer-accumulated information controlled to meet a certain level of quality. However, information generated on computers and provided over the Internet has increased and diversified at a tremendous speed, and the range of information

producers has broadened widely, from the conventional information service providers to private individuals. As things stand now, "Web information" created in digital format does not always maintain a certain level of quality.

Under such circumstances, the concept of a database is undergoing some changes. For example, the Japanese Copyright Law sets forth: "'database' means an aggregate of information such as articles, numericals or diagrams, which is systematically constructed so that such information can be searched for with the aid of a computer." (Article 2, Paragraph 1, 10-3). In the EU, the European Parliament and the Council of the European Union in 1996 adopted "Directive 96 on the legal protection of databases". This directive sets forth: "'database' shall mean a collection of independent works, data, or other materials, arranged in a systematic or methodical way and individually accessible by electronic or other means." In the United States, Bill Number H.R. 3261, the "Database and Collections of Information Misappropriation Act", was introduced in the 108th Congress, 1st Session (October 2003). Section 2 of the bill gives the following definition: "The term 'database' means a collection of a large number of discrete items of information produced for the purpose of bringing such discrete items of information together in one place or through one source so that persons may access them". Excluded from this definition of the database is "a work of authorship, other than a compilation or a collective work". The above definitions clearly indicate that the range of meanings given to the term "database" varies considerably at present.

3. Database Quality

Database quality is viewed from various aspects, but the generally accepted definitions are "conformance to requirements"[1] and "suitability for purpose".[2]

The Western countries began to recognize the importance of database quality from the late 1980s to the early 1990s. One of the major activities in the early period is the work done by the Southern California Online User Group (SCOUG).[3] SCOUG proposed an extremely detailed list of criteria for evaluating quality that affects the use of a database. The proposed criteria include 1) Consistency, 2) Scope/coverage, 3) Timeliness, 4) Error rate/accuracy, 5) Ease of use, 6) Integration/harmonization (with other like databases), 7) Output, 8) Documentation, 9) Customer support and training, and 10) Value to cost ratio/charging. Further details are given in each of the criteria, and in addition, SCOUG listed detailed requirements that bibliographic databases, full-text databases, and directory databases should satisfy.

The European Association of Information Services conducted a survey[4] of database users and asked them to rank the general headings of the SCOUG criteria. The survey result shows that "Coverage" is ranked top, followed by "Accessibility," "Timeliness," "Consistency," and "Accuracy", in that order.

In the United Kingdom, the Centre for Information Quality Management (CIQM) was set up as a clearinghouse at which database users can report quality problems. CIQM collects statistical data

on database quality and conducts activities to identify any improvements needed. As a result of studying methodologies of quality assurance in databases, CIQM proposed "Database Labelling".[5] The basic concept of this labeling method is the "Quality Flag" of a database. More specifically, this method presents the latest database specifications as a set of parameters by which users can accurately evaluate the attributes and content of each database. The label specifications vary with the types of database, such as bibliographic, full-text, numerical, or image database, and with the subject area covered by a database. Database Labelling makes it possible to "map users' expectations of databases and avoid unreasonable expectations". Users can use a database label as a key to determining the extent to which they can trust the database and make sure of its performance measurements and quality assurance. The concept of this labeling has led to the recent movement of using metadata to describe the quality attributes of Web information. Furthermore, the development of automatic quality labeling and evaluation methods is progressing, based on the results of research and development of technical infrastructures as exemplified by PICS (Platform for Internet Content Selection) and RDF (Resource Description Framework).

As the volume of Web information increases rapidly and its significance as information resource grows further, there is greater interest in the evaluation of Web information. Information service organizations, such as libraries, have taken a central role in establishing evaluation criteria from the standpoint of users.

4. Research on Evaluation of Databases and Web Information

The foregoing research and surveys were conducted mainly for proven databases because of the social climate of the time. Databases were also evaluated from the users' point of view. Since users were normally restricted to certain groups of people, such as researchers and information experts, it was relatively easy to determine user requirements and measure "conformance to requirements", which is one of the definitions of "quality" mentioned earlier. Nowadays, however, databases have expanded to include a variety of information resources and serve a broader range of users. More importantly, Web information is increasing and diversifying dramatically, and at the same time its importance as an information resource is growing: Web information is now part of our daily lives. On the other hand, with the World Wide Web (WWW) spreading swiftly, Web information producers have increased in number and variety, making Web information a jumble of good and bad. What is most needed now is to develop a new evaluation method by which users can select reliable material from the "universe of information", and to establish criteria for producers in order for them to create information to a certain quality standard.

In full recognition of such demand, the Database Promotion Center in Japan set up in 2002 the "Research Committee for the Evaluation of Databases". Following its three-year plan, the Committee has been conducting research and surveys with the aim of first establishing quality evaluation criteria from the standpoint of users, and furthermore, formulating guidelines with

selected criteria for creating Web information of a certain quality from the standpoint of Web information producers.

In 2002, the Committee conducted studies to identify evaluation criteria from the viewpoint of database use, based on the preliminary research conducted in the preceding year on the methodologies of database evaluation. More specifically, the Committee analyzed its earlier works in the light of the Dublin Core, Smith,[6] and others[7], and as the result, worked out hypothetical evaluation criteria consisting of 12 elements and studied them in further detail. A questionnaire survey was also conducted to verify the validity of the criteria. Table 1 shows the proposed evaluation criteria.

Table 1　Database Evaluation Criteria

Evaluation Criteria	Elements
Title:	Name given to the information resource
Author/creator:	Person or organization responsible for the content of the information resource
Subtitle/keyword:	Subject of the information resource
Summary:	Text description of the content
Responsible person/organization:	Person or organization that processed the information resource in the current format
Person/organization with supplementary roles:	Person or organization that made indirect contributions to the information resource
Date of creation/update:	Year, month, and day of creation/update
Source:	Relation to other information resources
Spatial and time characteristics:	Period, coverage, volume of information
Rights:	Copyright and other information related to access restrictions
Accessibility:	Ease of use and view, user friendliness
Usefulness:	Completeness, exhaustivity

Discussions by the Committee focused on whether evaluation criteria should vary with database types and characteristics, and while most of the earlier studies targeted scientific information, the Committee should try to develop an evaluation model for a certain type of database in the hope of developing further a general-purpose model for a greater variety of databases. Accordingly, the Committee chose newspaper databases available on the WWW as the object of research, and studied in more details the afore-mentioned 12 elements of the criteria, and identified 171 properties. These properties were examined together with the results of the Web sites analysis of four national newspapers and two municipal organizations. As the result, the Committee selected 70 properties as the evaluation criteria of newspaper databases provided on the Web. In order to verify this result, the Committee grouped together the 70 properties into 25 for the purpose of a questionnaire study and carried out a questionnaire survey at DATABASE 2002 TOKYO. These results demonstrated that the 70 properties are valid as the evaluation criteria of newspaper databases.

In 2003, the Committee conducted research and surveys from the viewpoint of Web information

producers, aiming to establish criteria for creating Web information that satisfies a certain level of quality. There are already some laws, regulations, and standards that regulate Web information. For example, the "Specific Commercial Transactions Law" on electronic commerce makes it mandatory for information producers to indicate 14 properties including 1) name of entity, 2) name of the representative, 3) geographical address, and 4) contact information (telephone number, e-mail address). Also, the Japanese Standards Association has published the "Guidelines for Designs in Consideration of the Elderly and Handicapped – Equipment and Software Service in Information Communications, Part 3: Web Content (JIS draft)". They are now working on the standardization of Web information accessibility as part of the Japan Industrial Standards. The "Online Shopping Trust Mark" is a certification program of the Online Mark General Center, sponsored by the Japan Direct Marketing Association (JADMA) and the Japan Chamber of Commerce and Industry (JCCI). This program verifies the physical presence of direct marketers, examines whether indications on their Web sites comply with laws and regulations concerning mail-order sales, and permits compliant marketers to use the mark.

The Ministry of Public Management, Home Affairs, Posts and Telecommunications is considering the introduction of a "Safe Content Mark" (tentative name). This system is designed for bulletin boards and other Web sites whose content changes frequently. The Ministry will examine the operating system of each entity operating a Web site and, when acknowledging the site is safe, allow it to display the Content Safety Mark.

The protection of personal information is another important issue related to the creation of Web information. The guidelines published by the Organization for Economic Cooperation and Development (OECD) set forth eight rules for the protection of personal information. In Japan, the "Law Concerning Protection of Private Information" took effect in 2003, and the Private Information Protection Guidelines were also established.

With these trends in view, the Research Committee for the Evaluation of Databases has been analyzing the ways in which various Web sites are created. The Committee is also studying actual Web information with the aim of identifying elements and establishing guidelines for creating Web information that is reliable and meets certain quality standards.

5. Conclusion

Research and development of Web information evaluation methodologies have been conducted in a variety of fields including biomedical science,[8] particularly in relation to gateway construction. The Cultural Content Forum, an international consortium of agencies engaged in the digitization of cultural information resources, analyzed the methodologies of about 90 evaluation projects and pointed out the potential of establishing a standard evaluation method, together with issues to be discussed in the process.[9] These evaluations are made mainly from the standpoint of users. On the other hand, information producers and providers are studying evaluation criteria in the hope of increasing access to their Web sites. Evaluation criteria for users and evaluation criteria for producers are, in many respects, two sides of the same coin. Research and studies of qualitative guidelines for Web information generation will contribute to the further development of databases and the sophistication of Web information in Japan.

References:

1) Aichison, T. M.. Aspects of Quality. Information Services & Use, 8 (1988) p.50.
2) Ibid.
3) Basch, r. Measuring the quality of the data. Database Searcher, 6 (8) 1990, p.18-23
4) May, N. A. Methodology for the measurement of quality of electronic database.
   Paper presented at Third International ISKO Conference, 20-24 June 1994, Copenhagen.
5) Armstrong, C. J. Database Information Quality. Library & Information Briefings, Issue 62, July 1995, p.6-8.
6) Evaluation of Information Sources
   http://www.vuw.ac.nz/staff/alastair_smith/evaln/
7) http://www.pace.edu/library/instruct/evaltech.htm
8) How to Evaluate an Internet-based Information Source
   http://biome.ac.uk/guidelines/eval/howto.html
9) Evaluation of Digital Cultural Content: Analysis of Evaluation Material
   http://www.culturalcontentforum.org/publications

*URLs are as of the end of April.

I. Trends in the Database Service Market in Japan

1. Sales of the Database Service Industry

In the "Survey of Selected Service Industries" conducted by the Ministry of Economy, Trade and Industry (METI), information service industry are classified into "custom-made software development service," "information processing service," "software products," "facility management service," and in particular, "database service."

The total sales value of the information service industry for 2002 stood at ¥13.9731 trillion (US$111.61 billion), which was a 2.0% increase over the previous year (total sales value for the previous year: ¥13.7039 trillion). The 2.0% rate of increase was the lowest since 1995, when the rate of year-on-year increase in total sales turned from negative to positive growth. Total sales of the industry, which reached a peak in 1992 then decreased till 1994, increased by as much as 29.2% in 1998. One of the main reasons was a review made of the total number of business establishments within the industry. However the pace of growth slowed down in 1999 and 2000.Although there was a significant increase in total sales of the industry in 2001 as a result of increases in demand from major users such as financial institutions, the yearly rate of increase in total sales declined in 2002, presumably reflecting stagnation in private equipment investment.

A breakdown by category or type of business of the total sales shows that sales of "custom-made software development service," which is the largest category with a 49.2% share of the total sales, increased by 1.5% in year-on-year terms. Sales of "information processing service," which ranks second in sales value, posted only a 3.5% year-on-year increase (Table I-1, Fig. I-1).

Sales of "Software products," which increased significantly in 2001, decreased by 3.4% in 2002. This was attributable to a marked decrease in sales of "basic computer software, etc."

On the other hand, the sales value of "facility management service," which has been increasing since 1996, continued to grow in 2002, posting a 7.2% increase year-on-year. In recent years, there have been moves to promote outsourcing in the information divisions of banks, insurance companies and securities firms, among others.

In 2002 average sales per business establishment hit a record high for the second consecutive year, increasing from ¥1.75 billion in 2001 to ¥1.828 billion in 2002.

Sales of "database service" (Fig. I-2), which declined in 1999, posted a year-on-year increase for two consecutive years, in 2000 and 2001, but in 2002 fell to ¥273 billion (US$2.18 billion), an 8.4% decrease from the preceding year. This was the first time in three years that sales of "database service" posted a year-on-year decrease. The ratio of sales of "database service" to total sales of the information service industry decreased slightly from 2.2% to 2.0%.

Average Sales per business establishment of "database service" also decreased slightly from ¥371 million in 2001 to ¥356 million in 2002.

According to the "Report on Selected Service Industries Statistics" by METI, which serves as a

source of leading indicators, the sales value of "database service" compared with the corresponding period of a year did not turn from negative to positive growth until 2001. Although it returned to negative growth in the fourth quarter of 2001, it began to increase in the first quarter of 2003. This corresponds to the trend in private equipment investment (GDP basis, nominal) during that period. In other words, it seems that sales of "database service" during that period were affected by measures to reduce costs and expenditure by private businesses. According to the forecast by the DI (Diffusion Index), it returned to negative growth in the fourth quarter of 2002, and the prospects for "database service" are not good. A breakdown by outlet of the total sales of "database service" shows that "Internet" accounts for 37% of the total, while "other" accounts for 63%.

Table I-1　Annual Sales by Various Information Service Sectors

Classification	2001			2002		
	Annual sales (¥100 million)	Ratio (%)	Change from the previous year (%)	Annual sales (¥100 million)	Ratio (%)	Change from the previous year (%)
Total	137,039	100.0	18.2	139,731	100.0	2.0
Information processing service	26,114	19.1	36.4	27,032	19.3	3.5
Custom-made software development service	67,634	49.4	8.4	68,682	49.2	1.5
Software products	14,827	10.8	49.1	14,316	10.2	▲3.4
Assigned system management	11,598	8.5	44.8	12,436	8.9	7.2
Database service	2,979	2.2	1.8	2,730	2.0	▲8.4
(through the Internet)	1,214	0.9	▲10.5	1,007	0.7	▲17.0
(Others)	1,765	1.4	12.5	1,722	1.2	▲2.4
Researchers	2,958	2.2	1.3	3,131	2.2	5.8
Others	10,929	8.0	3.3	11,405	8.2	4.4

Source: "Survey of Selected Service Industries," METI

(¥ 100 billion)

□ Information processing service (*1)
▨ Order-made software development and software products (*2)
□ Other (other than the above)

88 89 90 91 92 93 94 95 96 97 98 99 00 01 02
(*1) Classified as "on-line and off-line data processing services" until 1999.   year
(*2) Classified as "software development/program production, etc." until 1999.
Source: "Survey of Selected Service Industries," METI
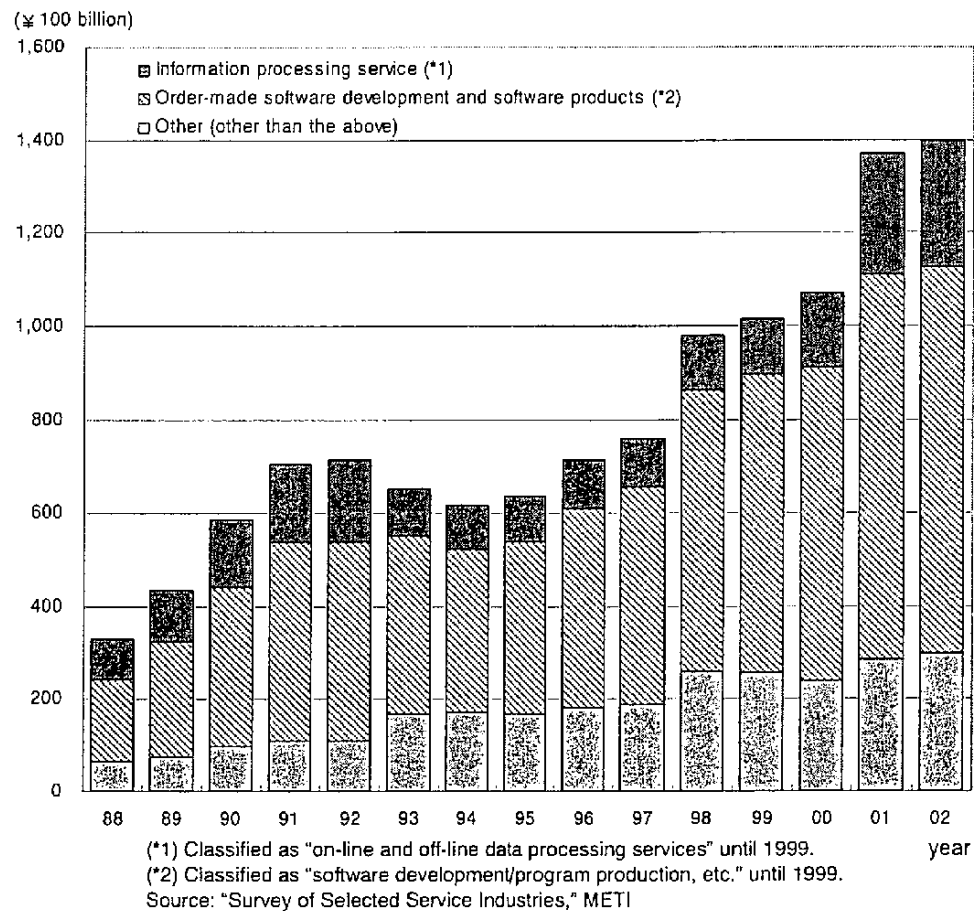
Fig. I-1   Trends in Annual Sales of the Information Service Industry (1988-2002)

(¥100 million)

	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	2000	2001	2002
¥100 million	523	787	957	1,008	1,143	432	1,063	1,576	1,886	2,160	2,141	2,115	1,988	1,973	2,354	2,578	2,910	2,683	2,916	2,979	2,730
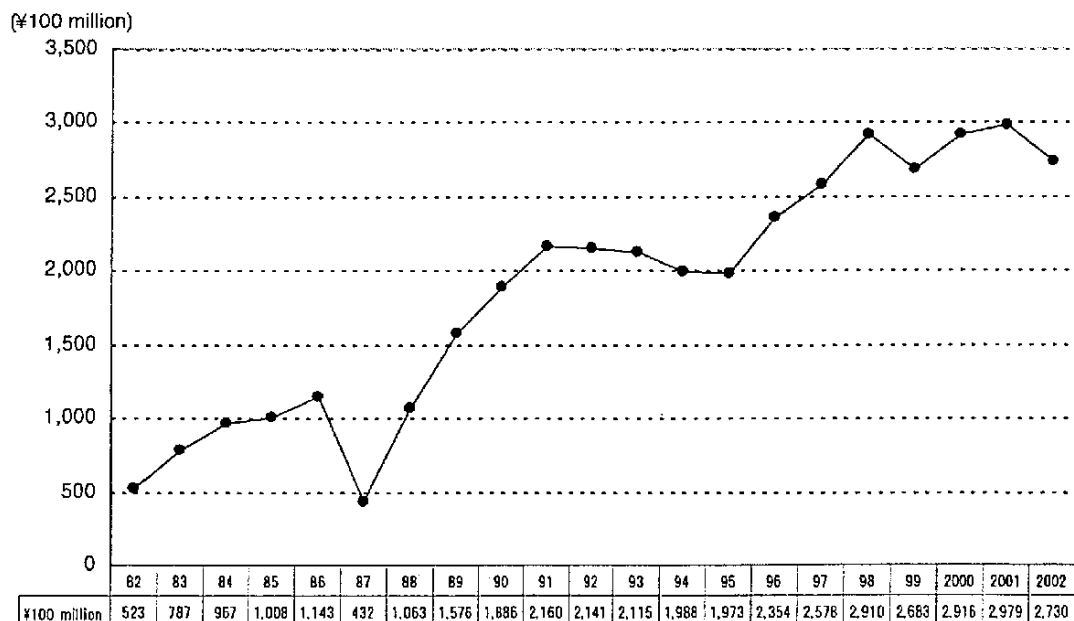
Note: Parts of the service classification were revised in 1987, and "Information supply service" was renamed as "Database service."
Source: "Survey of Selected Service Industries," METI

Fig. I-2   Trends in Annual Sales of Database Service Industry

9

2. Distribution of Commercial Databases

2.1 Number of Accessible Databases and Participating Companies

The "Directory of Japanese Databases", which has been published annually by the Ministry of Economy, Trade and Industry since FY1982, provides indices by field, medium, and service corporation, as a directory of commercial databases that are available in Japan. The entries in the directory are based on reports submitted by service companies. In FY2002, a total of 179 companies submitted their respective reports; the corresponding figure for FY2001 was 186. As stated in "1. Sales of the Database Service Industry," the number of reporting companies is lower than that of database service companies included in the "Survey of Selected Service Industries." This is because for the latter companies, the number of business establishments increased with the change of service classification to include "database service provided through the Internet." It is a cause for concern that the number of business establishments specializing in the provision of database services is decreasing gradually. A breakdown by type of business of the total number of these business establishments shows that the three main categories, namely, producer (the number of companies is 61, compared to 60 in the previous year), producer/distributor (51, compared to 48 in the previous year), and information search agency (18, compared to 22 in the previous year), account for approximately 70% of the total. When the number of companies which offer producer service concurrently with their main service are added to the number of producers, there is a total of 144 companies providing producer's service, representing approximately 80% of the total number of reporting companies.

The total actual number of commercial databases reported by the 179 companies was 2,379, a 6.8% decrease from the previous year. The number of databases was on the increase until FY1996, decreased in FY1997 and remained flat until FY2002 (Fig. I-3).

An analysis of the changes in the combined total number of overseas databases and domestic databases shows that the ratio of the total number of overseas databases to the combined total, which was nearly 80% in earlier years, began to decline gradually in FY1996 and was surpassed by that of domestic databases in FY1999. The trend continued into FY2002.

2.2 Diffusion of Databases by Field

A breakdown by field of the total number of accessible databases shows that 695 databases cover the field of "general" (a 22.4% decrease from the previous year), 743 cover "natural science and technology" (a 1.6% increase over the previous year), 186 cover "social science and cultural science" (a 19.2% increase over the previous year), 706 cover "business" (a 1.1% decrease from the previous year), and 49 cover "other" (a 10.9% decrease from previous year). There was a sharp decrease in the number of databases covering "general" (Table I-2). An analysis of the changes in percentage distribution by field of the total number of databases over the past 10-year period shows

10

that, while the percentage of databases covering "business" has been decreasing, that of databases covering "general" has been increasing. From FY1999 till FY2001, the percentage of databases covering "general" was larger than that of databases covering "business." In FY2002, however, the percentage of "general" decreased drastically, and as a result, the respective percentages of "general" and "business" became nearly equal(Table I-3).

As regards the databases covering "general," the total number of those which cover such subfields as "recreational, leisure and facility guides," "health and sports" and "law" has decreased sharply. Many of the databases covering other subfields of "general" have also suffered year-on-year decreases. While sales of databases for business use are not growing due to the stagnant business conditions, sales of databases for the use of general users are also being affected by the weakness of the economy.

Fiscal Year	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Actual number of DBs	1,795	1,964	2,128	2,354	2,666	2,799	2,980	3,061	3,308	3,350	2,598	2,478	2,581	2,458	2,552	2,379
Number of reporting companies	137	194	211	213	222	228	203	201	187	176	193	197	189	176	186	179
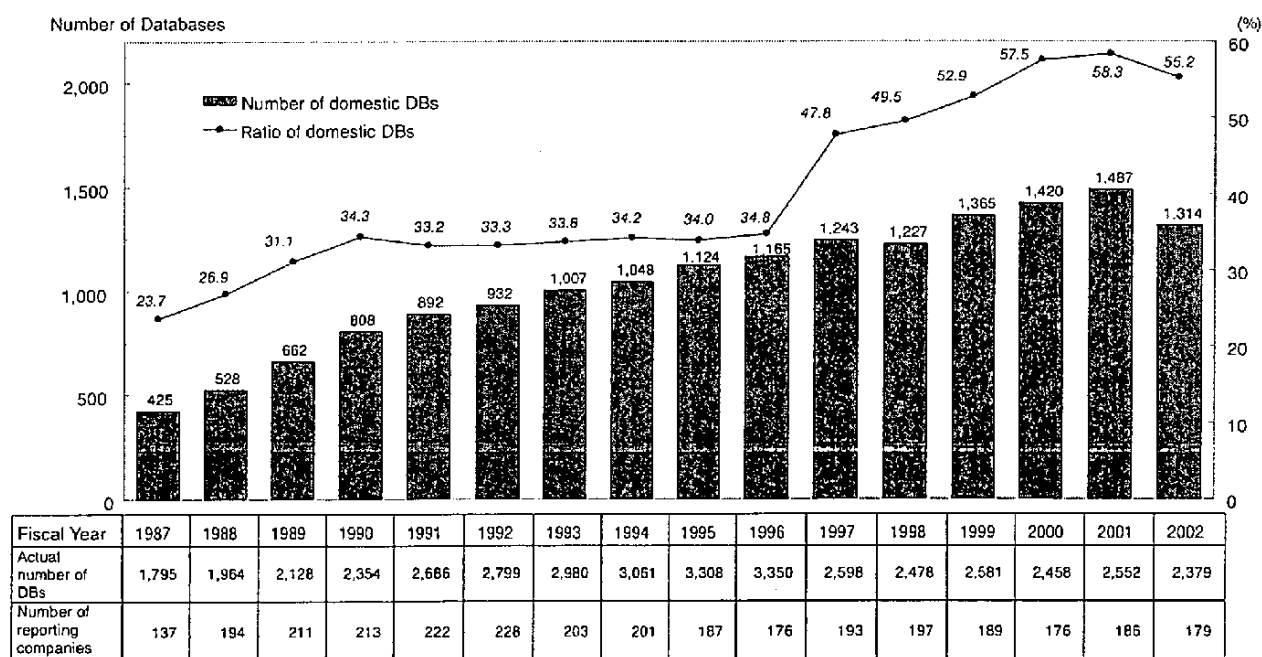
* The actual number of DBs in the table includes DBs compiled overseas. The ratio in the graph indicates the ratio of domestic DBs.

Source: "Directory of Japanese Databases," METI

Fig. I-3   Trends in the Number of Domestic Databases (actual numbers)

Table I-2-a Changes in Distribution of Databases by Field (actual numbers)

FY / Subcategory	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
General
General	69	69	75	82	78	76	62	68	72	70	75
Newspapers, journals and news	322	373	412	469	444	393	390	401	348	364	328
Who's who/Organizations	68	75	73	82	80	73	65	67	70	60	66
Government administration	38	46	48	54	59	46	36	36	34	37	33
Laws	63	67	80	76	77	46	49	49	51	50	26
Politics	43	46	40	36	34	6	5	5	2	1	2
Health and sports	15	27	27	33	35	27	24	38	36	36	9
Travel and scheduling	8	9	8	8	8	12	10	10	9	8	2
Recreation, leisure and facility guides	39	49	52	63	64	68	41	100	97	95	4
Domestic culture and home life	12	20	24	29	31	45	48	50	55	53	40
Dictionary and supplementary files	33	26	30	29	28	24	22	23	23	23	23
Place names, maps and addresses	11	7	5	5	8	19	20	31	48	66	68
Others	42	57	53	62	55	42	35	37	32	33	19
Subtotal	763	871	927	1,028	1,001	877	807	915	877	896	695
Natural science and technology
General	73	61	58	68	68	68	62	59	59	63	75
Patents	82	93	95	112	137	118	120	124	138	140	143
Medicine, pharmaceuticals, biotechnology, biology	164	178	152	196	216	191	185	184	178	213	228
Chemistry	101	106	106	121	121	96	77	69	79	78	79
Physics	6	7	7	7	7	7	6	6	8	8	7
Mathematics	4	3	3	3	3	4	4	4	4	3	3
Electricity, electronics and information	130	128	128	134	138	36	39	55	41	58	41
Machinery	18	15	19	21	22	14	9	9	9	10	9
Construction (civil engineering, architecture)	18	19	18	22	24	18	21	24	32	35	29
Space, earth and marine	28	31	32	35	32	12	10	12	11	11	10
Nuclear power	4	4	4	8	7	7	5	5	5	5	5
Environment and pollution	48	55	51	55	53	21	22	16	14	13	13
Energy and resources	41	48	61	62	61	17	17	18	18	12	13
Agriculture	13	12	14	14	14	15	13	13	11	13	18
Meteorology	12	12	5	4	5	9	7	4	3	4	6
Metals and materials	38	38	33	38	38	31	25	25	30	35	32
Food	8	12	12	15	13	7	8	8	8	8	11
Fibers, lumber and pulp	5	4	4	4	4	5	5	6	8	6	6
Others	23	19	17	21	22	14	12	12	13	16	15
Subtotal	816	845	819	940	985	690	647	653	669	731	743

Source: "Directory of Japanese Database," METI

12

Table I-2-b Trends in Distribution by Category of Databases Accessible in Japan (actual numbers)

	Subcategory / FY	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Social and Cultural Science	General	9	10	6	12	13	12	10	24	23	26	39
	Education	12	12	10	11	9	5	4	3	2	3	4
	Sociology	7	7	6	6	7	6	6	6	6	7	6
	Population statistics	32	34	39	40	43	50	55	60	76	84	100
	Law	1	2	2	2	2	4	2	1	3	7	5
	History and political science	2	4	5	5	6	2	2	2	2	4	7
	Arts (movies, music)	8	10	8	11	10	6	6	9	8	6	6
	Psychology	2	2	1	1	1	5	5	6	5	7	8
	Linguistics	2	2	2	2	2	4	4	5	5	5	3
	Philosophy	1	1	1	1	1	1	1	0	0	0	1
	Religion	8	8	7	6	6	1	2	2	2	2	2
	Others	6	5	7	8	7	4	4	4	3	5	5
	Subtotal	90	97	94	105	107	100	101	122	135	156	186
Business	General	93	109	110	105	118	118	116	114	89	91	88
	Marketing and products	132	120	121	133	139	116	97	105	67	68	59
	Economy (overseas)	73	71	68	71	70	48	38	33	20	16	16
	Economy (Japan)	89	87	92	84	79	101	120	112	120	125	122
	Corporate finance and profiles (overseas)	235	256	267	257	257	165	161	142	108	104	108
	Corporate finance and profiles (Japan)	88	88	92	82	79	80	87	85	90	89	98
	Accounting and management	15	16	17	31	29	19	17	17	11	12	11
	Finance, securities and foreign exchange	156	154	161	168	196	153	166	151	145	128	113
	Energy industry	15	15	13	17	19	5	4	4	2	2	2
	Communications and broadcasting	28	38	53	65	62	4	4	4	3	3	1
	Agriculture, forestry and fisheries	9	5	4	4	4	4	3	2	2	6	6
	Chemical industry	10	15	13	12	12	10	9	8	8	10	16
	Labor	6	9	13	13	17	11	12	16	18	23	24
	Education	-	-	-	-	-	0	0	1	1	2	7
	Architecture and construction	6	5	5	8	7	3	0	1	0	0	1
	Transportation	23	27	27	27	26	8	7	10	10	9	12
	Sales and services	15	15	16	15	13	1	2	5	5	6	8
	Distribution and real estate	9	7	5	4	5	5	5	5	5	5	1
	Others	57	69	83	78	71	24	19	17	15	15	13
	Subtotal	1,059	1,106	1,160	1,174	1,203	875	867	832	719	714	706
Others		71	61	61	61	54	56	56	59	58	55	49
	Grand total	2,799	2,980	3,061	3,308	3,350	2,598	2,478	2,581	2,458	2,552	2,379

Source: "Directory of Japanese Database," METI

13

Table I-3   Trends in Distribution of Databases by Field (actual numbers)

(%)

Field	FY1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
General	27.3	29.2	30.3	31.1	29.9	33.8	32.6	35.5	35.7	35.1	29.2
Natural science and technology	29.2	28.4	26.8	28.4	29.4	26.6	26.1	25.3	27.2	28.6	31.2
Social and cultural science	3.2	3.3	3.1	3.2	3.2	3.8	4.1	4.7	5.5	6.1	7.8
Business	37.8	37.1	37.9	35.5	35.9	33.7	35.0	32.2	29.3	28.0	29.7
Others	2.5	2.0	1.9	1.8	1.6	2.2	2.3	2.3	2.4	2.2	2.1
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Total number of databases	2,799	2,980	3,061	3,308	3,350	2,598	2,478	2,581	2,458	2,552	2,379
Business (revised)	51.8	52.1	53.7	52.1	51.6	51.6	53.3	50.4	46.3	44.6	46.2

Source: "Directory of Japanese Database," METI

Note: "Business" was revised because "Newspapers/journals/news" and "Who's who and organizations" are frequently used in the business field, and so these were added to the business field and the figures were revised and tabulated accordingly.

3. Comparison of Database Market Conditions in Japan and the United States

This section compares database market conditions in Japan and the United States. Figure I-4 shows a comparison of the size of the Japanese database market with that of the U.S. The outside graph represents the situation in the United States and the inside graph that in Japan. The following six items were used for comparison.

[1] Number of producers

[2] Number of participating companies

[3] Number of accessible databases

[4] Number of domestic databases

[5] Number of business databases

[6] Number of fact databases

Of the six items, items [1] and [2] concern the size of the database service industry, indicating the strength of the database industry. And items [3] through [6] concern the distribution of databases, indicating the size of the database market.

The United States maintains an overwhelming edge over Japan in terms of strength and distribution.

14

□ United States (outside)
□ Japan (inside)

① No. of producers (U.S. companies only)

1,567

⑥ No. of fact databases
11,378

② No. of participating companies and organizations (U.S. companies only)
1,642

144

540

1,342

706   1,314

2,379

⑤ No. of business databases
1,636

③ No. of accessible databases
16,417

④ No. of domestic databases
7,049
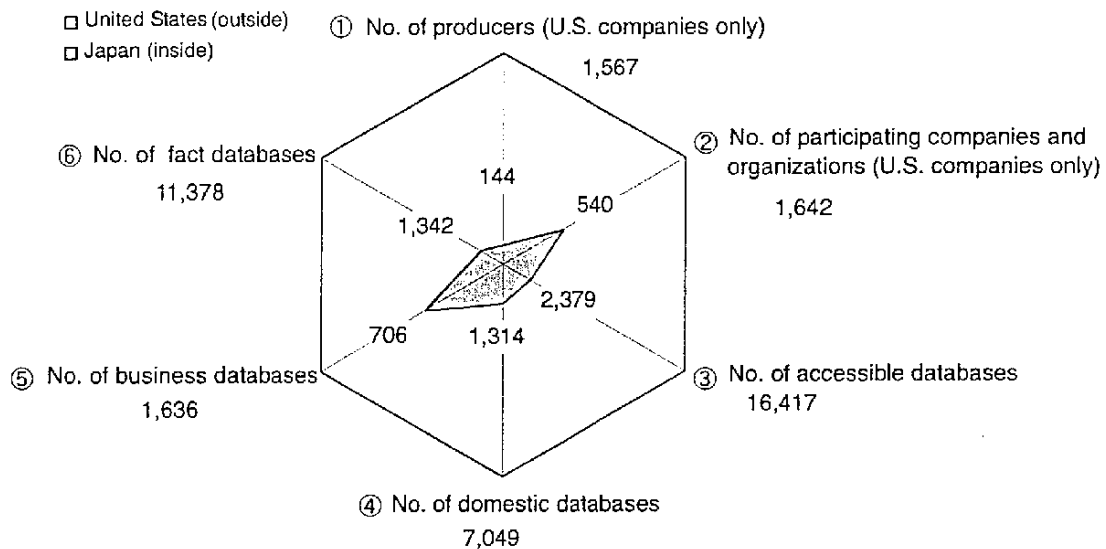
Note 1:   The figures for the United States for [1] through [6] were tabulated from the 2002 edition of "Gale Directory of Databases" Vol. 1 (Online) and Vol. 2 (CD-ROM), based upon an article by Martha E. Williams. The figures for Japan were tabulated from the 2001 edition of "Directory of Japanese Databases" and the 2001 edition of "Survey of Selected Service Industries."

Note 2:   The figures for [3] through [6] are based upon actual numbers.

Fig. I-4   Graphical Comparison of Database Situation in the United States and Japan (2002)

3.1 Comparison Between Japan and the United States in Strength of Database Industry

There is a wide difference between Japan and the United States in the strength of the database industry, and the difference is increasing gradually.

The U.S. economy remained sluggish due to the increases in crude oil prices and the slip in consumer confidence since the autumn of 2002. In the second quarter of 2003, however, it began to improve as a result of the large-scale tax reduction/low interest measures adopted by the Bush Administration. Hence, the above-mentioned difference between the two countries will likely continue to widen. In any event, the development of the U.S. economy will greatly affect the database markets in both countries (Table I-4).

Table I-4　Strength of the U.S. Database Industry Relative to Japan (2002)

(Japanese database industry = 1)

Criteria	United States	
	2002	2001
(1) No. of producers	10.9	10.6
(2) No. of participating companies and organizations	3.0	2.9
(3) No. of accessible databases	6.9	4.7
(4) No. of domestic databases	5.4	4.6
(5) No. of business databases	2.3	2.3
(6) No. of fact databases	8.5	5.7

Note:　The source and·tabulation method are the same as for Fig. I-4.
　　　The number of participating firms in (2) is tabulated for Japan by using the sum total of
　　　"independent business establishments" and "head office."

In comparison with the United States, where the foundation of the database service industry has been established over a long period, the database industry in Japan has not yet been firmly grounded. For this reason, it is difficult for a firm to maintain a viable business solely with the database services, as evidenced by the fact that most Japanese database service firms are concurrently engaged in other fields. In addition, with the economic deflation of recent years, fewer firms are entering the market today. Given the balance between investment cost and demand, overall financial tightening, and general business stagnation, the future of the Japanese database industry does not now appear to be so bright, and it will be a very difficult task to narrow the gap between the two countries. The U. S. database industry constantly grows faster than Japan's, and the widening gap in content should be seen as a gap in the ability to make use of information.　Thus, efforts must be redoubled to develop Japan's database industry.

Databases compiled in the United States are distributed not only to meet domestic demand but also throughout the world, and this fact is reflected in the figures that express the strength of the database industry. This should be kept in the mind.

3.2　Comparison of Distribution of Databases in Japan and in the United States

There is also a wide difference between Japan and the United States in the distribution of databases, a difference which widened in 2001. As stated in the preceding section, the U.S. economy is growing. How the Japanese economy manages to recover on the strength of the recent upturn in corporate performance will be key to narrowing the disparity.

The difference in the "number of domestic databases" between the database industries in the two countries has fluctuated over the past few years, so it is difficult to identify a clear trend. This may be because while most of the databases developed in Japan are utilized only in Japan, those developed in the United States are utilized throughout the world on the strength of the language used, namely, English which is now the global standard, and the U.S. economy which is the center of the world economy.

16

II. Trends in the Database Service Industry

1. Introduction

This chapter deals with the supply situation of database services and future trend on the basis of the findings of the "Survey of the Status of Japanese Database Services: Vendors Edition," which the Database Promotion Center, Japan, conducted in October 2003. This survey was carried out by sending a questionnaire to a total of 179 companies, of which 105 responded (response ratio 58.7%).

2. Business Mode of Database Service Providers

According to this survey, 38.1% of the database service providers fall into the category of "information processing/supply," 21.0% "newspaper and publishing," and 19.0% "public services." The category of "public services" includes research organizations, while "other services" (9.5%) includes restaurants, hotels, providers of services for business establishments, think tanks, and advertising agencies, etc. The fact that, although limited in number, commercial businesses and companies in other manufacturing responded to this survey is evidence that companies in various industries are engaged in the database service business (Fig. 10-2-1 in Reference Data).

According to the "Directory of Japanese Databases" (FY2002), in terms of the business mode of database service providers, the full-time producer companies make up 34.1% (vs. 32.3% in the previous year), and the producer companies that concurrently serve as distributors make up 28.5% (vs. 25.8% in the previous year). These two types of companies comprise 62.6% (vs. 58.1% in the previous year), or well over half of the total. In addition, companies that are fully engaged in providing information search broker services are placed third, with 10.1%.

3. Positioning of Database Service

The respondents, many of whom are providing database service in addition to their main operations, were asked to rate the importance of database service within the scope of their operations. Of the total number of respondents, 40.4% replied, "Database service is our main business now," while 46.5% replied, "Database service will be our main business in the future," suggesting that the number of companies which regard database service as their main business will increase. Compared with the previous survey (conducted in October 2002), the percentage of respondents who consider database service as their "main business now" decreased by 5.7 percentage points while the percentage of those who expect database service to become their "main business in the future" decreased by 0.5 percentage points. Is this because many database service providers are uncertain about the social and economic outlook or because there are some problems with the database service business? Although it is difficult to determine which is the case, as more

companies are now providing database service free of charge and in view of the concerns over security, many providers are surely finding it increasingly difficult to profit by specializing in the provision of database service.
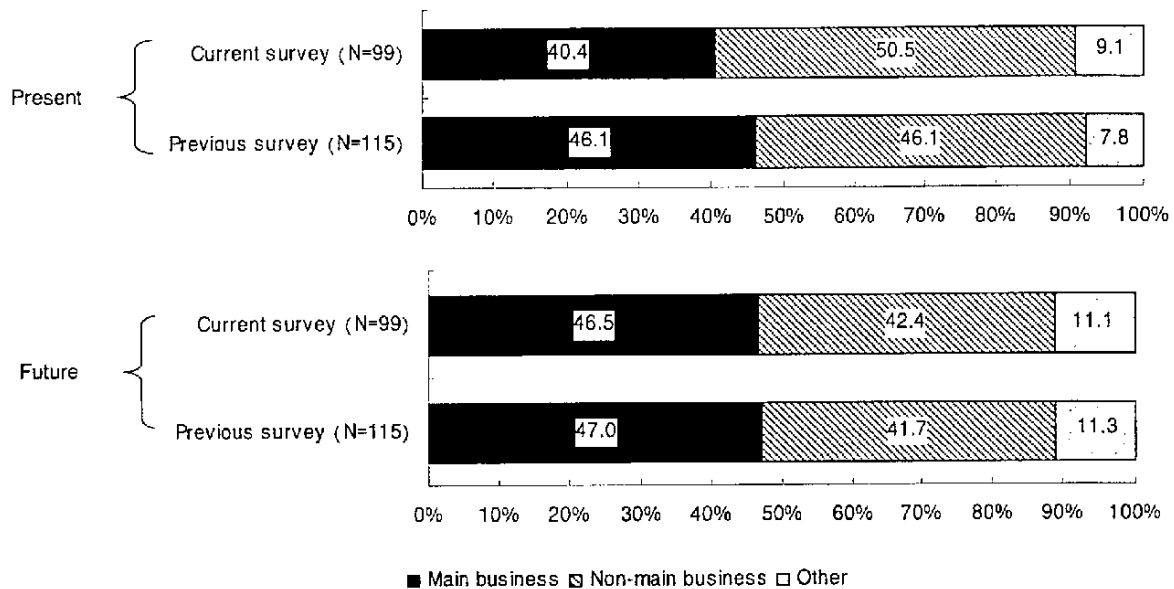
Fig. II-1 Positioning of Database Business

4. Sales of Database Services

4.1 Ratio of Sales of Database Service to Total Company Sales

If the question as to whether database service is the main business, which was mentioned earlier, indicates database service providers' subjective view of the viability and growth potential of the database service business, the percentage of sales of database service is a more objective and important indicator.

As shown in Fig. II-2, the average ratio of sales of database service to total company sales is 28.0%, a decrease of 0.8 points from the previous survey. The percentage distribution of groups classified by percentage of sales of database service shows that the largest group is the "under 1%" group, representing 34.1% of the total, followed by the "over 90%" group, representing 14.1%, the "5-10%" group and the "15-50" group, each representing 11.8%, and "the 10-15%" group and the "50-70%" group, each representing 8.2%. The combined total percentage of the groups whose sales of database service account for less than 10% is 51.8%, which is an increase of 2.4 percentage points from the previous survey. The percentage of sales of database service as a whole declined slightly and the number of providers with a percentage of sales of database service of less than 10% increased slightly. Overall, the database service business has neither grown nor shrunk.

18

Overall (N=85)
(Average: 28.0%)

34.1	5.9	11.8	8.2	11.8	8.2	2.2	14.1

0%   20%   40%   60%   80%   100%

☒ Less than 1% ☐ 1-5% ☒ 5-10% ☐ 10-15% ▥ 15-50% ▨ 50-70% ☐ 70-90% ▤ 90% or more
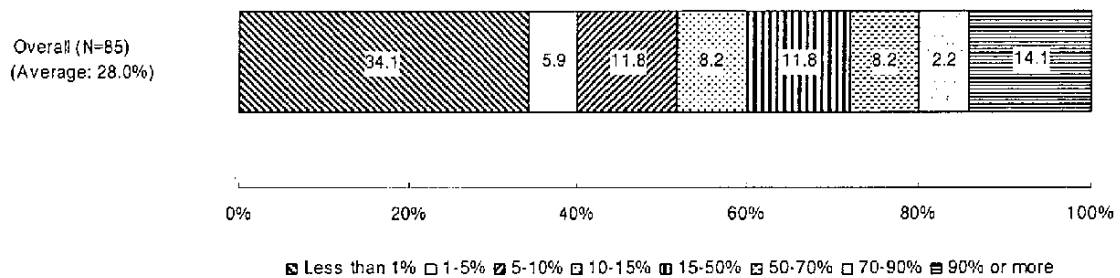
Fig. II-2   Distribution of Ratio of Database Sales to Company's Gross Sales

4.2 Ratio of Domestic and Overseas Database Sales

The ratio of sales of domestic databases to the total sales of databases is 85.6%, almost the same as in the previous survey (84.8%)(Fig. II-3). On the other hand, the percentage of sales of overseas databases is 14.4%. However, looking at the providers show that providers for whom database service produce less than 10% of sales account for 76.7% of the total. This means that few providers are enjoying large sales. Since 1988 when the survey was started the percentage of sales of domestic databases has remained at more than 85%, so domestic databases are still the main source of revenue for the companies included in the "Directory of Japanese Databases."

85.6	14.4

0%   20%   40%   60%   80%   100%

▨ Domestic databases ☐ Overseas databases
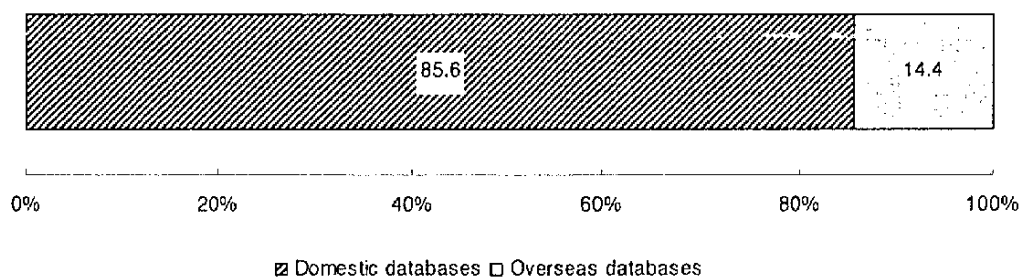
Fig. II-3   Ratio of Database Sales (Domestic and Overseas)

19

4.3 Ratio of Sales of Databases by Method of Supply

This section examines the ratio of sales of databases by method of supply.

As shown in Fig. II-4, "the Internet accessed through personal computers" remains top, as in the previous survey. The combined total percentage of this medium and "mobile telephone and PDA" is 43.0%, an increase of 1.3 percentage points from the previous survey. The Internet has thus become the main medium of supply of database service. "Online" represented nearly 50% of the total up until a few years ago, but with the spread of the Internet, it now accounts for only 10.0% (16.8% in the previous survey) of the total. This confirms that "online" is in decline. On the other hand, the percentage of "CD-ROM" increased by 4.8 percentage points from the previous survey to 29.2%, and still ranks second. The combined total percentage of "CD-ROM" and "DVD" was approximately 30%., so "CD-ROM" is still a main "offline" medium.

On-line ■ Internet (for mobile phones and PDAs) ▤ DVD ☐ Internet (for personal computers) ▨ CD-ROM ☐ MT ▨ Others
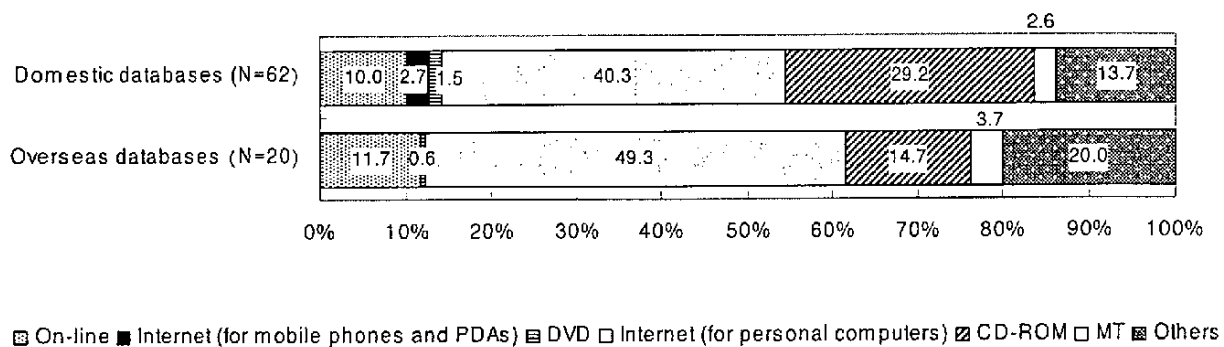
Fig. II-4 Ratio of Database Sales by Method of Supply

20

4.4 Annual Rate of Growth of Sales of Database Service

Sales of databases at the 78 companies that responded increased by 17.1% on average, which compares favorably to 12.7% shown in the previous survey (Fig. II-5).
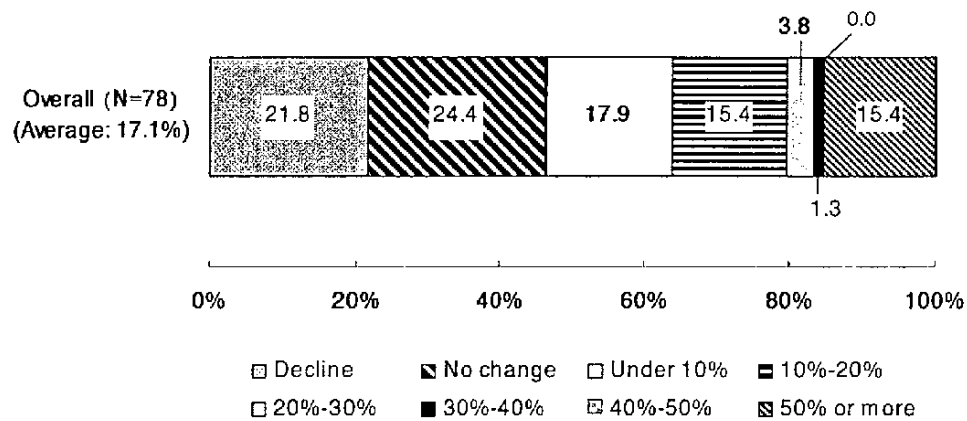
Fig. II-5   Distribution of Year-on-Year Changes in Database Sales

5. Promising Media and Types of Data

5.1 Media of Supply

The most promising medium is the Internet, which has already passed the first stage of diffusion. In this survey, the percentage of the medium increased by approximately 5 percentage points to 93.2%, outdistancing the second ranking medium, and expectations for this medium remain very high (Fig. II-6). "Mobile phone" ranked second as in the previous survey, with high hopes also for this medium.

On the other hand, "mobile computer and PDA" (percentage decreased from 31.0% to 27.2%) and "car navigation services" (from 13.8% to 8.7%) did not fare so well. Although in widespread use, these media's growth potential as database service supply media seems to be limited now.

As to offline media, "CD-ROM" has passed the growth period, but the percentage of "DVD" increased from 25.0% to 32.0%. Increasingly high expectations are placed on this new medium, which is better poised to meet the demands of the multimedia age.
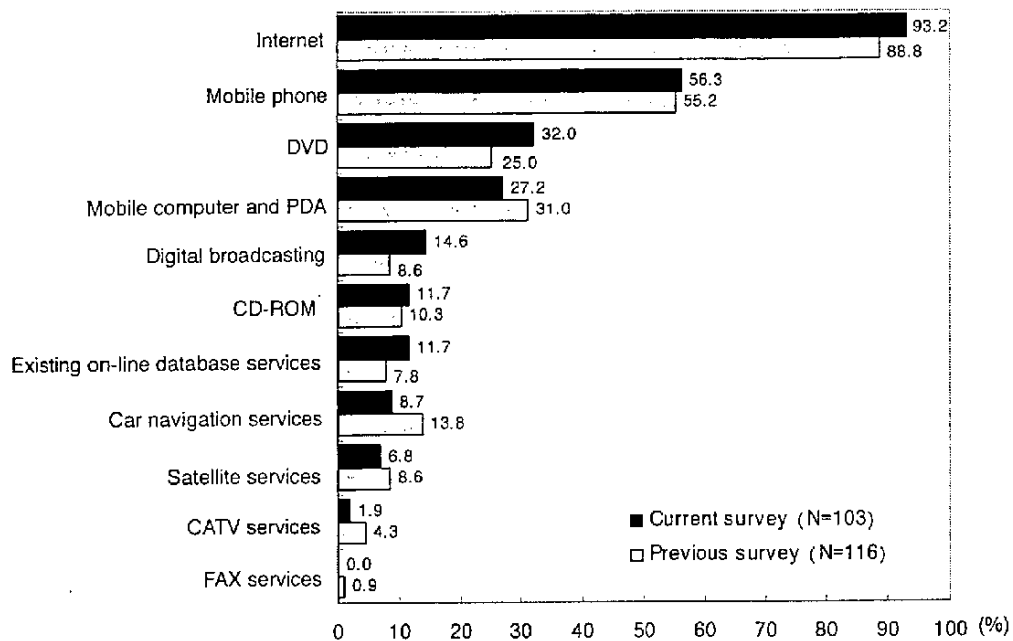
Fig. II-6 Media with Higher Growth Potential (multiple replies)

5.2 Types of Data to Be Supplied in the Future

Of the total number of providers who responded to this survey, 28 replied to the question about promising types of database service, giving detailed descriptions of promising types and contents of data. They pointed out a total of 57 promising types of data. of which "text data" ranked first (60.7%), followed by "multimedia" (53.6%), "graphics/images" (46.4%), and "numerical data" (42.9%).

6. Distribution of Databases

With regard to the method of distribution of databases, producers develop and distribute their respective databases or rely on other companies for their distribution. The former method is called "producer/distributor" within the industry.

Of a total of 73 companies which responded to this survey, 91.8% replied, "We distribute our database by ourselves," and the remaining 8.2% replied, "We rely entirely on other companies for distribution of our database." On the other hand, 31.5% of the respondents replied, "We service our customers, but also ask other companies to service our customers." These figures show that many database service companies in Japan are operating both as producers and distributors.

In the case of those companies which rely on other companies for distribution of their respective databases, the average number of companies which are commissioned by these companies to

22

distribute databases was 8.6 companies, which compares favorably to 5.6 companies in the previous survey.

(Average number of companies requested to distribute: 8.6 companies)
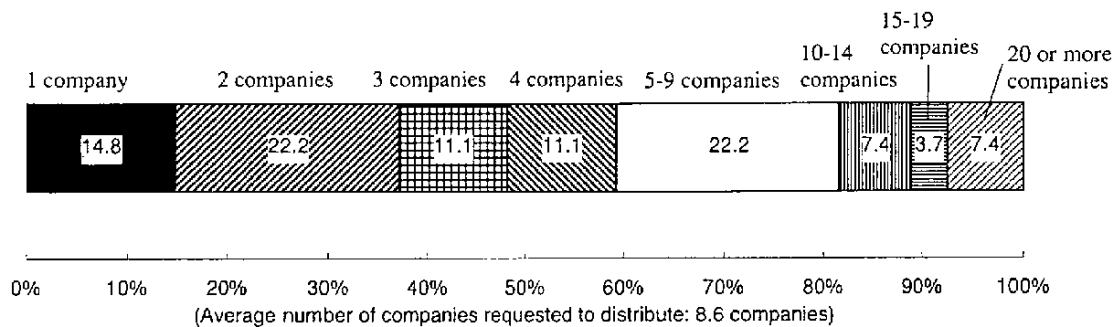
Fig. II-7   Number of Database Distributors Consigned by Producers (N=27)

7. Problems Encountered in Database Construction

The producer companies were asked to comment on the problems encountered in database construction. As shown in Fig. II-8, "cost and time which are necessary for data collection/entry" and "high cost of maintenance of the database constructed" are their major concerns. During the present economic slowdown, it is essential for producer companies to strengthen their capability by focusing on service cost. On the other hand, many respondents pointed out "shortage of system developers, including programmers," so vendor companies must secure human resources, especially experts. as well as deal effectively with cost problems. The percentage of respondents pointing out "concerns about data security" increased from 23.1% to 30.3%. If they are to comprehensively ensure data security, they will have to tackle the problems of human resources and increases in cost, which are indeed difficult problems for vendor companies.
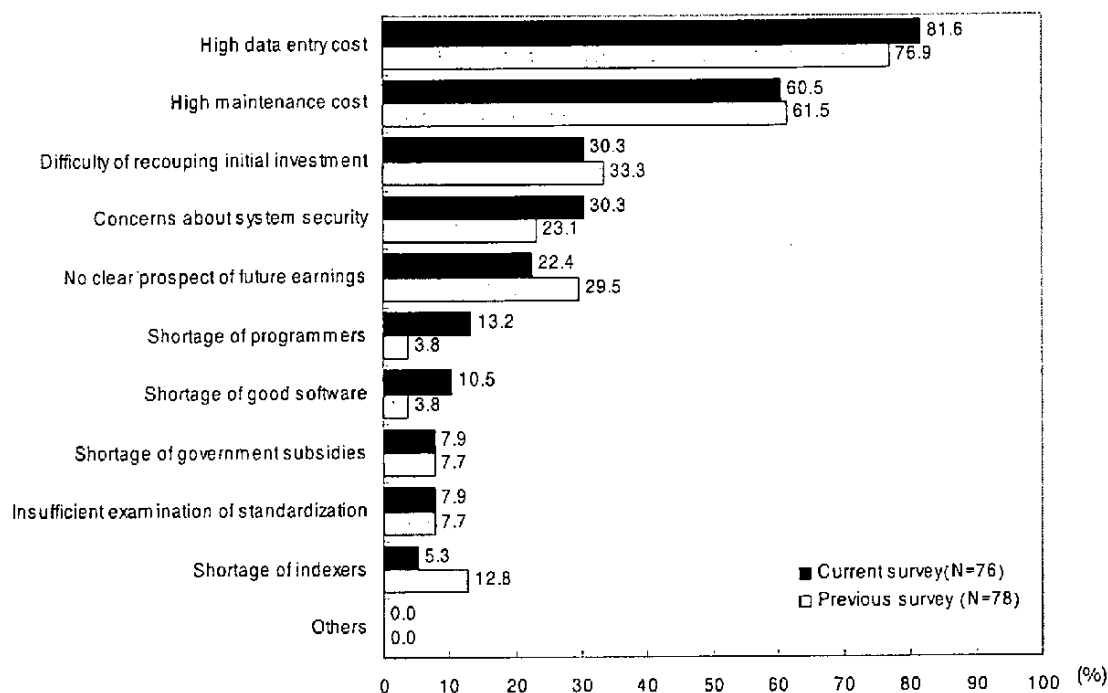
Fig. II-8 Awareness of Problems in Database Construction (multiple replies)

8. Making Use of Public Data

Statistical data/information and bibliographical information which the central government and other national public bodies compile and make public (white papers, reports of advisory councils, statutes, etc.) are highly reliable and therefore are important sources of information. Of a total of 77 respondents, 38, or 49.4%, replied that they are utilizing these public data as source data.

To the question about the types of public data utilized, in the area of text data 41.7% (35.1% in the previous survey) cited "Kampo (the Official Gazette of the Japanese Government). It should be noted that an equal percentage of respondents (16.2% in the previous survey) cited "local governments' data," indicating that needs have grown for data developed by local governments as well the central government. The percentage of respondents who cited "government white papers, commissions and council reports" decreased from 32.4% to 30.6%. As a whole, however, the needs for government agencies' data/information remain high (Fig. II-9). As shown in Fig. II-10, "general business/industry" tops the list of numerical data, being cited by 36.1% of respondents, followed by "population, employment and labor" (30.6%), "domestic economy, business situation, finance" (27.8%), "corporate information" (27.8%), and "international accounts and trade" (22.2%). Over the past few years, "general business/industry" has been at the top of the list of numerical data, and numerical data that deal with economy-related matters rank higher in the list (first to fifth ranks).

Official gazette 41.7
Local government data 41.7
Government white papers, commissions and council reports 30.6
Government press releases 22.2
City planning data (including map information) 19.4
Patent information 11.1
Judicial precedents, etc. 8.3
Others 13.9

0 10 20 30 40 50 (%)
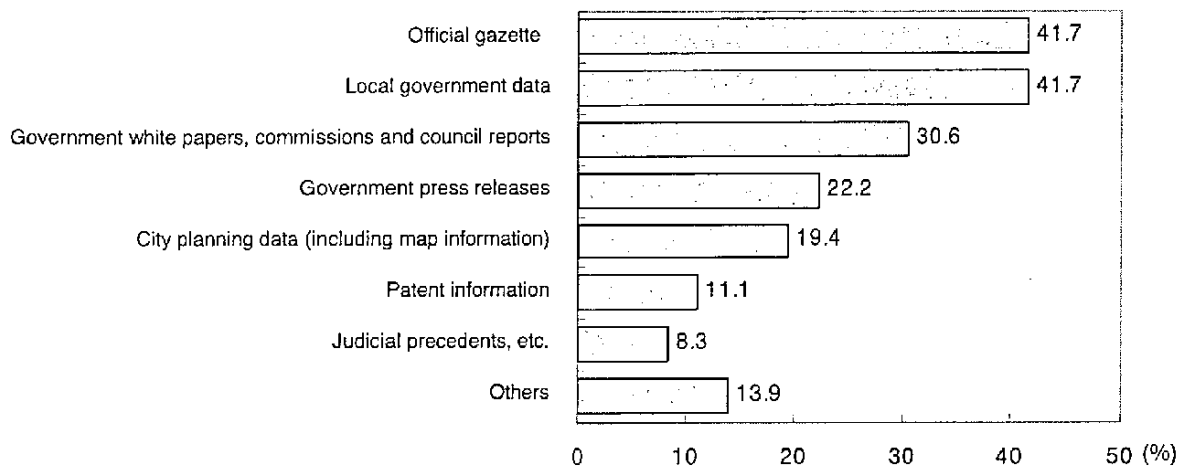
Fig. II-9 Types of Public Data Used Now [Text Data] (N=36, multiple replies)

General business/industry 36.1
Population, employment and labor 30.6
Domestic economy, business situation, finance 27.8
Corporate information 27.8
International accounts and trade 22.2
Others 2.8

0 10 20 30 40 50 (%)
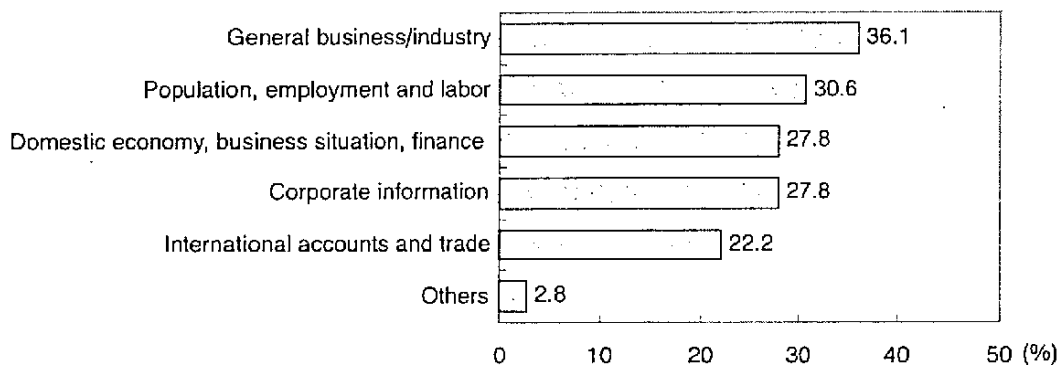
Fig. II-10 Types of Public Data Used Now: [Numerical data] (N=36, multiple replies)

9. Internet and Database Services

9.1 Status of Provision of Services

As shown in Fig. II-11, of a total of 89 respondents, 74, or 83.1% (76.3% in the previous survey) replied, "We are providing database services through the Internet." The fact that more than 80% of respondents are providing database services through the Internet confirms vendors' full-fledged entry into the Internet market. The percentage of respondents who are "planning to provide" internet-based service was 3.4%, a slight decrease from the previous survey. Most vendors are thus keen to provide services through the Internet.
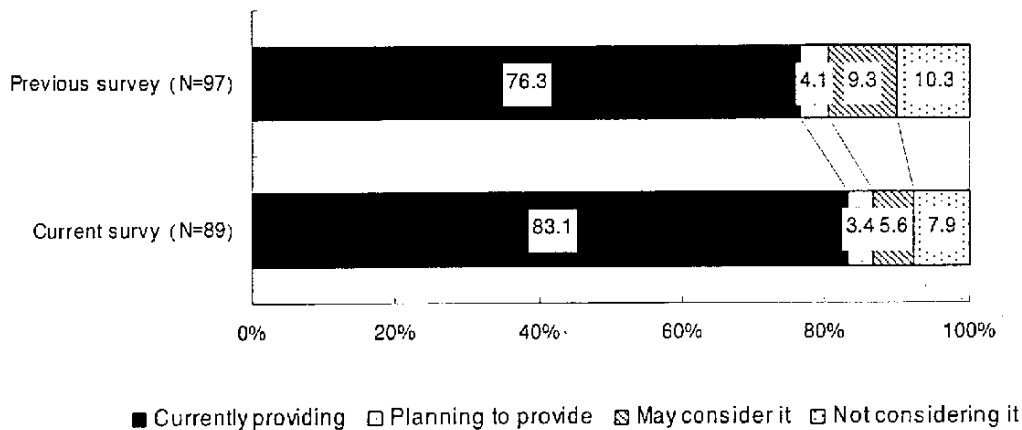
Fig. II-11 Current Utilization of Internet Database Services

9.2 Problems Involved in Provision of Service

With more than 80% of database service vendors providing services through the Internet, many vendors are facing various concerns and problems. (This is true for both vendors who provide database services through the Internet and those who do so through dedicated lines.) Figure II-12 shows a summary of the concerns and problems which were cited by 103 respondents. The most commonly cited concern was: "Homepages and other database services now can be accessed free of charge, which is hampering the growth of commercial database service" (50.5%). This was followed by "feel threatened by illegal access and virus invasion from outside" (45.6%), "concerned about protection of database copyrights (intellectual property)" (44.7%), and "possibility of free public database services affecting commercial services" (40.8%). In the case of vendors who are seeking to develop their database services making use of the Internet, "free database services and free homepages on the Internet" are threatening the foundation of their businesses. In addition, the rate of response to this survey may have been affected by frequent, detailed news reports on virus invasion and illegal access to computer systems.
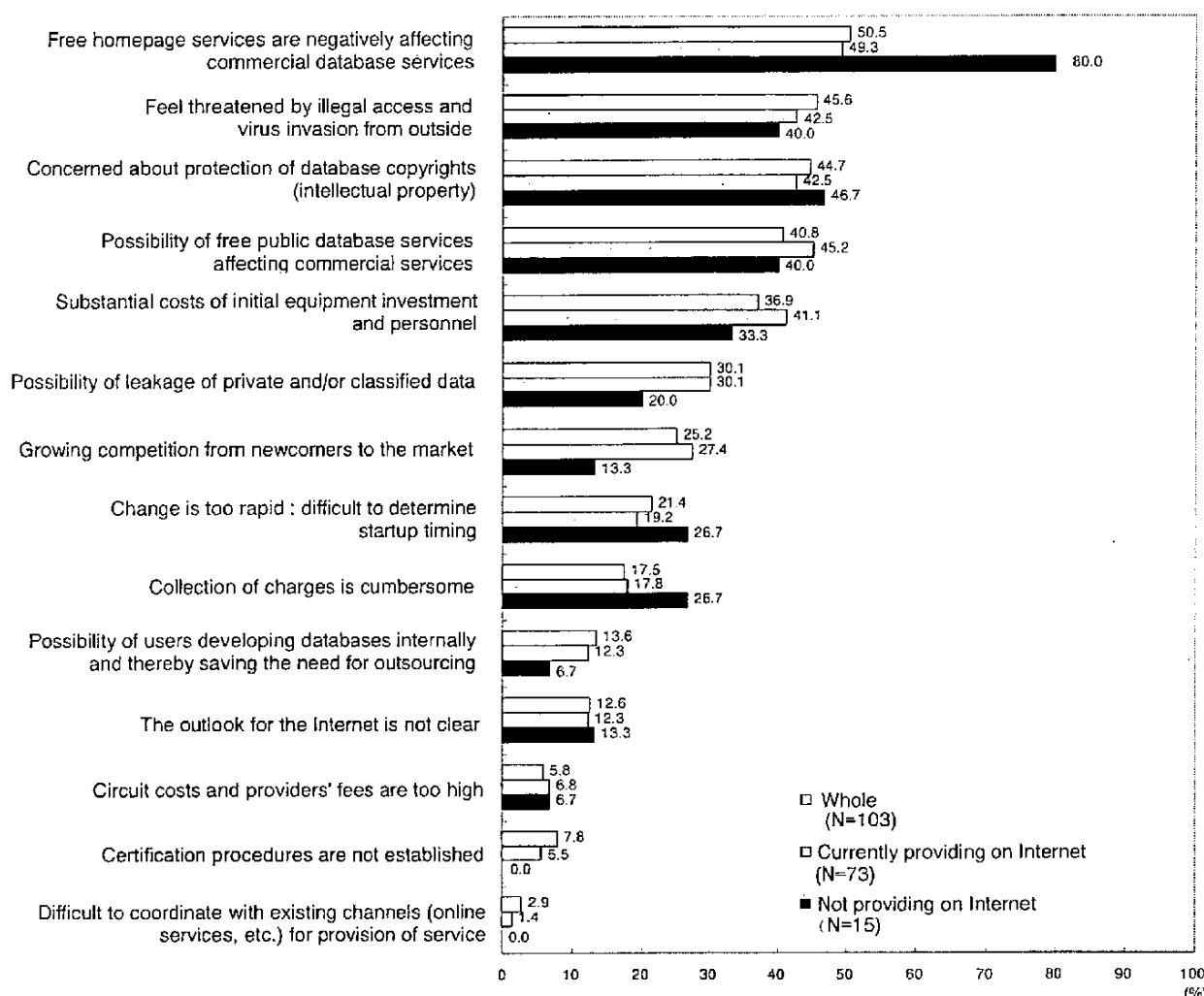
Fig. II-12 Concerns and Problems Relating to Database Services Provided Through
the Internet (multiple replies)

10. Supply of Databases Services on CD-ROM and DVD

10.1 Percentage of Providers of Databases on CD-ROM and DVD

As shown in Fig. II-13, at present 59.8% of database vendors are providing databases on CD-ROM and DVD, a slightly higher figure than in the previous survey. It is concluded that the percentage of such database vendors has leveled off.

10.2 Breakdown by Field of Total Number of Providers of CD-ROM and DVD Databases

A breakdown by field of providers of databases on CR-ROM and DVD shows that 39 providers are supplying "business" databases, 20 "general" databases, 17 "natural science and technology" databases, and 1 "other" databases. No providers are supplying "cultural and social sciences" databases. Compared with the previous survey, there was a uniform decrease in the number of

27

providers of these databases. The number of providers of "business" databases decreased by 4, that of providers of "general" databases decreased by 1, that of providers of "natural science and technology" databases decreased by 2, that of providers of "other" databases decreased by 1, and that of providers of "cultural and social sciences" databases decreased by 5.

	Supplying companies	Non-supplying companies
1992(N=99)	26.3	73.7
1993(N=91)	30.8	69.2
1994(N=104)	36.5	63.5
1995(N=93)	41.9	58.1
1996(N=79)	46.8	53.2
1997(N=74)	56.8	43.2
1998(N=82)	53.7	46.3
1999(N=95)	54.7	45.3
2000(N=82)	59.8	40.2
2001(N=77)	58.4	41.6
2002(N=95)	57.9	42.1
2003(N=87)	59.8	40.2

■ Supplying companies   □ Non-supplying companies

Fig. II-13   CD-ROM and DVD Database Supply

11. Present State of Japanese Databases Supplied Overseas

The Japan Database Industry Association (DINA) was commissioned by Database Promotion Center, Japan (DPC) in June 2003 to conduct the 17th survey on their members, DPC supporting members and other organizations involved in database services. Survey forms were sent to 313 organizations, and replies were obtained from 138 (response rate: 44.1%).

Seventy-two of the respondents made their databases accessible from overseas and 8 were planning overseas services. The past twelve years had shown a steady increase in the number of Japanese databases available to overseas users, but the figure for July 1999 stood at 403, down 48.0% from the previous year. The decrease was attributed to NIFTY's narrowing of its operations from the pre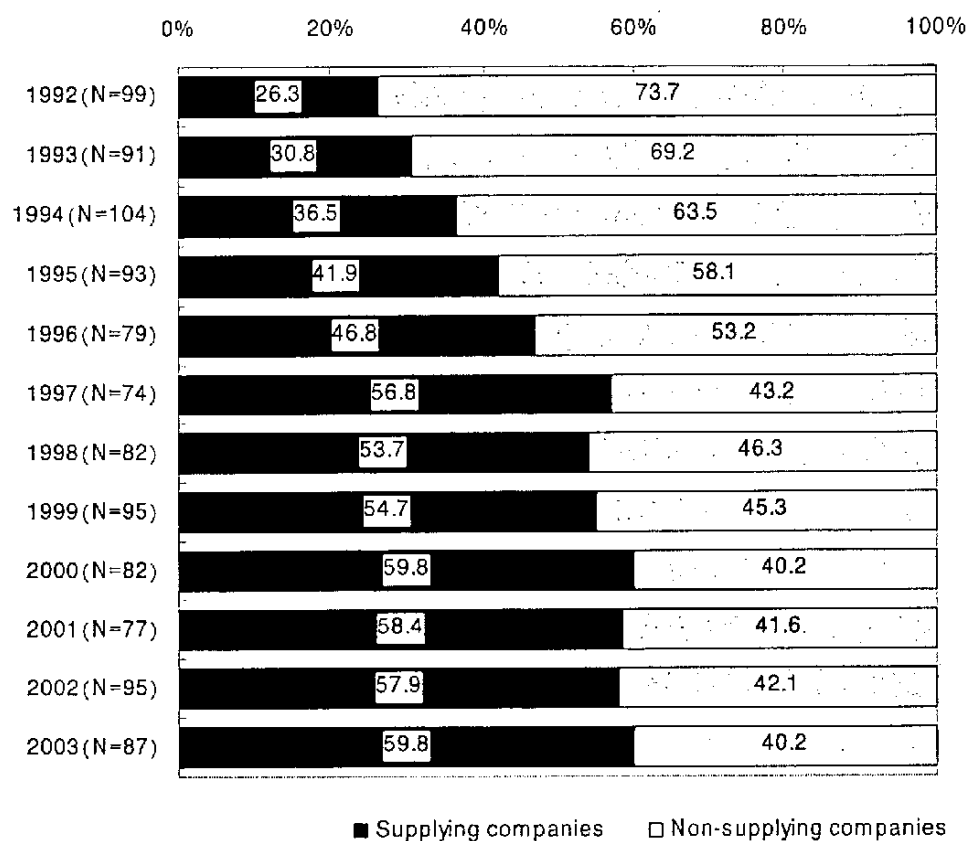vious year's 452 databases to 3. In the 2003 survey, the number of Japanese databases available for overseas users stood at 350 (based on actual numbers), a decrease of 12.7% from the previous year (Table II-1 and Fig. II-14).

By category, 350 databases are classified into the following: 71 (20.3%) in science/technology; 127 (36.3%) economy/business/finance; 140 (40.0%) general; 3 (0.9%) science/technology + economy/business/finance; 1 (0.3%) science/technology + general; 6 (1.7%) science/technology + economy/business/finance + general; 2 (0.6%) economy/business/finance + general.

By language, Japanese accounts for 196 (56.0%) databases, English 47 (13.4%), others 6 (1.7%), Japanese + English 85 (24.3%), Japanese + English + others 13 (3.7%), Japanese + others 1 (0.3%), and English + others 2 (0.6%).

By distribution media, 44 (12.6%) of the databases were distributed on the Internet, 136 (38.9%) online, 7 (2.0%) MT/FD, 23 (6.6%) CD-ROM/MO/DVD, 57 (16.3%) Internet + online, 5 (1.4%) Internet + MT/FD + CD-ROM/MO/DVD, 5 (1.4%) Internet + CD-ROM/MO/DVD, 3 (0.9%) Internet + online + CD-ROM/MO/DVD, 8 (2.3%) Internet + online + MT/FD + CD-ROM/MO/DVD, 17 (4.9%) online + MT/FD, 16 (4.6%) online + CD-ROM/MO/DVD, 21 (6.0%) online + MT/FD + CD-ROM/MO/DVD, and 8 (2.3%) MT/FD + CD-ROM/MO/DVD.

Table II-1   Number of Japanese Databases Supplied Overseas

Year of Survey / Item	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
No. of databases and proportion	353	374	430	739	833*	403**	437	417	401	350
supplied (%)	(33.7)	(33.3)	(36.9)	(59.5)	(67.9)	(29.5)	(32.0)	(29.4)	(27.0)	(26.6)

*Including 452 NIFTY databases
**Including 3 NIFTY databases
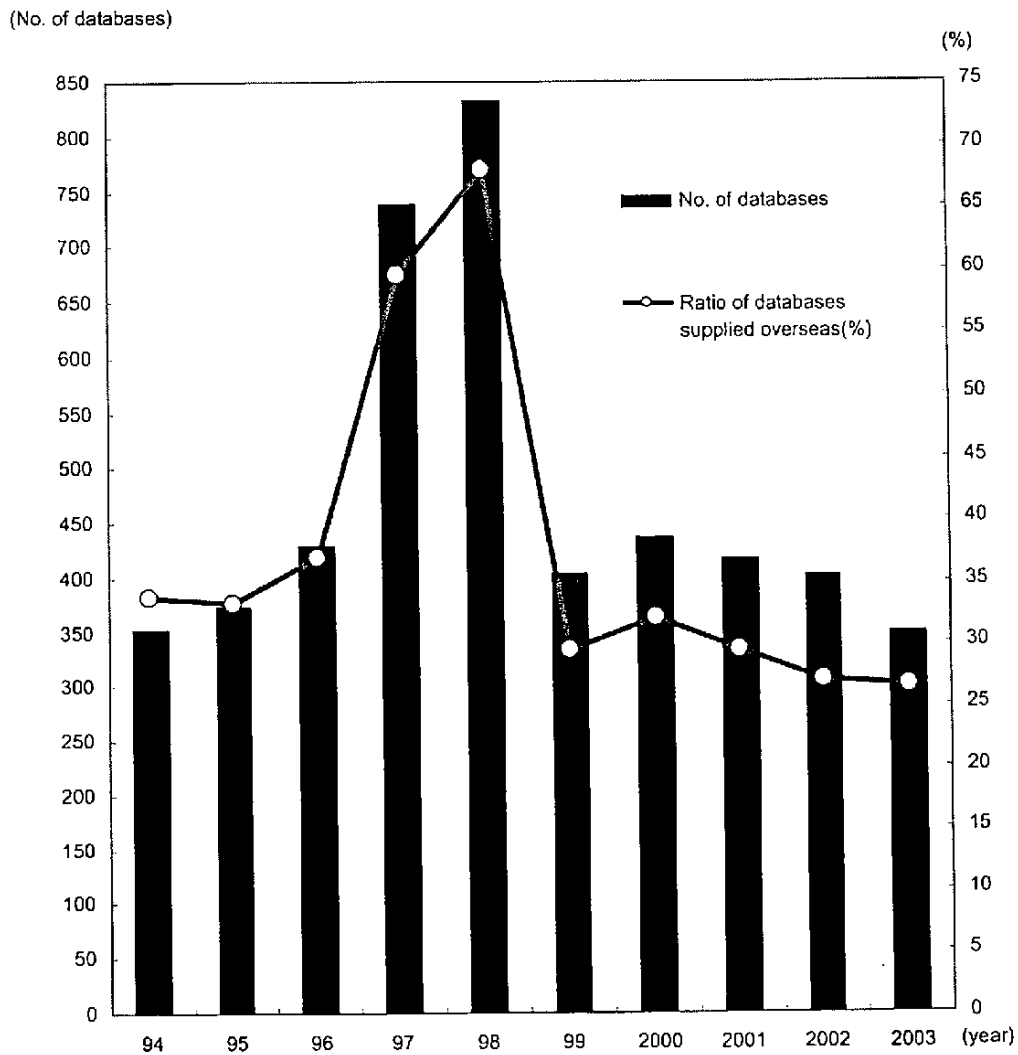
(No. of databases)

(%)

Fig. II-14 Growth of the Number of Databases Supplied Overseas

III. Current Utilization of Commercial Database by Corporations

1. Introduction

On the basis of the findings of the "Survey on the Status of Japanese Databases: Users Edition" which was conducted by the Database Promotion Center, Japan, in October 2003, this section deals with the present situation of the use of commercial databases and users' awareness of commercial databases. A questionnaire was sent to a total of 4,500 companies, of which 594 replied (response rate of 13.2%).

2. Current and Projected Use of Commercial Databases

2.1 Status of Use of Commercial Databases in FY 2002

Of the responding companies, 65.7% use commercial databases, 3.5% used to use commercial databases, and 30.8% have no experience of using commercial databases. As in the previous survey, the percentage of companies using commercial databases fell below 70%, but the percentage itself increased by 2.2 percentage points. Approximately two-thirds of the total number of responding companies are database users, which indicates an upward trend of the use of commercial databases. The percentage of companies which have no experience of using commercial databases decreased. Thus, commercial databases have established themselves as a business tool in many companies.

Of a total of 390 companies which are users of commercial databases, 382 gave descriptions of the degree of use of commercial databases in terms of fees paid. A comparison of the use of commercial databases in fiscal 2002 with that in the previous fiscal year (FY2001) shows that 66.2% of the users mentioned "about the same as last year" (60.6% in the previous survey) and 23.8% mentioned "increased" use (27.1% in the previous survey). The percentage of companies which mentioned "decreased" use was 9.9% (12.2% in the previous survey). As to the percentage distribution of the three differing degrees of use, the percentages of "increased" use and "decreased" use decreased, with only the percentage of "about the same as last year" increasing. This pattern is the same as in the previous survey. In other words, the use of commercial databases is becoming common practice in many companies (Fig. III-1).

Of the 91 companies whose use of database services "increased" and which were asked the reason for the change, 88 replied (Fig. III-2). The most common reason was "we increased the number of contracts with commercial databases that we can use" (44.3%). The percentage was almost the same as in the previous survey. The second-ranking reason was "more personal computers and other terminals, searching has become easier" (38.6%, compared to 43.9% in the previous survey). Next came "business has grown and the information search volume has increased" (22.7%, compared to 31.6% in the previous survey). The percentage decreased by 8.9 percentage points. The decrease in the number of companies where "business has grown" seems to be the main

31

reason for the slowdown in the rate of increase in the use of databases.

On the other hand, all of the 38 user companies whose use of database services decreased and which were asked the reason for the change replied (Fig. III-3). The most common reason was "we now use free homepages more frequently" (44.7%). Although the percentage decreased by 14.4 percentage points from the previous survey, this was again at the top of the list. Next came "reduction in commercial database service contracts" (23.7%), a decrease of 8.1 percentage points. On the other hand, the percentage of "other" reasons (31.6%) increased by as much as 18.0 percentage points. These figures clearly indicate the diversification of the reasons for "decreased" use.

Figure in ( ) indicates the number of respondent companies.

| Whole (382) | 23.8 | 66.2 | 9.9 |

0%   20%   40%   60%   80%   100%

□ Increased   □ Remained at same level as last year   □ Decreased
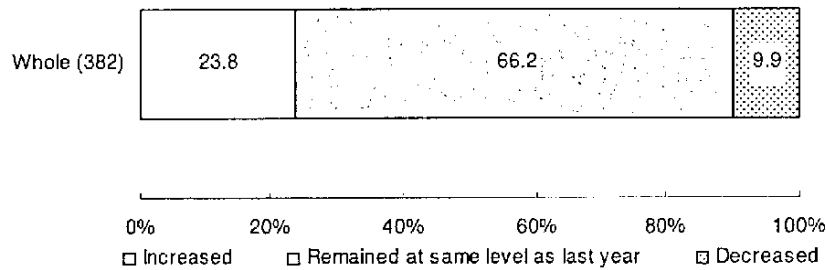
Fig. III-1   Utilization of Commercial Databases in FY2002 Compared to FY2001

	Current survey N=88	Previous survey N=98
Increase in contracts for available commercial database services	44.3	42.9
Increase in PCs and terminals facilitated data retrieval	38.6	43.9
News business increased volume of data retrieval	22.7	31.6
Others	11.4	13.3

0   20   40   60 (%)

□ Current survey N=88
☒ Previous survey N=98
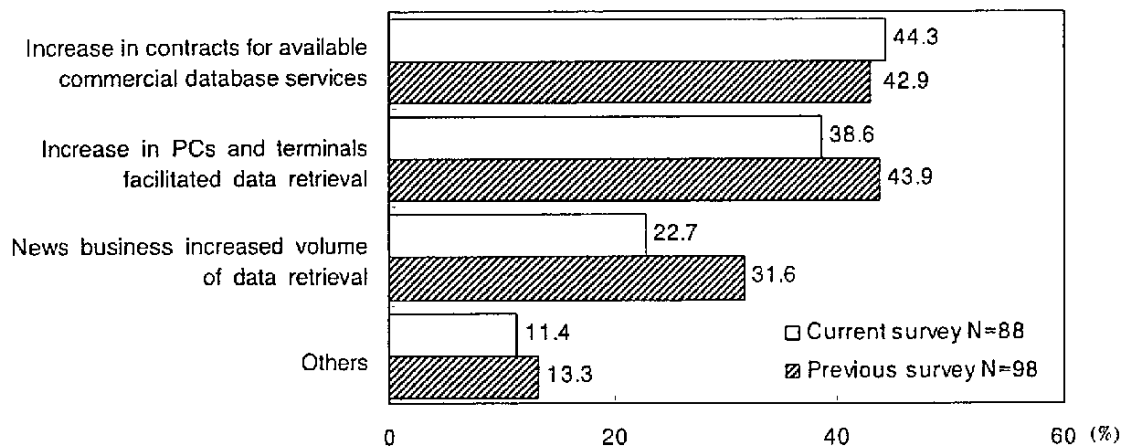
Fig. III-2   Reasons for Increase in Commercial Database Utilization in FY2002

32

We now use free homepages more — 44.7 / 59.1

Reduction of commercial database service contracts — 23.7 / 31.8

Reduced work leads to reduced database use — 21.1 / 18.2

Others — 31.6 / 13.6

□ Current survey N=38
☑ Previous survey N=44

0   20   40   60   80   (%)

Fig. III-3   Reasons for Decrease in Commercial Database Utilization in FY2002

2.2   Projected Use of Commercial Databases in FY2003

Fig. III-4 shows that the prospects for database use in FY2003 as compared with the previous year. 73.5% of users mentioned "almost the same as last year," and the percentage is somewhat higher than in the past two fiscal years. "Will increase" was cited by 20.1%, almost the same as the 20.4% in the previous year, but far higher than the 6.4% for "will decrease." A trend common to the majority of users is that, as the continuing high percentage of "almost the same as last year" implies, the rate of use of commercial databases is expected to increase slightly from the previous fiscal year.

Fig. III-5 shows that the percentage distribution of the reasons for the possibility of increased use of commercial databases in FY2003, while Fig. III-6 shows that of the reasons for the possibility of decreased use of commercial databases. Of the 75 companies that mentioned "will increase", 74 enumerated the reasons (multiple replies). The most common reason was "will increase the number of database service contracts" (39.2%). The percentage decreased slightly from the previous year, but remains around 40%. This shows that many companies are highly motivated to use commercial databases. Next came "increase in PCs and terminals facilitates information retrieval" (32.4%), followed by "new business will increase volume of data retrieval" (28.4%). The percentage of the latter reason increased by 5.1 percentage points, suggesting that many of those companies which replied that the use of databases will increase are planning to expand new business.

By contrast, of the 24 user companies which replied that the use of databases will decrease, 16, or 66.7%, replied, "We intend to use free homepages more frequently." The percentage of this reply decreased by 2.1 percentage points, but the use of "free homepages" remains important. The increase in the use of the Internet, resulting from cost reduction and increased use of homepages will continue to be a major factor for the changing manner of use of commercial databases. Noteworthy is the sharp increase in the percentage of "reduction in the use of commercial databases as a result of reduced business," which jumped from 6.3% to 20.8%. The future developments in the economy will have a great effect on the use of commercial databases.

33

Figure in ( ) indicates the number of responding companies.
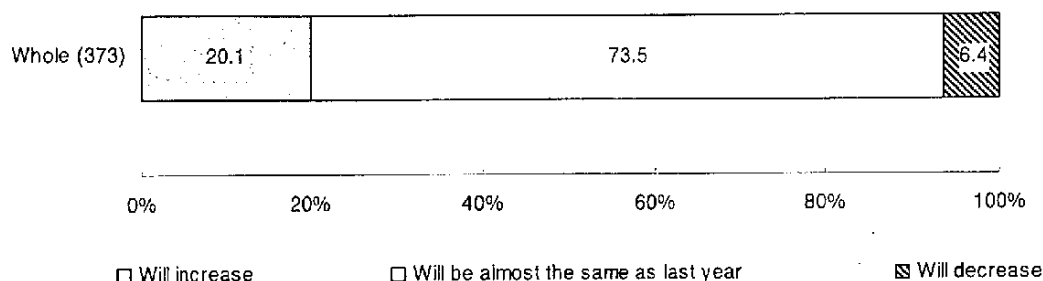
Fig. III-4   Prospects for Commercial Database Service Distribution in FY2003
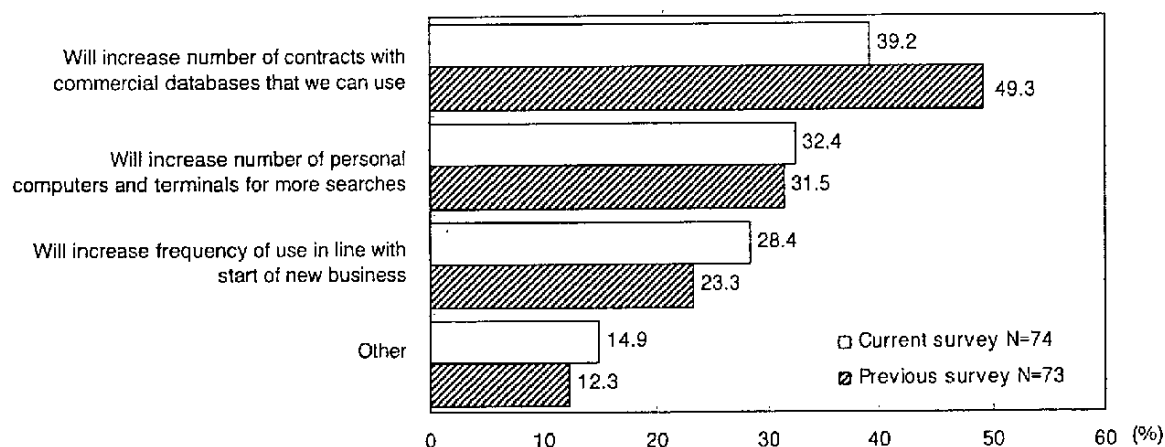
Fig. III-5   Reasons for Projected Increase in Use of Commercial Databases in FY2003
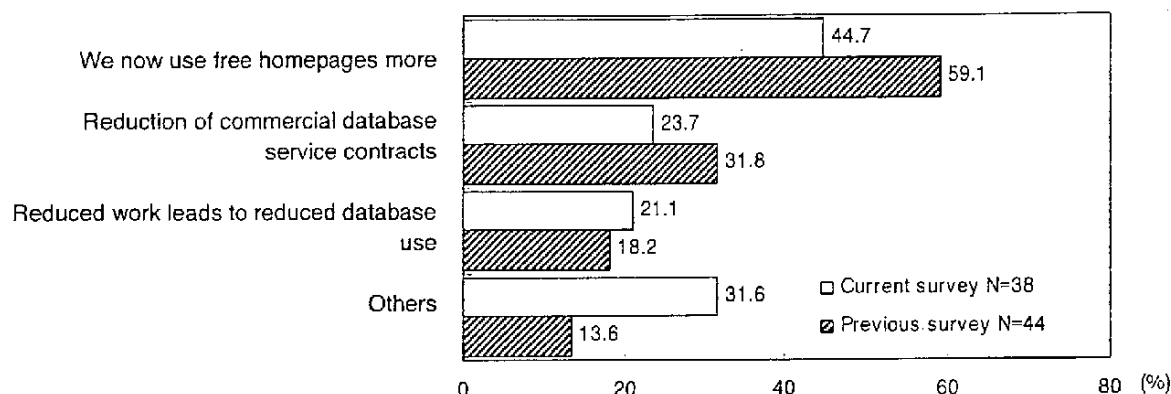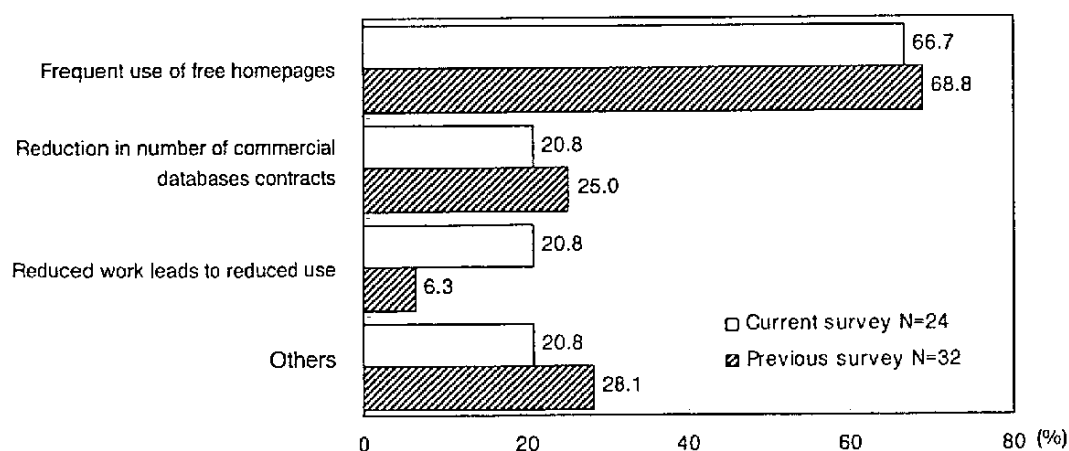

Fig. III-6   Reasons for Projected Decrease in Use of Commercial Databases in FY2003

34

2.3 Use of Databases by Division

The findings of this survey were examined by type of business in order to see which corporate divisions use commercial databases frequently.

Respondents were asked to select the top three divisions in database use (in value terms), and Table III-1 shows the results. For all types of business, "survey division" and "research division" ranked first, both cited by 20.1% of respondents. They were followed by "planning division" (12.5%), "sales division" (12.4%), and "patent division" (9.7%). An analysis of these divisions and their ranking shows that only the percentage and ranking of "research division" changed, increasing by 3.7 percentage points to rank alongside "survey division." In the case of the other divisions, there were no noteworthy changes in percentage and ranking.

Table III-1 Database Use by Company Division

	Total / average	Total/average for previous year
No. of responses	377	373
Planning division (%)	12.5	13.8
Survey division (%)	20.1	22.7
Research division (%)	20.1	16.4
Patent division (%)	9.7	10.9
System engineering division (%)	2.8	2.9
Production division (%)	1.7	1.9
Environment division (%)	1.3	0.7
Sales division (%)	12.4	13.0
General affairs division (%)	4.9	4.8
Legal affairs division (%)	1.2	1.9
Accounting/finance division (%)	6.5	4.7
Other divisions (%)	6.7	6.5

3. Types and Contents of Commercial Databases with High Utilization Rate

3.1 Systems and Databases Used Frequently

Respondent companies were asked to select the top five database service systems that they use heavily in terms of usage charges and the top five databases that they use most frequently. The respective rankings are as shown in Tables III-2 and III-3.

The database service system that is used most heavily (in terms of charges) was "JOIS," followed by "Nikkei Telecom," "COSMOSNET," "PATOLIS," and "G-Search," in that order. "NACSIS-IR," which ranked twelfth in the previous survey, rose to tenth place. Not listed in Table III-2 were "BIGLOBE," "JOIN," and "Data Star."

The database (file) with the highest frequency of use was "COSMOS" as in the previous survey, followed by "TSR," "Nikkei Shimbun Article Database," and "PATOLIS Patents and Utility Models File," in that order. There were a number of changes in ranking. "TSR" rose from fourth to second place, "PATOLIS Patents and Utility Models File" rose from fifth to fourth place, "JSTPLus" dropped from second to fifth place, and "WPI" rose from tenth to ninth place. As in the previous survey, "COSMOS" maintained the first place. In addition, there was a wide gap between "COSMOS" and the second ranking "TSR."

A breakdown by field of the top 15 databases shows that many databases cover economics, including corporate/financial information and newspaper articles. Databases on patents and trademarks, as well as chemicals and pharmaceuticals, both of which are closely related to patents, etc. were also used frequently.

Table III-2   Service Systems with High Utilization Rate (N=353, multiple replies)

Rank	Name of service system	Number of responses	Response rate (%)
1	JOIS	131	37.1
2	Nikkei Telecom	128	36.3
3	COSMOSNET	89	25.2
4	PATOLIS	75	21.2
5	G-Search	58	16.4
6	DIALOG	57	16.1
7	@nifty	50	14.2
8	STN International	46	13.0
9	TSR-VAN	35	9.9
10	NACSIS-IR	27	7.6
11	NICHIGAI/WEB (NICHIGAI-ASSIST)	25	7.1
12	ELNET	16	4.5
12	JACIC NET	16	4.5
14	NEEDS	15	4.2
15	QUICK	12	3.4

36

Table III-3 Databases with High Utilization Rate (N=320, multiple replies)

Rank	Database name	No. of replies	Reply ratio (%)
1	COSMOS (Teikoku Databank Corporations' Information)	138	43.1
2	TSR (Tokyo Shoko Research Corporations' Information)	92	28.8
3	Nikkei Shimbun Article Database	84	26.3
4	PATOLIS Patents and Utility Models File	80	25.0
5	JSTPlus	77	24.1
6	CA/CA plus CA-SEARCH	47	14.7
7	Asahi Shimbun Article Database	36	11.3
8	MEDLINE	24	7.5
9	WPI	23	7.2
10	Nikkei BP Magazine Articles Database	21	6.6
11	PATOLIS Trademark Files	20	6.3
11	Nikkei Who's Who	20	6.3
13	JMEDPlus	19	5.9
14	REGISTRY	18	5.6
15	BIOSS	15	4.7
15	WHO	15	4.7

3.2 Fields of Databases That Are Used Frequently

This section examines the fields of databases which are used frequently (Fig. III-7). In the case of domestic databases, "corporate finance/profiles" tops the list with 53.4%, followed by "newspaper/journal/news" (39.2%), and "patents" (31.3%). There is a wide gap between these three fields and "who's who/organization information," which ranks fourth. Of the top four fields, "patents" rose from fourth to third place, while "who's who/organization information" dropped from third to fourth place. Of the lower-placed fields, "electricity/electronics/ information engineering" rose from eleventh to seventh place.

With regard to databases developed overseas, the fields of databases that are widely used are "patents" (48.1%), "chemistry" (44.2%), "medicine/pharmacy/biotechnology/biology" (35.1%), "electricity/electronics/information engineering" (22.1%), and "newspaper/journal/news" (20.8%). "Corporate finance/profiles" fell from fourth to sixth place, while "patents" rose from third to first place.
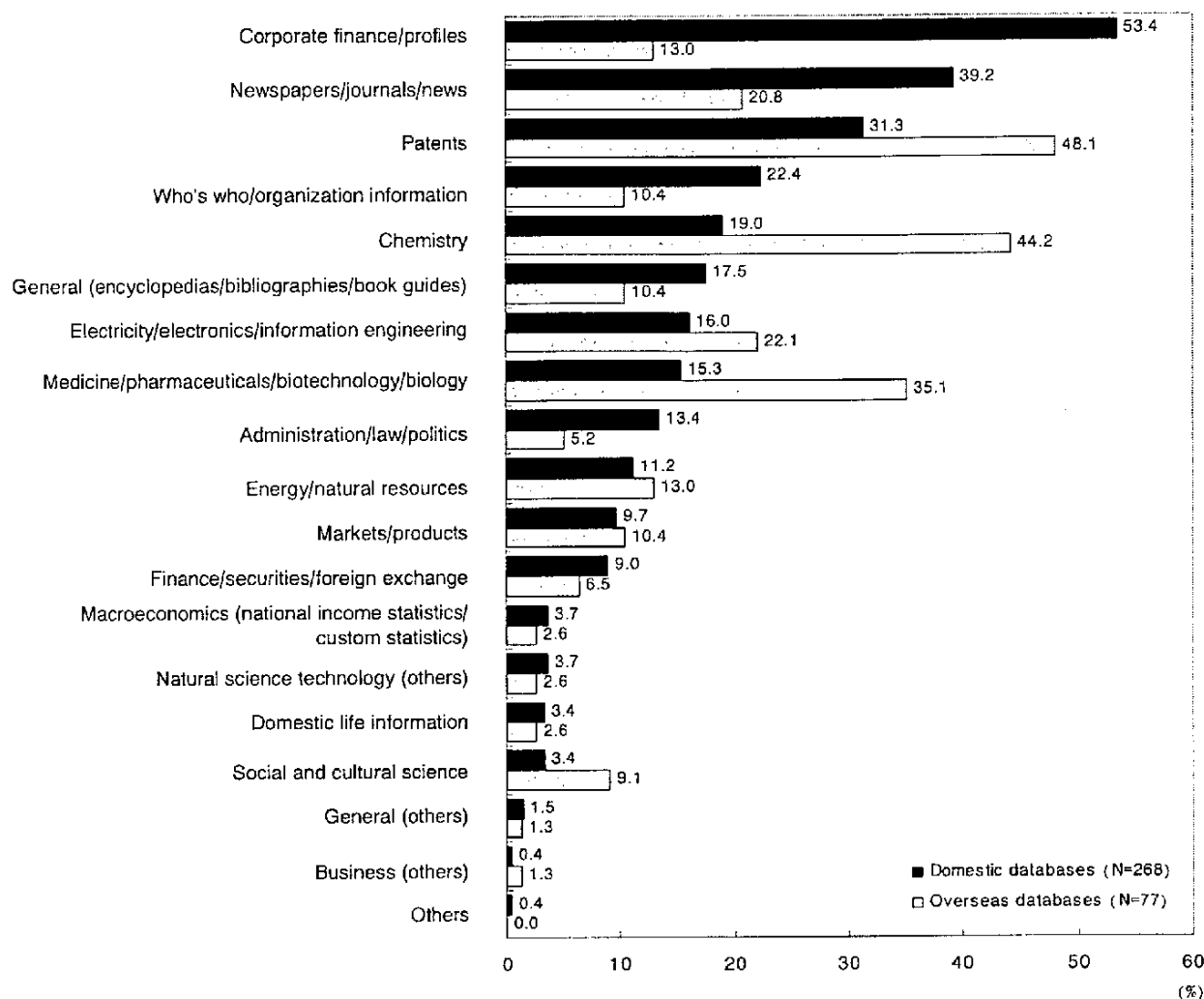
	Domestic	Overseas
Corporate finance/profiles	53.4	13.0
Newspapers/journals/news	39.2	20.8
Patents	31.3	48.1
Who's who/organization information	22.4	10.4
Chemistry	19.0	44.2
General (encyclopedias/bibliographies/book guides)	17.5	10.4
Electricity/electronics/information engineering	16.0	22.1
Medicine/pharmaceuticals/biotechnology/biology	15.3	35.1
Administration/law/politics	13.4	5.2
Energy/natural resources	11.2	13.0
Markets/products	9.7	10.4
Finance/securities/foreign exchange	9.0	6.5
Macroeconomics (national income statistics/custom statistics)	3.7	2.6
Natural science technology (others)	3.7	2.6
Domestic life information	3.4	2.6
Social and cultural science	3.4	9.1
General (others)	1.5	1.3
Business (others)	0.4	1.3
Others	0.4	0.0

■ Domestic databases (N=268)
□ Overseas databases (N=77)

Fig. III-7   Breakdown of Popular Databases by Field (multiple replies)

3.3   Utilization of Public Data

The use of different types of data that are provided by public bodies (namely, data developed and made public by government agencies, etc.) was examined (Table III-4). The total number of companies that use databases containing public data is 282, which accounts for 72.3% of the total number of companies using commercial databases (390), a decrease of 1.3 percentage points from 73.6% in the previous survey. But the percentage itself remains high. When data are broadly divided into "text data" and "numerical data," the percentage of "use now" and that of "want to use in the future" were both higher in the case of "text data."

Table III-4   Utilization of Databases Covering Public Data (multiple replies)

		Currently using (N=282)		Want to use in the future (N=176)	
		No. of replies	Reply ratio (%)	No. of replies	Reply ratio (%)
Text Data	Government white papers, commissions and council reports	145	51.4	40	22.7
	Government press releases	90	31.9	44	25.0
	Kampo (The Official Gazette of the Japanese Government)	124	44.0	68	38.6
	Judicial precedents, etc.	86	30.5	59	33.5
	City planning data (including map information)	39	13.8	73	41.5
	Patent information	165	58.5	43	24.4
	Local government data	91	32.3	64	36.4
	Others	14	5.0	2	1.1
	Total	268	95.0	154	87.5
Numerical Data	Population, employment and labor	108	38.3	47	26.7
	Domestic economy, business climate, finance	89	31.6	62	35.2
	International accounts and trade	64	22.7	49	27.8
	General business/industry	96	34.0	75	42.6
	Corporate information	123	43.6	67	38.1
	Others	5	1.8	3	1.7
	Total	182	64.5	125	71.0

4. Methods of Using Databases

Methods of using databases were classified into the following four types, and the ratio of use based on fees paid was investigated (Table III-5).

[1] Use of online database services other than those which are provided on the Internet

[2] Use of database services through the Internet

[3] Use of database services provided on CD-ROM and DVD

[4] Use of other offline database services

The results for FY2002 show that in the case of method [1] overseas databases enjoy a high rate of utilization, while in the case of method [2] domestic databases enjoy a high rate of utilization. Many methods of use of database services other than those which are provided on the Internet are used to access overseas databases. Compared with the results for FY2001, method [1] was used less frequently to access both domestic databases and overseas databases. By contrast, method [2] was used more frequently to access both domestic databases and overseas databases.

In FY2003, it is projected that database services will be used through the Internet more frequently. This trend has been continuing for the past few years, as Internet usage spreads from year to year.

There was no significant change in the use of offline databases, and the rate of utilization has remained at about the same level for both domestic and overseas databases.
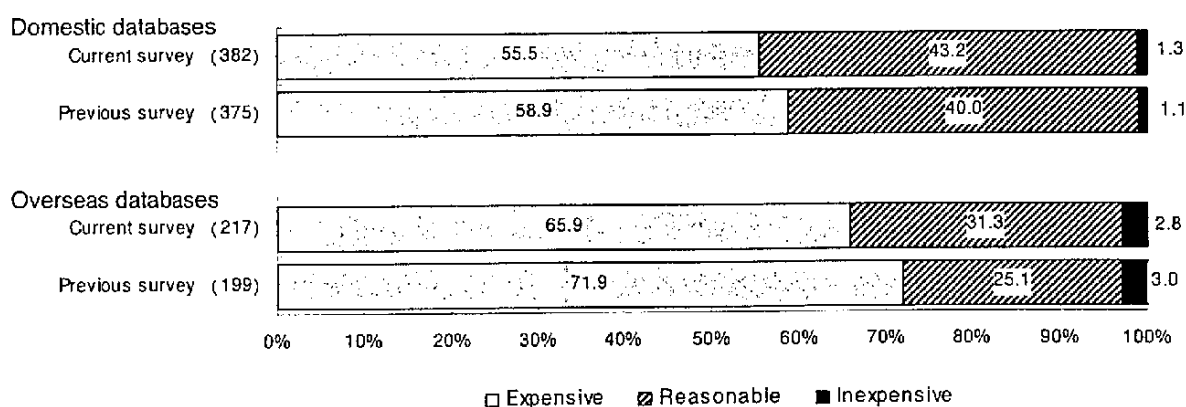
Table III-5   Average of Ratio of Method of Using Commercial Databases

	2001 results (%)		2002 results (%)		2003 projection (%)	
	Domestic (N=304)	Overseas (N=108)	Domestic (N=324)	Overseas (N=100)	Domestic (N=318)	Overseas (N=101)
[1] Commercial on-line (except by Internet)	27.6	42.0	22.7	37.6	20.6	33.0
[2] Commercial on-line by Internet	53.7	44.3	59.1	49.5	61.8	53.3
[3] Commercial off-line: CD-ROM and DVD	15.7	12.8	15.0	11.8	14.4	12.7
[4] Commercial off-line: others	3.0	0.9	3.2	1.1	3.2	1.0

5.   The Problem of Charges and Utilization

5.1   Problem of Charges

User companies were asked to comment on the overall system of charges for commercial databases. For domestic databases, 55.5% of responding companies replied that charges were "expensive" and 43.2% replied that they were "reasonable," only 1.3% replying that they were "inexpensive." For overseas databases, on the other hand, the percentages of responding companies mentioning "expensive," "reasonable" and "inexpensive" were 65.9%, 31.3% and 2.8%, respectively. Compared with domestic databases, user companies tend to think charges are rather heavy (Fig. III-8).

Domestic databases
Current survey (382)   55.5   43.2   1.3
Previous survey (375)   58.9   40.0   1.1

Overseas databases
Current survey (217)   65.9   31.3   2.8
Previous survey (199)   71.9   25.1   3.0

0%   10%   20%   30%   40%   50%   60%   70%   80%   90%   100%

□ Expensive   ☒ Reasonable   ■ Inexpensive

The figures in ( ) represent the number of responding companies

Fig. III-8   General Comments on Commercial Database Charges

5.2 Comments on Recorded Information

User companies were asked to comment on the information recorded in databases. With regard to domestic databases, 22.1% of responding companies complained that "information is not updated frequently." Other comments included "time between release of information and updating is too long" and "full text is not covered." These findings show that user companies want the latest, detailed information. In this respect, domestic databases still fall short of fully meeting users' needs in terms of frequency and speed of information updating.

With regard to overseas databases, on the other hand, "information is not updated frequently" did not top the list. As in the previous survey, the most frequently cited complaint was "information is sometimes inaccurate" (22.3%), which was followed by "it is not clear whether needed information is available" (21.6%). It seems that many users are discontented with overseas databases in terms of accuracy of information.

6. Users of Databases

This section examines the results of a survey on how companies actually use databases internally. As shown in Table III-6, the most common reply was "the person who needs information searches for it himself or herself" (71.9%), followed by "request the database search department " (27.0%) and " request in-house searchers only for specialized searches " (16.3%). The percentage of respondents who mention "the person who needs information searches for it himself or herself" is increasing year after year, attributable to the increase in the number of access points due to the progress of networking and the improvement in individual users' data search capability as well as the database use environment. In addition, the fact that information which individual users need is increasingly diversified and that the need to obtain necessary information quickly is growing, are also responsible for this trend. The least common reply was " request outside information brokers " (3.7%), and so this practice is in decline.

Table III-6 Database Utilization Method by Companies (multiple replies)

	Person who needs the information searches for it himself		Request another person in the same team		Request the database search department		Request in-house searchers only for specialized searches		Various patterns		Request outside information brokers		Other	
	No. of items	Ratio (%)	No. of items	Ratio (%)	No. of items	Ratio (%)	No. of items	Ratio (%)	No. of items	Ratio (%)	No. of items	Ratio (%)	No. of items	Ratio (%)
Total (N=381)	274	71.9	43	11.3	103	27.0	62	16.3	47	12.3	14	3.7	7	1.8
Total in previous survey (N=378)	268	70.9	51	13.5	95	25.1	63	16.7	53	14.0	14	3.7	4	1.1

7. Utilization of the Internet and Databases

7.1 Services on Which High Expectations Are Placed

Users were asked to comment on what they expect of database services provided through the Internet and vendors were asked to comment on what kinds of database service markets they consider promising. Both users and vendors were asked similar questions. Figure III-9 compares their comments.

The most common comment was "information search service using the Web" (73.4% of users and 67.9% of vendors), followed by "information service by text, numerals, voice, animation, still pictures, etc. on websites" and "news distribution service by e-mail." There was no significant difference in terms of percentage between users and vendors. Obviously, both users and vendors consider these services promising.
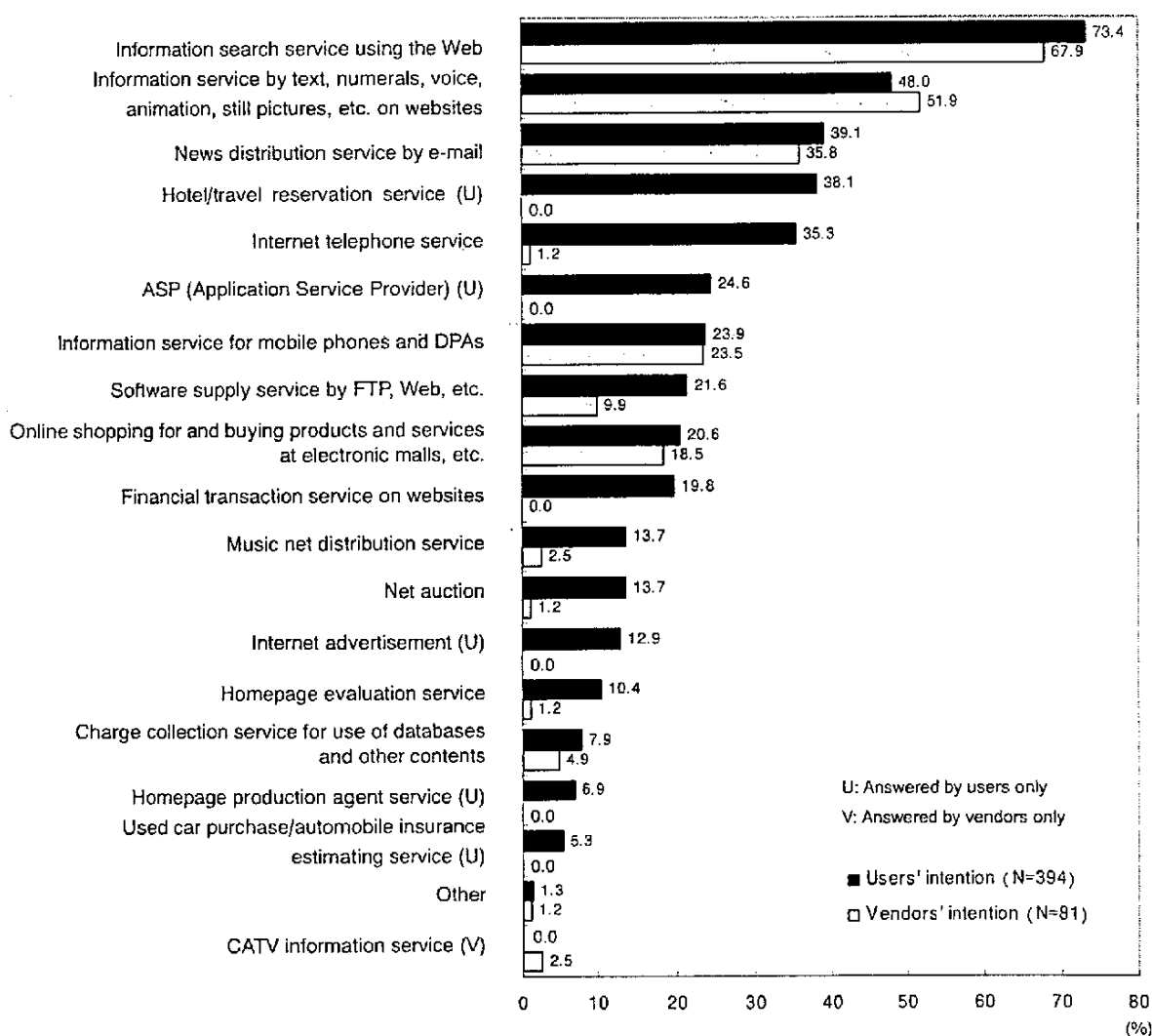

Fig. III-9 Comparison of Promising Internet Services as Seen by Users and Vendors (multiple replies)

42

On the other hand, there were some discrepancies in other respects between users and vendors. While a certain percentage of users mentioned "Internet telephone service," "financial transaction services on websites," "music net distribution service" and "net auction," the percentage of vendors who made such comments was less than 5%. This represents a wide difference between both sides. It should be noted that many of these services are already in widespread use. It appears that vendors, namely database operators who responded to this survey, have not been highly motivated to enter these database service markets, as they are preoccupied with other database services.

7.2 Problems With the Use of the Internet

The most common problem with the use of the Internet was "feel threatened by inadequate security such as illegal access and virus invasion" (83.1%), followed by "danger of virus contamination and data destruction" (82.6%), which topped the list in the previous survey; these two problems are the most important to users. They were followed by "possibility of leakage of personal/classified information" (61.7%). The high percentage of users citing this problem shows that achieving data security including means for adequate protection of privacy is the greatest challenge facing users.

8. Use of CD-ROM and DVD

One of the major media of offline use of commercial databases is CD-ROM. In view of the fact that DVD, among other large-capacity media, has recently come into widespread use, DVD was placed in the same group as CD-ROM and was included in the survey items. Of the 364 responding companies, 150, or approximately 40%, were users of commercial databases on CD-ROM or DVD.

As regards the fields of use of CD-ROM databases and DVD databases, "dictionaries, encyclopedias, directories" (24.6%) topped the list, followed by "science, technology/patents" and "medicine, pharmacy, biotechnology, biology." The fact that "dictionaries, encyclopedias, directories" or "biographies, books, periodicals," which do not have a very high rate on online services, are used frequently on CD-ROM and DVD reflects the characteristics of offline media. In contrast, "corporate finance/profiles" and "newspapers/journals/news" are used far more frequently on online services because speed is critical for this type of information. They are not used as frequently on CD-ROM or DVD, which are advantageous for recording and storing information.

9. Types of Database Services That Users Want to Use in the Future

With regard to the fields of commercial databases that companies want to use in the future, Table III-7 shows the fields currently used by 289 companies or fields which 47 companies which currently do not use commercial databases intend to use in the future.

The ranking in descending order of future information needs of those companies who currently use commercial databases was "corporate finance/profiles" (55.0%), "newspapers/journals/news" (41.9%) and "patents" (39.1%). The top two fields were the same as in the previous survey. Except for "patents" and "macro economics," all other fields retained the same ranks as in the previous survey. In the case of the group of companies that want to use commercial databases in the future, "corporate finance/profiles" and "newspapers/journals/news" ranked first and second as in the case of the former group of companies. "Patents" and "market/products" ranked relatively high. It appears that expectations are high for databases dealing with economic and social problems.

Table III-7   User-preferred Fields of Commercial Databases (multiple replies)

(%)

	Currently using databases (N=289)	Want to use in the future (N=47)
Corporate finance/profiles	55.0	61.7
Newspapers/journals/news	41.9	46.8
Patents	39.1	40.4
Who's who/organization information	31.1	21.3
Administration/law/politics	31.1	31.9
General (encyclopedias/bibliographies/book guides)	28.7	44.7
Market/products	24.2	34.0
Chemistry	24.2	12.8
Medicine/pharmacy/biotechnology/biology	23.2	19.1
Electricity/electronics/information engineering	22.8	19.1
Energy/natural resources	17.3	10.6
Macro economics (national income statistics/customs statistics)	13.8	17.0
Finance/securities/foreign exchange	13.1	19.1
Home life information	10.7	23.4
Social and cultural sciences	7.6	6.4

44