

**Economic and Social
Council**Distr.
GENERALCES/2000/17/Add.3
20 April 2000

ENGLISH ONLY

STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE**CONFERENCE OF EUROPEAN STATISTICIANS**Forty-eighth plenary session
(Paris, 13-15 June 2000)Theme 2**THE USE OF ADMINISTRATIVE SOURCES FOR STATISTICS AND
INTERNATIONAL COMPARABILITY**Supporting Paper Submitted by the Statistical Office of Slovenia 1/**I. PRESENTATION OF SLOVENIA AND HER RECORDS AND STATISTICS THAT ARE
REGISTER-ORIENTED**

1. It is widely known that the Slovene statistics is register-oriented. In the former state, i.e. the Socialist Federal Republic of Yugoslavia (SFRY existed until 1992), Slovenia made use of its status of being an independent federal entity also in order to produce a tailor-made multi-purpose statistical data model and a data model of records. With its proposal Slovenia in part managed to exert influence on the development of similar or related registers within the SFRY. In that period, in which there was predominantly social property, both the understanding of and the method on how to approach to the records and statistical data were greatly different and not such as are defined in the (10) UN Fundamental Principles of Official Statistics of 10 April 1992. In those times, the valid principle was that all records and administrative data were public, except in cases when they were (by law or otherwise) determined as confidential.

2. The Slovene statistics really used the available possibilities and therefore, within the scope of the principles of the then valid social system of informing, prepared numerous initiatives and itself took action in organizing the 3 basic registers in Slovenia, which it set up, kept and maintained until the

1/ Prepared by Mr. Tomaž Banovec, Director of the Statistical Office of Slovenia.

complete, well-organised administrative registers were transferred to more adequate performers. Nevertheless, Slovenian statistics retained (by law) the right to make use of all the data - merely for statistical purposes.

3. **The set-up of registers.** In the new legal settlement, following the year 1992, the National Statistics Act (1994) laid down that the national statistical office (hereinafter: SORS) were to give over part of the administrative registers to other institutions/Ministries, into their competence and to become their field of work. This applies to the **Central Population Register**, which was transferred to the Ministry of Interior in 1998, and to the extremely detailed **Register of Territorial Units**, which was slowly taken over by the Geodesic Office. The third one of the basic registers, i.e. the **Business Register**, is still in the domain and authority of the SORS. It is anticipated, however, that upon the introduction of some supplementary legislation it shall be transferred to a new state organisation (Agency for services and records under public law, which is being established), but SORS shall retain the right to make use of the compiled data for statistical purposes.

4. **In common questionnaires, requests referring to both: records and statistical demands for data, are to be linked.** In those times it was possible and a usual task to compile and add statistical requirements into accounting reports. This is still done at present, as the current information technology facilitates also other kinds of linkages (not only horizontal ones) and take-over of data from numerous other sources. Yet, then it was essential that the national statistics had a right to acquire all the necessary administrative data and this right still applies in Slovenia - except in cases when this is explicitly forbidden by law.

5. **Examples and national & economic rationalisations.** On the basis of identifications and data from records, contained in such registers and when making links to other data, SORS can **design its statistical registers**. Several times SORS has proposed a common **rationalisation of the administrative matters** in the country. Among other possibilities, it is possible to enter into the population register with all its attributes via the PIN (i.e. personal identification number); then into the register of health insurance and into the register of pensioners, where in both cases SORS has access only to some of the attributes. From this common database a **statistical register of permanent population** is derived and currently maintained. From a variety of numerous other registers, the SORS is able to set up other lists and registers for its use as well. Yet, it is essential to make use of the uniform identification of the observation unit in a specific domain and certain relations among identifications in various administrative lists are important. Similarly, the understanding of these functions and the support that is given to them by way of records are also vital.

6. **Basic presumptions for such an orientation were and predominantly are:**

- The desire to rationalize the administrative work (via records, that spread over all the country and are linked, and registers) which has emerged already in the 1970s, when informatization spreader to the field of statistics as well.

- The possibility that in those time (from 1973 onwards) it was possible to apply the legal provision on the **uniform classification of activities** (1973) and that, in case of central planning, this classification which was founded on the basis of MPS (Material Product System) is used on the level of enterprise or commercial company **for the statistical referential purposes and simultaneously for planning purposes as well.**

- **An adequate state of prepare(d)ness of the bodies (predominantly Offices) who cooperate/d in Slovenia and have taken over their shares of responsibility and have cooperated in**

implementation phase with the SORS for over 20 years and have been setting up registers, collecting data for statistical needs and other multi-purpose objectives.

- Slovenia and her public administration are relatively well equipped with the information technology from those times and have a good concept of informatization of the then times.

- Consistent and general use of administrative and linked registers for various administrative purposes (personal matters, elections, passports, permits, et al., claims for /right to/ land, etc.). This even nowadays presents a warranty that attributes are currently up-dated (in the population register, containing 2 million entities, 250.000 events are registered annually).

7. **From the beginnings of informatization to an information society**

Even then, from the 1970s period until the declaration of Slovenia as an independent state in 1992, and nowadays SORS gave and gives a majority of incentives for the organization of its own and a common administrative environment. Thus, even at present we maintain in the old and new registers together new classifications and new identifiers. Yet, nowadays, this task proves to be more difficult, taking into account the explosion of use of the Information Infrastructure and Electronic Commerce. Nevertheless, we have some means and possibilities, which we describe and which at least in part enable and facilitate exertion of influence on the data models of the entire country. The basic problems can thus be broken down into two large categories of questions:

- How to determine the information society as a macro-economic category, how to measure it statistically and what does introduction of information infrastructure in statistical observation units mean, what meaning do the reactions of the statistically observed observation units have in view of our relatively slow statistical harmonisation to this new occurrence?

- How can we enter into the records and into the transactions of electronic commerce? How to enter **into networks for value added**, how are we to determine the occurrences in networks that are important for statistics and by what statistical infrastructure are we to enter into the networks?

8. We anticipate, that at least some of these questions, that were set in November 1999 at the International Statistical Conference (Radenci '99), entitled Electronic commerce and statistics, shall not remain unanswered.

II. **LEGISLATION AND OTHER SETTLEMENTS, SUPPORTING LINKAGES**

9. **Legal and other legislative means** for the co-ordination of contents of data model of the Republic of Slovenia not only co-ordinate the contents of the data model of the Republic of Slovenia, but also facilitate the **horizontal and vertical linkage of records and statistical linkage of the contents of records and data**. The SORS and other authorised performers of the National Programme of Statistical Surveys (hereinafter: NSPP) have adequate experience herein.

10. Some principles and orientations for our current work have been based on:

- the past and current co-ordination and adjustment of administrative records and their contents which is required for about 30% of those statistical surveys (statistics) that are listed in the programme until 2002,

- the register-orientation of national statistics whose foundations represent administrative and other horizontally linked administrative sources (there are 35 major ones and deviations of national data bases or multi-purpose registers),
- the principles of protection of personal data and of data on enterprises, whereas both are linked to records and statistical data,
- an adequately well-structured national information infrastructure and support in the majority of authorised organisations,
- the requirements related to European integration (Harmonograms and screenings or analytical examinations, National Programme for Adoption of Acquis - May 1999, the EU Response - June 1998, NPAA - October 1998, et al.).

11. In view of requirements to harmonise the **contents, techniques and data dissemination with the international** environment, especially with the EU, **the Slovenian national data model** has found itself to be in a new and somewhat more difficult situation, as new tasks in the field of records and new statistical requirements have started to emerge implicitly and vertically in various sectoral harmonizations (there are 31 negotiating groups) at screenings and, last but not least, under various new titles and definitions and synonyms, which are sometimes being used differently in various institutions.

Terminology and the national data model

12. Seeking for the European, worldwide and other definitions and the determination of needs for data on such bases is a difficult task. When reviewing and making preparations for screening exercises it becomes obvious that the terms are not harmonised. In Slovenia there is a richness or abundance of such synonyms. The problem of terminological inconsistency, emerging in the contents of records and statistics, becomes even more rigid when these terms are translated into the Slovenian language. In case the Slovenian Government determines the wording of an act in which it uses two different Slovene words (i.e. "trajnostni razvoj" or "sonaravni razvoj") in two different Articles to express the meaning of "sustainable development", and it fails to give an explanation of the different terms, then such an inconvenience arises. **What is to be measured statistically, what are the components parts of records and how to express in words what a national data model is?**

13. It is the proposal of SORS that the translation services use also all the **definitions** that are already valid in our **classifications** which were transposed and are EU and internationally settled (there are cca 20 major ones). An instance from the **Combined Nomenclature** - can these names (Latin: nomen) be used also when determining the negotiating positions (and documentation) and legal texts? We would certainly gain a wider degree of harmonization and transparency among the records and the statistical part of the databases and their partial correspondence to the legal textual descriptions.

14. The expressions and terms of the EU, definitions, tasks and harmonisation have been taken into account also:

- prior to final enforcement of the Accession Partnership Agreement in 1998 (Article 95 thereof determines the commitments of statistics and there are many tasks related to records as well - TARIC, CN, NACE Rev.1);
- upon the completion of the Screening exercise in July 1998 (held for statistics and some others) and upon having received the new critical message of the EU to Slovenia, as for instance the Self-assessment Reports and Response (July 1998, autumn 1998 and both assessments from 1999);
- in the working papers and orientations of the Government and Government Office for European Affairs (hereinafter: GOEA) in making due preparations for the National Programme for Adoption of (EU) Acquis and in the documents of the in-line Ministries;

- in the anticipated and planned reforms of some development policies in the Union (agriculture from year 2000 onwards, State Aids, Structural Funds and Assistance, SAPARD, etc.);
- in those definitions that were merely in part harmonised with other national documents, providing orientation (strategies, national programmes, etc. - about 35 in all) and
- in requests for data of other international organisations (UN, OECD, FAO, ILO, UNICEF, WTO, IMF, MBOR, etc.)

Harmonisation with the European and thus with world-wide statistics - co-ordination

15. The Slovenian national statistics is to be harmonised in contents and in view of other commitments with that of the EU by 31 Dec. 2002. The Slovenian commitments in relation to other institutions (UN) are less laborious and are usually handled simultaneously - the only exception being our commitments to IMF, due to whose dissemination standards the reporting and contents are becoming increasingly tense. Yet, in our opinion, IMF thus stresses the importance of the correct timing of data presentation.

16. A special question is, whether in case of **such pressure we shall managed not only to keep and maintain the statistical national data model but also to simultaneously deal with the national data model of records**. At least there, where it is viable or has been viable to have vertically maintained records, linked to the current National Programme of Statistical Surveys, it would be advisable to maintain such a state also in the future.

17. The selected legal provisions that are to be applied in Ministries and their Offices and Public Institutions list more tasks for these than for the physical persons and the private sector. The latter anticipate such rational operation. For instance: merely one of the public administration bodies actually made their request to SORS (in a 3-year time period) to obtain from SORS its the legally compulsory opinion on the methodological bases of their data collection - whereas others failed to contact SORS until everything has already been decided. The predominant part of co-ordination regarding the contents, which is made in line with the stated acts, is achieved **via Advisory Committees** at the SORS. Namely, it is at the meetings of the Advisory Committees that both: the statistical approach and that in view of records are discussed in view of how to solve such a certain problem or occurrence which links the record (micro data) and the statistical standpoints (macro data) and prepares the administrative source for multi-purpose use of these records and for the needs of national statistics.

18. It is essential to strengthen the required knowledge on **use of records in our relatively young Slovenian local self-government**. Therefore it is vital to settle and properly arrange the contents of these registers and adapt them to the newly determined functions of the state on local level. The standpoint of records and that of statistics in principle include observation **for at least that part of identical entities**, characteristic of this occurrence, and in this way the municipality majors can acquire for their needs not only statistics (on demography, economy, employment) but also the print-outs (from the Central Population Register) for their territory, the print-outs of all business subjects on their territory and an extract from the Register of Territorial Units. From another register they can obtain the address of the seat of the enterprise which gives employment to the citizen of a certain municipality (the Parliamentary Procedure has been commenced).

III. ACTING AND ELEMENTS OF STATISTICAL AND REFERENTIAL STRATEGY FOR THE FUTURE

19. Hereinafter we point out some possibilities and strategies, which can be used in the developed information society for set-up of a linked **administrative and statistical data model** which can

strategically meet the requirements of **statistical time series**. Yet, there is subsequent need to give them a new definition and contents, to widen or re-shape the old, electronically archived, still not adequately arranged data from records - to be used for new statistical tasks.

20. **Possibilities of subsequent inclusion of still not included referential data into statistical time series.** This task is a general one and viable in case we have (at the right time) arranged and saved those referential data that could be inserted subsequently and we have a solution which is to facilitate linkage of contents (attributes) and technical linkage (informatization) via or with the assistance of those identifications that were linked in time (SAJPI).

21. The kernel of linkage are **common and linked identifications** from several registers or lists (PIN - Personal Identification Number, BIN - Business Identification Number, geocode - Centroid, MID, etc.) Matters or records are thus sized down to the shaping of lists and their linkage via current numbers or one's own identifiers.

22. Adequate care is required in view of definitions of common or comparable attributes (in records and statistics), besides maintenance of linkage of common identifiers and SAJPI (i.e. Common administrative kernel of linked identifiers).

23. Timely **breakdown of data and additional, complete and timely linkage of informatization and archiving of a large group of current data for** important occurrences for statistics in future.

24. During informatization and collection of statistical reports on occurrences it is possible not only to provide for statistical purposes, but we can also provide for complete informatization and for dynamic archiving of the administrative part of data, collected for such an occurrence and on the same entities (transaction). This may not be so interesting in the present, in view of forming a statistical aggregate, but it shall certainly be of importance in near future (in 2005).

25. **Instances: Customs simple documents**, completely informatized tax reports, daily occurrences from registers, transactions, etc. Complete accounts on use of energy per each business entity, et. al. These are linked to an adequate, timely normalisation of contents of data at least on the level of entities **among records and statistics** and the elaboration of adequate software. In this view the geocode centroid (real estate or postcode) is extremely useful, as the majority of linked occurrences can be linked to the location which is almost the only one with the physical basis for identification.

26. **The register-orientation of statistical offices, "diminishing" of statistical requirements and tools** of the national statistics in administrative records. Certainly it is difficult to insert purely statistical requirements into various data collections, especially if these are made by SNA or ESA criteria which are not comprehensible or are deemed to be unusable by the vast majority of administrative users. But, on the other hand, this is an opportunity for the future of national statistical institutes - in case they wish to assist the national decision-makers and provide them also with statistics, based on the **statistical requirements of SNA and ESA** and similar ones, including the European and world requirements. In the phase of its informatization, the majority of acts on national level cannot directly include the ESA or SNA statistical requirements, but they can include certain parts and adequately fragmented groups of data that are useful not only for the institutional but also for the macro-economic and statistical purposes.

27. **The kernel of poly-functionality and of linkage** and "a timely and in entity harmonised model" of data and the understanding of the occurrence. We are **building a model of an entity that is settled and arranged in its contents, space and time**, and is to link also common and uniform identifications of statistically observed units, which are observed administratively as well and are

adequately informatized. Here belong the compulsory calculations, made in line with world accounting standards, tax statements and records and informatized annual reports of commercial companies and statistical reports... In fact this refers to everything that can be linked immediately and regularly, daily, weekly, annually or on-line to an observation unit with different identifications in various records and statistics - eventhough we perhaps do not need all these data immediately and just now. Yet, in order to understand the methods and kinds of non-statistical observations of a single occurrence, the methodologist needs an adequate statistical identification which should be linked to the administrative identification if it refers to the same entity.

28. **Longitudinal successive linkages of recorded and informatized interesting occurrences**, or documentation of contents that were acquired via monitoring and are to serve for subsequent, possible insight. In 1988, in Slovenia we linked the statistical unit - inhabitant from the Central Population Register (CPR) - with the identification from the register of the pension insurance company and the identification for the same entity from the register of health insurance. In such a way identifications were linked to and from the BRS, the register of spatial units and some others. Via CPR we obtained the address of permanent residence and via insurance companies we obtained data on the employer and all the locations he works on. Therefore it is now possible to inquire where a person worked, what kind of work he performed and for how long, where he lived (registered residence), etc. This gives substantial proof of the possibilities of various usage of records - made on the model of the state for such purposes.

29. **Monitoring and use of formalised and informatized non-statistical sources and anticipation of future statistical requirements via registers and other sources.** This task is derived from the above stated items, presuppositions and ideas. Anticipation means a methodological task on what has to be (with some wisdom) anticipated for the future consensus of those who engage in collection of data from administrative records and of the state and in-line statistics. We monitor administrative tasks of a certain field also from the view of future statistical demands and we strive to anticipate or presuppose which data shall be determined as required ones in the future.

30. **Analyses and already in advance provided informatization of legal acts** and insertion of statistical needs in the proposed legislation. Thus it is presumed that the NSIs can actively be included in the preparation of those parts of in-line legislation that refer to data collection for administrative purposes. In Slovenia this is possible and legally provided for, eventhough it may not be fully carried out as planned. The modelling of the consequences or results of legislative acts, taking in advance full account of anticipated and unforeseen effects (Los test in Germany in Law model in Denmark) can be considered to be a special form of cooperation of the statisticians.

30. **Harmonisation, re-engineering and generation of new administrative sources.** Sometimes it is difficult to merely observe and wait until a Minister proposes an act. Therefore it is suitable that the NSI has the possibility to itself propose a legal regulation - and this is especially the case when classifications are proposed. An example is: the **Business register for statistical and administrative purposes**, the Census Act, the National Programme of Statistical Surveys, etc. The initiative of NSIs is vital in case of classifications and registers which are to have coverage of the international and national data model and extend also to the national and international administrative environments. Yet, the legislation is not only a constraint, but also an advantage. More should be transferred to standards and authorisations.

31. **The data model of a state, the statistical model, links to the data model of the international environment and occurrences that are monitored.** The idea of the model is in principle transparent, but it is difficult to describe it. Nevertheless, some statisticians, especially SNA experts, master it (in their heads). Little that is statistical belongs merely to one country. How do we understand such complex

models? It is possible to master **meta data bases**, but discussion is still required on occurrences that are measured, on functions (Model of functions), which are performed upon these occurrences, items, and data that are required in this view. The explosive development of informatics has facilitated informatization to spread everywhere, but it has not increase the understanding of these occurrences and their causal relations in time, by contents and space - in our heads.

32. Permanent deliberation of **future tasks of statisticians**, openness for the new, acceptance of new ideas and trespassing of the narrow professional mind, expert discussions held in and among statistical circles and circles of other professional associations, dialogues between statisticians themselves and with users. It is required that we are widely-oriented and deal with many linked occurrences and are specialists with deep in-sight into occurrences and their most important entities, able to determine entities and attributes for statistical monitoring and monitoring of records, that we understand the development that has been completed in this field in the world, and also other domains and forms of the occurrence and needs for the future.

IV. CONCLUSION

33. Take-over of data from administrative sources was and shall be the strategy, tactics and everyday implementation task of the Slovenian statistical office. At present we have adequate legal and theoretical bases and many experience. Introduction of the statistical identification in each entity, which is monitored via records, and linkage of various identifications, which pertain to such an entity in other domains, facilitates linkage and take-over of administratively determined attributes. We require only the explanation of the differences among the records and classifications and the statistical use, which is based on them.

34. The Office (SORS) can itself make proposals for its classifications and definitions for administrative work, as standard classifications for the territory of the entire country. In some cases, classifications shall have to be made for the past. Many entities shall be monitored statistically and by records - in the future and for the past, but this is still not exactly known. This is why we should save some of the currently horizontally linked and overtaken attributes, linked to these entities, for the future and archive them together.

35. In a global world and in network economy, **electronic commerce** shall become internationally standardized on the level of records, just as statistics is in the SNA framework. Institutional national solutions shall retreat, national legislation shall be substituted by world standards - and it is an advantage to be present at this stage - even in view of statistics and records.

36. Yet, the question remains, whether such a diversified school system can by itself educate and train people for such tasks. Probably not. Additional training is required, to again shed light on the understanding of the occurrences and entities which it explains and to acquisition of data, referring to these entities, from other domains which are in part related to the observed occurrence.
