

STATISTICAL COMMISSION and
No. 9
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS

Working Paper

English only

Work Session on Statistical Metadata
(Geneva, Switzerland, 18-20 February 1998)

Item 3 of the provisional agenda

**SYSTEM OF DEFINITIONS OF SOCIO-ECONOMIC TERMS
AS AN ELEMENT OF THE STATISTICAL METAINFORMATION SYSTEM**

Submitted by the Central Statistical Office of Poland ¹

¹ Prepared by Anna Paradowska.

I. INTRODUCTION

1. Statistical Metainformation System (SMS) provides the information needed to interpret and analyse statistical data. It includes definitions of socio-economic indicators and specifies their component elements.
2. The **definitions** of socio-economic indicators used in statistical surveys are one of the most important elements of SMS. They identify a statistical indicator as a subject and its descriptive components (e.g. observation units, price).

II. LEGAL BASIS OF THE MANAGEMENT OF SOCIO-ECONOMIC DEFINITIONS IN CSO OF POLAND

3. The law on official statistics currently in force, was issued in 1995. It adapts the statistical system of Poland to the conditions of a democratic state and market economy. It is in accordance with the Fundamental Principles of Official Statistics in the European Region, adopted by the Economic Commission for Europe at its 47th Session in 1992.
4. The law on official statistics concerns, among other things, statistical metadata issues. Accordingly, the Central Statistical Office of Poland is responsible for "developing standard classifications, nomenclatures and **definitions of basic categories**, relationships between them, and their interpretation" as well as "maintaining the official state registers: business register and the register of territorial division of the state".
5. Registers contain metadata describing statistical units. Classifications are a collection of profiles in which statistical information is presented, and **definitions** explain the terms used in statistical surveys.
6. The President of the CSO of Poland is, according to the regulations, obliged to publish methodological principles and definitions of terms used in statistical surveys in the Official Journal of Poland Republic.
7. Following the law on official statistics, CSO maintains an integrated information system "Definitions". The aim of the system is to ensure a uniform and unambiguous interpretation of socio-economic terms used in statistics. It enables information to be compared over time and internationally.
8. Definitions are given for general terms of particular fields of statistics, and for socio-economic indicators and their components.

III. THE RELATION BETWEEN THE SYSTEM OF DEFINITIONS AND THE GENERIC MODEL OF A STATISTICAL INDICATOR

9. According to the "Standardisation of statistical indicators and metadata²", definitions specify the meaning of a statistical category and describe its component parts. The component

2

UN/ECE Work Session on METIS, 22-25 November 1994.

elements of an indicator alone are often not sufficient to give an exact description of an indicator. The definitions contain additional information which explains the socio-economic phenomena and makes their interpretation and analysis easier.

10. Let us consider the generic model of a statistical indicator:

$$I = [C, S, T, V]$$

where: I - statistical indicator
C - content of the indicator
S - space, i.e. the observed statistical units
T - time of the observation
V - value of the indicator

11. The attribute C can be described as a pair:

$$C = [A, MA]$$

where: A - attributes of the statistical category
MA - measurement attributes

12. The parameter A can be described as a vector:

$$A = [B, A_1, \dots, A_n]$$

where: B - basic indicator
 A_1, \dots, A_n - subject-matter attributes of the basic indicator:

$$MA = [MU, KP, MM, TC]$$

where: MU - measurement unit
KP - kind of price
MM - method of measurement
TC - time characteristic

13. The “Definitions” system describes, first of all, element **B** of statistical indicators, i.e. the **basic indicator**. But it also includes definitions of socio-economic categories as general concepts, superior to basic indicators.

14. For specific fields of statistics, where data often appear in profiles by different classifications (e.g. foreign trade), the descriptions of those classifications are included as a source for subject-matter attributes. The measurement attributes often include definitions of kinds of prices (**KP**). Definitions are also given for different types of **statistical units**, which constitute the element **S** (space of objects).

15. Definitions concern both primary and derived statistical indicators. The data for primary indicators is obtained directly from statistical surveys. The derived indicators are calculated according to special algorithms. In this case, these algorithms are included in the definitions.

IV. DEVELOPMENT OF THE SYSTEM OF "DEFINITIONS" AT THE CSO OF POLAND

4.1. Establishing definitions

16. Definitions are created and developed by subject-matter departments of CSO in consultation with scientists, and in agreement with respective Ministries. The work is co-ordinated by the CSO Department of Standards, Registers and Informatics.

17. The operational scheme of the "Definitions" system is shown in the Annex.

18. Official normative documents are the source material for the definitions: e.g. laws, resolutions and regulations of the Parliament and Council of Ministries, and other documents such as notices, conventions, publications, encyclopaedias, etc.. We try to a maximum degree to take into consideration the recommendations of international statistical organisations which concern unification of statistical methodology. We often use publications of UN Statistical Division, Eurostat and others.

19. The sources of definitions belong to 8 main groups:

- 1/ Laws, decrees
- 2/ Regulations of the Council of Ministries and Ministers
- 3/ Resolutions of Parliament and Council of Ministries
- 4/ Dispositions of the Prime Minister and Ministers
- 5/ Dispositions of the President of CSO
- 6/ Orders, notices, guides, circular papers, statutes
- 7/ Publications, encyclopaedias and other issues
- 8/ Other documents.

20. Definitions are grouped according to the following fields of statistics:

- 01 Population and demography
- 02 Organisation of state economy, jurisdiction
- 03 Living conditions
- 04 Health services
- 05 Natural resources and environment protection
- 06 Education
- 07 Science and technical progress
- 08 Culture, sport and recreation
- 09 Organisations and associations
- 10 Housing, municipal infrastructure
- 11 National accounts
- 12 Prices
- 13 Labour market, wages
- 14 Public finances
- 15 Money- and capital market
- 16 Non-financial firms effects
- 17 Proprietary and structural transformations
- 18 Investments and fixed assets
- 19 Material and fuel-energetic economy

- 20 Agriculture and forestry
- 21 Industry
- 22 Construction
- 23 Transport and communications
- 24 Trade and services
- 25 External trade
- 60 Registers
- 70 Information standards

21. At present the system contains nearly 2,300 definitions belonging to 17 statistical fields. The official definitions are also published in "Methodological Bulletin".

22. Each Bulletin consists of 3 parts:

- 1. Definitions, arranged into thematic **groups**
- 2. List of source materials
- 3. Alphabetical index of socio-economic terms defined in the given Bulletin

23. The thematic groups often include a group of general terms for the given area of statistics, and a group of derived indicators calculated on the basis of primary indicators. Derived indicators do not necessarily form a separate group; they can be components of other thematic groups.

24. The groups of terms which are specific for a given statistical field often include definitions of particular measurement attributes (MA), e.g. the kind of price. In the area of foreign trade, there are definitions of respective classifications and nomenclatures, including international ones.

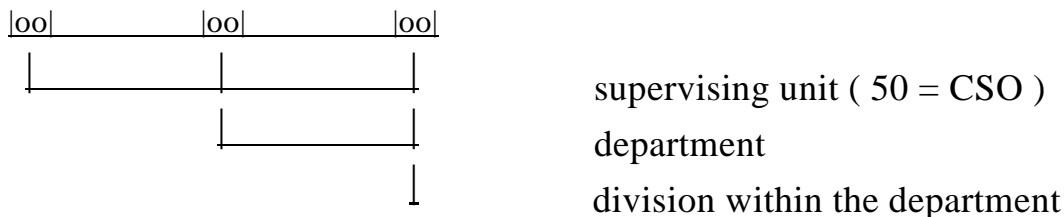
25. For each socio-economic term, sets of key-words and sources of the definitions are presented.

4.2. Information system "Definitions"

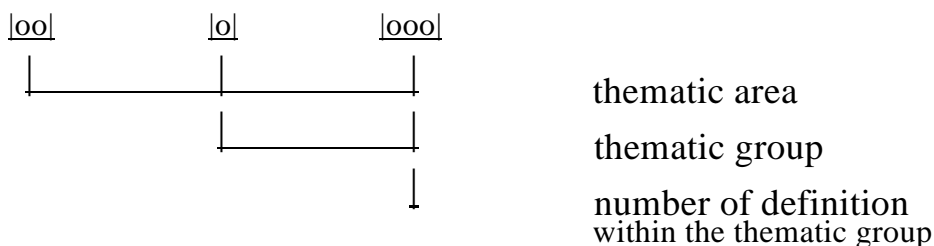
26. The information system enables access to the definitions in interactive mode using IBM PC and LAN. Each definition in the system is described by a number of attributes. Generally, this description consists of:

- 1. symbol of the unit in CSO who is responsible for the definition
- 2. unique identifier
- 3. title of the definition (i.e. name of the defined term /expression)
- 4. text of the definition
- 5. key-words
- 6. identifiers of synonymous terms
- 7. identifiers of sources being the basis for the definition.

27. Symbol of the responsible unit consists of 6 digits as follows:



28. The identifier has the structure:



29. The thematic areas are listed in section 4.1. of the paper.

30. The breakdown into groups is made by the responsible unit in CSO, which also provides the serial number of definition within thematic groups.

31. There are two kinds of users of the system:

- active users, i.e. authors who can update the definitions;
- passive users, who can only read the data.

32. Users can search for definitions by:

- 1) number of the term
- 2) name of the term
- 3) key-words
- 4) thematic groups.

33. The information included in "Definition" system can be obtained:

- through CSO LAN, with the possibility to retrieve definitions and make printouts in CSO departments and regional statistical offices;
- from the Official Journal "Polish Monitor";
- in the form of "Methodological Bulletins";
- in the form of printouts from the system for specified groups of users;
- on diskettes.

34. Information system "Definitions" works under DOS and is programmed in CLIPPER language. It is planned to develop a new version of the system working in Windows environment.

Operational scheme of the "Definitions" system

