

# Foundations of Ethics in Statistics

## in the Light of the ISI Declaration on Statistical Ethics and the UN Fundamental Principles of Official Statistics

Józef Oleński, professor  
Department of Economics, University of Warsaw

### Introduction

For providing the citizens' rights to information in *information society* and in *information economy* the United Nations statistical Commission and the International Statistical Institute elaborated and "codified" the rules of quality criteria and standards for information produced and disseminated with the use of modern IT. Official statistics may and should play an important and active role in defining those criteria and standards. Professional ethics of statisticians and of official statistics seems to be proper basis for that. Ethical principles of statistics are formulated in two basic documents: *The fundamental principles of official statistics* adopted by the UN Statistical Commission in 1993 and the *ISI Declaration on statistical ethics* adopted by the ISI in 1985.

This paper is an attempt to identify specific problems of the impact—both positive and negative—of modern IT on the realization of the citizens' rights to information. The advantages and threats generated by modern IT and by the globalization of information processes and systems are discussed. The duties of the democratic state, which should be responsible to protect the citizens' right to information in the light of those two documents, are analyzed. Some proposals concerning the role of statistical ethics formulated in *The fundamental principles of official statistics* and in the *ISI Declaration on statistical ethics* in the implementation of quality standards for social and economic information to protect citizens' right to information are formulated.

### 1. Information prerequisites of human rights in democratic society

1.1. **The right to truth is the fundamental human right** and the prerequisite of social order, freedom and democracy. The right to truth as the human right was verbalized inter alia in the Universal Declaration of Human Rights of the United Nations.

1.2. In the Latin civilization<sup>1</sup>, the *human right to truth* is expressed and concretized in laws proclaimed by the states, as the **citizens' law to information**. The *citizens' law to information* is the basis of democracy. The better this

---

<sup>1</sup> See: Koneczny F., *Wielość cywilizacji*, Gebethner i Wolff, Warszawa 1932 (reprint: Warszawa 1996).

law is obeyed and employed in practice, the more democracy the citizens enjoy and the more democratic are the state and the society. In democratic society and in democratic state any citizen has *the right to true, verified, understandable, accurate, relevant, pertinent, timely, available and practically accessible information*.

1.3. Social and political order of democratic states is based on the *integrity of rights and duties* protected by law. From the point of view of the citizens' right information this integrity of rights and duties means, that the governments, businesses, any other organizations and individuals are obliged to respect the law to information by providing respective information relevant to the rights to information. If any organization or any individual produce, deliver or disseminate the information, which does not meet the quality requirements specified above, they violate the citizens' right to information and the human right to truth.

1.4. The implementation and compliance of the human right to truth and the citizens' law to information in modern societies and economies require specific approaches. Conceptual frames developed do describe new processes in the societies and the economies of today, seems to be helpful for that. However, some terminological noise and redundancy may be confusing. Numerous concepts like *the new economy, the information society* as well as such terms like: *the digital economy, electronic economy, virtual economy, virtual business, or e-commerce, e-trade, etc.*, become nowadays very popular. It is the result of the fact, that journalists and marketing experts introduce most of those concepts, not the scientists. Some serious economists are of the opinion that this terminological *embarrass de richesses* is caused by "the seasonal fashion" of early stage of aggressive invasion of modern telecommunication and information technologies rather than by the real need of scientific considerations and professional activities.

1.5. I share that opinion to some extent. I think, that from scientific and practical point of view some terms used for new features in social, political and economic life, are rather pre-matured<sup>2</sup> and should be replaced well defined scientific concepts. First of all the term "new economy," which is obviously "seasonal" term, has to be changed as soon as possible<sup>3</sup>. Official statisticians may contribute to stop the terminological chaos and to help establishing the conceptual order by introducing official concepts and terms to the statistical classifications of activities and products.

<sup>2</sup> In particular, the terms: *new economy, digital economy, electronic economy*, are probably good for marketing of new communication services, however they do not meet the requirements of scientific concepts. E.g. *the new economy* seems to be the temporary term only, because after some time it will become the old one. Terminological order should be introduced to eliminate redundant terms like *the digital economy, electronic economy or e-commerce*.

<sup>3</sup> We do not have to much time to do that, because the term "the new economy" started to be used widely even in economic sciences and official statistics (vide: SINE—*Statistical Indicators for the New Economy* research program of the EU). The authority of IAOS and IARIW may help to propose and to get common approval for good scientific terms.

1.6. However, because we do not have well established terminological frames for those new phenomena of the economy yet, I will use in my paper—*nolens volens*—the existing terms. For the needs of this paper the following the most important aspects of some of those terms are specified below.

*Information society is a social system, in which:*

- a) level and dynamics of social and economic development depends on existing *information resources* and their growth, much more, then on the production of material goods and services;
- b) *minimal information resources* necessary for proper activities of companies and individuals are very large (i.e. larger then the natural ability of storage of information by individuals), and must be permanently and continuously updated;
- c) because of that the information resources of companies and individuals are not sufficient and their activities must be supported by specialized *external information systems* providing data and information services;
- d) *natural demand* for information from the part of individuals and companies is very high;
- e) the share of *expenditures for information* in the budgets of households and companies is also high;
- f) global information systems and resources play dominating role in providing information;
- g) end-users of information are not able to evaluate and verify the quality of information themselves; they must rely on the external systems, which deliver or disseminate the information.

1.7. Some scientists use also the term *new economy* to describe not very new—in fact—phenomena of post-industrial economies. This concept and term, invented in the USA, became popular in the European region, i.a. thanks to the IST program and its sub- program SINE (*Statistical Indicators for New Economy*). By this concept it is understood the complex of information-dependent sectors and processes of modern economies (knowledge industry, financial sector, information sector of the economy etc.). It seems however, that the expression “the new economy” will become obsolete and outdated in few years time, because of the adjective “new.” Therefore, good scientific term for this concept is needed.

1.8. Other concept used now is the *information economy*. This concept seems to be useful from scientific and practical point of view. By this term it is understood is the post-industrial economy, which has following new, specific features:

- a) knowledge and information is an indispensable and most important factor of economic growth;
- b) the resources of knowledge of population, its contents and quality, are playing the key role in the economic and social development;

- c) technological progress is based on very large resources of information, the results of technological progress are—first of all—new information;
- d) information sector is the largest sector of the economy;
- e) information infrastructure of the economy and society is decisive for the level of economic development and growth.

1.9. Rather confusing seems to be the term *digital economy*. It seems to be rather marketing expression, then scientific or professional. It may occur, that it is also the temporary, provisional concept, representing the complex of tools and activities connected with the implementation of modern IT and telecommunication in the economy in its pre-matured phase of aggressive invasion in the economy. Providers of Internet services and software developers use it for marketing purposes. Its synonym is *the electronic economy*. One should not forget the *e-commerce*, *e-trade*, *e-business*, *e-... e-etc.* in this noisy chorus. It seems to be rather useful technical and marketing term representing the trade via *internet*.

1.10. It seems to be necessary to end this terminological noise as soon as possible, before official statistics is made to invent statistical measures of those “phenomena”. Official statisticians are not only able, but also authorized to elaborate scientifically based conceptual and terminological proposals and to implement them via statistical classifications and surveys.

## 2. Functions of information modern society and economy

2.1. In modern societies and economies (so called “information economies”) the information plays the following functions:

- representation of real world,
- creation of the resources of knowledge,
- control function,
- decision function,
- consumption function.

2.2. **The representation of real world** is the fundamental criterion of quality of any information. It is also fundamental criterion of quality of statistical information. Statisticians use scientific methods and specific language of statistical data to describing complicated phenomena and processes of technology, of social and economic life. *Differentia specifica* of statistics are statistical methods and statistical language. The prerequisite of the representation of real world by statistics is the adequacy and coherence of statistical methods and of semantics and pragmatics of statistical language with real phenomena and processes.

2.3. **Function of the creation of the resources of knowledge** is dominating in information systems, which the objective is to register real social, economic or natural processes or to extend the knowledge about real world objects, phenomena and processes. The function of the representation of the real world is dominant in scientific research, in technology sector, in book-keeping, in official statistics etc. Dissemination of information, which real-

izes the function of representation of real world, is realized by education, by scientific publications, by scientific and technical information systems. Main function of statistics is to produce information, which realizes this function.

**2.4. Control function of information** means, that the control, governing or guiding is realized using the information as main instrument of influence on peoples' behavior. This function is dominating in political information systems, in marketing and advertisement, in some segment of information management systems, in the army (orders). The jurisdiction is also an example of control function of information. *The control by information in the information society is very effective.* It is very easy and cheap to provoke expected behavior of individuals or large groups of population disseminating little information, which can not be personally verified and "will be taken for granted." The efficacy of influencing the behavior of large groups by dissemination of information is higher, when people are convinced, that it is delivered by the information systems specialized in the representing the real world<sup>4</sup>. Agencies organizing the advertisement, electoral campaigns or social actions know that very well. In non-democratic social systems the control by information is widely used and protected by information monopoly of the state. In global information society these forms of control are also implemented, e.g. by creating stereotypes of some regions, nations, ethnic groups, economies, products and producers. Modern mass media (in particular: TV, radio and popular press) are very powerful instrument for realization of control function of information in global scale.

**2.5.** *There is no direct relation between the function of representation of the real world and the control function.* The information, which does not represent any real phenomenon, may be even more effective for control of other people, then the true information honestly representing the real—world phenomena.

**2.6. Decision function of information** is realized, when an individual or an organization is taking any decision on the basis of information collected. For decision making the information should be in the disposal of an individual or an organization in *proper content, language, time and place*. In modern societies there are huge information systems providing information necessary for taking decisions. Some of those systems, because of their infrastructural character, are managed or financed by the governments and international organizations. Those systems compose wide and fine information web, covering all areas life. The decision-makers expect that information received from those information systems meets the quality requirements and is well representing the reality. However, it may not always be necessarily true.

**2.7. Consumption function of information** in information society seems to be the dominating function. In many countries *huge industry is developed to*

---

<sup>4</sup> E.g. the advertisement is more effective, when some advertisement slogans are read by a "professor," a "doctor" or by popular actress.

*produce information for consumption only.* The major flows of information produced for and disseminated by TV, radio, by books, magazines and newspapers, is appropriated for consumption only. It may have nothing common with the real world, with the decision-making processes or any other requirements of the users, except the consumption itself. The scale of production of information for the consumption is growing dynamically. Global market for that kind of information was created and is developing very fast. The higher is the level of economic development of the country, the higher is the consumption of information (*panem et circenses*). The market of information for consumption operates on global scale.

2.8. One single set of information may play several functions at the same time: to represent the reality, to help to take decision, to control some processes and to be tasty cake for consumption. E.g. well elaborated, honest advertisement should:

- 1) to inform the customer on real technical characteristics of the product,
- 2) to influence the customer and create the demand for the product,
- 3) to help the customer to take proper decision of buying or not buying the product,
- 4) to satisfy esthetic needs of the perception of nice pictures and sweet music.

The same requirements may refer to political electoral campaigns, to the education processes as well as to scientific conferences. Scientific lectures supporting results of careful research (extension of the resources of knowledge), should not necessarily be boring. They may deliver scientific information (extension of knowledge) in attractive form (nice and easy consumption of information), convince people to some theories (control functions) and help them to take decision to read or to read not the paper of the lecturer (decision function).

2.8. From users' point of view the best is the situation, when a given set of information fulfils all four functions. However, much more often there is the "functional specialization" of information and of information systems. For example, let us ask ourselves, how often our very important official statistical information, which represents the reality and contains important knowledge is presented to businessmen or politicians in the form almost unreadable for them, in a language, which they hardly understand? How often the content of statistical tables is not relevant and redundant? Is the timeliness acceptable or the data are delivered to late to take any reasonable decision? We should not be surprised, that many end—users prefer to read our statistics published in colorful magazines or presented in TV, in the form of nice graphs, surrounded by pictures of glamorous girls (consumption effect).

2.9. In social and economic information systems the priority should be given to the function of creation of knowledge. It should be accompanied by the dissemination of relevant and pertinent knowledge to the users. This is necessary for proper realization of the decision function of information.

2.10. Relatively new phenomenon in modern economy is the *formula—driven policymaking*. The formula—driven policy means that social or economic decisions are taken on the basis of formalized algorithm fed by statistics. Politicians rather eagerly accept arithmetical formulas replacing their own political thinking and decisions. Those formulas are often introduced by law. The adoption of the formula—driven policy in practice requires statistical data of very high quality, timeliness, accuracy and relevance. It causes also strong pressure on statisticians to producing data expected by politicians. To helping the governments and statisticians to respect statistics as science and as the attribute of democracy important was an important reason of adoption of the UN Fundamental Principles of Official Statistics.

### **3. The Fundamental principles of official statistics as the instrument of protection of the citizens' right to information**

3.1. *The Fundamental Principles of Official Statistics* adopted by the United Nations Statistical Commission in 1994, stresses that statistics should “**honour citizens' entitlement to public information**” and that “**the essential trust of the public in official statistics depends to a large extent on respect for the fundamental values and principles which are the basis of any democratic society.**” Fundamental values and principles are universal and should be respected by all legal, political and institutional levels of the information infrastructure of countries and societies. Below we try to interpret the commandments of the Fundamental Principles as the universal rules of protection of citizens' right to information in democratic society.

The document *Fundamental Principles of Official Statistics* consists of two parts:

- a) the Preamble specifying general political and social foundations and objectives of the Principles,
- b) 10 rules, which are colloquially called “commandments;” it is the reason why the Principles are often called *The Statistical Decalogue*.

#### 3.2. The Preamble

3.2.1. The declaration of *Fundamental Principles of Official Statistics* was adopted by the United Nations Statistical Commission in 1994. It was preceded by the resolution of the Economic Commission for Europe on Fundamental Principles of Official Statistics in the Region of the ECE, adopted in 1992. The draft of the resolution was elaborated by the Conference of European Statisticians, by statisticians for statisticians and for governments.

3.2.2. The world of today is a global village. The prerequisite of globalization of politics, economy and society is the global information system. This globalization of information is a real thing now. Powerful communication systems, in particular satellite TV, are collecting and disseminating informa-

tion all over the world. Globalization of information systems is a chance for the sustainable development of the global economy. However it is also a threat for the mankind. Information, if misused, is an extremely dangerous weapon and poison for humanity. In a global information village there should be common, accepted and adopted rules and laws for information and information systems. All democratic societies understand very well the need for such rules and laws.

3.2.3. The universal declaration of human rights defines the indispensable attributes of freedom. One can deduce that the right to information and the right to privacy are those which determine the space of freedom of a human being. The rights to information and to privacy constitute the basis of a democratic society.

3.2.4. In a democratic society an individual has the right to privacy and confidentiality. The space of privacy is determined and defended by law. That space should be a compromise between the needs of an individual and the needs of the society for information. This compromise is based on the “**minimax information principle**”: **minimum information from an individual—maximum information for the society**. In a democratic society the individuals have the right of access to information needed for their life as active citizens. The same principle refers to all forms and institutions, in which social and economic life is organized. The information infrastructure of any democratic state should be based on precise laws regulating the co-operation of all systems, which compose the information infrastructure of the state and the economy.

3.2.5. The declaration of the United Nations Statistical Commission: **Fundamental principles of official statistics**, creates the frames for legal basis of official statistics, as an important part of information infrastructure on the levels of countries as well as on global level. It is addressed to official statistics. Anyhow, the spirit and the letter of the declaration are universal. It contains the rules, which may and should be used in the development of all infrastructure information systems. In this paper I will try to stress those aspects of the principles, which have a universal, general character.

3.2.6. Official statistical services are the institutions of public trust. The position and authority of official statistical services is based on the mutual trust of statisticians, data sources and data users in:

- a) Statistical deontology,
- b) Statistical methodology,
- c) Statistical efficiency.

The trust must be mutual. To achieve that, statisticians must prove every day, in every survey and data, which one can rely on their professional ethics, methods and skill.

3.2.7. “The essential trust of the public in official statistical information depends to a large extent on respect for the fundamental values and princi-



ples which are the basis of any democratic society which seeks to understand itself and to respect the rights of its members.”

3.2.8. Mutual trust of statisticians, data sources and users is the prerequisite of the quality of official statistics. Statisticians understand (or should understand) very well that without trust and confidence it is not possible to expect co-operation of citizens, enterprises and other respondents in providing appropriate data needed for necessary compilations. Most of the users do not know or do not understand the conventions and details of methods used by statisticians. The transparency of standards and methodology is also the prerequisite of mutual trust.

3.2.9. The human right to truth, to good, verified, honestly selected information should be realized by all infrastructural information systems. Fundamental principles of official statistics show how official statistics as well as all other infrastructural information systems should act for the benefit of that universal human right.

### 3.3. Commandment I—Statistical “Ockham’s Razor” for information

3.3.1. **“Official statistics provide an indispensable element in the information system of a democratic society, serving the government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honor citizen’s entitlement to public information.”**

3.3.2. The commandment contains two contradictory requirements:

- a) Statistics should provide to the public **all necessary** information;
- b) Statistics should provide **nothing but necessary** information.

The question is, who—in a democratic society—is entitled and has the authority to decide, what information is necessary, and what is not? How to test practical utility of official statistics? What is the minimum of information necessary to honor citizen’s entitlement to public information?

3.3.3. In a democratic society there is a need for social control over all information systems which create the information infrastructure of the country and society. This control may and could be achieved by establishing—by law—representative social bodies (councils, committees etc.) entitled to monitor respective infrastructural information systems. Such bodies are an integral part of all official statistical systems in democratic societies. Their members are the representatives of governments, economic and social organizations, businesses and scientists. They represent both the users of statistical data and the respondents. As a rule, they concentrate their attention on the program of surveys and methodological issues. They also monitor the activity of official statistical systems and offices.

3.3.4. Social control established by law should be the pattern for all infrastructural information systems. The scope and forms of control may be different. If an information system is static, i.e. the scope of information col-

lected and stored is stable or fixed (e.g. stabilized social security records, tax records, etc.), the contents of information, rules and organization of a system may be determined by law and controlled within the frames of monitoring of law. If a system is dynamic, (e.g. official statistics, scientific and technical information systems etc.), there is a need of continuous monitoring by special bodies.

3.3.5. The practice of official statistics, i.e. control of programs and methodology by special committees or councils representing the society and the government, is adaptable for other infrastructural information systems.

3.4. Commandment II—End-user’s dilemma: “To trust or not to trust”

3.4.1. **“To retain the trust in official statistics, the agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.”**

3.4.2. Official statistics and statistical offices are the institutions of public trust. The principle of public trust should be the “credo” of all infrastructural information systems in democratic society.

3.4.3. Public trust is based on the confidence, that all data are collected, stored, processed and disseminated according to professional ethics, on best scientific basis, according to the standards coherent with international recommendations, scientific methods, national laws and real economic and social conditions. Professionalism and respect for law are the base of public trust.

3.4.4. Why often the public does not trust official statistics?

One reason is that statistics uses the scientific language of economics, sociology, technology etc. while the users expect to obtain data represented in their own (colloquial) language. The “linguistic gap” between users and producers of statistics is a real threat for statistics. As long as the users of statistics were mainly professional economists or politicians, the official statistics could speak to them in scientific, professional language. Now, in global information society, statistics should learn to speak to the public in the language of the public. Statisticians should develop such communication language as a tool for transfer of statistical data to the public and semi-professional users.

3.4.5. In modern information infrastructure different system use different languages for the same or similar events, processes or things. This modern “Babel Tower” is also a cause of in confidence and trust. Semantic standards for infrastructural information systems are the prerequisite for social trust in statistics as well as in other systems.

3.4.6. All systems of the information infrastructure of the country should define their professional ethics. The professional ethics of statisticians is codified in the ISI Declaration on Professional Ethics. The globalization of information processes and systems requires the “codification” of general principles of professional ethics for any public activity in information.

3.5. Commandment III—Statistics should “make sense” not only for statisticians, but also for end-user

3.5.1. **“To facilitate a correct interpretation of data, the statistical agencies are to present information according to scientific standards on the sources, methods and procedures of the statistics.”**

3.5.2. Statistical data must be understandable for all users. The language in which statistical data are presented should be user-friendly and user-adjusted. End-users interpret statistical data in the semantics of their own language. To facilitate the interpretation statisticians should present and disseminate “user-tailored” chunks of statistical knowledge. Modern information technologies enable to develop statistical knowledge based systems, which meet the requirement of the Commandment III. This is true for any infrastructural information system.

3.5.3. Chunks of statistical knowledge should be self-contained. They ought to cover relevant data, respective metadata, all necessary methodological explanations and references to pertinent knowledge. The contents of chunks should be adapted to the specificity of particular classes of users.

3.5.4. The language of official statistics is the language of science. Representation of statistics information in the language understandable for end-users can not change, reduce or “primitivize” the contents and semantics of statistical information. If the complexity of statistical data can not be represented in the categories of end-user languages, it would be better to offer the data in “original” professional language of statistics rather, than to mislead the users by erroneous simplifications.

3.5.5. Training of users and wide dissemination of general statistical knowledge is the duty of official statistics. Training and teaching of those who co-operate in data dissemination to the public (in particular journalists) is of utmost importance. Ignorance of journalists may cause severe losses of information in the process of dissemination. It may also destroy the trust in statistics and mislead the users.

3.5.6. Many scientists and different organizations produce data, which have the same form as official statistical data. The odium of poor quality, errors and mistakes of those data is put on official statistics. That is why official statistics is authorized to develop standards for representation of statistical data and of any other economic and social data which are presented in the form of statistical indicators. Those standards should cover: lexical rules, syntax, semantic rules and generic standards for representation of different classes of chunks of statistical knowledge and metadata.

3.6. Commandment IV—Statistics is the job for courageous men ready to tell “the truth, all the truth and nothing but the truth”

3.6.1. **“The statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.”**

3.6.2. This commandment is a real test of democracy and independence of statistical agencies. Statistical data and metadata are used by many different classes of users: scientists, politicians, journalists, businessmen, “ordinary citizens”. Politicians and mass media use statistics to prove their political and social ideas and theorems. Such use of statistical data coincides with “purpose-oriented” interpretation, erroneous interpretation or even misinterpretation of statistical data.

3.6.3. Most frequent forms of improper interpretation of statistical data are following:

- a) Change of names of statistical indicators. Numerical values are correct, but the name of an indicator is reduced, changed, reworded. Even small modifications of names of indicators may cause significant changes of the meaning of data, of their economic and social sense;
- b) Teleological selection of data, leading to the misinterpretation of economic or social phenomena;
- c) Use of incorrect definitions of concepts;
- d) Mixing of data from different sources: official statistics, results of scientific research, information transferred by mass media, data from other unverified sources;
- e) Falsification of data.

3.6.4. Statistical offices should react if erroneous interpretation of data may cause misuse of official statistics or if it may lead to the questioning of the quality of official statistics and destroy the trust in official statistics. The reaction should be adjusted to the place, form and consequences of errors and misuses. Most difficult and serious cases of misinterpretation are committed by mass media. The consequences of misuse of official statistical data in mass media are long lasting, they influence very large number of respondents and users of statistics. The result of erroneous interpretations and misuses in mass media is not only the dissemination of incorrect data, but also the damage of reputation and position of statistical offices. The more powerful are the media, the more long lasting and wide are the consequences of such misuses.

3.6.5. Statistics should demand the rectification of any misinterpretations or misuses in mass media, even if the correction is printed may be non equivalent. In many democratic countries the law obliges the media to publish rectification.

3.6.6. Erroneous interpretation or misuse of official statistical data by scientists, in reports or books is not to be ignored. There is a feedback between research and statistics. If there are misinterpretations done by scientists and researchers, official statisticians should take such events as signals of some gaps, fails or imperfections in statistical methodology, documentation, ways and methods of presentation of statistical data and metadata.

3.6.7. Erroneous interpretation and misuses of data by other governmental institutions, officials and offices are very sensitive issue. Although such

misinterpretations are often delivered to the public and disseminated by mass media, the reaction of statistical offices is—as a rule—addresses not to the public, but to the respective government bodies. In critical situations however the reaction should be public, if it may destroy the trust in statistics. Statistical offices are parts of government structures, so the reaction on the errors and misuses done by other governmental bodies should be well-balanced and long term oriented. It is better to teach and to explain than to get into conflicts.

3.6.8. Standardization of representation of data and metadata may be helpful in eliminating most serious errors and misuses of statistical data. Wide dissemination of metadata: catalogues of indicators, definitions of terms, methodological explanations, and implementation of standards, including generic standards for representation of data, are effective in long term.

3.6.9. Harmonization of basic infrastructural information systems is also an effective approach to eliminate potential threats of errors and misuses of statistical data. The harmonization should cover:

- a) Unified definitions of common concepts;
- b) Harmonized classifications and nomenclatures;
- c) Common identification systems (e.g. of businesses);
- d) Common standards for representation of data and metadata;
- e) Sharing data and metadata according to the rules and possibilities defined by law.

3.7. Commandment V—Do not victimize your respondents = collect once—use many times

3.7.1. **“Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents.”**

3.7.2. Use of administrative records for statistical purposes is the future of modern statistics. Also the use of administrative records as the sources of data is the future of modern information infrastructure of any democratic society. Mutual transfer of data between infrastructural information systems must be determined by law and controlled by the law. Possibilities and limits of matching registers should be explicit precisely by law according to the legal system of particular country.

3.7.3. We want to stress that the postulate of wide access of official statistics to administrative records is the consequence of the principle of reduction of response burdens and costs of statistics for the society and for the state. In a democratic society information stored in and by any public organization, for the money of citizens, is the property of the society. The most expensive phase of any information system is data collection, capture and editing. Therefore the duplication of data collection and edits should be elimi-

nated. Existing and available data files and data bases should be used not only for operational purposes of organizations which collect and store the data, but also they should be used for economic and social analyses, research and for better information of the society.

3.7.4. Use of administrative records for social and economic analyses, in particular for the purposes of official statistics, assumes that the certain level of integrity and coherence of infrastructural information systems is reached. One can easily define the minimal level of integrity of that infrastructure. It was specified in par. 4.9 above.

3.7.5. Who should take the responsibility and whom law to co-ordinate should authorize the information infrastructure of the country? It seems that official statistics is predisposed to play the role of impartial co-ordination of that. There are several important arguments for that:

- a) Statistical offices do not collect data for themselves; their duty is to inform all the users according to their rights and needs; all other organizations which create the infrastructural information systems collect data for their operational tasks (e.g. for collection of taxes, to pay pensions, to collect customs, to license some activities etc.). Official statistics in not taking any decisions. It is obliged to supply data to the others.
- b) Statistical offices cover all subject matter areas of social, economic and political life, while all other organizations of the information infrastructure cover very specific frames of reference only.
- c) Official statistics covers all the territory of the country; it also collects data from abroad to enable international comparability. The fields of interests of the other services cover—as a rule—a part of territory of the country.
- d) Official statistics uses widely internationally approved standards and scientific methods, while other information systems are based on administrative regulations and principles valid, as a rule, for one single country.

3.7.6. Statistical data should be produced timely. “Classic” collection of input data, editing, quality control are time, costs and labour consuming. Administrative data sources update and control their data much more frequent, for their administrative purposes. Also quality control of administrative records is very detailed. However, the criteria of updating and quality control of administrative data differ—sometimes very much—from those used for statistical purposes.

On the basis of comparative analyses of several administrative data systems (social security system, tax system, customs, cadastre etc.) it was concluded, that the integrity of administrative information systems with statistics could be reached rather easily. It is rather the problem of good will and of mutual understanding of advantages, which may be achieved thanks to the integrity. E.g. it would be quite easy to adjust the contents of tax declarations in such a way that it could be used as the frame for many surveys; the same could be done with the cadastre. Social security records in some countries

need only minor adjustments of contents and classifications and implementation of common identifiers of official statistics to become a rich source of data for statistics of employment, wages and salaries, conditions of work etc.

3.7.7. Use of administrative records for statistical purposes meets all the requirements of the commandment V. It helps to get timely good quality data at minimum costs and with very little burden for respondents. International statistical organizations should use all their authority to explain the governments how important, beneficial and how advantageous is the co-ordination of information infrastructure from the point of view of the use of administrative data for statistical purposes. It is necessary to repeat all the time, that the use of administrative data for statistical purposes does not mean the access of statistical offices to administrative data. Official statistics does not collect data for its own purposes but for the users.

3.7.8. It should be also stressed that infrastructural information systems can not limit their responsibilities to the internal functions of the organizations, which manage those systems. They must contribute to the development of the whole information infrastructure of the country and society. So they should meet some general requirements and provide some information services for the whole infrastructure.

3.8. Commandment VI—Statistical “black hole”—the prerequisite of trust and quality

**3.8.1. “Individual data collected by statistical agencies for statistical compilations, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.”**

3.8.2. Individual data are the data, which can be identified with concrete natural or legal person. Non aggregated data, which can not be associated with a person, by any existing method of identification, may be used not only by official statisticians. Those data can be supplied for external researchers, analysts etc. The critical issue is, what methodological solutions should be adopted in a survey to enable the disclosure of non-aggregated data when identifiers (e.g. personal identification number, business register number etc.) and other unique characteristics (e.g. postal address and code) are separated.

3.8.3. For any survey the list of data, which may cause disclosure of individual data, should be specified. This specification should be an integral part of the methodology of a survey. Those data should be kept together with other data only when aggregated statistical information is produced or in the files strictly confidential, with no access to external users.

3.8.4. Matching administrative records and registers may also cause the danger of disclosure. Having in mind the commandment V (access to administrative records), the rules of preventing disclosures should cover all individual source data, also those taken from administrative records.

3.8.5. Modern information technologies: data bases and knowledge based systems, direct remote access to the data, possibility of matching data from different statistical sources by the users themselves, enlarge the area of potential disclosure. For example having the access to the business registers and to the aggregated data by municipalities one may for smaller municipalities identify concrete businesses. Statisticians should carefully analyze all input data available in a statistical system and in other co-operating infrastructural information systems, all relevant aggregated data produced, respective available metadata, including registers (of businesses, persons, territorial lots, addresses etc.). Statisticians should simulate all potential “case studies” of disclosure and propose measures (organizational, methodological, technological) to reduce the possibility of disclosure.

3.8.6. If the test of disclosure is negative, i.e. there are no practical possibilities of disclosure; the non-identifiable individual data may be given to the disposal of other users. However it seems that official statistics should select customers and applications for whom and for what the non-aggregated data are delivered from the official statistics. In my opinion the following rules should be obeyed:

- a) The data from administrative records stored in statistical systems can not be re-distributed without special request or special permission of the “owners” of those data.
- b) Data should be used for scientific research, professional analyses and teaching; if the customer does not justify the transference of data, non-aggregated data should not be delivered.
- c) Statistical offices should offer the elaboration of statistical compilations for individual customers rather than the transfer of data outside the statistical system.

3.8.7. Individual non-identifiable data delivered by official statistics to other customers can not be redistributed or disseminated by those users. The purpose, for which non-aggregated data are sent outside the statistical system, should be known and controlled. Any overuse in that issue should be treated very seriously.

3.8.8. Any data collected on the mandatory basis by statistical offices are confidential. Statistical confidentiality has no exclusions. It is the kind of confidentiality close to that of lawyers, medical doctors, priests etc. Statistical confidentiality is the basis of statistical ethics and trust of the respondents and users to statistics. It must be protected by law. In a democratic society the government, public organizations and all users should respect statistical confidentiality.

3.8.9. Statistical systems should be developed in such a way as if they were the cybernetic “black boxes” for individual data.



3.9. Commandment VII—Statistical metainformation convert digits to statistical information

3.9.1. **“The laws, regulations and measures under which the statistical systems operate are to be made public.”**

3.9.2. Democratic statistical system is based on the principle of equilibrium:

- a) Confidentiality of individual data—open access to aggregates;
- b) Mandatory supply of data by respondents—explicit legal basis for statistical duties;
- c) Access to administrative records—prohibition of redistribution;
- d) Impartial statistical duties—impartial access to statistical data
- e) Confidentiality of identifiable data—transparency of metadata.

3.9.3. Statistical data without all relevant, transparent, available metainformation do not represent any information at all.

3.9.4. Statistical metainformation covers metadata related to statistical data, as well as information about official statistical system, including the laws, regulations and measures under which official statistics operates.

3.9.5. In a democratic society citizens and businesses are entitled to know, what are laws and regulations of statistical activities, to know their duties and their rights toward the statistical system.

3.9.6. National statistical offices should facilitate the access of respondents and users to the metainformation on laws and regulations of official statistics. Dissemination of that information is the task of statistical offices. The respondents should be informed, what are legal basis and regulations of any survey in which they are expected to co-operate, is it mandatory or optional, what are the consequences of participation in a survey, and what are the consequences of refusals.

3.9.7. The users should be informed what data, where and when are or will be available, how to access the data, which statistical services are free, which can be delivered on commercial basis, what is not available.

Education of respondents and users in laws and measures regulating statistical activities is very helpful for better co-operation of those persons and units with statistical offices. Respondents and users should be treated like partners of professional statisticians. “Partnership for better statistics” of all actors of statistical processes may be helpful in improving the quality of statistical production and in increasing its efficiency.

3.9.8. In advanced information technology environment the commandment VII is implemented by the development of complex metainformation systems. The concept of METIS (statistical metainformation system) elaborated by the group of experts of the UNDP/CES research project in the 80th is an ideal model of that kind of metainformation systems.

3.9.9. There are two models of regulation of statistical activities:

- a) Centralized model: one basic legal act, voted by parliament or introduced by highest governmental authorities plus sub-ordinate legal acts dealing with more detailed matters or regulating specific areas of statistics
- b) Decentralized model: regulations concerning statistics are integral parts of legal acts regulating selected areas of social and economic life, e.g. in the legal act regulating medical care issues there are specific paragraphs defining statistics of health etc.

Centralized model is often adopted in the countries in which statistical activities are concentrated in one central statistical office. There are however situations where decentralized model of legal regulations coincides with the centralized model of the national statistical office.

It is the duty of national statistical services to take care of extraction of statistical issues from different legal acts. Information on legal basis should also refer to legal regulations concerning other infrastructural information systems and data sources. Statistical offices should be entitled to take active part *ex officio* in all legislative processes relevant for statistics, even indirectly.

3.9.10. In democratic societies statistical data are often used as the basis for administrative decisions, e.g. revalorization of pensions is based on the CPI. All users should know indexes (algorithms, methods of compilation, definitions of concepts) all meta-information related to such data, and available in statistical publications and in accessible statistical knowledge based system (if they exist).

### 3.10. Commandment VIII—How to co-ordinate the co—ordinators?

3.10.1. **“Co-ordination among statistical agencies within countries is essential to achieve consistency and efficiency in the statistical system.”**

3.10.2. There are three models of the organization of statistical services:

- a) Centralized model: one central statistical office responsible for all official surveys;
- b) Decentralized model: many statistical offices, often integrated with different ministries;
- c) Mixed model: a part of surveys conducted by the central statistical office and some areas of statistics organized by other institutions (e.g. national bank, some ministries).

3.10.3. To avoid information noise, misunderstandings, duplication of surveys and data, excessive burdens of respondents, there is a need for effective central co-ordination of all statistical activities in the country.

3.10.4. Tools of effective co-ordination are the following:

- a) Joint, integrated program of statistical surveys prepared by one agency (e.g. central statistical office or one selected office) containing all surveys conducted by all offices or agencies entitled by law to conduct official statistical surveys. The office responsible for preparation of the joint program of surveys should be equipped with the competencies of elimination

of duplications and inconsistencies ex officio or in the process of negotiations. Joint program of work may be also helpful in co-ordinating the participation and co-operation of other infrastructural information systems with official statistics.

- b) Common registers and identification systems of persons, businesses, other organizational units, territorial units etc. Those tools of co-ordination are of importance for other infrastructural information systems too.
- c) Common or coherent classifications, nomenclatures, definitions of terms;
- d) Co-ordination of data dissemination policy (very important postulate, however a sort of wishful thinking).

3.10.5. This commandment, if accepted by all actors, may help in developing very effective informational infrastructure of the country. Co-ordination should not be limited however to statistical activities only, but to all information systems of infrastructural character.

3.10.6. Co-ordination of statistical activities and infrastructural information systems within countries creates the effect of positive synergy of information systems. This synergy may be observed by:

- full information of all existing data resources for all classes of authorized users;
- more effective, cheaper and quicker access to data;
- wide transfer of data between systems;
- better control of security, confidentiality and privacy;
- elimination of excessive duplication and redundancy of data flows and files;
- better control and “self-control” of data quality;

Positive, creative competition on the market of information services. “Splendid isolation” and poor coordination of statistical surveys data causes negative synergy within statistical system and within information infrastructure of countries and economies..

3.11. Commandment IX—Standardization—way of progress or stagnation trap?

3.11.1. **“The use by statistical agencies in each country of international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels.”**

3.11.2. In the preamble we put the stress on the globalization of information infrastructure. Without standardization of data, metadata, organization and technology global information system would collect and produce noise. Some consequences of the noisy global information systems may be observed in the sphere of mass media. One can easily imagine, what would have mean such a “noisy situation” for statistics and for any other global information systems.

3.11.2. Official statistics is a scientific information system. The sciences, which are the basis of good statistics, are of the global nature. The achieve-

ments of sciences in watching and measuring social, economic and environmental phenomena and processes are represented in the form of statistical standards: classifications, nomenclatures, methods, concepts.

3.11.3. All national statistical systems must observe internationally approved standards. It does not mean that the offices should adopt international standards directly, inside the country. Adoption of international statistical standards may be done in different ways:

- a) direct implementation of international standard for national purposes;
- b) creative adoption of an international standard for domestic needs, following the approaches of a respective standard (e.g. HS—CN—PCN (Polish Combined Nomenclature) in such way that no “gateways” between international and national standards are necessary;
- c) Developing “gateways” (e.g. correspondence tables between classifications or nomenclatures) between international and national standards;
- d) Developing duplicated data flows and data processing—one for national purposes, second—in accordance with international standards.

Strategies (a) and (b) are recommendable. Strategy (c) is often of necessary, however in long terms one should consider the replacement of (c) by (b). The strategy (d) shows the gap between national statistical system and “the rest of the world”.

3.11.4. Statistical international standards are important not only for official statistics. They should be respected or considered by other statistical agencies (“non official”) and by all infrastructural information systems.

3.11.5. Often direct implementation of a standard is very difficult, sometimes it is not possible at all. Therefore special attention should be given to the elaboration and implementation of **generic standards**. GESMESS—generic standard for statistical tables is a good example how efficient that approach could be. The work done and continued by the CES expert group on METIS in the field of development of generic standards for statistical metadata (classifications, correspondence tables, registers etc.) is also very promising.

3.12. Commandment X—“You are not alone”—the role of international co-operation of statisticians and statistical offices

3.12.1. **“Bilateral and multilateral co-operation in statistics contributes to the improvement of systems of official statistics in all countries.”**

3.12.2. In most of the countries there is one central statistical office. Official statisticians do not have equivalent partners for solving different problems within countries. Our situation is similar to that of the scientist working in a specific branch of science. International co-operation, exchange of experiences, joint research work are necessary to stimulate the progress in methodology and technology of statistical work. Basic problems of surveys development, methodology, statistical standards, statistical production and informatics the statisticians must solve by themselves.

3.12.3. Open, friendly international co-operation of official statisticians and statistical offices is an element of statistical ethics. Statistical know-how is the part of statistical scientific knowledge. It should be accessible to all members of statistical community as it is in other branches of science. The exchange of statistical knowledge, know-how, ought to be done on the non-profit basis. ISI and co-operation of national associations of statisticians show how to integrate international statistical community.

#### **4. The ISI Declaration on professional ethics in statistics and its role in the identification and protection of the citizens' right to information**

4.1. The *ISI Declaration on Professional Ethics* was adopted by the General Assembly of the ISI in Amsterdam in 1985. This declaration is addressed to both statisticians, any organizations involved in statistical surveys and official statistical agencies. The objective of this declaration, from the point of view of the citizens' rights to information, was to codify deontological principles relevant to statistics to protect the rights of users of information and the rights of suppliers of information in statistical production process. Although the *Declaration* is addressed to statisticians, it contains many statements important for politicians, government officials, researchers, journalists and any subjects involved in conducting statistical surveys and in the use and dissemination of statistical data to the public. Below I will try to point out those commandments of the Declaration, which may have direct impact on the proper understanding and implementation of the citizens' right to information.

4.2. The *Declaration* determines four layers of obligations, which should be fulfilled in any statistical research, inquiry, survey or statistical analysis:

- **obligations to society**
- **obligations to sponsors and employers**
- **obligations colleagues**
- **obligations to subjects**

##### 4.3. Obligations to society

The layer of the **obligations to the society** defines the principles, which should be obeyed by statisticians as the producers and suppliers of information and knowledge to the society. In this sense the statisticians are those professionals, who inter alia realize the right of a citizens to be well informed. The following rules may be specified in this layer:

- a) Statistical survey is predicted on the belief, that greater access to well grounded information is beneficial to the society. Therefore the statisticians should use any possibilities open to them to extend the scope of statistical surveys and to communicate their findings for the benefit of the widest possible community.

- b) At the same time the statisticians should be aware of any possible misconstruction, misuse, or misinterpretation of statistical data. They should conduct the surveys and disseminate the data in the forms and scope minimizing the possibility of manipulating and misuse.
- c) Statisticians should forecast and consider any possible conflicting interests of different groups of users and founders. They should guard the data against predictable misinterpretation and misuse.
- d) Statistical inquiries and surveys must be impartial. So, while statisticians operate within the value system of their societies, they should attempt to uphold their professional integrity without fear and favor. They should not engage or collude in selecting methods designed to produce misleading results, or in misinterpreting statistical findings by commission or omission.

It seems, that the rules specified above should be obeyed not only by statisticians, but also by any subjects producing and disseminating information to the society. Producers and disseminators of information should not be made to deliver to the society the information, which do not meet the requirements specified above. They should have the right to refuse to produce and disseminate the data, which do not meet the quality requirements and which may cause misinformation. Statisticians and all other producers and disseminators (e.g. journalists), of information are also the citizens. Their “active” right to information as the citizens is to refuse producing the information, which does not meet professional and social quality criteria.

#### 4.4. Obligations to sponsors and employers

Relation between the statisticians and the sponsors, founders of surveys or employers is very sensitive area. Founders and employers may expect from statisticians, journalists and other producers of information to receive the information adjusted to their needs. Sometimes the requirements of founders may be in conflict with the quality criteria (representation of the reality). The statisticians and other producers of information may be pressed to adjust the methodology of surveys and analyses to receive “expected results,” which are in line with political or economic expectations of founders.

To avoid this, the statisticians (and other producers of information) should:

- a) clarify in advance the respective obligations of employer or founder, to inform them on relevant rules of professional ethics,
- b) consider all available statistical methods and procedures and to choose those, which fit the best to produce good quality data relevant for the founder, but—first of all -representing social or economic reality,
- c) to accept contractual conditions that are contingent upon a particular outcome from a proposed survey,
- d) keep the confidentiality rules of the information received from the founder or employer as confidential.

#### 4.5. Obligations to colleagues

In information society, where a lot of poor quality information is produced and disseminated, the confidence of the society to statistics as the reliable source is the essential value. All statisticians should maintain the confidence in statistics by strict obeying the scientific level and rules of their activities.

Within the limits of confidentiality requirements they should exchange information on scientific methods, procedures, techniques and other findings which may help the colleagues to improve their tools.

#### 4.6. Obligations to respondents and sources of information

All citizens have some obligations to deliver information for different purposes. The ISI declaration has formulated important principles protecting citizens against misuse in the area of the collection of information for statistical purposes. The most important are following:

- a) avoid undue intrusion,
- b) collect only the pertinent information, necessary from the point of view of the objectives of the survey,
- c) if the survey is voluntary, obtain freely given consent of the subjects; they should be clearly informed on all aspects and consequences of their participation in the survey, and they should be aware of the fact, that they may refuse to take respond,
- d) if the supplying of information by the subjects (individuals, households or companies) is obligatory, they should be aware of that and of the consequences of the refusal.

In modern IT environment there are technical possibilities of accessing the information on individuals without informing them. Government administrations, the units offering social services units, businesses store and maintain huge amount of information on individuals. Accessing those administrative records without explicit legal authorization one brakes the citizens' law to privacy and to the protection of information.

The ISI declaration clearly says that such practices are in contradiction with the principles of professional ethics of statisticians.

Serious violation of the citizens' right to information is the overuse of the trust of individuals in collecting information, especially the trust of children, of people not acquainted with what may happen with the information they provide if they are asked.

Very important deontological commandment is the obligation to protect the interests of their respondents by implementing respective methods and tools to keep the confidentiality of data, maintaining confidentiality of records and inhibiting disclosure of identities.

4.7. It seems to be clear, that the ISI declaration on statistical ethics may, and in my opinion should be extended to any information producers, collectors and disseminators. In modern IT environment the threats for citizens' rights to information and privacy are much more dangerous then in the past.

The similar codes should be elaborated and adopted by all other professions in the information sector of the economy.

## 5. Conclusions

5.1. In modern IT environment the need for codification and introduction by law the professional ethics of all professions involved in the collection, storage and dissemination of information is the prerequisite of protection of citizens' right to information.

5.2. In democratic society the governments are obliged to control the implementation and to execute the observance of the laws. Codes of professional ethics of statisticians and other professions are helpful in proper identification of the cases and interpretation of laws.

5.3. In modern IT environment many information systems operate on supranational and global scale. International organizations, both official and professional, should actively react, within the frames of their mandates, to monitor, control and influence the protection of the citizens' right to information.

5.4. Fundamental principles of official statistics and the ISI declaration on professional ethics should be included to the education programs of statisticians. All students who are thought statistics should be aware of the professional ethics related with statistics.

5.5. ISI may encourage other international organizations of professions involved in information processes to elaborate respective deontological codes. The *Fundamental Principles of Official Statistics* and the *ISI Declaration on Professional Ethics* may be used as good example and effective pattern.

## References

- Fundamental Principles of Official Statistics*, The Statement of the United Nations Statistical Commission, New York, 1994.02.10.
- ISI Declaration on Professional Ethics*, *ISI International Statistical Review*, (1986) No. 54, pp. 227–242.
- Oleński J., *Elementy ekonomiki informacji (Foundations of economics of information)*, Uniwersytet Warszawski—Wydział Nauk Ekonomicznych, Seria: Studia Informatyki Gospodarczej, Wyd. Nowy Dziennik, Warszawa 2000.

### **A b s t r a c t** Foundations of Ethics in Statistics in the Light of the ISI Declaration on Statistical Ethics and the UN Fundamental Principles of Official Statistics



1. Every information activity is based on the information users' trust in the sources of information and on the presumption of professional reliability of the information producer and disseminator. It is so because the user of information has no means to verify its quality. Thus, in many areas of the information activity, codes of professional ethics are functioning. Statistics are one of such areas. In 1985, at the session of the International Institute of Statistics (ISI) in Amsterdam, the declaration called in the statistician circles "Code of Statistical Ethics" was adopted. In 1992, the UN ECE adopted in the form of resolution the "Fundamental Principles of Official Statistics" called—because of the number of 10 rules—"The Statistical Decalogue". In the paper, the fun-



damental ethical principles are discussed that, in the light of the above-mentioned documents, are to be observed by the persons and institutions involved in statistical survey or being users of statistical data.

2. The paper also takes into consideration the results of the proceedings of the special ISI committee on the code of statistical ethics that is presently working on a new version of this document. The amended version takes into account the ethical implications of modern information technologies for statistical surveys. Such implications, including the threat to the information quality and statistical confidentiality and the danger of misinformation, arise from the development of information and telecommunications technologies, the ever broader use of the so-called *formula driven policy* (direct linkage of social and economic instruments with the value of statistical indicators) and the technically easy access to the conduct of statistical surveys by insufficiently methodically prepared teams.

3. The author suggests the need to include the ethical principles set down in the document of the International Institute of Statistics and the UN into the high-school and university programs of statistics. Also, these principles are to be observed in planning and carrying out statistical surveys as well as in the quality management in statistics.