STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

ECE Work Session on Statistical Output for Dissemination to Information Media (Geneva, Switzerland, 28-30 October 2002)

Topic (b): Methods, Policies and Techniques of Statistical Dissemination

DISSEMINATION OF INTERNATIONAL LABOUR STATISTICS ON INTERNET

Submitted by International Labour Organization¹



http://laborsta.ilo.org Copyright International Labour Organization

Prepared by Patrick Cornu and Sophia Lawrence.

La Diffusion des Statistiques internationales du Travail par le BIT sur Internet Résumé

Le BIT et son Bureau de statistique diffusent des statistiques du travail pour les Etats membres depuis plus de 80 ans. Cet article décrit les différentes étapes du développement de la dissémination des informations statistiques, par exemple la plus grande disponibilité des extraits de base de données rendue possible par la technologie des dernières années. Le récent site Internet statistique du BIT, créé en 1998 afin de diffuser avec la meilleure technologie les données et les métadonnées significatives, y est présenté au lecteur. Ceci représente une innovation importante dans le domaine de la diffusion des statistiques du travail mise au bénéfice des utilisateurs. L'opinion sur le site et sur les statistiques qui s'y trouvent ainsi que des renseignements sur l'utilisation qu'en font les divers utilisateurs du site BIT est sollicitée pour pouvoir améliorer les développements futurs. En dernière page, les lecteurs sont de ce fait invités à réfléchir aux questions qui s'y trouvent et nous renvoyer leur réponses par courrier électronique à l'adresse: laborsta@ilo.org. Nous les remercions d'avance.

La Diseminación de Estadísticas internacionales del Trabajo por la OIT en el Interne Resumen

La OIT y su Oficina de Estadísticas distribuye estadísticas del trabajo para los Estados Miembros desde hace más de ochenta años. Este artículo presenta las etapas diversas de desarrollo de esta diseminación de informaciones estadísticas, incluyendo la mayor disponibilidad de bases de datos que se debe en gran parte al avance tecnológico de los últimos años. Se presenta también el sitio Internet de la OIT sobre estadísticas del trabajo creado en 1998 para difundir los datos y meta-datos más significativos, utilizando la mejor tecnología. Este sitio es una innovación importante en el área de la diseminación de estadísticas del trabajo. Para poder mejorar el desarrollo de esta diseminación en el futuro, nos interesa conocer su opinión sobre este sitio y sobre las estadísticas incluidas; también nos interesa contar con información acerca del tipo de utilización que hacen los usuarios de este sitio. Es por esto que se invita a los lectores a reflexionar acerca de las preguntas que se hacen en la última pagina del artículo y a enviarnos sus respuestas por correo electrónico a: laborsta@ilo.org. Muchas gracias de antemano.

Introduction

The ILO and its Bureau of Statistics have been disseminating labour market statistics to an international audience on behalf of ILO member States for over 80 years. This article presents in three parts the main components of these long-term activities. Part I is the history of data dissemination which describes the various stages in the development of statistical information dissemination. Part II covers the statistical and methodological information currently available in the Bureau, mentioning the greater electronic availability of data sets made possible in the past few years. Part III introduces the ILO statistical Internet application which has been developed to disseminate data and significant metadata, and represents an important innovation in the area of ILO statistical dissemination. Feedback on the service provided by the web site as well as on the uses made of the labour statistics made available by the ILO is requested in order to facilitate future improvements. Readers are therefore invited on the last page of this article to answer a few questions concerning their impressions and their use of the web site and respond by email to laborsta@ilo.org for which we are very grateful.

PART I: History of data dissemination

Since the early 1920s, an important role of the International Labour Organization in the international arena has been to make available to the widest possible audience the results of national statistical inquiries into the characteristics of their working populations and their conditions of work and life. This "distribution of information on all subjects relating to the international adjustment of conditions of industrial life and labour" started when price statistics and unemployment began to be published in the International Labour Review in 1921 using the technical means of the times. In the years which followed, the Review supplemented its coverage of "labour statistics" by adding employment, and increasingly to include other topics such as statistics on wages and hours of work, and information on industrial disputes and collective agreements. The Review was eventually produced as a "monthly" for the price of 60 cents per issue or on a yearly subscription basis! It could only be obtained from the publishers or from distribution points in Canada, Switzerland, the USA and the UK, but, according to the publication itself "constitute[d] a unique source of information, since only the Office is in a position to secure all the relevant data."

A next step in the ILO's fulfilment of its mandate to carry out statistical dissemination was the publication of the *Year Book of Labour Statistics*, starting in the early 1930s. The official statistics presented then covered the "gainfully occupied population", employment and unemployment, hours of work, wages, cost of living and retail prices, family budgets, international migration, industrial accidents, production and wholesale price indices, and exchange rates. Already at that period, the statistics "were presented in time-series, in a systematic way with a view to assisting users in their interpretation, especially for international comparisons" ⁴. Contemporary printing and mailing methods ensured distribution of the approximately 200-plus page *Year Book* to users around the world. This impressive publication next became a trilingual (English, French, Spanish) edition and presented the most complete data obtainable "for some 50 countries in all parts of the world". It was truly a major reference work for intellectuals of the day. Today the latest edition includes 194 countries and territories. Over the years, the selection of subjects covered by the *Yearbook* (as it is now known⁵) changed for a number of reasons, some related for example to the altered division of labour between the United Nations family of agencies, to changes in the world of work and official statistics and some intrinsically related to the physical restrictions of paper and the printing press.

Other publications of data joined the *Yearbook* in later years: the *Bulletin of Labour Statistics*, a quarterly update of time-series presented annually, its *Supplements*, and the special edition of the Bulletin presenting the *October Inquiry Results*. The statistical tables in all of them give time-series, the longest being ten years in the *Yearbook* so that the 2000 edition covers the years 1990 to 1999. All of these ILO products have been prepared by the Bureau of Statistics using established practices. The Bureau is committed to this responsibility and is required by the ILO Governing Body as part of its mandate to serve national statistical agencies and to benefit researchers and professional statisticians and the public at large. It has continued to seek to improve upon the dissemination of statistical information, applying the techniques of the times. Thus, as the technical means evolved so have the Bureau's approach and strategies to information dissemination.

Use of recent information technology

In the early 1980s, newly available information technology enabled the Bureau to begin developing a database system for the storage and management of its collection of statistical information gathered from national statistical agencies. The bulk of the data from 1969 onwards was computerised. The database also contains

² Constitution of the ILO, Article 10 (1).

³ Advertisement for the International Labour Review, in Studies and Reports, Series N (Statistics) No. 25 (Revision of No. 19), Montreal, 1943.

⁴ Idem.

⁵ The change in spelling to one word "Yearbook" took place in 1995.

statistics on the total and economically active population derived from population censuses going back as far as 1945. The database on monthly, quarterly, etc. statistics published in the *Bulletin* start in 1976. The computerisation of a vast number of files was initially and primarily to prepare and produce more easily (by photo-composition) the regular statistical publications. The database facility contributed to reducing the time required for processing the data. It also served other significant functions in the Bureau, such as its analytical work, and lent itself to a swifter and more personalized satisfaction of external requests for statistics. As the facilities were built up and became better known, requests from users and their requirements increased in number and intensity, which led to a further evolution of the Bureau's activities as it strived to meet the higher demands.

PART II: Information available in the Bureau

LABORSTA and users of labour statistics

"LABORSTA" is the acronym which designates the whole set of ILO statistical databases covering all subjects for which the ILO is the custodian within the United Nations system. LABORSTA is managed and operated by the Bureau of Statistics. The contents of these databases have their origin in the commitment of the countries which are the member States of the International Labour Organization. By adhering to the organization and ratifying the ILO Core Conventions, Convention 160 on Labour Statistics, or the earlier Convention 63 concerning Statistics of Wages and Hours of Work, they thereby agree to comply to reporting on relevant information to the ILO, and in this case to the Bureau, on the statistics they collect. This reflects a commitment to sharing information in all domains - one of which only is statistics and constitutes a fundamental strength of the ILO. Other non-member countries and territories also cooperate in supplying their national statistics to the ILO by request, reflecting their recognition of the importance of the ILO and LABORSTA as a vehicle for international dissemination of their official statistics.

LABORSTA comprises nine major individual databases covering: 1) total and economically active population, 2) employment, 3) unemployment, 4) wages, 5) hours of work, 6) labour cost, 7) consumer price indices, 8) occupational injuries, and 9) strikes and lockouts. Custom-made extracts of any part of LABORSTA in printed or machine-readable form have been made available to the public for many years. Most requests are for long

time-series for one or several countries and often for more than one subject. Recently, the Bureau's own basic statistics on the type and origin of such data requests and how they are responded to show that users are interested in all topics and all areas of the world. Breakdowns of the different subjects by industry, occupation and/or sex are most frequently required. Many users want information in the form of rates rather than as absolute numbers or levels. Unfortunately, some frequently recurring requests cannot easily be answered based on the official national statistics in LABORSTA, such as worldwide aggregates ("What is the world-wide unemployment rate ?") or requests for combined results such as the share of wages in total production costs. Other queries difficult to answer take the shape of very detailed

Milestones

in

ILO Labour Statistics Dissemination

1921: International Labour Review

publishes statistics

1935: 1st Year Book of Labour Statistics

1964: *Technical Guide* series appears1983: *LABORSTA* database developed

1998: LABORSTA on the Web available on

Internet

disaggregates ("number of employed dentists"), and for statistics which do not form a part of the core labour statistics topics covered in the Bureau (labour turnover rates, job flexibility, training statistics, etc.) and/or because few countries produce such statistics. Some requests can be redirected to other sources such as the United Nations or other international organizations or to national statistical offices themselves. Sometimes we

can offer no solution or only approximations: a request for "number of blue collar workers" may be answered by suggesting the use of combined sub-groups of the International Standard Classification of Occupations. The Bureau's user-request statistics are, however, insufficient to tell us enough about customers' (changing) needs which, again, is why readers of this article are invited to provide answers and comments to the questions on the last page.

The labour statistics we can supply from LABORSTA allow users to perform various types of analysis over time as well as between countries and regions. All of the topics are important when depicting or quantitatively studying the world of work. Apart from the information they bring, the results of the research performed also contribute to enhancing the visibility of individual countries which might not otherwise receive as wide a review. Other than requests for statistics, numerous methodological and concept-related queries are also received. These are predominantly concerned with questions of definitions and issues of scope and coverage, or inquiries about ILO-recommended procedures, calculations, etc. Most of this type of information may be grouped into what is now referred to as metadata, the information surrounding the data.

Metadata

Metadata are any set of methodological information which describes and explains the statistics to which they relate. The ILO has regularly undertaken, in addition to publishing statistical tables, to provide users with detailed footnotes to tables as well as some basic methodological explanations of the statistics themselves. In the past, the latter took the form of introductory texts and later a series of *Technical Guides* which started in 1964 and continued up to 1980. After that year, the Bureau began disseminating information in a comprehensive series, now entitled *Sources and Methods: Labour Statistics* (see box below).

A novelty of the series is its organisation by type of data source, rather than by the subjects presented in the *Yearbook*, as in the past. Information is obtained through special questionnaires directly from countries and/or from national publications. There are currently ten separate volumes with various editions. Each volume gives standard, structured descriptions for data sources in the countries which provide the statistics published in the *Yearbook*. Volume 9 is a "one-of-a-kind" edition covering all the areas of labour statistics for countries in transition to take into account the major changes affecting labour statistics in their emerging labour markets.

The texts of the methodological descriptions represent the biggest portion of stored meta-information related to the data.

In addition to these presentations of national sources and methods, other metadata available in the Bureau are the underlying concepts and definitions for subjects established by resolutions of the International Conferences of Labour Statisticians, as well as abbreviated versions of the international standard classifications used to present disaggregated data (ISIC, ISCO, ISCED, ICSE ⁶). Together with the notes and the footnotes associated with the data tables, these metadata are essential for the proper understanding and

Sources and Methods: Labour Statistics

Vol. 1 Consumer price indices

Vol. 2 Establishment surveys

Vol. 3 Household surveys

Vol. 4 Administrative records

Vol. 5 Population censuses

Vol. 6 Household income + expenditure surveys

Vol. 7 Strikes and lockouts

Vol. 8 Occupational injuries

Vol. 9 Transition countries

Vol. 10 Estimates and projections of the economically active population 1950-2010

analysis of the statistics, their scope, coverage and comparability.

⁶ International Standard Industrial Classification of all Economic Activities (ISIC), International Standard Classification of Occupations (ISCO), International Standard Classification of Education (ISCED), International Classification by Status in Employment (ICSE).

Advantages/Disadvantages associated with the use of the technology

The latest capacities for data dissemination have resulted in heightened user awareness of the ILO itself and its role in the development of new and improved labour statistics. Because the Bureau accompanies its dissemination with the natural warnings and explanations concerning the data and their limitations, interest in the underlying sources and related international statistical standards and guidelines is generated. However, the increase in the volume of requests and the subsequent electronic distribution requires time and personnel to satisfy it. Some of the advantages gained by the rapidity of computerisation also created new requirements for the Bureau in order to best serve the interest being generated.

PART III: The ILO statistical Internet application

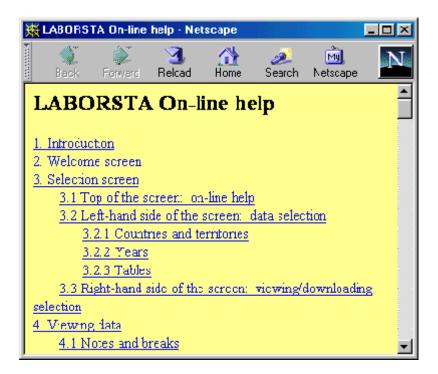
Dissemination through Internet

In the last ten years, several factors have entered into the Bureau's consideration of its role as disseminator of international labour statistics. One factor was the regular increase in user requests for electronic files containing extracts from LABORSTA, coming from a much wider array of users. This larger audience would conceivably require more guidance and information concerning the use and interpretation of the statistics being provided than the traditionally concerned experts and experienced statisticians we dealt with primarily in the past. The increased burden on the Bureau's staff, resources and functioning was another factor. The overriding consideration was the advent of new technologies such as CD-ROMs and the Internet which opened up enormous potential for innovation. Thus, while maintaining its primary programme of printed publications (necessary to many users in the world who still may not have easy access to electronic media), as well as its electronic data distribution service, the Bureau of Statistics began to develop an Internet application associated with its data bases. The aim was to give access to as much of the available statistics as possible to users of the World Wide Web. We baptised it: *LABORSTA on the Web*.

The Web application

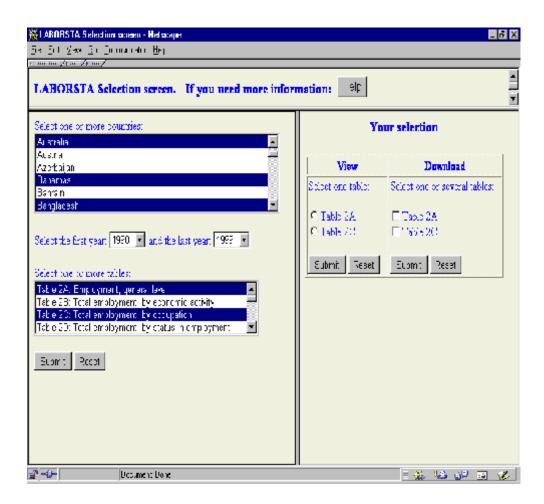
The application was designed during the summer of 1998 within the framework of an ILO interdepartmental project called "ILO Flagship Databases". The purpose of this project was to facilitate the availability of and improve access to all types of ILO information, through one common Internet entry point. The Bureau chose to develop its application using the computer software SAS: Statistical Analysis System. Originally a statistical package which was used in the Bureau to develop its data base systems and to perform statistical data analysis, SAS has transformed itself over the past decade into a fourth generation computer language comprising a specific module for Internet development work. Our application takes advantage of the particular possibilities of the user's Internet navigator and allows for the incorporation of numerous "hyper links" in order to offer a great quantity of metadata with the data themselves. In this way, *LABORSTA on the Web* is a dynamic application which combines in electronic form the multitude of statistical data and metadata previously contained in several full series of printed publications, namely the *Yearbook*, the *Bulletin, Sources and Methods*, etc. mentioned above. The Bureau presented its application to the public for the first time in October 1998 on the occasion of the 16th International Conference of Labour Statisticians.

Around 70,000 time-series on the main topics of labour statistics are accessible to public users of Internet. The main features of the application are that it allows users to select, view, print and download data as well as the relevant metadata. The combined statistical and metadata enhancement enables users to better understand the statistics and how they can be used for their needs. The welcome screen (shown on the front page of this article) opens the application and allows users to: i) display the "on-line help" screen (below) which contains the information required to select, view and download data, ii) print the help contents or, if users so wish, iii) minimize the "help" in the toolbar which can be invoked as needed during use of the application. From the welcome screen we continue to the "selection" screen, which is the heart of the system.



Accessing the data

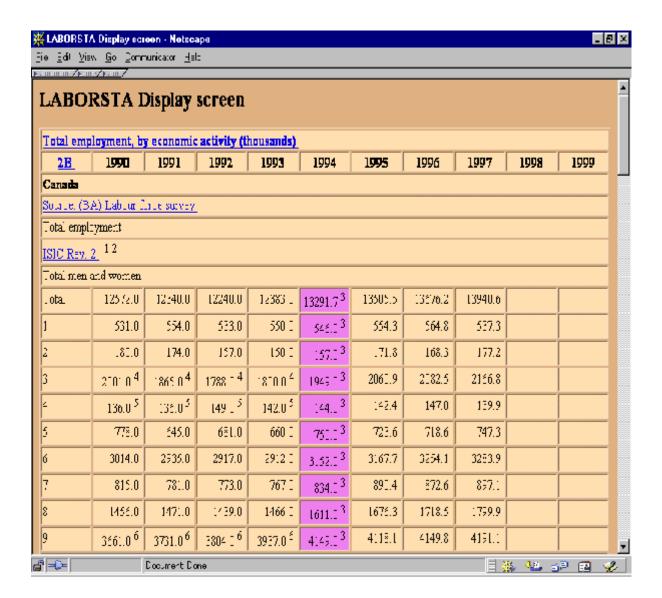
The prevailing idea during the development and design of the application was to make the data accessible to users as simply as possible, with the least amount of navigation or, in other words, a minimum number of mouse "clicks" to avoid frustration due to too many manipulations before you are actually able to obtain the desired information. Directly following the "welcome" screen, therefore, users are taken to the "selection" screen (below).



The selection screen consists of three different parts or frames. The top frame enables the user to invoke the "help" facility at any time. In the left-hand frame, users can specify the three criteria (country, year, subject) needed to process a data request. The right-hand frame allows users to "view" or "download" data, according to their selection. Some 200 countries and territories are listed in alphabetical order by continent. By clicking in the selection window and typing a letter, the system is automatically positioned at the first listed country starting with the letter typed (for example "N" goes to Namibia). The possibility exists to select more than one country at a time. Several countries in a row may be selected: contiguous countries = "Shift+Click". For a non-contiguous selection of more than one country, use "Ctrl+Click". Start and end years may be specified or the default selection gives the 10 latest years. Following selection, the pending or waiting message is replaced with the confirmed selection as submitted to the Web server. The two-column table which then appears is either for "on-screen viewing" by users or to download to their own computer (the client).

Viewing the data

To "view" or display the data selected by the user, the application opens up the "display" screen (below) where the data are presented, one table per country. Within multi-country selections, each new country is indicated by a change of colour. Notice the table is multi-coloured where the blue titles indicate the hyper-links available, which when clicked take users to different selections of meta-information. Purple coloured cells signal a break in series.



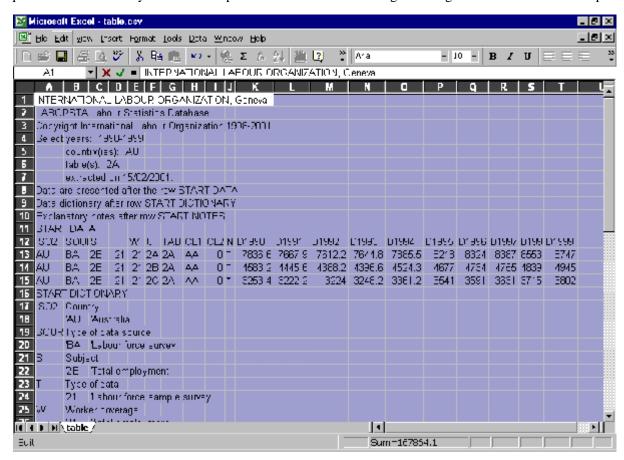
To "read" the table, proceed from the top to the bottom of the screen by clicking on the relevant hyper-links as follows:

- 1. Total employment, by economic activity is the subject of the table and links to the employment concept from the International Conference of Labour Statisticians' resolution on employment, etc.;
- 2. 2B is the table number and links to specific explanations associated with the table;
- 3. Source (BA) Labour force survey links to the country's Labour Force Survey methodological description from *Sources and Methods*; the purpose of the two-letter source code (BA) is to provide users with a rapid comparison criterion;
- 4. ISIC Rev.2 links to the text of the relevant International Standard Classification, at the level of detail presented. Data displayed according to updated versions of a classification are shown, one following the other. This example is the Rev.2 version of the industry classification;
- 5. Note Call Superscripts 1, 2, 3, etc. refer to explanations at the end of each country table (not shown here):
- 6. The purple cells in the year 1994 indicate a break in the time-series as compared to data for earlier years the text of note 3 at the end of the table explains the break.

Along with the country and source-specific information associated with a table, users can at any time also browse the rest of the metadata (concepts, definitions, classifications, methodological descriptions) associated with any of the statistics contained in the Web application. Updating of the statistics in the application is done annually at the same time as the distribution of the latest *Yearbook* edition.

Downloading the data

Once one or more tables have been selected and the request has been submitted, a dialog box appears. Users must specify whether they wish to open the downloaded file or save it on a disk. Submitted tables can be downloaded in a CSV (comma separated values) file specifying the file name, which can be stored on disk or imported directly into any software data facility such as Excel, Access, etc. This process works automatically if the list separator of the operating system is a comma and the computer is set up to open files with the CSV extension into the desired application - Excel in the screen example (shown below). Settings depend on users' requirements and will vary from one computer to the next. The defining and design elements of tables for export



in electronic formats have been programmed behind the scenes to accommodate an automated functioning system to prepare a complete output for users.

The downloaded data set is structured in four blocks, in the following manner: 1) "SELECTION" provides the identification summary of the request; 2) "START DATA" presents the data in a block of contiguous data cells in time-series which enables direct graphical analysis; 3) "START DICTIONARY" provides textual information on the meaning of codes used, and the contents of each field, and 4) "START NOTES" displays all of the relevant notes to the data (entire series and/or particular years) in full text. The data are automatically accompanied by this metadata in conformity with the Bureau's responsibility to provide information which is as complete as possible and to draw users' attention to any differences between statistics for different years.

Users may choose not to retain such information when dealing with the data after they have downloaded it must be received by them.

Purposes served

Up to the present time, all of the information made public by the International Labour Office through the Bureau of Statistics' application is on a complimentary basis. In this respect, LABORSTA is different from other statistical sites on the Internet, as is the close interface between the statistical data and the metadata making them available simultaneously, on-line. The Bureau considered but rejected the production of a CD-Rom for dissemination of the LABORSTA data electronically, primarily because of the lack of an efficient means to regularly update the statistics. Development of the Internet application takes advantage of the superior potential to reach out to the public provided by the World Wide Web within the cost-frame of the budget provided by the ILO constituents. Inside the ILO, the application is available at every single work station which is linked up to the Internet, thus also giving all (statistical and non-statistical staff) in the Office access to *LABORSTA* on the Web for their work.

Future outlook

Following the experience gained in making the application available on Internet, several steps are envisaged next:

1) making the site trilingual as are all ILO printed statistical publications;

2) extending the coverage to all statistical subjects included in the *Yearbook of Labour Statistics*;

3) introducing other regional groupings apart from individual countries, such as worldwide, more developed, less developed or others; and

4) incorporating the Bureau's second largest data base, LABOCT (results of the ILO's annual *October Inquiry*, from 1983 on wages and hours of work and from 1985 on food prices) into the site.

Under the umbrella of statistical information and technology management, the Bureau is considering documenting a quality control system associated with the statistics which we disseminate. Together with the systematic information concerning the sources of data, this would provide users with an indication of the quality features and reliability of the national statistics made available.

Linked to the use of the Internet technology is the development of new, standard electronic reporting formats, expected to assist countries to respond to ILO information requests in a way which is harmonised with the site's specificity. Further hyper-links to the Web sites of other ILO programmes and/or departments⁷, for example, international and other United Nations agencies etc. will also be implemented where possible.

Additional changes could also depend on the findings obtained through the feedback to our inquiry below, thanks to your interest and replies.

Feedback from users

After having visited the site *LABORSTA* on the *Web*, you may have ideas concerning other information you would wish to find there, or on different improvements you would like to see take place. If this is the case, kindly provide us with your responses to the information requested below. Our email connexion is available at the address *laborsta@ilo.org*.

-

⁷ One example is the Key Indicators of the Labour Market (KILM) programme to which LABORSTA provides some 60% of the data underlying the indicators or used to derive them.

Questions to users of LABORSTA on the Web:

- 1. Please tell us what kind of user you are (e.g. government policy maker, student, worker or employer representative) ?
- 2. Are you interested in our data for an individual or institutional reason?
- 3. What is your purpose in viewing /downloading our data?
- 4. How many times approximately have you visited the site or is this your first time?
- 5. What suggestions to improve the site or additions to the information presented would you make?
- 6. How does this site compare with other statistical Web sites that you have used?
- 7. What other comments would you like to make? (Please feel free.)