

Chapter V: techniques and technologies for improved information flow between statistical offices and the media

TABLE OF CONTENTS

5.1 (COMMUNICATION VEHICLES COMPLETED) Techniques and technologies for improved information flow between statistical offices and the media.....	3
5.1.1 Introduction.....	3
5.2 (TIME GAP MORE AND MORE NARROW) Improving information flow between statistical offices and the media.....	3
5.2 Impact of Internet dissemination on the structure, functioning and resource requirements of public relation offices in statistical offices.....	4
5.2.1 Internet distribution of time sensitive reports.....	4
<i>First reports on-line.....</i>	<i>5</i>
5.2.2 Web activates passive users	
<i>Modifying organizational build-up.....</i>	<i>6</i>
<i>To serve Internet society.....</i>	<i>7</i>
<i>New burdens, new patterns.....</i>	<i>7</i>
5.2.3 Managing web-communication.....	8
New services require new skills.....	8
<i>Technical skills mix.....</i>	<i>9</i>
<i>Communications skills mix.....</i>	<i>10</i>
5.2.4 The “advanced disseminator”	11
5.3 (DISPERSION IN CHARGING FOR SERVICES) Pricing of Internet based services.....	11
<i>Creating demand in Netherlands.....</i>	<i>11</i>
<i>Corporate account at Eurostat.....</i>	<i>11</i>
<i>Basic data free in Belgium and Luxemburg.....</i>	<i>12</i>
<i>Answering economic necessities in Italy.....</i>	<i>12</i>
<i>Finland’s problematic small clients.....</i>	<i>13</i>
<i>Users’ contracts in Germany.....</i>	<i>13</i>
<i>Charging by constraint in Czech Republic.....</i>	<i>13</i>
<i>Information as freeware in Denmark.....</i>	<i>13</i>
<i>Strictly revolving funds in Turkey.....</i>	<i>14</i>
5.4. Measurement and evaluation of Internet use/use of Internet by the media..	14
5.4.1 Assessment of Internet use by hits, visits.....	14
Overview of measurement concepts.....	14
<i>Hits.....</i>	<i>15</i>
Page View/Impression.....	15
<i>Visit/ Total user sessions.....</i>	<i>15</i>
View Time.....	15
<i>Proxy problems.....</i>	<i>15</i>
<i>Cache server evasion.....</i>	<i>16</i>
<i>Results of measurement.....</i>	<i>16</i>
5.4.2 User sessions in focus.....	17
<i>Press releases neglected.....</i>	<i>17</i>
<i>Unvisited “What’s new” page.....</i>	<i>17</i>
5.5. Making use of new technological developments to better meet the needs of the media and get their feedback.....	18
5.5.1 Census: full use of technologies.....	18

<i>Role of the Public Information Office</i>	18
<i>Embracing Emerging Technology</i>	19
<i>Census 2000 Media Operations</i>	19
<i>Broadcast</i>	20
<i>Video</i>	20
<i>Radio</i>	21
<i>Photography</i>	22
<i>News Conferences</i>	22
<i>Internet/Intranet</i>	22
<i>Accomplishments at a Glance (October 1, 1999 - July 31, 2000)</i>	22
5.5.2 Vision of statistical libraries' mission	23
<i>Software's abilities</i>	23
<i>Multiple agents, multiple innovations</i>	23
The Architecture.....	24
<i>The Client</i>	24
The Library.....	24
The Compute server.....	24
<i>The Data server</i>	25
<i>Actors</i>	25
<i>Implications of using</i>	25
5.6 Experiences in evaluating the quality of communication with media ..	25
5.6.1 Reports about the press	25
Press-clippings.....	25
Interviews with journalists.....	26
General data on Statistics.....	26
Problems.....	26
5.6.2 Exploitation of customers' feedback	27
Poor TV coverage.....	27
Exploiting Internet.....	28
Database: flexible output.....	28
<i>Good old telephone</i>	28
5.6.3 Do statistics need news media?	29
Possible commutation of functions.....	30
<i>Free databank access</i>	30

5.1 (COMMUNICATION VEHICLES COMPLETED) Techniques and technologies for improved information flow between statistical offices and the media

5.1.1 Introduction

While the time delay for data dissemination is getting shorter and new tools become available, very few agencies have been able to increase funding to cover all additional costs needed to comply with the changes. In the meantime traditional ways of dissemination have not disappeared but had to be adapted to the new tools and possibilities – while the use of regular means were to be reviewed in order to be improved. At the same time, it would be desirable to have certain points of reference about the dilemma whether to charge for Internet services or not? And this had to be solved in the middle of a whirlwind when a real reincarnation evolved: printed messages became displayed on the screen, web pages were printed, video and audio tapes could be played on PCs, databases tapped and figures manipulated according to users wishes... There were constant appeals to unify the different kinds of releases through Internet. And there were huge organisational efforts to use all tools in unison to promote the greatest venture of statistics, the population census. Not surprisingly, after thorough investigation by National Statistical Offices of their efficiency in using old and new tools of communication some voices occurred predicting that the role of news media in disseminating statistics would be gradually replaced by statistical agencies themselves.

5.2 (TIME GAP MORE AND MORE NARROW) Improving information flow between statistical offices and the media

Internet had a large **impact on financial resources** as well as **on organisational matters**.

Timeliness in dissemination is no more limited on the side of its vehicles: web seems to develop more rapidly than abilities to servicing it. Faster distribution of data required from statistical offices considerable investment to resources and a **reshuffle** of their **data production** as new technical skills and knowledge were needed to cope with maintenance and development of Internet-based services. In addition existing tools of dissemination also grew user-friendlier, raising further demands.

The time between availability of data and their dissemination has become very short, thanks to Internet. There were, however a number of issues involved in this development. What and **how much information** should be placed on the Internet? Should the information be made available **free of charge** or should the users **have to pay** for the data? What information should be free and what should be chargeable? If the users should pay then what price – and how to keep track of the use of the data? (see also 2.2.)

The situation varied from country to country. Some National Statistical Offices made their data available on the Internet entirely free of charge, others planned to charge but considering a combination, i.e. **some data sets would be free of charge** while other data sets would have a **price attached** to them. In countries where **revenue generation** is also a mandate of the NSO, the Internet raised the very important issue of how to strike a **balance** between revenue generation and **the public's right to the information** as a public good as well as the need to promote data dissemination? In this regard, it was felt that national statistical offices clearly needed guidance from the Governments. On the other hand, with the rapid development and expansion of the Internet, some countries were of the view that it was not advisable to take decisions now on the pricing policy of data on the Internet, particularly when most of the information currently available on the web is entirely free of charge.

The work group on dissemination to media also noted that there had been **benefits** to be gained from providing data on the Internet other than in money terms. Data dissemination can be expanded at a very marginal cost; the public could have access to the data quickly and freely; more use of the

data meant more scrutiny of the data quality; and statistical staff could spend less time responding to requests for data.

The whole **interaction between statistical products and communications is deeper** than it seems at first sight. This is demonstrated by the daily presence of statistics in the newspapers, on the radio, and on television. It can be explained among others by both the constant qualitative and quantitative development of the NSOs' activities and the improvement of the diffusion techniques.

The positive effects of the communications and diffusion policies adopted by NSOs were visible: the **involvement of the media has grown**; the quality of the released data has improved; the information has been conducted underlining the importance of the data diffused; the public's information and the assistance in its decisions had grown.

5.2 Impact of Internet dissemination on the structure, functioning and resource requirements of public relation offices in statistical offices

5.2.1 Internet distribution of time sensitive reports

*Several agencies of the **U.S. Department of Agriculture (USDA)** produce statistical reports containing information and **data critical to the agricultural marketing system**. These reports are prepared under **strict security** conditions, and are released at specific times published in an annual calendar. Distribution of electronic versions of these reports on the Internet was one of the important steps in ensuring that all market participants have **efficient, timely and equitable access** to U.S. crop and livestock production estimates, world supply and demand estimates, and commodity market analyses.*

*The National Agricultural Statistics Service (NASS), the Economic Research Service (ERS) and the World Agricultural Outlook Board (WAOB) are responsible, within the U.S. Department of Agriculture, for preparing a variety of **sensitive reports**. These reports include estimates of crop acreage, yield, and production; grain stocks; livestock, dairy and poultry production, etc. This information and **data may affect market prices**, often in ways unanticipated by traders, and the reports are **very time-sensitive** because the value of this information degrades quickly. For both reasons, the preparation and **release** of sensitive reports is **strictly controlled** to ensure that all market participants have equal access to the information they contain.*

*Until the mid-1980s, all **reports** were released only as **printed materials**. On schedule, members of the media would gather in the USDA pressroom and watch a precision clock move toward the release time. Meanwhile, copies of the reports were placed face down at telephone stations around the room. The tick of the clock to release time was the signal for reporters to rush to telephone stations and uncover the reports. Progress in computer-to computer communications, however, led to the development of the USDA Computerized Information Delivery system (CIDS) in 1986. By 1994, customer concerns about the cost and effectiveness of computerized delivery and the rapid development and adoption of Internet technologies prompted the three agencies to explore Internet alternatives for electronic distribution of reports in preparation for a phase-out of computerized delivery in 1997.*

First reports on-line

Till 1994 USDA did not have the Internet communications, servers, technical infrastructure, and personnel, but a new project enhanced their implementation. It became a natural vehicle for the distribution, and all three agencies' reports were available on the Internet beginning in March 1995. A web interface to the system was introduced at that time as well as an automated e-mail delivery system. Due to failures of routers and interconnections at that time, the service standard for Internet release was set at the **release time plus three hours**.

Internet delivery, however, could not completely satisfy the media's need for certain, instant electronic release. This need was met in a unique way beginning in 1996: **reporters**, with their

laptop computers, were **admitted to the “lock-up”** 1-2 hours before official release of a report. Lock-up is a secure suite of offices at USDA carefully isolated from all forms of communication outside the suite (see also 4.1.3.1. and 4.2.1). No one admitted to the lock-up is permitted to leave until the report being prepared is officially released.

Like the USDA staff preparing the report, journalists could not leave the lock-up before the release time or communicate with anyone outside the suite in any way. But they were **given an electronic copy of the report** and permitted to use the new report to write stories, prepare data tables and otherwise prepare files on their laptops for electronic transmission through a modem as soon as they, and report, are released **simultaneously from the lock-up**. At that moment, telecommunications to the suite are enabled, and transmission of the reports and news stories begin. USDA experience showed that **user access** to reports using FTP or HTTP protocols on the Internet could be provided reliably **within 5 minutes of release**. From a user perspective, the services provided were remarkable, almost magical.

5.2.2 Web activates passive users

When the Government's Press Office started to set up the **first Internet sites** for **Luxembourg's** State Department, Statec was one of the first to make an important contribution. This was in 1996.

*As at that time, Statec was **not prepared to invest in Internet** (either human or financial resources), the Government Press Office decided to sponsor the initial website, developed by private enterprise. The first site contained basic information. The updating was **maintained within a low budget** and with basic human resources investment.*

But growing demand pushed the office to improve the website and to offer more statistical information, especially short-term economic indicators. Most of users requested all encompassing statistical information on an annual and monthly basis and notably download facilities of EXCEL files.

*Statec's budget and human resources did not allow a substantial development of the site at the time. For this reason since the beginning of the year 1999, a new Internet policy had been discussed with the aim of convincing political authorities to increase the **financial resources dedicated to Internet**, to increase the staff dealing with Internet and, at the same time, to find possibilities to facilitate the job of the dissemination staff by offering a large amount of information on Internet. The best way to **rationalize the dissemination** work seemed to be to concentrate all efforts on putting a maximum of information on the website. The result of this effort was seen in the **growing number of requests reoriented** towards website and thus great **time saving** for dissemination by other means. Redirecting the requests for statistical information towards the Internet has the great advantage of **converting passive users into active users**. The time saved in this way could be reinvested in the continuous improvement, updating and management of the entire website. One important problem to be solved in order to optimize dissemination work under the constraint of the new technologies remained: the **training of staff**. Most of the staff dealing with Internet was **self-taught**, and professional training involved time and expense if no internal courses were available.*

Modifying organizational build-up

The then new-born Internet policy anchored on four principles:

- to put a **maximum of information** on the Internet as quickly as possible
- to give **basic information** (statistical or other) **for free**
- to **charge for** downloadable files and **detailed statistical data**
- to redirect users to the Internet site to facilitate the task of dissemination staff and to minimize printing and mailing costs.

All these efforts had a large **impact on the organization** at the different levels of the statistical institute.

The set up of an Internet server is quite expensive and the charge for the budget might be rather high, especially for a small institute. A common problem for many statistical institutes is the fact that, due to budgetary techniques, **they cannot reallocate receipts of dissemination**. As these receipts usually flow in the central public treasury, the resources cannot be directly reinvested in Internet matters.

Updating the site or transferring files via FTP generates far **fewer costs than printed** publications or electronic products like diskettes or CD-ROM. Notably the mailing costs for printed publications are extremely high compared to electronic transmission.

The setting-up of an **Internet site demands substantial financial resources**, which might not be foreseen by an Institute's budget. In this case, new resources have to be found. Once a Website is running, it saves money compared to traditional dissemination. The existence of a well running Internet site appears at nearly every level of an institute's organization, but in different ways.

Obviously the dissemination and the computing units have the heaviest workload. But the **statistical departments** (production units) are also involved and **have to "think Internet"**.

Internet adds a new task to the daily work of a dissemination unit. During the setting-up of an Internet site, the development of a good concept with a user-friendly navigation and a relevant content is very labor-intensive and time consuming. Innovation and follow-up of the newest evolutions and technologies are a full-time job. The site also has to be supervised.

Links and addresses have to be checked regularly. The staff in charge of the updating of the site have to be trained and must follow the evolution of Internet technology.

On the other hand, a good running **Internet service leads necessarily to a reorganization** of the dissemination unit. If well-structured basic information is available on the Website, with detailed tables that can be downloaded, the information officers normally receive fewer requests by phone, mail or fax. They will have to send fewer printed publications, photocopies, diskettes and CD-ROM's to the customers. This advantage could usefully be **reinvested in Internet-related work**.

To serve Internet society

The main "conflict" between distributors and producers of statistical information remains. Shared experiences of people dealing with dissemination of statistical products always conclude with the same reproach: 'the producing units do not think "user-oriented", only "product-oriented"'! They require their output data to be published, but do not take the users' needs sufficiently into account. Their **output is incomprehensible**, concepts are difficult to understand, **tables are too large**, too many figures, too many details!

These problems are not specific to Internet dissemination, but they do represent a new challenge for distributors to convince their colleagues **to produce "on-line friendly" information** and to organize their files in a way which permits an easy transfer to a web server.

New burdens, new patterns

The great advantage of Internet is its ability to provide information in many ways (basic statistics, detailed figures, short-term indicators, databases). Therefore one is dealing with different groups of customers: those who are seeking on-line data which fits the computer screen and is easy to understand, and those searching for more scientific data, at a detailed level, notably historic series. There are private customers, students and, most of all, professional customers. But on one point they are all agreed: the **required data must be immediately available and up-to-date**.

More than ever, the Internet society puts a burden on information providers. The pressure to provide latest figures is increasing. That's why data producers have to understand that they must **organize their work in a new and more disciplined way**, because the time between availability of data and their dissemination has become very short. In the past, the time needed to capture, lay out, print and send the information allowed correction, modification and updating of that information. This "grace period" is now extremely short.

Internet obviously has a large **impact on financial resources** and on **organizational matters** in a statistical institute. The set-up of a website for the first time is certainly a challenge; it is also a great opportunity **to disseminate more, better and faster**. This requires an important **investment**

and concerns more or less every level of the organizational chart. If **statistical institutes** want to be perceived as **serious competitors** on the information market, they must consider all these facts and organize their data production and dissemination accordingly.

5.2.3 Managing web-communication

The advent of the World Wide Web technology, which has made the Internet so popular, has changed – probably forever – the ways in which statistical agencies communicate with their audiences.

In the past, public affairs and media-relations sections in statistical agencies did not have to concern themselves with items such as the computer skills of their clients, or the capacity of their computer equipment. Traditionally, communication with the news media and the public was done by print. Before Web technology arrived, a limited number of clients had direct electronic access to a selection of statistical products and services. Statistical agencies developed **specific electronic dissemination systems** to meet the needs and preferences of these clients. Now, an increasing number of people just “log on” to the Internet to get the statistical information they need. It’s impossible to know who all these people are. Like it or not, media and public relations offices in statistical agencies have had to adapt themselves to the technology of the Internet.

New services require new skills

The increasing use of the Internet by the media and other clients has resulted into the development of **new Internet-based communications programs**.

*At **Statistics Canada**, Communications Division underwent **some re-structuring**. Some new positions were created and the skills mix of its staff was updated to respond to the needs generated by implementing new Internet programs. This was viewed as necessary to manage in a cost-effective manner the rapidly growing “public good” portion of the Agency’s website.*

*In the **United States** where the nature and extent of Internet based media/public relations programs vary due to decentralization of its the U.S. statistical system.*

*Statistical offices **developed some new tools**, guides, and materials for the media using the Internet **without any major structural change** or additional staff for the most part. Many media or public affairs offices did develop operational links with other parts of the agencies' communications structure and formed new partnerships to offer and support Internet services. Virtually all of the U.S. statistical offices offer internet-based communications services and some have a wide array of products and services. With those services and features such as Extranet (password access to embargoed materials) these **agencies have expanded the media's access** and use of the data. All agree, however, that it is essential that **staff be knowledgeable** about, and adapt to, evolving Internet technologies, and be able to acquire the new skills quickly.*

Technical skills mix

Experience has shown that a **combination of technical skills and knowledge** is required to develop and maintain Internet-based statistical services and products on a daily basis that meet the needs of users and the Agency’s goals. This is the case at Statistics Canada’s Communications Division. Such skills include extensive experience with and knowledge of:

- **Operation of a NT server or a Unix server:**
 - . Operating Systems (DOS, NT, Unix)
 - . Knowledge of Client/Server
 - . Networking (TCP/IP, FTP, Security, Data sharing)
 - . Hardware
- **Office tools:**
 - . Word processing, Spreadsheets
 - . Database management systems
 - . Web publication tools

- **Programming:**
- HTML, Dynamic HTML, JavaScript, Perl
- MS Visual Basic
- Macros, Scripts
- **Other:**
- Internet experience
- Graphic design

On-going training and development are also essential to keep abreast of latest software and hardware developments.

*At **Statistics Canada**, Communications Division obtained **mixed success by training** existing public and media-relations staff to acquire the required technical knowledge and skills to develop and maintain its Internet-based program. Some limited additional staff had to be hired to perform the new duties associated with the growing Internet program. So it was decided to hire staff to help achieve the desired goals. Notably, these **goals included the daily dissemination** of timely statistical products and services meeting both the needs of users and their technical abilities and capacity.*

*The **technical work** is performed **under the guidance** and supervision of public and **media-relations staff**. Their communications skills are key to ensuring that the information needs of target audiences and the goals of the Agency are met in a cost-effective fashion.*

*Another **option** could have been to **contract out** the initial technical development and implementation work and subsequent trouble-shooting, and assign the maintenance of the service to regular public/media relations staff, after some basic training. However, it was viewed as **more cost-effective to integrate** the necessary technical capacity into the organizational structure.*

In the American statistical agencies, increasing use of the **Internet has had a different impact on staff members**. **Requests** for additional personnel to do web work have generally been **rejected on budgetary grounds**. It was rare that new personnel resources were added, at least not directly in the media operations. In most cases, agencies **current staff took on new functions** and obtained new skills, shifted functions, or formed **new interdisciplinary teams** to bring together the skills and information needed. Staff received formal training or literally picked up the new skills as they performed the new duties. In some cases the additional skills were acquired on an evolutionary basis as the Internet itself evolved with changes and development occurring steadily. In some cases, the fact that there was perhaps a little less emphasis on printed publications has cushioned the extra-Web related work. In other cases, agencies still provide a full array of published as well as electronic products and services.

Communications skills mix

When developing communications plans or media strategies, communications advisors must have an extensive **knowledge of a variety of elements**. These elements have included the target audiences, the applications and uses of traditional communications tools, the concepts related to the presentation and distribution of information, and the tailoring of products. Communications advisors must be familiar with the same elements as they apply to the Internet. Specifically, they should ensure that the structure and design of their agency's website allow for easy navigation and maximum use.

The Internet can improve service to the public, but it can also have the opposite effect. It depends on whether the organization has put in place the adequate structure, mechanisms and resources. **The Internet can work against the organization** if clients do not obtain the requested information within a reasonable time frame.

Statistics Canada's Communications Division organized a **training program** aimed at improving the knowledge and competency of its media and public relations staff with regard to the use of the Internet as a communications tool. The course provided an overview for communications

practitioners who must be able to give advice on the **strategic use of the Internet**, and to work with technology professionals to develop suitable Web information packages.

*In the **United States**, many agencies have news centers or **news buttons for the media** to offer specialized information or services on their websites. In addition, several interagency organizations have been formed in the United States to help statistical and other government programs improve their websites.*

5.2.4 The “advanced disseminator”

Internet is seen as an “advanced disseminator” medium by statistical institute of **Turkey** but it was assumed that besides its advantages, Internet had some **disadvantages**:

- needed expert skill and equipments
- expensive when is used to full capacity
- could be slow and complicated
- contained the risk of incorrect interpretation and
- risk of inappropriate comparisons

Web service of the institute became operational in December 1995 in Turkey.

5.3 (DISPERSION IN CHARGING FOR SERVICES) Pricing of Internet based services

As it is extremely difficult to find general rules governing the revenue generating of statistical offices (as it is clear from 2.2. and also from 5.1.), the same goes for the data presentation on the web. A great variety of approaches could be mirrored by early ponderings on how to handle the new vehicle of dissemination.

Creating demand in *Netherlands*

In Netherlands, since the office is mainly funded by tax money, it had to be seen as a public duty to inform the public and disseminate its information as widely as possible.

Budget cuts occurred in early nineties had made realize that Statistics Netherlands was dependent on wide popular support. The office started to take a look at the demand for its figures, and how can increase it. It was thought demand might be stimulated informing the public through the media that new figures were available.

Initially there was a hesitation about providing free access via Internet to publications and the output database Statline. Departments were afraid of “losing income and jobs for publishing activities and client services”. However, the experiment in 1996 with free access to the weekly Statistical Bulletin and to the StasLine output database (with limited download facilities) had no negative income effect and more information was published through the Internet site. In spring 1998, management decided to lift the restrictions: all statistical information for the public should be retrievable through the Internet in a user-friendly way and free of charge. However, customers are charged for individual information services.

Corporate account at *Eurostat*

*One of the objectives of the Eurostat site is to **sell statistical products**. Traditional payment methods are still used. On the Eurostat site, the users will also have new methods of paying for their purchases. The main difference with the existing method of payment is that a user can provide his credit card number on-line.*

The site allows the visitor to group its purchases in a basket and he should be able to get the total of his purchases and to change the products in the basket at any point in time. The site makes it easy

to market products. It has the possibility to market products and manage prices dynamically, for example by temporary and targeted promotions and reductions given after a certain amount of purchases.

The electronic commerce tool should have standard functions or integrated modules for:

- registration, subscription and ordering
- billing
- electronic catalogue for soft goods as well as for hard goods
- customer care
- help desk (for accepting credit card numbers by phone, fax or e-mail and for questions from users)
- clearing (with banks and credit card companies)
- alternative electronic payment method (bank transfers and cheques)
- corporate account (with discount possibilities for specific users)
- customer self management (for customisable views)
- fraud and bad debt detection and delinquent account management
- EDI interface for professional users

Anyone who frequently orders products on Eurostat site can have a **corporate account**.

This account is especially meant for companies or organisations, which have several employees who can access the Eurostat site and want to order products, which must be paid for. The corporate account should be a module of the selling tool.

Basic data free in Belgium and Luxemburg

The National Bank of Belgium decided to make available on the web all the statistical information from BELGOSTAT database, **free of charge** as far as the **latest data** were concerned. For historical series customers were advised to use Belgostat CD-ROM, issued quarterly.

The Luxemburg's STATEC site adopted different approach. It contained two separate sections:

1. basic statistical data for the "general public" accessible free of charge,
2. detailed data available in the form of files which can be downloaded, accessible on payment of a subscription.

Answering economic necessities in Italy

In Italy Istat, with its bulletins, volumes, data banks, is surely the country's leading supplier in this sector. The boundary between free information and on payment information is very difficult to delineate. What is the position to be held as to the expected boom of the customized information demand and to the commerce of electronic data? From the viewpoint of the **institute's global image** this is not a negligible problem, since the institute is seen both as a **subject competing with private companies** (with the risk of being accused of monopolistic privileges) and as a **public institution** with the risks of being too closely related to the State in a country where public administration does not have a good image.

On the one hand an institute like Istat is expected, from a political viewpoint, to **self-finance on the market** at least a part of its activities. On the other hand the institute is to fulfill its **mission of diffusion of statistical knowledge** and culture even though these activities are not profitable.

Therefore, the data collection, processing, analysis and diffusion policies answer economical necessities, but these necessities are only partially measurable in terms of publications or floppy disks sold, while they partially correspond to a task required by the State along for the granted public funds.

Differences between free products and on payment products are becoming clearer, but the process is still in progress.

In 1998 on an experimental basis, **on-line services were all free**. For the future in that time it was considered to supply:

- free data via Internet;
- on payment BBS services
- on payment floppy disks and CD ROMs with macro-economic data

Generally the price is connected to the supplying cost, while the contents are free because they are connected to Istat's institutional task.

Finland's problematic small clients

*For Statistics Finland charging when using Internet as a delivery system did not pose problems with big customers – the agency made, as before web, a comprehensive agreement on the statistical service to those clients. But small and ad hoc clients were more problematic. There was no universal way to use money in the net. It was thought that the possibility to buy **small pieces of information** straight away would **increase the sales** of statistical material.*

Statistics Finland made a sales catalogue of publications on the Internet pages, where a client could tick the publications he or she wanted to have and send the order to sales office.

Users' contracts in Germany

*With the exception of time series of the Statistical Information System, the data the Federal Statistical Office of Germany offered in the Internet belonged to the basic statistical information provided **free of charge to the general public**.*

*The time series service was rendered against payment on the basis of a user contract. Users were provided with individual user codes and passwords after having had annual subscription. The subscription included the annual provision of 200 time series. Additional time series can be ordered by additional payment. Accounts that were made at quarterly intervals, could be recalled on-line. The conditions specified above refer to users of rather small quantities of data. **Large-scale users** would be **granted quantity discounts**. Besides, these payments did not refer to commercial information distributors for whom prices and license fees were presently being specified. The time series service encompassed a documentation and search system to be used **free of charge** and an order mode provided **against payment**.*

Charging by constraint in Czech Republic

*The Czech Republic's Statistical Office's incomes are planned each year according to the current budgetary regulations and they are sent back to the state budget. Despite efforts to provide via Internet as much as possible of standard statistical information free of charge, by 1998 it **became inevitable** to provide some of them in **paid form**. There were several possible ways how to pay for the provided information:*

- *in cash or with invoice after closing license contract: this method has been utilized up to now, but has some limitations, the main one being time circumscription (first of all when providing information into abroad)*
- *payment cards; but the payment cards are not so much spread over in our country, although for the foreign countries the method would be suitable*
- *payment by phone cards: this method would be the best under our conditions, but it is not still utilized in our country and it also depends what attitude will be taken by the telephone company, also the purchase of a new software for the contact with a customer is expected.*

Information as freeware in Denmark

*In Denmark on-line services were looked at as a field that had to **balance economically**. Therefore all subscribers were **charged according to their use** of data, which was widely accepted both by private enterprises, ministries and other organisations from the beginnings. One of the user groups, which were not attracted by the idea of payment were journalists. The principle of information as a "freeware" for news agencies and newspapers was seen to be fundamental.*

Strictly revolving funds in Turkey

*The crucial point for the statistical institute in Turkey was that information should be given as possible at a **minimum price** (and even free of charge for some categories of users). In practice, the charge included the dissemination costs but not the information itself.*

*As to the **structure of the budget**, activities of the institute were covered through the general budget of the government. Annual grants from government were supposed to finance basic activities as*

well as the cost of compiling, controlling and editing the data to document information or databases. Pricing policy depends on the nature of products. To make easy pricing some general calculations are made, giving a fixed price per page, per diskettes regarding to their size, per characters varying with circulation.

The prices calculated and offered by the Publication, Communication and Public Relation Division were analyzed and approved by Revolving Fund Administration (RFA) Management Council comprising five members. After their approval the prices were announced to users and updated with respect to current inflation rate yearly.

To provide information and to allocate all publications to users, **statistical information requests** from both domestic and foreign users were **regularly gathered** and met by the communication division. Firstly it was analyzed from where the information requests can be obtained. If the requested information were published, the request would be met by photocopying from the related publications. If not published, it would be taken from the related division. After preparing an **official letter** and having signed by the head of department, the vice president or the president depending on where the information were requested or on the characteristics of the information, requested information was allocated to the users.

If private companies, researchers, unions and associations applied for the information, they were required to deposit the **amount determined by the General Directorate** of the revolving fund administration to its account (or directly to the administration) as the cost of dissemination.

5.4. Measurement and evaluation of Internet use/use of Internet by the media

5.4.1 Assessment of Internet use by hits, visits...

For a user-oriented design and further development of their data stock offered online, it is crucial for statistical and other professional information providers to know user profiles. There are various methods to measure Internet use: their possible pitfalls are presented experiences gathered by **Germany**' Federal Statistical Office by 1998. *On the basis of the web server log entries made by default, requested documents and user data are stored in "log files". There are no uniform and generally accepted standards yet with regard to the quantitative and qualitative evaluation of the data of those using web site. Compared with traditional media, the logging of the "web site traffic" opens up new opportunities of directly measuring client contacts. The inaccuracies occurring in such measurement are largely due to the underlying network technology and logs.*

Overview of measurement concepts

To analyse the structure of demand as well as user profiles of online information offers, the following measurement criteria were applied or tested by 1998:

Hits

Any object on a web page is counted, irrespective of whether it is a text file, a graphics file, a logo, a picture, a navigation button, or a control file for frames. As many web pages contain more than one such element, accessing one web page may produce quite a number of hits, i.e. of all files transmitted. Recording just the **hits** would result in an **inexact measurement**; it would lead to very impressive request numbers, but, at the same time, be of almost no information value as far as the real interest in the data offered is concerned.

Page View/Impression

A tool much better suited for measuring the response to the information offered is using Page View/Impression as the variable measured; it **counts sight contacts with HTML pages**, irrespective of the number of elements generating hits on that page, i.e. irrespective of layout and graphics. This provides the measure for the use of individual pages of the data offer. However,

using frames – one screen page contains several HTML pages – again poses the problem of **multiple counts** (similar to measuring hits), even when Page View is the variable measured.

Visit/ Total user sessions

A Visit refers to an uninterrupted use process of a user accessing a web site. The term “user” refers to **users of a URL** (web address) rather than “real” person because, for reasons of data protection, the latter cannot be allocated to individual visits.

In 1998 the measurement criteria Page View and Visit were used in the German online branch mainly to perform comparative measurements of the transmission range of advertising media.

View Time

Is to be used for measuring the time and duration of individual uses (time per visit).

Proxy problems

Defining measuring criteria does not guarantee their correct measurement. Employing proxy cache servers and firewall computers are the main reason why measuring actual Internet use is difficult. Proxy servers are used by Internet providers, companies and institutions in order to better utilize the net bandwidths available. This is achieved by **using cache servers**, which locally and temporarily **store pages** that are frequently demanded. When a user who is connected to a proxy server requests a web page, the server will first of all check whether that page is really stored locally. If so, it will transmit the page to the user, while the server of the page provider will not get a page request and thus no entry in the server log, although the user gets the page on his screen. Only if the page is not available locally, will the proxy server pass on the request to the provider service. Although, in this case, a regular entry in the log file of the provider will be made, the **Internet address stored will be that of the proxy server**, which leads to a bias in Visit numbers.

A similar effect is caused by local cache memories available on the user’s computer.

Proxy/cache mechanisms lead to significantly smaller figures when counting the number, kind, and duration of accesses. This is true especially of data offers, which are used both intensively (many users) and extensively (long average connection times). Determining user numbers is difficult: behind one computer name or one IP address, there may be several users with their own web browsers (e.g. at large Internet service providers and enterprises with firewalls). This has to be taken into account when interpreting the results of the measurement criteria.

Cache server evasion

To **avoid the problems of proxy/cache servers**, the Association of German Periodical Publishers) developed a method for measuring the performance of advertising media. The method works as follows: on the one hand, the HTML pages of the respective information offer are extended by a transparent graphic through the above method: due to the operation of a so called CGI (Common Gateway Interface) script program, a **one-pixel graphic** is necessarily **transmitted** from the web server of the provider for **each case of browser access to a URL** (i.e. a web page) even if the page is already stored in the local cache memory of the browser or the proxy server. The page itself is subject to all cache mechanisms. Consequently, the network load per page request will increase by not more than the size of the mini graphic (43 bytes). Since all requests for pages are recorded, an easy and precise determination of the number of Page Views is possible. On the other hand, the method classifies user actions by “click streams”. A click stream is generated by a sequence of page requests of a given web browser over a specific period of time. The **method identifies individual browsers** by analysing supplementary information, which is not recorded in normal log files.

Results of measurement

Federal Statistical Office of **Germany** had been present on the Internet since March 1996 and **first measured requests** for its data between November 1997 and February 1998. In order to obtain results for total number of requests only an “analogue” method was applied.

In the reference period the number of requests for data reached a total of 862,603 pages, with 60 % of the relevant users coming from Germany. An evaluation with regard to the **days of the week**

showed that **requests** had been made primarily in an **occupational or scientific context**: an average of about 26,000 requests were made on working days, compared with 12,000 requests on weekends. Pick hours were largely within working hours in Germany. By selected chapters of the time on the site, “Press services” (including press releases issued since 1 January, 1996) accounted for 144,990 (16,8 %) requests. From this, 117,913 pages of the German-language offer were visited. Among these the “press home page” was the first (10,815 requests), followed by the service page “breakdown by subject fields” (7,257), the page “current press releases” (5,280) and press releases, chronological” (2638). “Indicators”, a section, which gave quick information about economic trends, had 38,691 requests, while “Time series” was available to 670 registered clients. **Comparison** of these **analogue results** to those obtained by **extension of HTML** pages revealed a moderate difference in requests. When surveyed by analogue method, the German-language home page counted 99,760 requests while measuring by extension method it showed 98,760 requests.

5.4.2 User sessions in focus

With the advent of the web, statistics about activity of websites had proliferated. Webmasters commonly threw out numbers about the number of hits their sites received. But by 1998, there was little understanding about what was being measured and the meaning of the measurement as the methods were in the earliest stage of development. Most of the measures resulted from the **need to manage web servers and not to manage their content** and simple counts did not provide real information.

Being aware of deficiencies, **U.S. Bureau of Justice Statistics** in 1997 undertook an attempt to sweep website activity by analysing its outside use during a 16-week period. In the course of survey, a total of 155,216 user sessions were logged and the site averaged 1,424 **user sessions** per day. The length of the average user session on the site ranged from 7.5 minutes to slightly over 8 minutes. Over 100,000 downloadable files were requested, including 19,000 press releases.

Press releases neglected

The analysis looked at the number of user sessions where a **particular web page** was **requested**. A total of 298 different HTML pages on the site, almost all, were requested at least once during 16 weeks. The least requested pages were either relatively old, or very new. **The pages most frequently requested** appeared to be in the topical data sections of the site. While on the top was the “Publications alphabetical listing” with 11,404 sessions, the press release page was requested in 3,060 user sessions and ranked 32nd in the total number of user sessions. Links to the press releases were on the report abstracts and the “Press release” page. BJS issued during 16 weeks 5 press releases, but they did not appear to greatly affect the requests for these files on the website.

Unvisited “What’s new” page

The feedback provided from analysis proved to be useful in confirming many decisions to be taken about the content of the site, suggesting changes to it and planning for new content. The survey found that only a **fraction of users** were **going to the “What’s new” page** where updates and notices had been placed. Since the site was changing constantly and the agency was offering many electronic only products, it became clear that it could not more rely on the “What’s new” page to alert users to changes. Therefore several **outreach activities** had been **initiated to notify users** of changes to the site. Amongst them was the announcement for major updates and releases in JUSTINFO, the National Criminal Justice Reference Service listserv and preparation of a semi-annual catalogue in paper announcing updates to electronic only files.

5.5. Making use of new technological developments to better meet the needs of the media and get their feedback

5.5.1 Census: full use of technologies

During the 1990's, the **U.S. Census Bureau's** Public Information Office had embraced **new technology** to help disseminate its materials to the country's media. The office established **video and radio production studios** to serve the broadcast media to the same degree as the print media.

As Census 2000 unfolded, the video production area turned out a steady stream of materials, including national **video news releases**, video news feeds aimed at specific markets, editions of stock shots for television stations, public service announcements, and satellite media tours. The same variety of products was created for radio stations. The Internet was used extensively to service the media.

The director's **periodic news briefings** were carried **live on the web**, and **archived** for later reference. At the same time, reporters from around the country were able to take part in these news briefings by special telephone line.

Census 2000 saw a 5 percent rise in households responding by mail over 1990, allowing follow-up operations to be conducted in a shorter time. While media relations were only one part in a five-point program, it made a sizeable contribution to this success.

PIO's operations saw a combined audience of 178.3 million for its video products, while print efforts produced some 20,000 newspaper clippings, twice those generated in 1990.

The 1990's saw an **explosion of new communications technologies**. The Internet had become a major factor in U.S. life. The three through-the-air television networks were no longer dominant. Cable News Network and Fox were at the forefront of greatly increased penetration of U.S. households by cable and satellite-delivered programming. The average American now had less time to devote on information and entertainment, even as their sources for these diversions multiplied. The **communications program** of the Census Bureau to support Census 2000 was developed amid **political controversy, major operational changes**, the need to complement a paid advertising campaign, increasing language and cultural diversity, decreasing levels of civic participation, and major growth in information technology,

Role of the Public Information Office

The responsibility for dealing with the print and broadcast media about all Census Bureau programs rests with the Public Information Office (PIO), which organizationally reports to an Associate Director for Communications. In turn, this person reports directly to the Census Bureau director. This **simple, direct structure** assures that major problems receive immediate attention.

To adequately represent the full spectrum of the bureau's work, PIO developed good **working relationships with subject matter divisions** and key officials throughout the organization. PIO worked with the media throughout the decade, and has a staff of approximately 40 for this effort, divided into **three teams**: One **researched and wrote all materials** on the bureau's ongoing demographic and economic data; the second produced **materials for radio and television**, as well as providing still photography support; the third is **administrative in function**. For Census 2000, this staff was augmented by a temporary team of 20. Its job was to coordinate all aspects of the bureau's media effort on behalf of the decennial census.

Embracing Emerging Technology

Media outlets used to complain about **receiving embargoed material too late** when it was distributed **through the mails**, particularly those located on the west coast. The solution was to switch to a computer-controlled fax machine, this assured almost simultaneous reception throughout the country.

While the office historically had done a **good job of reaching the nation's print media**, its broadcast efforts were purely **reactive** in nature. Noting that an increasing majority of Americans receive their news and information via radio and television, it was decided to **establish internal radio and television production capabilities**.

A temporary video production group, formed to support the 1990 census, was made permanent. Beginning in 1996, a series of video news releases (VNRs) was produced and distributed, **highlighting major data releases**.

Shortly after the video facility was established, a **radio production studio** was built. A weekly, magazine-format program in English and a companion one in Spanish were produced from 1993 to 1997. At that time, the programs became daily, 60- second features, keyed to each specific calendar day. The **digital editing equipment has been upgraded**, and the most recent addition is an ISDN line, allowing studio-quality audio to be fed to any station in the country with similar capability.

Still photography support has seen a shift from 35 mm film to digital images. While both formats remain in the inventory, digital images are preferred by the print media.

Census 2000 Media Operations

While PIO operations changed during the 1990's to accommodate new developments in technology and in the way the media operates, the pace picked up significantly in 1999 and into the census year of 2000.

One cause was the increasingly active political framework in which the census was planned and executed.

Changing from a largely reactive organization onto one, which has hard news operations, and is taking the **initiative** to use the media's daily news cycle meant paying an increasing amount of attention to our media contacts. The top 25 or so newspapers, wire services, and broadcast outlets had been identified and good working relationships established. But analysis of the hard-to-enumerate areas in 1990 showed that a much wider group of media outlets needed to make aware of the census as an ongoing story that affected their local area. To this end, a number of initiatives were undertaken:

- Vendors were contracted to supply accurate, complete media lists, broken down by type of media, audience, and geography. These services provided terminals and proprietary software to allow the preparation of customized lists.
- Workshops at media organization conventions were conducted to familiarize reporters and editors with the complexities of the census.
- Tours were arranged for groups of reporters at census headquarters and at local offices, to acquaint them with the census and its potential as a news story.
- Every effort was made to answer media queries quickly, whether by phone or e-mail. Staff was trained to find out reporter's deadlines.
- A cadre of top officials was given media training and made available for interviews. These often occurred in hectic circumstances and were conducted via cell phones.
- A steady stream of timely material was made available on the Internet, both the general website and the special PIO embargoed site for media.

To ensure flexibility and speed, a Rapid Response Team was established.

Broadcast

The PIO Broadcast Service Team normally is made up of three video producers, a radio producer, and a radio writer, aided by three support staff. For 2000, more than 40 contracts were awarded for

various radio and video professional services, and a rotating cadre of some 15 to 18 contract video producers, writers, editors and researchers worked continuously at the Census Bureau. The regular staff concentrated on planning and contract supervision.

Video

A full range of video products was brought into play, forming one of the key elements in the communications campaign:

Video News Releases (VNRs). These are video building blocks, offered so TV- stations may produce their own stories. They are fed via satellite, so any station may pick them up and use them. Typically, they run from 5 to 12 minutes, contain several information slates, sound bites with one or more interview subjects, and stock shots, called B-roll. The key to success with VNRs is the notification process, a combination of advance faxes and telephone calls. Since VNRs are fed nationally, they are

used mostly for large-scale stories.

Video News Feeds (VNFs). These are much like VNRs, except they are fed to selected markets and even single stations. Again, notification is the key. While VNFs are fed by satellite so any station could pick them up, only those stations selected to solve a particular communications problem are notified of the feed.

Satellite Media Tours. These are feeds to single stations, in which a Census Bureau official is offered for an exclusive interview. They may be live or live-to-tape. Because these feeds are exclusive to a single station, they generate a strong interest by the station concerned: they also take considerable staff/vendor time to set up.

B-roll (stock shots for television). PIO produced a new edition of b-roll approximately every four to six weeks, usually keyed to a new phase of Census 2000 operations. These were used extensively by stations and networks since they illustrated operations that could not be covered by local crews because

of confidentiality laws. It was very important that the b-roll be of professional broadcast quality to ensure its usage.

Public Service Announcements (PSAs). These are spots the same length as commercials (:10, :20, :30 and :60) offered to stations to play as a public service in support of a non-profit undertaking. PIO produced a number of sets of PSAs, several of which featured “doughnuts” or holes in which a soundbite from a local dignitary could be inserted, increasing the local appeal of the spot.

Specialized Event Videos. These are short productions (3 to 8 minutes), usually used in conjunction with a live presentation. They may be for a one-time only use, such as a specific speech, or general in scope for widespread use.

In addition, monitoring services routinely recorded news programs from several hundred stations in key markets and offer air-checks for a fee. This is a valuable service, because it allows an evaluation of how census stories are handled by the media.

Radio

The use of radio by the **Census Bureau** saw great growth during 1999 and 2000. There are some 10,000 commercial radio stations in the U.S., compared to less than a 1,000 television stations.

The following key elements comprised the radio communications effort:

Radio Media Tours. These are the audio equivalent to Satellite Media Tours for television, described earlier. These interviews may be carried live, or recorded for later airing. Because of the local angle and the exclusivity, stations are usually eager to take part.

A typical Radio Media Tour will see the Census Bureau director in PIO’s radio studio, being interviewed by a succession of radio stations at pre-arranged times. While the interview is actually conducted on the phone, the phone line is patched through the control board, and the subject speaks into a microphone and wears headphones.

These tours are a useful tool to project a message into a specific situation, such as countering a negative newspaper story or editorial in the same market. Again, these feeds require considerable effort by staff/vendors to arrange.

Audio News Releases (ANRs). These are the radio equivalent to Video News Releases. Audio material is supplied to stations to be used in producing their own news stories. The ANR may be fed by satellite across the country, or to individual stations, depending upon the audience it is desired to reach.

Daily Radio Programs. PIO produces two daily, 60-second radio programs, one in English and one in Spanish. Beginning in January, 2000, the normal tag line: "Profile America is brought to you as a public service by the U.S. Census Bureau," was modified to read: "Profile America is brought to you as public service by the U.S. Census Bureau, now counting down to Census 2000." This line was used until the end of March. From April through July, the tag line read: "Profile America is brought to you as a public service of the U.S. Census Bureau, now conducting Census 2000."

Public Service Announcements (PSAs). Key census advertising and public relations messages were reinforced with PSAs produced by PIO. These were sent to stations directly, and also sent to field staff to distribute to stations in their area. Some PSAs were produced in both English and Spanish.

Photography

The office also provided a variety of still photographs for print media. These shots were usually taken at the same time simulated operations were being videotaped for the next edition of B-roll.

Photographs were made available **in two formats** – on Compact Disks or through Internet/Intranet.

News Conferences

The idea of holding news conferences by officials for the media is well established. However, for Census 2000, it was decided to stage an ongoing series of such events, termed "Operational Press Briefings." These were held by the director every two to three weeks on average, but occurred more often at the height of taking the census.

These briefings used advanced technology to go beyond the normal parameters of a Washington news conference attended by a cadre of local reporters. As a result, they became national events. Each briefing was broadcast live on the Internet.

Following the briefing, the video of the event was archived on the website for a period of 90 days. A verbatim transcript of the event was also put up on the bureau's website.

Internet/Intranet

The Internet proved to be a valuable tool in conducting media relations on behalf of Census 2000:

- It allowed treat all media equally. A password-protected server was established, allowing three-day advance access to news releases and report data.
- It also allowed make accessible a wide range of reference material.
- Another key use of the Internet was receiving e-mailed queries from reporters.
- The establishment of an Intranet site allowed PIO to share in advance with Census Bureau staff in the field policy documents needed to prepare the staff for local media queries.

Accomplishments at a Glance (October 1, 1999 - July 31, 2000)

Video. 202 productions reached audience of 205,886,202.

- ☐ 23 Video News Releases
- ☐ 11 Regional Video Feeds
- ☐ 2 Satellite Media Tours
- ☐ 63 Target City Feeds
- ☐ 17 B-roll editions
- ☐ 31 Sets of Public Service Announcements
- ☐ 17 Webcasts of events
- ☐ 38 Video Titles Produced and Distributed

Radio. 108 productions, broken down as follows:

- ☐ 24 Radio Media Tours

- ☐ 8 Sets of Public Service Announcements
- ☐ 32 Audio News Releases
- ☐ 31 Soundbite Assemblies
- ☐ 13 Interviews Conducted From Studio

News Conferences.

- ☐ PIO arranged for 31 news conferences held by the director.

Internet Queries.

- ☐ PIO staff handled more than 8,000 e-mail queries.

News Releases, Statements, Advisories Issued.

- ☐ Over 140.

Press Clippings.

- ☐ There were 24,446 clippings gathered on Census 2000, compared with 10,662 during comparable period in 1990. Heaviest day was April 3 afternoon edition (clippings are issued twice a day), with 308 clippings.

5.5.2 Vision of statistical libraries' mission

“How can we provide an environment where users, novice and expert alike, can publish, on the Web and elsewhere, statistics of a demonstrably high quality and transparency to permit their use in a global context?”

An R&D funded project envisaged in 2000 one of possible **systems** in which existing statistical **data and metadata can be combined** and delivered to the desktop of any authorized person. **MISSION** (Multi-agent Integration of Shared Statistical Information over the [inter] Net) had the vision of having a number of independent organizations publishing their data within a framework that makes comparisons and harmonization possible. The project aimed to provide a **modular system of software**, which would enable providers of official statistics to publish their data in a unified, and unifying, framework, and will allow consumers of statistics to access these data in an informed manner with minimum effort.

Software's abilities

- ☐ Allows suppliers of statistics to subscribe to an integrated network of data-stores via an interface to their existing data, while retaining control over all aspects of access to their data. This includes their level of involvement; the data they supply; the users who can access it; and the level of resources to commit.
- ☐ Allows users to make declarative requests, with a minimum of understanding of statistics, or the domain area, and still retrieve meaningful results from our internal routines or through an interface with external statistical packages.
- ☐ Gives the user a range of options for automatic harmonisation of statistical data, with clear indication on the interpretation of the results.
- ☐ Provides audit trails of data manipulation and analysis, so that methods can be retained, re-used and published.
- ☐ Maintains libraries of metadata that can be made available to other users.
- ☐ Provides a flexible architecture that allows third parties to act as Independent Metadata Providers, thus encouraging the free exchange of knowledge.
- ☐ Allows users to build up individual profiles, accessing data and methods most relevant to their needs.
- ☐ Offers a number of independent, interoperable systems that can run on different hardware platforms and access heterogeneous data storage systems.

Multiple agents, multiple innovations

The innovation of MISSION is twofold. First, it adopts emerging technologies and combines them together in an imaginative way, and second, it proposes an innovative organizational structure, which will overcome some of the difficulties statistical offices have in integrating new technologies into their current practice.

Four innovative strands combine in the MISSION software: agent technology, XML based descriptions of metadata, harmonization and data merging techniques, and information self-organization.

Agents are used to provide an intelligent interface for the user. The user supplies the agent with a rough description of the information that he/she is interested in. Using XML descriptions of statistical data, the agent tries to locate the relevant data that satisfy the user's interests. Once this is located, a query can be formed with (or without) interaction with the user. These query agents can interact with the interface agent for a query answering mechanism to supply the answer. Query agents take a declarative approach to statistical queries. A user can state his needs in terms of the goal that is a statistical (macro) **table, plus its related metadata** (footnotes).

Mediation agents provide the mechanism for harmonising metadata without the need for a predefined global ontology (a set of shared concepts to which all data definitions can be mapped). Instead of a universally shared ontology, each user will carry his own personalised ontology that can be used for the mapping. Alternatively, **third parties can act as repositories** for shared ontologies in a particular domain or country. When a new data provider is accessed, there is a need to negotiate a modified shared ontology – the combination of the old shared ontology with the new provider's ontology.

The Architecture

The architecture of the system comprises four basic logical, or conceptual, units or building blocks, which can be deployed in different scenarios. The components are:

- The Client, which provides the interface to the system.
- The Library, which is a repository for statistical metadata.
- The Compute server, which is a statistical processing engine and stores no information of its own.
- The Data server, which is the unit that gives access to the data.

The Client

The Client component is a Web based user interface that connects a user to all sites participating in the architecture. Obtaining a request from the user, and sends an agent to search for a Library that can satisfy the request.

The Library

The Library software supports a repository for statistical metadata. Different Libraries can communicate with each other. A Library holds three different kinds of metadata. The most basic type of metadata is *access* metadata, which is the physical, and logical information required to access statistical data. The second kind is **methodological metadata**, which is the information required to process that data in order to satisfy requests for statistical analysis.

The third kind of metadata is **contextual metadata**, which supplies background information and explanatory notes for the user. This kind of information includes, for example, the purpose of a survey, or an explanation of a break in series for a time series. The first two types of metadata are machine understandable. The last is machine readable and human understandable.

When a **Library** receives a request, it **decomposes** it, and, if necessary, it can send to other Libraries in the system for any metadata it requires. On receiving the reply to the request, it returns the answer to the Client.

The Compute server

The Compute server is a statistical processing engine, which stores no information of its own. Based on the query it receives, it obtains the necessary data from various data servers, performs the request, and returns the result to the Library that made the request. It may also make a request to third party statistical packages.

The Data server

The Data server is the unit which gives access to the data. It holds the data itself, management tools for registering and maintaining the system and a gateway module. The gateways hold the minimum amount of metadata necessary for the safe use of the data.

Actors

Main Actors in the system could be identified:

- The User: the user is **anybody with access** to a Client who wishes to make a query on statistical data. Users will have different levels of statistical expertise and domain knowledge, and the Client will cater for their differing needs.
- The Statistical Data Provider (SDP): the SDP is a **site that has data**, which it would like to offer for statistical analysis.
- The Third Party Provider: This is a site which is neither user nor SDP, and which houses one or more of the basic units.

Implications of using

The architecture offers a **variety of configuration models**, so that private and public libraries of expertise can be built up. A private library may be local to a department and used to hold the processes needed for their day-to-day operation, while the public library would hold the processes needed to map the collected data into a format or aggregation needed by other actors in the business. The library will have registration formalities, which ensure that only valid users have access to sensitive or controversial information.

The **third party libraries** are maintained to a quality standard. This is of equal importance in a system that is entirely internal to an organisation as the quality of the management information available is controlled by the quality of the metadata of the poorest library.

5.6 Experiences in evaluating the quality of communication with media

5.6.1 Reports about the press

*There were several ways in which **Statistics Finland** acquired **feedback information** on communication activities aimed at the media. These were as follows: 1) press-clippings service and the related analysis, 2) survey interviews and free discussions with journalists and 3) general data on public opinion regarding Statistics Finland. The feedback system was fairly comprehensive, while the analysis of the acquired data was not thorough enough.*

Press-clippings

Press clippings were collected daily for Statistics Finland by a private organisation. The number of clippings is about 15 000 a year. The papers followed amounted to 400, whereof about 50 were daily papers.

The clippings were immediately mailed to Statistics Finland. They were then piled by the release they related to or, if there was no such release, by the statistical unit, which was the source of information used in the story. The clippings, which needed to be commented on, were then quickly sent to the statistician in charge of the problem.

*The **incoming data was then reported statistically** once a month by organisational unit and release in internal bulletin board system. Statistics were based on the number of clippings. The number of so-called negative clippings was also calculated. The level of 'negativity' of a story was evaluated in the press office of Statistics Finland.*

In order to obtain more information out of the data a sample of about 1 000 clippings was picked out.

*The sample was **analyzed** more thoroughly: the size of the story as well as the use of pictures, graphics and tabulations were counted. The type of the paper and that of the story were also coded.*

The sample study was reported annually with more qualitative analysis of the press-clippings.

Interviews with journalists

Representatives of the most important journalist organizations in Finland would be invited to visit Statistics Finland. Groups of about ten journalists visited about once a month; more groups in wintertime and less in summertime. Journalists were normally interested in getting acquainted with Statistics Finland. The program of the visit would begin with a general presentation of Statistics Finland, continued with special themes of interest or demos and ended with free discussion. Most often the visits would lead to some kind of co-operation afterwards. These visits were very useful from the point of view of the changes we planned in our releasing system or other services. Statistics Finland had also participated in a survey research conducted among Finnish journalists.

General data on Statistics

A general survey on public opinion concerning Statistics Finland was conducted every other year. There are about 1 000 respondents in the survey, which was conducted by a private survey organization. The survey had been performed since the 1970s, which means that there are great time series based on it.

In a way this survey gave us the end result of our PR-activities: an indicator of how many people knew Statistics Finland and an overall profile of attitudes connected with the office. The profiles of attitudes towards other Finnish organizations were also included for the sake of comparison.

The results of this survey seemed to correlate reasonably in time with the press clipping data: the growth in the number of clippings seemed to connect with the rising knowledge of Statistics Finland. The attitudes of the stories seemed to be reflected in the attitudes of the public.

In addition Statistics Finland conducted annually a survey among its clients. Although news media had been allowed a considerable amount of free service, they were often among the paying clients, too. The evaluation in the survey did not make distinction between different kinds of services. Thus the feedback was obtained on the news service as well.

Problems

The number of press clippings was so immense and their processing such a tedious job that there was not enough time to analyse the material.

Another problem is that it had not received a good and clear feedback from electronic media.

Good comparability between organizations is of utmost importance in survey research.

Comparisons could be made with other Finnish organizations, but owing to the special nature/role of a statistical agency that was not enough. It would be useful to have internationally comparable indicators on how many people know the name of the statistical agency and what kind of attitudes people (or journalists) have towards statistics in general.

When working with journalists the lack of time is always a problem. Thus journalists are not very eager to take part in evaluation surveys. That is why it is often difficult to get systematic and reliable data on their attitudes. On the other hand, they give their criticisms very freely, when they see concrete problems in their work.

5.6.2 Using customers' feedback

With an increase of the statistical information available in the output database StatLine, improved accessibility through the Internet free of charge, **Statistic Netherlands** has encouraged wider use of its statistical information during nineties.

In the context of its wish to become a responsive public information provider, the office has progressed from being a highly product-centered organization to a more client-oriented entity.

On the basis of this more client-oriented focus, the wishes of important customers and Customer-groups were increasingly being taken into account when new statistical activities and new information supply operations were set up. However, this approach had to be balanced with other considerations, one of them being the paramount cornerstone of Statistics Netherlands' policy not to sacrifice objectivity to business considerations.

Poor TV coverage

The main statistical results are published and disseminated through **press releases**. Various steps had been taken **to improve coverage** by the news media: the quality of the press releases had been improved and they were better planned throughout the year, preventing the publication of more than two press releases on the same day. The news media were also informed well in advance. These activities had resulted in an **increased coverage** in newspapers and journals; many summaries of press releases were published directly verbatim. In 1998, 254 news items were released to the press; the national newspapers published 1,380 articles based on these news items, 10% more than in 1997. There was also a **rise** in the number of **newspaper interviews** to 200 in 1998.

Television, on the other hand, had **almost no coverage** of the press releases issued at 9.30 a.m., because their newsrooms did not have time to prepare a visual news presentation. As TV current affairs and news programs are the main source of topical information for the Dutch public. Following a successful try-out in 1997 the office offered the radio and television news programs on a contract base:

□ To discuss every Friday the press releases planned for the following week and the type of information expected to be included. Program producers could thus obtain an idea of what might be of interest for them. All this information was confidential and might not be used before the press release. Information that might effect financial market prices is not given at all.

□ If requested, human-interest information (health, traffic, environment, tourism and demography) may be given strictly confidentially 18 hours before the press release, in aid of a film report.

The effect of this policy was a substantial rise in the number of interviews and in the presentation of statistics on radio and television.

Exploiting Internet

The form chosen on the Internet in 1996 followed the familiar presentations of information on paper, and was at least consistent with the other media Statistics Netherlands used. However, the Internet has its own rules and dynamics and one came to the conclusion that simply translating existing products to an Internet version is not enough.

Graphic design also became an important issue: how to present the information in such a way that users can find what they are looking for? The most popular products on the site were key figures, StatLine, and the weekly Web Magazine with respectively 17,103, 14,636 and 6,610 hits in November 1998.

The use of **Internet** in the Netherlands had showed a very **high growth rates**, although it was slowing down: 40% up in 1998 compared with 85, 68 and 53% in 1995, 1996 and 1997 respectively.

The **average age of the users** was 36, and most were men and boys (78%). Over half of the users exploited the Internet privately, although 36% also used it for business. **The favourite activities** on the Net were surfing (33%) and e-mail (30%). The most popular surf subjects were specific information (39%), news (23%) or entertainment (14%).

Database: flexible output

To provide for a consistent, timely and efficient publication process, flexible standardized medium-neutral output database StatLine has been developed. This output database must be situated between the internal statistical process and the dissemination and must create conditions for 'tailor-made' statistical information for many different media. (see 4.2.) Besides 'one stop' on-line access to the database, such a system could also **automatically provide information for e-mail** subscriptions, faxes, CD-ROM and paper publications.

If all available statistical information is accessible on the Internet, free of charge, millions of users can access it. The implications in terms of accessibility might be dazzling when compared with the few users reading printed publications. In 1999, for instance, the annual sales of about 15,000 Statistical Yearbooks with the 86,433 key figures users or the 123,222 StatLine users through the Internet in 1998 (user is defined as a specific customer of the tables of that publication on a certain day).

*With the improvement of accessibility through the Internet, Statistics Netherlands decided to **terminate paper publications** with fewer than 100 subscribers.*

Good old telephone

*As stated above, visitors to StatLine on the Internet or CD-ROM often need help from staff to retrieve, or to be sure to retrieve, the right information for their purpose. The **telephone information service fulfills this function.***

Statistics Netherlands market research in 1994 showed that clients greatly appreciated a telephone information service which could be reached directly and which could directly, or in one step at the most, connect them to someone who can provide the statistical information required.

In 1995, besides the official call service for price index figures, a general information service was started where staff was trained to provide direct information about published statistics.

*Moreover, throughout Statistics Netherlands, **subject-matter specialists were available** as a 'second line' to answer more specialized questions and provide 'tailor-made' information and analysis from 'publication (micro) databases' as well. This general information service had been a great success, with a growth rate in client requests of 68% during the time period from 1995 to 1997. The service for price index figures went up by 10% in the same period. And, at the same time, a fall was observed in requests for information from specialists. For the bureau as a whole, the rise in information requests in those three years was 14%.*

*This increase led to capacity problems. Although pricing policy in general is that individual client services should be cost recovering, in practice the **telephone service as a whole is not cost recovering**, and therefore no extra staff could be taken on.*

*To meet these problems, management decided to re-organize the telephone service towards the year 2000, putting it more on a **cost recovery basis** and more efficiently applying new technologies such as computer telephone integration (CTI) and automatic call distribution, with customer contact servers and, where possible, case-based reasoning.*

*Statistics Netherlands market research in the spring of 1997 and of 1998, i.e. before the renewal of the site, showed that, of all products and services, **the telephone service was the most appreciated by clients.** It got a better rating than paper publications and the Internet site on all values.*

5.6.3 Do statistics need news media?

If increasing numbers of users of statistical materials retrieve their information electronically in the future, according to **Statistics Denmark** one may ask questions such as:

How will the changing behaviour of statistical users affect the roles of statistical offices and news media as distributors of statistical information in the future? How will this affect the **relationship between news media and statistical offices**? And finally: may statistical offices grow to an independent media?

With more and more people accessing data directly from the source, the roles of news media might change further from being distributors of objective information towards being critics, seeking the story behind the figures. At the same time, statistical **agencies** seem to **become their own** (objective) **media** communicating directly with users and the public, diminishing the need for attracting the attention of other media.

The introduction of Denmark's largest Internet Databank, on 1 March 1999, proved to be a great success, showing that many potential users welcomed an easier access to data.

*In Denmark between 1993 and 1998 the number of "**tables**" **retrieved and paid** for by external users of the main online databanks **increased** roughly by 35%, while the number of paper publications sold have decreased by some 14%.*

*However, the type of products that have increased in terms of items sold, were not necessarily obvious substitutes for the type of products that have decreased in terms of items sold. Not only the number of electronic products had increased, but also the number of users obtaining free data from website. Thus, total **electronic dissemination had shown an increase** and was expected to grow even further, with more potential Internet users and more sophisticated products to be offered in the future.*

Possible commutation of functions

The way of **searching for information was changing** and affected those offering the information. With easier access and much to offer, the Internet gave users the possibility to play an **active role in accessing** information.

The web is the natural place to search for statistical information, as it can be obtained from the original source, at the exact time of release. This means that the **role of the traditional news media will also change**. It will be difficult to sell newspapers containing such information, as it is easily accessible elsewhere. Thus, **quality news media** will have to write serious articles, where they **analyze the information** and put it into perspective, **rather than distributing the bare objective facts**.

At the same time, the **distribution channel** from statistical offices to the final users is **becoming more direct with Internet dissemination**, and less importance placed at the news media channel. Thus, the role of statistical offices as being their own independent (news) media will be highlighted in future years, making it easier to fulfill the obligations of serving users and the public.

Categorising users of statistics into one group that actively seeks data and one group obtaining the information more occasionally, e.g. from news media, it seems to be the first group that will be easier to come into contact with via the Internet in the future. However, if the obligation of NSOs is to inform the public in general, the news media still seems to be an important channel, in order to reach the second group.

Although there seems to be a tendency that NSOs will rely less on other media for dissemination purposes in the future, there are other aspects of the relationship between news media and NSOs. News media play an important role in expounding the public opinion and setting the agenda for relevant issues in the debate of the development of society. These are also important issues for NSOs when deciding what statistics they should produce. This role, and the role of news media as critics in relation to NSO products will not diminish with more Internet dissemination. Thus, statistical offices will have to deal with negative press coverage also in the future, but it might be easier to retort through own distribution channels.

At the same time, news media will continue to be a distribution channel for statistical agencies in case of reaching a broader public audience, and as reviewers of new statistical products.

Free databank access

*On 1 March 1999 **Statistics Denmark** introduced a new Internet databank and a new and improved layout of its website. The number of visitors at website almost doubled during the first month. With only 250 paying customers on "old" online databanks, the expectations for enrolment to the new system were modest. However, more than 8,000 applied for access to the system within the first two months, and the enrolment seemed to continue.*

*A larger share of the new users could come from the thousands of users, with specific interest in statistics, visiting the website every day. Thus, a qualified guess is that the effects of **newspaper reviews had been limited** in proportion to the promotion via the Internet itself. If this conclusion holds, the role of **National Statistical Offices as independent media** could be even stronger in the future.*

The apparent change of behaviour in the way people search for information will also influence the **relationship between statistical offices and the news media**. Although NSOs could expect to strengthen the dissemination channel, they will still have to deal with **negative press coverage** from, what could be even more critical news media, in the future.

PAPERS USED:

- 5. (Techniques...) None, G5 is on lock-ups in the U.K= 4.1.1., P ECE Secretariat, C20 (Istat),
- 5. 1. (Impact...) C8, C10, C11 outdated, specific, none of them used; instead C room paper: USDA, P4, P11, P9
- 5.1.1. Guidelines on selecting content for Internet and web dissemination (=3.4. Effective website)
- 5.1.2. (Pricing...) C6, P2, Eurostat, V8, C20, Ot8, Ot14, C10, O13.
- 5.1.3. (Measurement and evaluation of Internet use...) C18 Germany, C19 BJS
- 5.2. (Making use of new...) G13 Put into Contacts with press in Norway? Transferred, V14 Census Bureau, V16 MISSION,
- 5.2.1. (How to take advantage of new technologies to increase the visibility...)V7 about Eurostat's website transferred to 4.2.1.
- 5.2.2. (How to benefit from new technological development...)
- 5.3. (Feedback from users...)
- 5.4. (Experiences in evaluating the quality of communication...) Ot7, P2, P13