

Explanatory Memorandum

Recommendation No.R (97) 18 of the Committee of Ministers to Member States concerning the protection of personal data collected and processed for statistical purposes

(Adopted by the Committee of Ministers on 30 September 1997 at the 602nd meeting of the Ministers' Deputies)

Foreword

Mission and action of the Council of Europe in the field of data protection

The Council of Europe was created in 1949 to enable governments of democratic European states ^{footnote 1} to co-operate "to achieve a greater unity between its members for the purpose of safeguarding and realising the ideals and principles which are their common heritage and facilitating their economic and social progress" (Article 1 of the Statute of the Council of Europe).

The Organisation is governed by the Committee of Ministers of Foreign Affairs of the member states, which is advised by the Parliamentary Assembly, and many intergovernmental committees of experts dealing with most aspects of the daily life of European citizens, except defence: human rights, harmonisation of law, culture and education, social affairs, public health and the economy. The Council of Europe's activities focus in particular on "topical issues" such as problems linked to drugs, terrorism, refugees and the prevention of torture.

One of the first and most important conventions concluded by the Council of Europe - provided for in Article 1 of the Statute - is the Convention for the Protection of Human Rights and Fundamental Freedoms, which was opened for signature in 1950. Article 8 of this Convention states that "Everyone has the right to respect for his private and family life, his home and his correspondence". This right can be restricted by a public authority only in accordance with domestic law and in so far as it is necessary, in a democratic society, for the defence of a number of legitimate aims. But the Convention also lays down, in Article 10, the fundamental right to freedom of expression. This right includes explicitly the "freedom to receive and impart information and ideas without interference by public authority and regardless of frontiers". The "freedom to receive information" set out in Article 10 is considered as implying the "freedom to seek information".

In the conceptual architecture of the Convention, Articles 8 and 10 are not contradictory but complementary. However, in practice, the exercise of one of these rights can be restricted by the exercise of the other. For this reason, the European Commission and Court of Human Rights have defined in case-law the limits to the exercise of each of these rights and, in particular, the extent to which public authorities have the right to interfere. This case-law has been - and still is - of great importance to the Council of Europe in its work on data protection as the source of criteria for the development of national regulations on data protection. Nevertheless, in the years following the adoption of the European Convention on Human Rights, it became apparent that efficient legal protection of privacy required more specific and systematic development.

From the beginning of the 1960s, rapid progress in the field of electronic data processing and the first appearance of mainframe computers allowed public administrations and large enterprises to set up extensive databanks and to improve and increase the collection, processing and interlinking of personal data. While this development offered considerable

advantages in terms of efficiency and productivity, it also gave rise to a clear trend towards massive electronic storage of data concerning the private sphere of individuals. In the face of this trend, the Council of Europe decided to establish a framework of specific principles and norms to prevent unfair collection and processing of personal data.

A first step in this direction was taken in 1973 and 1974, with the adoption of Resolutions (73) 22 and (74) 29 which established principles for the protection of personal data in automated databanks in the private and public sectors. The objective was to set in motion the development of national legislation based on these resolutions. However, during the preparation of these texts it became apparent that comprehensive protection of personal data would be effective only through further reinforcement of such national rules by means of binding international norms. The same suggestion was made at the Conference of European Ministers of Justice in 1972.

In 1981, after four years of negotiation, the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data ^{footnote 2} - known as ETS Convention No. 108 - was concluded. Contracting parties to this convention take the necessary measures, in their domestic law, to implement the principles laid down in it with regard to personal data of everyone resident on their territory. These principles concern in particular fair and lawful collection and automatic processing of data, storage for specified legitimate purposes, the non-usage of data for ends incompatible with those purposes, and the keeping of data for no longer than is necessary. They concern also the quality of the data - specifying in particular that they must be adequate, relevant and not excessive (proportionality) - their accuracy, the confidentiality of sensitive data, information on the data subject and his/her right of access and rectification.

The convention provides for free flow of personal data between states party to it. This free flow may not be obstructed, for personal data protection reasons, unless parties derogate from this provision, which they may do in two specific cases: where protection of personal data in the other party is not "equivalent", or where the data are transferred to a third state which is not party to the convention.

The convention establishes a Consultative Committee, consisting of representatives of parties to the convention, which is responsible for interpreting its provisions and for ensuring improvement of the implementation of the convention. This committee took the initiative to study to what extent contract law could facilitate transborder data flows between parties to the convention and states not party to the convention and, together with the European Community and the International Chamber of Commerce, drew up a model contract.

Inasmuch as Article 4 provides that states must have enacted adequate legislation before becoming party to the convention, only eighteen member states have ratified the convention ^{footnote 3}. Other states have signed the convention ^{footnote 4} and some of them have passed data protection acts and are preparing to ratify the convention. In addition, Article 23 provides for states not members of the Council of Europe to accede to the convention.

Since the conclusion of the convention in 1981, society has been computerised to such an extent that the personal computer is now a common instrument which permits any individual or organisation to carry out "automatic processing of data". In the meantime, social and economic development has led to even more complex forms of organisation, management and production, based on powerful processing systems. In this context, the individual becomes an active agent of the "information society" - while, at the same time, his or her privacy is

subjected to even greater interference by the information systems of numerous public and private services, such as banks, credit services, social security, social assistance, insurance, police, medical care.

This evolution constitutes an enormous challenge in terms of data protection. Today an ever increasing number of new problems and practical questions is submitted to national data protection authorities - in most countries the national data protection commissioner ^{footnote 5}. These authorities, which, like the ombudsmen, have become an integral part of the control system in a democratic society, must interpret the principles of the convention and apply them to these new problems and questions. Nevertheless, experience has shown that neither the principles of the convention nor national regulations on data protection can regulate exactly every situation in which personal data are collected in different sectors: medical care and research, social security, insurance, banking, employment, police, telecommunications, direct marketing, etc. Of course, in each of these sectors data must be collected and processed in accordance with the basic principles of the convention, but the ways and means may be different. In some sectors conditions may be more flexible than in others, and self-regulation may be more advanced in one profession than in another.

For each of the different sectors, therefore, the principles of the convention must be further elaborated. Rather than amend the convention, or add protocols to it, the Council of Europe prefers to use another tool for this purpose: that of recommendations to governments. Such recommendations have the advantage that they are easier to draw up, to adopt and to implement: instead of signature and ratification by each of the member states, they only require unanimous adoption by the Committee of Ministers. It is therefore simpler to adapt them to changing circumstances than to amend conventions; and, above all, although they are not legally binding, they contain real standards of reference for all member states, whether they are parties to the convention or not. A recommendation constitutes therefore a request to consider in good faith the possibility of elaborating and implementing domestic law in conformity with internationally agreed interpretation of the principles laid down in the convention.

In order to draw up these recommendations - a process which requires, in addition to legal experience, specific knowledge of the subject covered by the recommendation - the Committee of Ministers set up in 1976 a committee of experts on data protection, which subsequently became the Project Group on Data Protection (CJ-PD). This committee is composed of experts from each of the forty member states who are responsible for data protection in their respective countries. When dealing with sectoral recommendations, these experts are sometimes accompanied by consultants who are specialists in the field. It is also traditional Council of Europe policy to invite to these intergovernmental meetings observers from European professional organisations and non-governmental associations working on the same subject. The European Union also participates in drawing up these recommendations in fields in which it is competent.

Although the two committees - the consultative committee and the project group - work in close collaboration, their activities are separate: the consultative committee is the guardian and promoter of the convention, while the project group draws up more technical and detailed guidelines in the various fields.

Over the years, the project group has not only drawn up a series of recommendations ^{footnote 6} but has also published studies on specific subjects in the field of data protection ^{footnote 7}.

Introduction

A. The issues

1. Conducting state affairs, planning the economy or undertaking many public or private activities can hardly be envisaged without statistical information on the activities and the situation of the population, enterprises and administrations. Often motivated by topical concerns or decisions which will have to be taken rapidly, the need for information is often directed towards a combination of data and the most recent developments. But statistical activities also aim at collecting materials for a better global understanding of our societies, their structure and their long-term evolution. They can focus both on large scale phenomena - such as the macroeconomic flows or the demographic tendencies of a state - and on less important issues, such as the productivity of workers in an enterprise, or the effects of a therapeutic measure on a group of patients. Statistical information can therefore be produced or commissioned either by the public sector, the private sector or the scientific community. It is a necessary instrument for planning, forecasts, decisions, management and research and is therefore of use to the legislator, the citizen, public administrations, the economy or science.

2. Statistics aim at analysing mass phenomena. Statistics allow, by means of a condensation process, the drawing of a general affirmation from a series of systematic individual observations. The results of this process often become available as information in figures, on a given phenomenon or a target population. In this way, although statistics are based on individual observations, their objective is not to acquire knowledge of the individuals as such, but to produce synthetic and representative information on the state of a population or of a mass phenomenon. Statistical activities can be distinguished from other activities, in particular by the fact that they are not directed at taking decisions or individual measures, but rather at gathering knowledge of large entities - such as economic cycles, the living conditions of a social group or the structure of a commercial market - as well as at the analysis of phenomena - such as epidemics, opinion trends, fertility or consumer behaviour of households - and therefore arriving at collective judgments or decisions.

3. If we were to consider only the purpose of statistics and the fact that they provide only anonymous and aggregate results, the conclusion could be reached that this sort of activity does not come under the data protection rules. However, this conclusion would be premature and would not take account of the whole procedure of producing and releasing statistical information, which is based, on the whole, on the possibility of collecting or processing microdata, that is to say personal data. There is a risk, therefore, of the data in question being diverted from the purpose for which they were collected and thus being used for purposes relating to individuals. This may be the case, for example, when statistics come into contact with the administration and the police and where one could be tempted to use these data collected for statistical purposes for individual judgements and decisions. Moreover, in spite of their anonymous and aggregate nature, statistical results sometimes can be such that their analysis or recomposition enable the individuals upon whose personal data the results are based to be identified. Finally, the considerable commercial interests sometimes at stake should not be ignored, nor the consequent danger that personal data collected for statistical purposes may be considered as simply commercial goods, to the detriment of the protection of privacy. Taking these considerations into account, statistical activities must be governed by data protection principles.

4. In our democratic states, official statistics are ruled by specific norms of public law which guarantee wide protection of personal data; but statistical activities carried out in the framework of scientific research or in the fields of marketing, private management or opinion polls are not necessarily subject to this protection, or only in some cases. In all these different fields of recourse to information on individuals, the techniques used and the risks to privacy are very similar. Furthermore, there is not always a clear line between private statistics and public statistics. Therefore, in response to the diversity of applications of statistics - and the data protection problems which arise as a result - recommendations should be formulated to apply in a general way to the whole range of public and private statistical activities.

5. Such recommendations must take into account the provisions which, in the national legal order, already ensure the complete or partial protection of personal data which are collected and processed for statistical purposes. Such recommendations must also take account of the specific purpose of statistical methods and of certain research procedures. In this respect, article 9.3 of the convention itself acknowledges that it may be necessary to make exceptions to general data protection laws in certain areas and expressly recognises this fact in the case of statistics. Consequently, it is necessary to decide, according to the purposes and methods of collecting and processing, whether all the principles governing data protection apply, or whether it is necessary to define specific principles.

B. Characteristics of statistics

6. Statistical research is at one and the same time:

- a. a scientific discipline, with its subject matter, its theory, its methods and its aims;
- b. an activity firmly based in society, with its agencies, its regulatory framework, its professionals, its material (data), its occasional or permanent operations and its users.

7. The purpose of statistics is to bring to light the collective features of a specific population and to do so by piecing together individual data. The "population" in question may comprise all the physical or non-physical objects that one wishes to consider: groups of persons, enterprises, commercial transactions, plants, physical particles, therapeutic processes, newspaper articles, events, etc. From the standpoint of this recommendation, however, the observation and analysis of human populations are the main focus.

Statistical purpose

8. While the point of departure for statistical activity is indeed the (so-called "elementary") information relating to the various individuals which make up the population considered, its purpose is merely to describe the characteristics of that population as a whole. This means that individuals are the necessary medium for the background information but that they are not regarded as significant in their own right. In fact, starting with the basic material represented by individual information about many different people, the statistician elaborates results designed to "characterise a collective phenomenon" (Article 1 of the recommendation).

9. Statistical results frequently take the form of numbers, such as totals, averages, percentages or indices; or they may take the form of numerical tables showing a breakdown of the population in terms of various individual characteristics (age, profession, place of residence, level of education, income, miscellaneous behaviour patterns, etc.). Totals or enumerations are sometimes referred to as "aggregates", and this term is also sometimes wrongly applied to

other numerical results. However, statistical results may also be non-numerical: they reply to a question such as "Is a particular characteristic linked to another?" (for example, "Is a particular illness caused by diet?"); or they may take the form of a typology, for example the identification of groups characterised by certain common situations or behaviour patterns, territorial location or collective professional classification, etc. In any event - leaving aside special situations which will be further considered below (see paragraph 35) - such results say nothing specific about any of the persons whose data have been used. In other words, the statistical result separates the information from the person.

10. It is therefore necessary to make a fundamental distinction between the individual use of personal data and their collective use. This is a differentiation of purpose, which is essential from the standpoint of the protection of individuals, their privacy and their rights and freedoms :

a. individual use depends on the personal data per se, provided that they relate to a specific person. Such individual use has a personalised purpose, which is to obtain specific information about a particular person or to take strictly individual decisions to his detriment or his benefit, such as administrative, medical, legal or other decisions ;

b. collective use relies on personal data only via their concentration into a statistical result. This is the distinctive characteristic of statistical purpose: personal data are collected and processed with a view to producing consolidated and anonymous information.

11. The question of purpose requires consideration to be given to the use that will be made of the statistical results. Statistical knowledge is not an end in itself. It usually serves other purposes. The following three categories of purpose are not mutually exclusive:

a. general information purposes: statistical knowledge is made available to a variety of persons without prejudging its benefit or importance in their eyes;

b. purposes of assistance with planning and decision-making: the aim is to give a public or private decision-maker information about his/her environment or field of action which will enable him/her to establish a strategy or optimise a decision. The same decision-maker, or a third party, is also enabled to assess the effectiveness of the action that he/she is taking;

c. scientific purposes: the aim is to provide researchers with information contributing to an understanding of phenomena in fields as varied as epidemiology, psychology, economics, sociology, linguistics, political science, ecology, and so on.

12. Thus, while at the input stage the statistical purpose is totally distinct from administrative, judicial, health, educational, commercial and other purposes, at the output stage it links up with other purposes. It is not therefore exclusive: when reference is made to the statistical purpose in this connection, it is an immediate purpose, underpinning intermediate purposes. It must be emphasised, however, that the statistical information furnished for such intermediate purposes always relates to mass phenomena and cannot, therefore, under any circumstances entail direct or individualised consequences for individuals.

13. Assistance with planning and decision-making is a purpose relevant to officials responsible for taking two types of decision: general decisions (laws, scales of assessment, vaccination campaigns, organisation of transport, design of blueprints, entry into production, etc.) and individual decisions (admission or exclusion, taxation, benefit, remuneration, etc.).

Personal data collected and processed for statistical purposes may be used only for the first type of decision. Such general decisions no doubt have particular implications (favourable or unfavourable) for certain persons; but there must be no connection between the fact of having provided personal information which is then added into the statistical mix, and the fact of being concerned by an individual decision. To forestall any such connection, prescriptive rules are frequently supplemented by an organisational division between the services which handle statistical data and those which deal with files and take individual decisions (principle of the functional separation of duties).

14. Scientific research uses statistics as one of a variety of means of promoting the advance of knowledge. Indeed, scientific knowledge consists in establishing permanent principles, laws of behaviour or patterns of causality which transcend all the individuals to whom they apply. Thus it is aimed at characterising collective phenomena, this being the very definition of statistical results. It could be said, therefore, that research becomes statistical at a certain stage in its development.

a. In the biological and human sciences, much of the research process involves experimentation. In this area, personalised intervention is basic to research, even though statistical analysis may come into play at a later stage. This type of research calls for specific ethical and legal rules which have no place in the field of statistics as defined here.

b. In contrast, the statistics with which the recommendation is concerned operate on the principle that there should be minimum interference with the individuals providing the basic information. On the one hand, the information concerned existed prior to the statistical inquiry, and involves no experimental alteration of the status, conduct or opinion of the data subjects. On the other hand, every effort is made to ensure the least possible disruption to the latter through the actual gathering of the information. The data collected in this statistical framework may be used for research purposes, albeit strictly on condition that there is no personalised feedback to the individuals concerned.

c. No parallels should be drawn between the designation of agencies and the way in which the information is used: the appellations "research institute" and "statistical institute" refer to functions and do nothing to detract from the fact that the research is carried out in part by statistical methods. To the extent that statistical and scientific purposes coincide, the use by research agencies of data gathered by statistical bodies has both scientific and organisational merit. The recommendation therefore envisages this convergence of purposes and co-operation between the specialised institutions; as long as a so-called research institute uses statistical methods to process data, such processing is for statistical purposes.

Data collection and processing

15. There are two main types of data collection for statistical purposes: the "primary" collection of data from individuals and the "secondary" collection of data from public or private agencies which hold information concerning individuals.

16. Preparation of collection: Collection is usually preceded by a phase in which the persons to be surveyed are identified: scope of the collection (set of persons in principle concerned by the planned statistical survey) and decision as to whether to interview all the persons concerned or only a sample.

17. Primary collection: Data may be collected either directly from the persons concerned ("direct" primary collection), or from their entourage (that is, people who are close to or

associated with them) who reply in their place ("indirect" primary collection), or again by means of "direct observation" (for example, counting of visitors to a museum or of vehicles on a road, remote observation by means of aerial or satellite photographs, etc.). Indirect collection is a common practice in household surveys: the person present replies on behalf of the other members of the household, which is convenient both for the latter and for the interviewer. Another example is that of health surveys, where information is gathered from the doctor or nurse.

18. Primary collection techniques: Both direct and indirect collection may involve different technical procedures:

a. Under the traditional procedure, the person surveyed himself fills out questionnaires or forms, or they are filled out by an interviewer on the basis of his replies.

b. Under a recently introduced procedure known as "CAPI", the interviewer records the replies, not on a form, but on a portable microcomputer. This facilitates a variety of immediate checks as well as the coding of replies.

c. Another technique is the telephone survey ("CATI"), in which the interviewer is assisted by a computer which records the replies. The computer may even be used to manage the directory of persons to be interviewed and to make calls. This technique has the advantage that the interviewer does not know the identity of the interviewees.

19. Checking and processing of data: The collection of data is systematically followed by a checking phase, and then by the actual statistical processing, which leads to the production of statistical results (scheme A). However, this linear approach is not always appropriate. On the one hand, the first statistical results sometimes display anomalies which lead to renewed checks at the input stage. Secondly, the collection process is sometimes staggered so that some results can begin to be produced before it is completed (scheme B). Finally, some surveys are one-off exercises or, if they are repeated, they are carried out independently, with no effort to use the same persons twice or to collate their two sets of answers; in other cases, however, the (so-called "multi-stage") survey has to be repeated among the same people (scheme C).

20. Identification data used as ancillary data: A distinction is made between two sorts of data in this technical process:

a. the information that will be used to produce the statistical results, for which the identity of the data subjects is irrelevant;

b. the information which identifies individuals for collection and checking requirements, but which is of no relevance for the calculation of results. Identification data may be collected either to facilitate the implementation of repeat surveys, or to supervise the interviewers' work, or again to facilitate the checking of unreliable data with the persons surveyed. Moreover, data from two or more sources (surveys or administrative files) are used for certain statistical operations; a common identifier is then needed to match the data relating to the same individual.

These two categories are not completely separate: for example, date of birth or place of residence may simultaneously form part of the identifier and be used to produce certain results. Otherwise identification data are helpful only as "ancillary data", for the acquisition of genuinely informative data which alone will be used to produce results.

21. Separation, conservation and destruction of identification data: There comes a point where the identification data may be separated or erased. However, bearing in mind the fact that the process is not a linear one, as was said above, this stage is not identical in all statistical operations. Moreover, there are cases where the identification data gathered from a census or from certain surveys are kept separate, with a view to initiating a subsequent survey.

22. Secondary collection: Statistics are not confined to the processing of data collected by means of surveys. Data may also have already been gathered through the activity of a public or private agency which thus has a set of documents or a ready-made file capable of being used to produce statistics. In particular, a full or partial copy of the documents or file in question may be supplied to a body of statistical experts. This method of collecting data, which is known as "secondary collection", clearly has a twofold advantage: it saves the time of the data subjects, who do not have to reply twice to the same questions; and it is more economical for the body of statistical experts. In addition, the latter reaps the benefits of the checks carried out by the first collector, and this increases the reliability of the data.

23. Use of registers for internal statistical purposes in firms: Secondary collection is an extremely widespread practice, especially in the internal administration of companies and other enterprises (that is, statistics for the management of staff, wages, productivity, etc.), as well as in connection with their commercial relations (that is, statistics on the "purchasing habits" of customers). In this area, the techniques are frequently defined and implemented by persons whose grasp of the scientific and ethical aspects of statistics is very rudimentary. As a result, tasks relating to individual checks on persons may sometimes be confused with "statistical" tasks. However, sound management and ethical professional conduct usually go hand in hand with measures for the protection of personal data, particularly as regards the use for statistical purposes of the internal registers of firms (staff and customer registers). In this connection, it is already common practice among large and medium-sized firms to replace identification data at the outset by a code or internal identifier, the meaning of which is known only to the controller (staff number, customer number).

24. Personal identification number (PIN): In some countries, the system of official statistics is based on a set of administrative files or directories rather than on a set of surveys. Statistical information is produced from data contained in such directories, often by combining the content of several of them. This requires the data subjects to be registered by means of a common identifier. This identifier or personal identification number (PIN) does not constitute a direct personal identifier in cases where:

- a. the structure of the number prevents direct access to the identity of the data subjects;
- b. appropriate technical and organisational measures prevent unauthorised physical or logical access as well as the decoding of such numbers.

In the Nordic countries in particular, this practice has developed to the point where the identification of individuals by means of the PIN is possible only by expending an unreasonable amount of time and resources. Moreover, access to the PIN in the field of statistics is strictly limited to a very small number of persons, whose work is subject to stringent monitoring.

Communication of data to third parties

25. In accordance with the principle of purpose (see paragraphs 8

to 14 and 68), the ethical codes of statisticians require that personal data collected for statistical purposes should never be communicated to third parties for non-statistical purposes. On the other hand, such data may, under certain conditions (see paragraph 100 ff.), be communicated to third parties in order to be processed for other statistical purposes, as the purposes in that case are compatible. In this connection, reference has been made to the communication of data for research purposes, but the communication of data between the statistical units of government departments could also be mentioned; data collected by an agency for the production of education statistics may be communicated to another agency to be utilised or to be matched with other data, but always only for statistical purposes.

26. Encryption and scrambling: Data held by a statistical body are protected by rules of access and by physical safeguards. When they are communicated for the purpose of other statistical uses, there is a choice between imposing the same obligations on the beneficiary of the transfer and disseminating only non-identifiable data: that is to say, either anonymous elementary data, not capable of being re-identified, or encrypted or scrambled data.

a. Encryption consists in transforming data which would permit identification by means of a key that is kept secret.

b. Scrambling involves the random alteration of data to make them individually meaningless while preserving the overall structures of the population considered by means of compensatory adjustments (for example it is possible to switch some of the data concerning two persons). It is then advisable to make it known that the data have been modified, so as to deter any personalised use and thereby contribute to the security of the data.

Guarantees of protection

27. The question of identification is obviously central to the protection of privacy. As long as data are used only to produce statistical results

and the results themselves are impersonal, there is no threat of infringement of privacy. In the circumstances, there should be no obstacle to statistical activity. However, this presupposes the effectiveness of the two above-mentioned criteria: exclusively statistical use and impersonal results. It is in this connection, therefore, that the recommendation lays down legal guarantees and stipulates technical precautions, including the following:

a. data accompanied by identification data should be protected by guarantees against any non-statistical use;

b. data accompanied by an internal identifier (specific to the controller) are considered anonymous only if the meaning of the identifier is not accessible to third parties;

c. individual data that appear to be anonymous (unaccompanied by any identification data) may nevertheless in some cases be indirectly identifiable. This means that, for at least some if not all data subjects, the piecing together of informative data (such as age, sex, occupation, residence, family status, etc.) makes it possible in fact to discover the identity of the person concerned. Where this is a possibility, the data may not be considered to be genuinely anonymous and must therefore be protected ;

d. where statistical results only aggregate data for a small number of individuals, they may also occasionally be used, by means of analyses and cross-checks, to trace - at least approximately - the data specific to some of the individuals concerned. This risk exists in particular where the statistical results relate to very small geographical units (districts, small

municipalities, etc.), where a combination of variables (age, sex, civil status, occupation, etc.) permits identification of the data subjects or, more simply, where a person or a small group of persons occupies a dominant position in a given population. These contingencies should be heeded, therefore, when publishing such results. However, it will be noted that mere knowledge of the persons covered by a result does not in itself constitute a lifting of anonymity, as long as the individual data are not traced.

28. Anonymisation, which, in the statistical sector, is largely achieved by the separation of identification data, is the principal means of ensuring protection. Clearly, however, it is not always either necessary or adequate:

a. it is not always necessary: there may be a case for not making data anonymous at the outset, or even on a longer-term basis, particularly in the context of repetitive surveys. But this is offset by extra precautions to prevent unlawful access to the data ;

b. it is not always adequate: there is sometimes a residual risk that theoretically anonymous data may be re-identified.

However, although this is not strictly a zero risk, it may be considered to be non-existent in practice when re-identification would require excessively complicated, lengthy and costly operations. No safe is completely burglar-proof; precautions are required to make safe-breaking highly unlikely, though not strictly impossible. This requirement may vary according to the nature of the data, depending on how sensitive they are.

29. Risk of abuse of purpose. In this connection, it should be pointed out that it is not statistical processing as such which presents a threat to data subjects, but merely the fact that the data on which it is based could be diverted to non-statistical purposes. As in the case of other uses in the public interest, therefore, the question is not one of striking a balance between statistical interests and the interests of the data subjects. Indeed, there are no grounds on which statistical interests could justify an infringement of the interests of the data subjects, since the latter are not in any case infringed by statistics, whether or not such statistics are significant. What is at issue is the risk arising from the holding of identifiable data. For that reason, from the standpoint of both professional ethics and the principles contained in the recommendation, emphasis is placed on the prevention of that risk.

C. Development of data protection regulations

30. Attempts to lay down safeguards and standards to protect personal data in the statistics field are nothing new. The very first framework regulations were drawn up by practitioners themselves. As far back as 1947 the European Society for Opinion and Marketing Research drew up the first code for its members. Since then various professional and scientific associations have drawn up codes or charters of principles primarily aimed at protecting data and laying down requirements for access to personal data.

31. Within the Council of Europe the Committee of Experts on Data Protection identified a need to look closely at the use of personal data in research and statistics in 1979. A year later it instructed a working party to draw up a draft Recommendation on the protection of personal data used for scientific research and statistics. The recommendation was adopted by the Committee of Ministers on 23 September 1983 as Recommendation No. R (83) 10.

32. Since then, the Council of Europe's endeavours to promote efficient protection of personal data have been largely taken over by the scientific and official circles dealing with statistics. These communities have taken important initiatives, at national and international level, aimed

in particular at elaborating deontology codes, at reinforcing national legislations on statistics and at promoting increased international co-operation in respect of data protection techniques and methods. Such initiatives bear witness to a significant evolution, which leads the way to catching up with the spirit of the convention.

33. On 20 August 1985, the General Assembly of the International Statistical Institute adopted the Declaration on professional ethics ^{footnote 8}, whose aim is "to enable the statistician's individual ethical judgements and decisions to be informed by shared values and experience, rather than to be imposed by the profession". This declaration underlines in particular the statistician's obligations in respect of the subjects of an enquiry. In this framework, it recalls that the statistician has a duty to protect as far as possible the interests of the individual and it formulates recommendations related in particular to the consent of the individuals who are to take part in the enquiries, the duty to guarantee confidentiality of individual data, as well as the obligation "to take appropriate measures to prevent statistical data being published or otherwise released in a form that would allow any subject's identity to be disclosed or inferred."

34. The orientations of this Declaration on professional ethics have been taken over and completed by more precise provisions in "codes of conduct", which have been adopted by national statistical institutes or national associations of statisticians in certain Council of Europe member states.

35. In the specific field of official statistics, many Council of Europe member states have over the last ten years enacted new legal bases which lead to strengthening data protection. Also in most states in central and eastern Europe, the process of democratisation and transition to the market economy had brought about a fundamental review of national legislations, in particular with a view to ensuring the protection of personal data collected and processed for public statistical purposes. Finally, at present most new independent states have adopted, or are about to adopt, analogous provisions in their national legislations on statistics.

36. Personal data collected for the production of official statistics are moreover subject to increased protection because of the introduction of legal norms, established by the European Union, which apply to some Council of Europe member states. In 1990 the Council of the European Communities adopted a regulation aimed at guaranteeing confidentiality of statistical data communicated by national statistics institutes and services to the Statistical Office of the European Communities (Eurostat) ^{footnote 9}. This regulation provides for a committee on statistical confidentiality, defines data protection norms and ensures that all regulatory, technical and organisational measures must be taken to guarantee confidentiality of data transmitted to Eurostat. In addition, on 17 February 1997, the Council of the European Communities adopted Regulation No. 322/97 on Community statistics. These legal acts constitute a set of specific rules governing the confidentiality of data collected for Community statistics. Personal data protection in the field of statistics will be further reinforced when the member states of the European Union adapt their national legislation and practice to the provisions of Directive 95/46/EC of 24 October 1995 ^{footnote 10}. This must be done by 24 October 1998 at the latest.

37. Following a proposal by the Conference of European Statisticians, the United Nations Economic Commission for Europe adopted a resolution on the fundamental principles of official statistics on 15 April 1992 ^{footnote 11}. This resolution, which refers to the convention, ensures, among other rules, that "individual data collected by statistical agencies for statistical

compilation, whether they refer to national or legal persons, are to be strictly confidential and used exclusively for statistical purposes". This requirement, together with the other principles in the resolution, have a universal scope and as such were adopted at world level by the Statistics Commission of the United Nations at an extraordinary session on 15 April 1994.

38. This two-fold progress, on both the legislative and professional fronts, is testimony of the concern that is shared by statisticians and data protection authorities. Parliaments and the statistics profession are pursuing the same objective, which is to ensure respect for the individual and his/her privacy. The rules and practices which each is developing are equally necessary and complement each other. Statistical activity cannot be left to develop without an appropriate framework of legal guarantees. Equally, there is better protection and better prevention of abuses if professionals subscribe to the necessary rules and regard them as integral to sound practice. Thus self-regulation and development of effective methods are an indispensable adjunct to legal provisions.

D. History of the recommendation

39. In 1980, the drafters of Recommendation No. R (83) 10 were instructed to draw up standards for the processing of personal data for the purposes of scientific research and statistics. The drafters, however, gave greater importance to the consideration of certain fundamental problems of data protection in the field of scientific research. The recommendation therefore only deals incidentally with the problems related to the collection and processing of data for statistical purposes. Moreover, the evolution of professional ethics, of national legislations and of the definition of fundamental principles of statistical activities is certainly in harmony with the work and the initiatives of the Council of Europe; however, this is not yet sufficient to fully guarantee strict protection of personal data in all fields of statistical activity.

40. In November 1990, the Project Group on Data Protection felt therefore that there was a need to define more precise principles on the protection of personal data collected for statistical purposes. It was agreed that Recommendation No. R (83) 10 should be examined as to whether or not it dealt adequately with the data protection problems which had been identified in the whole field of statistical operations.

41. A working party was set up for this purpose. It was chaired by

Dr Jean-Philippe Walter (Switzerland) and consisted of experts from Austria, Denmark, France, Greece, Hungary, Portugal and Switzerland. It met seven times between December 1991 and January 1995. Experts from Germany, Luxembourg, the Czech Republic, Slovenia, Sweden and the United Kingdom, as well as observers from Croatia, the United States of America, "the former Yugoslav Republic of Macedonia", the Russian Federation, the Commission of the European Communities (Eurostat) and the International Chamber of Commerce attended the meetings.

42. On 22 and 23 March 1993, an exchange of views was held in Strasbourg between the Committee on Statistical Confidentiality of the European Union and the Council of Europe's working party. The European Union contributed to the preparation of the recommendation from its field of activity.

43. The draft recommendation on the protection of personal data collected and processed for statistical purposes was approved by the Project Group on Data Protection at its 32nd meeting (19-22 November 1996), and later by the European Committee on Legal Co-operation.

44. On 30 September 1997, the Committee of Ministers adopted Recommendation No. R 97 (18) on the protection of personal data collected and processed for statistical purposes, and authorised publication of this explanatory memorandum.

Commentary on the recommendation

Preamble

45. The preamble sets out the considerations which prompted the Committee of Ministers to issue the recommendation to governments of member states :

a. among those considerations the Committee of Ministers points out the importance, both for the authorities and public services in performing their functions and for private-sector bodies and undertakings in carrying on their activities, of having reliable statistics available to them. To analyse and understand modern day society and work out policies and strategies requires that such statistics be producible, which in turn calls for basic information necessitating the collection and processing of data, including personal data, with the help of ever more powerful data-processing technology, but subject to the requirement that the data subject's anonymity be guaranteeable ;

b. the Committee of Ministers acknowledges that statisticians themselves were among the first to take privacy into account in the production of statistics based on personal data and that statisticians' concern has been reflected in codes of conduct, basic principles and legal standards. As far as Council of Europe member states are concerned, the authoritative instrument in data protection is the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS 108), which was opened for signature in January 1981 and which developed out of the European Convention on Human Rights, which all the member states have signed. Article 9 of the 1981 convention in fact gives statistical work the benefit of an exception whereby data subjects' exercise of certain rights is restricted ;

d. the Committee of Ministers states the purpose of the recommendation to be to establish appropriate procedures reconciling the interests of the various parties concerned in the production of statistics: on the one hand, the producers' interest in being able to produce representative statistics and, on the other, data subjects' interest in being protected against unnecessary intrusion into their private lives ;

e. lastly, the Committee of Ministers is aware that, since its adoption fourteen years ago, Recommendation No. R (83) 10 on the protection

of personal data used for scientific research and statistics with regard to the collection and automatic processing of personal data for statistical purposes has ceased to reflect progress both in statistics and data processing.

Operative provisions of the recommendation

46. The Committee of Ministers firstly recommends that governments of member states take steps to ensure that the principles in the appendix are reflected in their law and practice. This recommendation is flexibly worded because it is also made to member states which are not yet parties to the convention and which consequently are not under any obligation to introduce the necessary measures to give effect, in domestic law, to the basic principles of data protection.

47. Secondly, governments are encouraged to ensure wide circulation of the appendix to the recommendation among all persons professionally engaged in collecting and processing personal data for statistical purposes or in data protection.

48. Thirdly, governments are asked to encourage the introduction, in the statistics sector, of codes of ethics modelled on the principles in the present recommendation.

49. Lastly, as regards collection and processing of personal data for statistical purposes, the Committee of Ministers withdraws Recommendation No. R (83) 10 on the protection of personal data used for scientific research and statistics. However, Recommendation No. R (83) 10 continues to apply to the collection and processing of personal data for other scientific research purposes.

Appendix to the recommendation

1. Definitions

50. The recommendation is intended to apply not only to governmental institutions of member states but also to persons, public authorities and bodies which collect and process personal data for statistical purposes and to data protection authorities. The drafters therefore considered it necessary to clarify the terms "domestic law", "law", "provided for by domestic law" and "permitted by domestic law", which are used in a number of principles in the recommendation:

a. "domestic law" denotes, according to the legal and constitutional system of the particular country, all substantive rules, whether of statute law or case-law, which meet the qualitative requirements laid down by the European Court of Human Rights, which include those of accessibility and foreseeability ^{footnote 12}. It covers all measures applying to an unlimited number of cases and an indeterminate number of persons. It encompasses rules that place obligations or confer rights on persons (whether natural or legal) or which govern the organisation, powers and responsibilities of public authorities or lay down procedure. In particular it includes member states' constitutions and all written acts of legislative authorities (laws in the formal sense). It also covers not only all regulatory measures (decrees, regulations, orders, administrative directives) based on such laws but also international conventions applicable in domestic law, including Community law in the case of European Union countries. It further includes all other acts whether of public or private law (including law of contract) together with court decisions in the common law countries ^{footnote 13} or which interpret a written law ^{footnote 14}. In addition it includes any act of a professional body under powers delegated by parliament and in accordance with its independent rule-making powers ^{footnote 15}. On the other hand, it does not include voluntary regulations (such as codes of professional ethics or good conduct), which are not binding ^{footnote 16} and to which the aforementioned qualitative requirements do not apply. Under the case-law of the ECHR institutions the terms "domestic law" and "law" are identical ^{footnote 17}. In some cases, however, the qualitative requirements, in particular those of accessibility and foreseeability must be even stricter and the degree of precision required is even greater ^{footnote 18}. The drafters of the recommendation thus intentionally used the word "law" (in the sense of a statute) where the qualitative requirements of the legal basis for collecting and processing personal data for statistical purposes needed to be stricter. In particular, in continental countries, those substantive requirements will only be met by written laws ;

b "provided for by domestic law" means that domestic law, as defined above, contains rules which expressly lay down provisions and safeguards required by, or exceptions to, the principles in the recommendation. To be "provided for by law", an interference with fundamental freedoms or rights and in particular the individual's privacy (for example, collection and processing of personal data for statistical purposes) must therefore have a legal

basis in domestic law and be performed in accordance with domestic law. In addition, the legal basis must meet certain qualitative requirements, including those of accessibility and foreseeability (see section a above) ^{footnote 19} ;

c. the words "permitted by domestic law" refer to any interference in accordance with the principles of the recommendation which are not explicitly forbidden by domestic law.

51. Principle 1 defines certain concepts central to the recommendation. During the drafting of the recommendation, it was suggested that it lay with the statistical profession to define the different terms used in this science. However, in the absence of universally accepted definitions, and to facilitate understanding of the recommendation, certain key terms should be clearly defined "for the purposes of this recommendation", without prejudice to any definitions formulated at a later stage by the statistical community.

52. Personal data. The definition is consistent with the one contained in the convention, as interpreted in the explanatory report attached thereto. It has already been used in a number of sectoral recommendations adopted by the Committee of Ministers in the field of data protection.

a. Person: the definition relates to a natural person. However, with regard to states which extend the application of the convention (under its Article 3, paragraph 2.b or of the recommendation (under its Prin-

ciple 2.4) to legal persons, the definition is understood to cover the latter also. Moreover, information concerning a natural person may relate to a unit with which he/she is connected, for example, a single-person enterprise, while at the same time constituting personal data (relating to the proprietor in the above example). In such cases, the information comes within the scope of the recommendation (see paragraph 64).

b. Households: the drafters agreed that, for the purposes of the recommendation, data relating to households should by analogy be considered as personal data.

c. Identifiable person: a person is said to be "identifiable" when he/she is not identified but could be by processing or cross-checking of his/her personal data (see paragraph 27.c).

d. Anonymous data: data are considered anonymous if identification requires an unreasonable amount of manpower, that is excessively complicated, lengthy and costly operations (see paragraph 28). Conditions for anonymity are relative, especially in relation to the technical means available for identifying data and taking away their anonymity. In this way, in view of the rapid progress in technological and methodological developments, the time and manpower required to identify a person, which would today be considered "unreasonable", might no longer be so in the future. However, the present wording is sufficiently flexible to cover such developments ^{footnote 20}.

53. Identification data. Such data in most cases consist of the name, first name and address, and serve to identify an individual directly. Identification data may be distinguished from the other personal data collected for statistical purposes (see paragraph 20).

a. In some states, identification may also be made by means of an officially allocated identification number used in the public services (personal identification number, see paragraph 24). To the extent that such numbers would not permit direct identification of the data subject, they are not identification data within the meaning of this recommendation ;

b. The drafters avoided using the term "anonymisation" in the text of the recommendation. They considered it worthwhile, however, for the explanatory memorandum to provide clarifications from the standpoint of the application of the recommendation. Anonymisation consists in erasing identification data so that individual data may no longer be attributed to the various data subjects by name (see paragraphs 21 and 28). The withdrawal of identification data does not in some cases provide total protection against re-identification: the risk of disclosure cannot always be a strictly zero risk. Leaving aside any attempt to define "acceptable risk", the effectiveness of anonymisation is somewhat relative (nature of the information at issue, conditions of security, date of anonymisation, stage of processing, etc.).

54. Sensitive data. The definition draws on the list set out in Article 6 of the convention. However, in accordance with Article 11, other categories (such as data concerning trade union membership or income) may be defined as sensitive in domestic law.

55. Processing. The definition relates to all automated operations necessary for the production and dissemination of statistics, apart from collection. Principle 2 specifies the range of such operations within the scope of the recommendation.

56. Communication. This term covers the process - whether active or passive - of making data accessible. It differs from "dissemination" in so far as communication relates in particular to the transmission of data between individuals or public or private agencies.

57. Statistical purposes. The definition applies to any operation in the collection and processing of data, that is to say their use not only when producing statistical results but also for the preparation and checking of surveys.

a. Article 5 of the convention stipulates that data must not be used in a way incompatible with the purpose of their collection. The drafters of the recommendation attached great importance to a precise and unambiguous definition of "statistical purposes". Individual data collected for statistical purposes, that is to say in order to visualise mass phenomena, are only raw material intended ultimately to lose their individuality in a statistical result. The statistical purpose is geared exclusively to this type of result (see paragraphs 8 to 14).

b. Statistical purpose as defined above excludes the use of the information obtained for decisions or measures in respect of a specific individual, particularly decisions of an administrative, judicial, fiscal or other such nature. Statistical purpose is the subject of Principle 4.1 (see paragraph 68).

c. Statistical purposes may include scientific research where the research in question consists of processing operations and analysis of personal data in order to characterise a collective phenomenon in a given population. However, statistical purposes are totally incompatible with any scientific research relating to individuals or capable of leading to decisions or measures in respect of such individuals (see paragraphs 14 and 102.c.)

58. Statistical results. These are not personal data, as they are not linked to an identified or identifiable natural person (see above, paragraph 9) :

a. Some statistical results may nevertheless permit the indirect identification of data subjects (see paragraphs 27.d and 109).

b. Intermediate results obtained during the process of producing statistics are not considered to be "statistical results" within the meaning of the recommendation as long as they are not at the disposal of a third party (see paragraph 35).

c. In the text of the recommendation, the drafters avoided the use of the term "aggregate", which is a form of statistical result (see paragraph 9), while expressing the wish that it be mentioned in the explanatory memorandum. In relation to a group of individuals considered without differentiation, it is therefore intrinsically anonymous.

59. Controller. The definition is based on the definition of "controller of the file" in Article 2 of the convention. The drafters nevertheless considered that, in the statistical field, the term "file" has been overtaken by new techniques for processing personal data. They preferred to make the definition of "controller" applicable to any person or any agency that determines the objectives, resources and organisational arrangements for the collection and processing of personal data for statistical purposes. Where the objectives, resources and organisational arrangements are defined by an act of parliament - as occurs in some member states - the controller is the person or agency responsible for organising the production of statistics.

2. Scope

60. The drafters discussed the status of statistical activities referred to in the recommendation, and more particularly the question whether a distinction should be made between "official" or "public" statistics and private statistics :

a. some experts considered that the recommendation should make this distinction since official statistics, unlike private statistics, are authorised by law, organised and monitored within an institutional framework and governed by public law standards guaranteeing extensive data protection. They also emphasised that the methodological bases of official statistics are made public and the results are accessible to all citizens. Finally, they referred to the regulatory framework for the "fundamental principles of official statistics" adopted by the United Nations, as well as the existence of common definitions, classifications and rules applicable to official statistics ;

b. on the other hand, other experts observed that the lines of demarcation between "private" and "public" statistics are not always clear. Statistical studies are sometimes commissioned from research institutes by the administration, and official statistics are sometimes produced by private agencies. The same experts also pointed out that the distinction between the two types of statistics may vary from one country to another in the light of national legislation, and that it would therefore entail inequality in the application of the principles contained in the recommendation in the different member states. Furthermore, they noted that neither the convention nor the relevant directive of the European Union makes a distinction between the public and private sectors. Lastly, they acknowledged that the introduction of data protection standards is necessary, particularly in view of the development of statistical activities in the private sector. However, such standards should not pose a major problem for the public sector which is already, in most member states, subject to legislation based on the principles of the convention.

c. the drafters finally agreed that the recommendation should apply to all statistical activity, whether carried out by official bodies or by firms, individuals, or private institutions ;

61. Principle 2.1 limits the scope of the recommendation to "the collection and automated processing of personal data for statistical purposes", as well as to "statistical results, to the extent that they permit identification of data subjects". This definition of scope covers not only strictly statistical activities but also activities based on statistical processes and including operations that involve collection and processing of personal data, such as opinion surveys or market research.

- a. In the statistical sector, data are for the most part collected manually on the basis of questionnaires. Accordingly, the recommendation applies to any instance of the collection of personal data for statistical purposes, whether carried out by automated means or by manual means with a view to subsequent automated processing.
- b. The term processing, as defined in Article 1, covers a wide variety of operations. The recommendation therefore applies to all these forms of automated processing of personal data.
- c. Although statistical results are not personal data, the drafters considered that it was necessary to ensure the protection of privacy with regard to certain results which would permit the relatively straightforward identification of individual data and enable a link to be established between such data and the identity of the data subjects (see paragraph 27,d). The recommendation therefore applies to the communication and publication of such results, which are the subject of Principle 14 (see paragraph 109).

62. Principle 2.2 encourages member states to "extend the application of the recommendation to the non-automated processing of personal data for statistical purposes". Collection is immediately followed by processing which is frequently carried out in part automatically and in part manually. In some cases (for example in censuses), personal data are stored or conserved manually. On the other hand, it is noted that, in some member states, data protection legislation does not apply to manual files. The recommendation therefore proposes an optional extension of scope to ensure full protection for personal data processed manually for statistical purposes. This principle is based on Article 3, paragraph 2.c of the convention, which stipulates that states may extend its scope to personal data files which are not processed automatically.

63. Principle 2.3 states that in no event should data be processed manually in order to avoid the principles in the recommendation. This provision is therefore designed to prevent evasion of the principles, particularly where national legislation on data protection does not apply to manual processing.

64. Principle 2.4 authorises an optional extension of the scope of the recommendation to the protection of data relating to entities and persons other than natural persons, when they are collected and processed for statistical purposes. The drafters wished, on the one hand, to ensure effective protection of the privacy of individuals and, on the other, to take account of the possibility, recognised in the convention, that states may extend the scope of the recommendation to include legal persons. The wording of this principle takes account of the considerations expressed by different experts:

- a. a principal purpose of data protection is to ensure respect of several human rights and fundamental freedoms and in particular of Article 8 of the European convention on Human Rights: the right to the integrity of private life. Some experts noted that, in the spirit of the convention, this is a specifically "human" right, as, in its Article 1, it attributes it exclusively to natural persons ;
- b. in the event of a too narrow interpretation of the scope of the recommendation, these experts emphasised that the protection of privacy would be ineffective if it were confined solely to the protection of data relating to individuals. In some cases, data on legal persons include information permitting direct identification of individuals. Furthermore, data relating to one-man firms should not be considered only as data on "legal persons", as they may constitute personal data in the "physical" sense, in which case they should be protected ;

c. some experts, moreover, remarked that there is no objective definition of "personal" data relating to human groups. In their opinion, if data on groups of persons with legal personality (such as couples or families) are easily considered personal data, this is not necessarily the case where the group data relate to broader aggregates (such as the occupants of a block of houses, a group of tenants or informal sports or cultural groups) ;

d principle 2.4 does away with any possibility of restrictive interpretation of the scope in so far as it lays down that the principles in the recommendation could apply to "the collection and processing of

data relating to groups of persons, associations, foundations, companies, corporations and any other bodies consisting directly or indirectly of individuals, whether or not such bodies possess legal personality". In this respect, the drafters recalled at this point that, in accordance with Article 3 of the convention, states have the option of extending the protection of legal personality to data subjects such as firms, which are not natural persons.

3. Respect for privacy

65. Principle 3.1 reiterates the purpose of the convention as set out in its Article 1: to secure for every individual respect for his or her rights and fundamental freedoms, and in particular, the right to privacy. This requirement is of vital importance at all stages in the process of producing statistics.

a. It is important first of all for the collection of personal data, particularly when the latter are still accompanied by identification data.

b. Respect for privacy must then be guaranteed in the context of the processing of personal data and, in particular, when such data are kept for future use (which is the subject of Principle 11), when there is any communication of such data to third parties (which is the subject of Principle 12) and when statistical results are disseminated or published (which is the subject of Principle 14).

c. Lastly, Principle 3.1 takes account of the fact that, in order to obtain reliable statistics, the statistician is sometimes required to correct and supplement the data collected. For example, he sometimes finds it necessary either to modify individual data by means of "editing" procedures or to correct replies which are incompatible with other replies from the same individual, or again to insert an assumed reply where one is missing. The statistician may also modify data to complicate reidentification of the data subject, and this is then a technical data protection measure. In fact, such modifications of data are not likely to infringe the fundamental rights or freedoms of the data subjects. In this connection, the drafters of the recommendation noted that the requirement concerning the accuracy of personal data, in Article 5 of the convention, is designed to prevent decisions or individual measures being taken with regard to data subjects on the basis of inaccurate or false data. However, this requirement is inapplicable when the processing is carried out exclusively for statistical purposes.

66. Principle 3.2 provides that persons involved in the production of statistics who have access to personal data must be subjected to professional secrecy, either by means of domestic legal rules or through codes of ethics approved by statisticians under the terms of voluntary regulations concerning good professional practice. Such codes of ethics (which are the subject of Principle 16) are based on corporate recognition of the principle of respect for privacy, and

on that basis they constitute a useful supplement to binding measures under domestic legislation.

67. Principle 3.3 provides for the anonymity of data as an immediate means of guarding against the risk of infringement of personal privacy. To achieve anonymity, identification data must be separated from personal data.

a. As was explained above (see paragraph 28) this does not provide an absolute guarantee of protection, as there may be a residual risk of re-identification of "anonymous" personal data, especially through more or less complex cross-checking and analysis. However, the separation of personal data from identification data is an elementary measure which serves at the outset to prevent the direct identification of data subjects. For that reason, Principle 3.3 requires such separation to be implemented as soon as the identification data are no longer necessary for processing operations (see paragraphs 21 and 53.b). Specific rules to that effect are elaborated in Principles 8.1 and 10.1 (see paragraphs 90 and 97).

b. The drafters considered that the principle of anonymity does not rule out the possibility of conserving personal data in an identifiable form for use in longitudinal studies, especially those based on repetitive surveys phased over a period of time.

4. General conditions for collection and processing for statistical purposes

Purpose

68. Principle 4.1 is based on a requirement in Article 5 of the convention: data collected for a specific purpose shall not be used for other incompatible purposes. Accordingly, where personal data are collected and processed for statistical purposes, the data in question must be used solely and exclusively for those purposes. This is a fundamental rule which entails three consequences:

a. where personal data have been collected and processed for statistical purposes, they must not in any event be used to take decisions or measures in respect of the data subjects. The decisions concerned include those of an administrative, judicial, fiscal or financial nature, as well as decisions by public authorities and measures affecting individuals in their workplace and in their community or corporate activities ;

b. data collected and processed for statistical purposes must not be used to supplement or correct data files used for non-statistical purposes. If the initial statistical purpose were to be distorted in this way, there would be no guarantee that the data concerned could not be used to take decisions or measures in respect of the data subjects. The same is not true of the use of data collected for statistical purposes in order to correct or update files established for other statistical purposes. In the latter case, the purposes are obviously compatible within the meaning of the convention, and this recommendation contains specific provisions on the subject in Principles 12.2 and 12.3 ;

c. if the same act of collection is meant to serve several purposes, statistical on the one hand and non-statistical on the other, the specific conditions of lawfulness regarding the non-statistical purposes should be determined by domestic law. In cases of multi-purpose collection, Principle 4.1 provides for separation of the processing operations in the light of their respective purposes, in order to ensure that each controller has access only to the data needed for the purposes (statistical on the one hand and non-statistical on the other) for which they are collected. This functional separation may be effected by means of technical (physical or logical) and organisational measures.

69. Principle 4.2 deals specifically with processing for statistical purposes of personal data originally collected for non-statistical purposes (secondary collection): the drafters considered it worth making clear that processing for statistical purposes was not incompatible with the non-statistical purpose or purposes for which the data had been collected provided there were appropriate guarantees to prevent the data's being used as a basis for decisions or measures affecting the data subject.¹ Principles 4.4 and 5.4 lay down the lawfulness and information requirements applying to secondary collection and processing. It should be repeated, however, that the lawfulness of secondary collection only applies in one direction: although processing for statistical purposes of data collected for non-statistical purposes is perfectly permissible, the converse (processing for non-statistical purposes of data collected for statistical purposes) is totally non-permissible.

Lawfulness

70. Article 5 of the convention requires that personal data be obtained fairly and lawfully. From the lawfulness standpoint collecting and processing for statistical purposes must have a legal basis or have either the data subject's consent or an appropriate justification. These three general conditions of lawfulness are alternatives in that they match the different degrees of coercion or intrusion to which the different types of collection may subject the data subject.

71. Principle 4.3 deals with the applicability of the general lawfulness requirements to different kinds of data collection and processing for statistical purposes.

a. Sub-paragraph a deals specifically with cases or situations where personal data are collected and processed for statistical purposes and this is provided for by law. This is intended mainly, but not exclusively, to deal with collection and processing of data for statistical purposes

where the collection and processing are performed in the public interest or by an official authority and where there is an obligation to provide data. The obligation may be a direct obligation on the data subject for the purposes of a survey (for example, a population census) or it may arise indirectly in secondary collection for statistical purposes (for example, where a national statistics institute collects data on income from the tax authorities). In such cases the compulsoriness of the collection immediately removes the requirement to obtain the data subject's consent; but in such cases the data collection and processing must be expressly provided for by law. Principle 4.5 lays down the condition which must be met for compulsory collection to be lawful when not performed on official authority in the public interest. Lastly, Principle 6.4 states that anyone refusing to respond may only have penalties imposed on them that are expressly provided for in domestic law.

b. Sub-paragraph b is specifically concerned with voluntary surveys and secondary collection in which data subjects may oppose processing of their data. In such cases collection and processing must be "permitted" by law - that is, there must be nothing in the law which prohibits them. In addition, Principle 4.3 lays down further lawfulness requirements which apply in three cases in particular:

- where the data subject or his/her legal representatives are questioned themselves (direct primary collection) one or other must express consent to collection and processing of the personal data. The characteristics of and requirements for such consent are dealt with in Principle 6;

- where, in secondary collection, the data subject is allowed to oppose processing of his/her data, such processing is only lawful on three specific conditions: firstly that the data subject

has been properly informed (see paragraph 81); secondly that the data subject has not opposed processing of his/her personal data; thirdly that no sensitive data will be processed;

- the drafters recognised that the circumstances and purpose of a survey may be such as to make it permissible for someone to respond on behalf and instead of others (indirect primary collection). This may occur in two rather different main sorts of cases. Firstly, the data may be collected from someone close to the data subject, in particular a member of the household: someone close to the data subject naturally also has the desired information by virtue of that closeness and can be assumed to have the same interests as the data subject (the person might be a member of the family or a teacher, for instance). In the drafters' view, collection from such a person was justifiable in, among other cases, statistical surveys of entire households given that it might be physically impossible for all members of the household to be present or available at the time of a survey. Any obligation to collect data direct from the data subject would impose a considerable and pointless burden both on the data subject and the controller - and all in all would make a good many voluntary household surveys impossible. Secondly, data are collected indirectly from a third person not close to the data subject. This mainly concerns surveys of shifting populations or of anomie or atypical groups (vagrants, nomads, gangs of young offenders, and so on). In such cases not only may it be physically impossible to establish contact with the data subjects, but there may also be considerable doubt as to the reliability of replies even if direct contact can be made and the data subjects themselves questioned. The drafters felt that, in such exceptional circumstances, it was legitimate to allow indirect collection of personal data from persons who, on account of their institutional role or social involvement, were familiar with the groups the surveys covered (social workers would be one example of such persons). However, they wished to lay down two lawfulness requirements for any indirect collection, whether from third parties remote from or close to the data subject: firstly, it was necessary for the person providing the data to give consent to collection; secondly, the collection and processing must in no way endanger the privacy of the data subjects and in particular must not be concerned with sensitive data. The drafters further considered that, where appropriate, it would be for the independent supervisory authorities provided for in Prin-

inciple 18 to say what type of data could be collected and processed in a particular survey. In both cases and in so far as this is reasonable and is provided for by domestic law, data subjects should be able to oppose processing.

72. Principle 4.4 lays down two further conditions which must be met if collection and processing, for statistical purposes, of data originally collected for non-statistical purposes (secondary collection) are to be lawful. As already explained, this type of collection has the considerable practical advantage of allowing data to be collected without repeatedly disturbing the data subject (see paragraph 22). In addition, from the data protection standpoint, secondary collection as such is consistent with the purpose principle (Principle 4.2, see paragraph 69). Principle 4.4 thus lays down two alternative requirements for lawfulness which correspond to two particular cases:

a. secondary collection for statistical purposes is permissible when it is "necessary for the performance of a task carried out in the public interest or in the exercise of official authority". As a rule, the type of work involved here is the work which official or public statistics services perform under public law and which statistics legislation specifically entrusts to them. However, secondary collection for statistical purposes may also be carried out, under a public-interest assignment, by public or private institutions which are not official statistics services

(universities, research centres, non-profit-making institutions, quasi-public enterprises, etc.). Similarly, secondary collection may be carried out "in the exercise of official authority" by bodies external to the official statistics system. In the drafters' view, such collection is lawful provided that "tasks performed in the public interest" and "exercise of official authority" are defined in domestic law and are covered by data protection law. Additionally, Principle 5 specifies controllers' information duties. It is therefore also for the independent authorities on data protection (Principle 18) to see to it that proper information about such secondary collection is provided and that such collection meets the requirements of domestic law, in particular as regards firstly legal powers to perform work in the public interest and secondly institutions capable of exercising public authority ;

b. secondary collection for statistical purposes is also lawful where "necessary for the purposes of the legitimate interests pursued by the controller". Such interests may be of various kinds: scientific (secondary collection may sometimes yield data of better quality than are obtained by direct collection), technical (secondary collection is simpler to organise), financial (the costs are lower), or courtesy to the data subject (the data subject is not caused the disturbance of having someone seek information he/she has supplied before). One or more of these considerations may prompt a decision to use secondary collection. Private firms may therefore use for statistical purposes the data in their records of staff, wages, clients or suppliers (see paragraph 23). More generally, in national statistics systems such as the Nordic countries have developed, there may be systematic use for statistical purposes of data in public authorities' files (see paragraph 24). In both cases the drafters felt that the controller's interests were legitimate "except where [...] overridden by the rights and fundamental freedoms of the data subject" - especially by ensuring compliance with Principles 3 and 4 ;

c. lastly the drafters wished to specify that the above conditions of lawfulness also applied where data collected for statistical purposes were processed for other statistical purposes.

73. Principle 4.5 lays down a lawfulness requirement applying to compulsory collection of data for statistical processing by a file controller not acting on delegated official authority or under a public-interest assignment. In such cases the data subject may not be compelled to provide information unless domestic law provides for it. As already explained, domestic law includes, in addition to statute law, all legal measures coming under public law or private law (see paragraph 50). In particular, it should be noted that law of contract forms part of domestic law and that it is therefore permissible for an obligation to provide information to derive from a contract entered into by the data subject or his/her legal representative as long as domestic law does not oppose this with regard to sensitive data. In such cases the obligation, by law, would continue to exist until the contract was terminated or its effects were extinguished. Under Principle 6.4, penalties may not be imposed on anyone refusing to provide data unless there is provision for them in domestic law.

74. Principle 4.6 lays down a requirement which must be met if matching or interconnection of personal data for statistical purposes are to be lawful: interconnection and matching may only be performed if domestic law lays down "appropriate safeguards to prevent [the data] being processed and communicated for non-statistical purposes". The safeguards concerned consist in particular in rules of professional secrecy (Principle 3.2), provisions governing communication of data (Principle 12) and technical or organisational data-security measures, dealt with under Principle 15.

Proportionality

75. Principle 4.7 is based on a requirement in Article 5.c of the convention: data collected must be adequate, relevant and not excessive in relation to the purposes which the controller seeks to achieve. In the statistical field this principle of proportionality requires that only those personal data are to be collected and processed which are necessary to achieve the survey's purposes.

a. The recommendation expressly mentions a direct consequence of the proportionality principle: identification data can only be collected and processed if that is necessary for the purposes it is sought to achieve. As explained, identification data are sometimes indispensable to perform technical checking or to assemble files for statistical purposes, but they are not used for the production of statistical results (see paragraph 20). Principles 8 and 10 require specific guarantees before collection and storage of such data for statistical purposes are allowed.

b. In the drafters' view the concept of "data necessary for the statistical purposes to be achieved" needed some clarification: in practice the statistician does not, in the first instance, decide what data it is necessary to collect, but rather what variables or combinations of variables are needed to characterise a collective phenomenon. It is only thereafter, during design of the questionnaire, that the statistician decides the type of data that need collecting to obtain the desired set of variables. In designing a survey, therefore, the statistician has a choice between different methods and approaches. The drafters accepted that it was inappropriate for the recommendation, tribunals or supervisory authorities to dictate criteria of objective internal necessity for making this type of choice since this was a matter of methodology and of statisticians' freedom to exercise scientific judgment. The drafters considered that the concept of "necessary data" must be viewed in terms not only of constructing a set of variables but also of how specific the object of the particular survey was. For instance, although it is easy to decide what data are necessary to establish school examination pass rates or produce statistics on household income and expenditure, it is much less so when it comes to statistics on, and indicators of, sex inequality or a social group's standard of living. In addition, some collection of personal data, such as population censuses, is not aimed solely at generating a predefined set of statistical results but also at amassing general-purpose data that can be used for a whole range of specific statistical purposes over a lengthy period (ten years, say). The drafters thus acknowledged that the statistical purpose of a collection could not always be defined in terms of an expected statistical outcome or a specifiable number of statistical results. They therefore agreed that the proportionality principle, while remaining a general reference criterion in the design of questionnaires and surveys, must be applied in such a way as firstly not to interfere with the statistician's scientific freedom to decide the methodological approach and secondly to make allowance for how specific the objective of the particular collection was.

c. Where the purpose of the particular statistical survey is well defined, the requirement to collect only "necessary" data has an important practical implication: if the desired data are already available, it should not be necessary to collect them again. Thus, provided the overall requirements concerning lawfulness and respect for privacy are met (Principles 3 and 4), the recommendation allows processing for statistical purposes of data originally collected for non-statistical purposes (secondary collection) as well as processing for statistical purposes of data originally collected for other statistical purposes. The drafters further took the view that, provided it was consistent with the principles in the recommendation, domestic law should encourage multiple use for statistical purposes of data already available.

d The proportionality principle also requires that the organisation and methods of data collection and processing be relevant and necessary to the statistical object pursued. However, the recommendation does not set out to lay down criteria of relevance or necessity, which are a matter for statisticians' independent methodological judgment and which, once again, depend on how specific the purposes of the particular collection are. The recommendation considers the proportionality of organisational arrangements and statistical methods purely from the standpoint of protecting the data subject's privacy. The drafters felt that the recommendation's provisions as a whole provided that protection, whatever the organisational complexity or the methods used to collect and process personal data. They noted, incidentally, that proportionality of method and organisation was imposed as a matter of course by the material, financial and institutional constraints of statistical work.

e. Lastly, although, from the proportionality principle alone, it was not possible to derive regulatory guidelines as to choice of type of collection, the drafters took the view that the data subject's interests should be a central criterion in any such decision. As far as possible, therefore, the controller should opt for sampling rather than full-scale surveys, for voluntary rather than compulsory surveys, and for secondary rather than primary collection.

Sensitive data

76. Principle 4.8 deals with collection and processing of sensitive data, which are defined in Principle 1. The drafters acknowledged that, in statistics, confidentiality measures and security measures sought the same degree of protection for all personal data and that, on the face of it, there was therefore no need for statisticians to draw any distinction between sensitive data and "normal" data. However, the convention does not allow any derogations from its Article 6, which requires that domestic law provide appropriate safeguards on processing of sensitive data. In Principle 4.8, therefore, the drafters laid down a rule which is additional to Principle 4.1, is based on the proportionality principle and deals specifically with collection of this type of data.

a. The rule is that, in principle, sensitive data should be collected in such a way that the data subject is not identifiable. Here, the drafters were of the view that collection of sensitive data without identification data was a safeguard within the meaning of Article 6 of the convention. Consequently subsequent processing of sensitive data did not require any further special protection measures since it was sufficient to then apply the protection measures laid down for all processing of data for statistical purposes (see Principles 11, 14 and 15).

b. However, the drafters took into account that there might be a legitimate need to collect sensitive data for statistical purposes in identifiable form. For example, it may be necessary to collect addresses or partial addresses for data subjects (post code, town of residence) so as to be able to produce statistical results at different levels of aggregation or for varying geographical units. One example is health statistics (statistics on medical and hospital care, for instance), which often need breaking down not only by administrative region but also by town or by area within the national health system. It may also be necessary to collect data in identifiable form so that a repeat survey can be carried out. To reconcile such legitimate needs with the privacy-protection requirement, the drafters agreed that the recommendation should allow simultaneous collection of sensitive data and identification data on two conditions:

- firstly, such collection must be necessary to the legitimate purpose of the collection. This is a particular application of the proportionality principle laid down in Principle 4.7 but requires that the particular purpose of collection be very carefully specified;

- secondly, domestic law must provide appropriate safeguards, including specific measures to dissociate sensitive data and identification data as from the stage of collection unless that is clearly unreasonable or impracticable. The drafters took the view that, to ensure that controllers processing sensitive data were unable to identify the data subject, domestic law must prescribe stricter protection measures than were contemplated in the general provisions of Principle 10 concerning collection and use of identification data.

c. In addition, Principle 6.2 lays down provisions regarding data subjects' consent to collection of sensitive data.

5. Information provided to the data subject

77. Principle 5 concerns information that the controller must give the data subject in order to comply with the requirements of fair collection as laid down in Article 5 of the convention. The drafters considered it necessary to draw a very clear distinction between this type of information and other types which the convention also dealt with:

a. under Article 8 of the convention everyone must be able to obtain information about any recording or communication of personal data and if necessary ask for the data to be rectified, and must have a remedy available to him/her. These are supplementary safeguards intended to enable people to protect their rights as regards automated data files, particularly in view of the risk of individual decisions affecting them being taken on the basis of the data in such files or data communicated to third parties by a controller. Such supplementary safeguards need not apply when personal data are collected and processed exclusively for statistical purposes as defined in Principles 1, 4.1 and 4.2 ;

b. on the other hand the drafters took the view that, under Article 5 of the convention, persons whose data were collected for statistical purposes must in principle be given appropriate information, in particular about the nature, characteristics, circumstances and purposes of collection. Giving this information is not just an essential ingredient of fair collection but is also a way of winning data subjects' trust so that they answer truthfully and the data are reliable. Fairness of collection and data quality are thus interconnected.

Information provided at the time of primary collection

78. Principle 5.1 deals with two types of information relevant to primary collection where the data subjects themselves are questioned (direct primary collection):

a. the first type (sections a to e) is information which must always be provided - by whatever means and whether before or during questioning - to data subjects. The drafters considered that, in any survey, data subjects needed to know whether their participation was optional or compulsory (in the latter case they had to be informed of the legal basis of the obligation). They must also be informed of the purposes of the collection and processing and of the name and status of the person or body responsible for the collection and processing. In addition they must be given explicit assurances that their personal data will be processed confidentially and will be used exclusively for statistical purposes. Lastly they must be told how to obtain further information on request, in particular the information items referred to in sub-paragraphs f to j of this principle ;

b. the second type (sub-paragraphs f to j) is information which must be provided at the data subject's request or in the manner laid down in domestic law. In a voluntary survey, data subjects must be able to obtain information on how to refuse to reply or how to withdraw consent (in particular where they do not wish to take part in repeat surveys). In a compulsory

survey, the interviewer or controller must answer any questions about the penalties the data subject may incur for refusing to reply. In addition, although the rights of access and rectification are very limited where data are collected and processed exclusively for statistical purposes (Principle 7), the drafters took the view that data subjects must be able to obtain information on those rights. They must likewise be able to find out to what persons or bodies their personal data may be communicated. Similarly, if they wish, it must be possible for them to obtain information about data protection measures and the various safeguards for ensuring confidential processing. Lastly, the interviewer or controller must answer any question from the data subject about the types of data it has been decided to collect and process, their significance or their relevance to the survey. The desirability, manner and extent of this information provision come under Principle 5.3 (see paragraph 80).

79. Principle 5.2 deals with the information which must be provided to data subjects when it is not they themselves who are questioned and someone else answers the questions in their stead (indirect primary collection). This form of collection, whose permissibility is established by Principle 4.3, clearly cannot be subject to the same information requirements as direct primary collection. The drafters agreed that the right to information laid down in Article 8 of the convention must be adapted to the particular circumstances of indirect collection for statistical purposes.

a. As a rule, data subjects who are not themselves questioned must be informed that their data have been collected from other persons. However, there is no obligation to provide this information where in practice it is clearly unreasonable or impracticable to do so.

b. When data subjects are told that their data have been collected from other persons, they must be able to obtain proper information about items listed in Principle 5.1, sub-paragraphs a to e. The information may be provided either by the controller or, quite simply, the person who replies to the survey if, at the time, he/she is given the relevant information orally or in writing in accordance with Principle 5.3.

80. Principle 5.3 states that anyone questioned in a survey, whether they provide their own personal data or other people's personal data, must be given the information items set out in Principle 5.1, sub-paragraphs a to e "at the latest at the time of the collection".

a. However, the drafters felt it needed specifying that the manner and extent of such information provision must be appropriate and adapted to the circumstances. That, indeed, is why they drew a distinction, in Principle 5.1, between two categories of information, taking the view that it would clearly be inappropriate to supply unduly detailed information or information which the data subject had not asked for at the outset. They therefore felt that the information need not be sent to all data subjects individually but that various forms and means of information could be used as required by the nature or scale of the survey. In addition, they acknowledged that not all data subjects would be capable of understanding the full details or implications of a survey. They likewise acknowledged that unduly detailed information might be a disproportionate and pointless burden both on the questioner and the data subject. They accordingly took the view that the information provided by the controller should be proportionate both to the data subject's interests and the circumstances, purpose and scope of the survey. In particular, the terminology used and the amount of detail in the information should be such that the data subject is able to grasp easily, but in general terms, the purposes and significance of the survey. Lastly, in voluntary surveys, sufficient information should be given to enable the data subject to give informed consent.

b. The drafters observed that, in certain circumstances, information about the nature and objectives of a survey might distort the data subject's answers, as in statistical surveys for psychological or social research projects on the characteristics and behaviour of certain social groups. The drafters took the view that, even in such cases, the right to information must be protected; they accepted, however, that provision of information or of part of it might be deferred where that was essential to a survey's legitimate purpose. If need be, it would be for the independent supervisory authorities on data protection (Principle 18) to give a ruling on whether deferral was warranted. The recommendation states that, as soon as the need to withhold information ceases, the information must be provided to the data subject unless it is clearly unreasonable or impracticable to do so. Where data were collected direct from the data subjects, all the information relevant under Principle 5.1, sub-paragraphs a to e, must therefore be supplied to them direct at some stage after collection.

Information provided at the time of secondary collection

81. Principle 5.4 deals with the individual's right to information during secondary collection of their personal data for statistical purposes. In this type of collection (the nature and permissibility of which are established in Principles 4.2 and 4.4), information cannot be supplied to data subjects in the same way or by the same means as in primary collection since it would be very expensive, and may even be impossible, to trace all the data subjects, who, in addition, might be surprised, or even needlessly worried, at receiving the information.

a. The drafters took the view that the right to information might here be met by appropriate publicity. In the case of company statistics (statistics concerning staff, clients or suppliers), the publicity may be a standard clause in the employment contract, internal circulars or notification either sent routinely to clients and suppliers or appearing on invoices and order forms. In the case of public statistics, appropriate publicity may consist in official publication of the statute or other measure of internal law authorising secondary collection.

b. In addition the drafters took the view that if the information items in Principle 5.1, sub-paragraphs a to e, did not appear clearly and in their entirety in the public notice of secondary collection, it was necessary for data subjects to be able to obtain the information concerned except where that was impossible or would require disproportionate effort of the controller or if secondary collection and subsequent processing of data were expressly provided for in domestic law.

The information rights of the legally incapable

82. Principle 5.5 deals with exercise of the right to information where data on the legally incapable are collected for statistical purposes. The legally incapable are persons whom domestic law treats as unable to act on their own behalf. The category includes both minors and persons who are mentally impaired and who consequently are unable to give or withhold consent in a free, informed manner.

a. When collection of personal data is concerned with anyone who is legally incapable, the relevant information must be given to the legal representative. The drafters noted, however, that in some member states domestic law allowed certain categories of legally incapable person to act on their own behalf, and in these specific cases they took the view that information could be supplied direct to the data subject.

b. Principle 5.5 encourages giving information direct to anyone legally incapable but able to understand it, but it provides that, in a direct survey, the relevant information need not be

provided before or during questioning but may be provided after questioning and before processing of the data: since the persons concerned are relatively fragile or mentally immature, the drafters took the view that information provided before collection might affect the reliability of the answers and cause significant distortion. They nonetheless felt that such persons were entitled to be given relevant information provided it was adapted to the circumstances. It could be given before processing of their personal data so as to allow them to say whether or not they agree to such processing. The possibility of expressing personal wishes is dealt with in Principle 6.3.

6. Consent

83. Principle 6 deals with consent where data subjects are able to allow or refuse permission for their data to be collected and processed for statistical purposes. This section is based on Article 5 of the convention and Principle 4.3 of the recommendation, which treat consent as one of the essential prerequisites for fair and lawful voluntary collection.

84. Principle 6.1 sets out the general methods and characteristics of exercising the right to give or refuse consent:

a. The data subject's consent must be free, informed and not open to doubt. This involves three prerequisites, which apply jointly:

- firstly, data subjects must not be subjected to any kind of duress, influence or pressure (for example, psychological pressure or pressure brought to bear through their work), whether direct or indirect, if consent is to be free;

- secondly, data subjects must be informed in the manner and by the means prescribed in section 5 of the recommendation if consent is to be informed;

- thirdly, there must be nothing to suggest that the data subject might be hesitant, apprehensive or in any other way reluctant to provide the requested information if the consent was not to be in doubt.

b. These three requirements are interconnected and together form the very basis of fair collection. The drafters were particularly anxious that the impression should not be given that a survey was compulsory when it was not.

c. When these three requirements are met, the fact of giving data amounts to clear consent and requires no further officialisation or confirmation: the view was taken - and in this the recommendation differs from some other recommendations on data protection - that written certification of consent would place an undue burden on the controller. The drafters accordingly took the view that this would be an unnecessary safeguard where data were being collected solely for statistical processing since, under Principles 1 and 4.1, such collection immediately precluded any use of the data for individual decisions directly affecting the data subject. They further considered that, in many statistical surveys (for instance, surveys of young people or social groups in difficulty), written certification of consent would put the questioner's relations with the data subject on too formal a footing, might be an obstacle to ease of communication and a climate of trust, which had a vital bearing on the success of such surveys, and in some cases might even upset or disturb the data subject.

d. The second paragraph of Principle 6.1 states that when data have been collected, the data subject must be able to withdraw consent to their processing. However, this is possible only if, when consent is withdrawn, the identification data have not already been separated from the

other data collected, since under Principles 3.3, 8.1 and 10.1 the controller must separate identification data from the other data as soon as they cease to be needed for processing purposes. This measure is basic to ensuring the anonymity of personal data but the result, when the measure is applied and someone wishes to withdraw consent, is that it is extremely difficult, or even impossible, to trace the data in order to erase them. The drafters therefore took the view that data subjects' right to withdraw consent to processing of their data ceased to have any meaning or effect when the data had already been made anonymous.

e. Principle 6.1 likewise states that data subjects must be allowed to withdraw their co-operation at any time in the case of a survey to be repeated over a period of time: if that possibility did not exist, participation in the survey would cease to be voluntary and would de facto be compulsory, which would contravene the principles of lawfulness and fairness. However, the drafters took the view that, although everyone had the right to stop participating in a survey conducted over a period of time, exercise of that right could in no case have retroactive effect. In the drafters' view that would be pointless as data collected in previous stages of the survey would in all likelihood already have been processed and incorporated in sequential series of data and intermediate results. Thus, even if it were still possible to trace and erase data which a data subject had provided on various past occasions, they felt that the work needed would be totally disproportionate to any interests a data subject might normally be able to invoke.

f. Lastly, the drafters deliberately confined themselves, in Principle 6.1, to laying down a general framework for consent. They acknowledged that consent might raise particular issues and necessitate special regulatory guidelines in the different types of collection which the recommendation allowed. However, they decided against laying down detailed recommendations on the ground that it was for domestic law and, where appropriate, the independent supervisory authorities (Principle 18) to lay down regulatory frameworks appropriate to particular cases. Such frameworks should reflect all the requirements laid down in the recommendation, in particular the statistical purposes allowed by Principles 1, 4.1 and 4.2, and allow differential application of data subjects' right to information as described in section 5.

85. Principle 6.2 deals with the data subject's consent to the collection and processing of sensitive data for statistical purposes.

a. The drafters took the view that, under the principle of proportionality, a requirement additional to those set out in Principle 6.1 was needed when consent was required for the collection of sensitive data. In collection of sensitive data it was not enough that free and informed consent was not in doubt: consent must also be expressly given. Here, the mere fact of providing data could not be regarded as amounting to consent: it was necessary for the data subject's explicit consent to precede the survey. According to the particular case and the arrangements laid down in domestic law, express consent can be given orally or in writing. In particular the drafters took the view that the method of consent might take the data subject's circumstances into account.

b. With reference to the provisions of Principle 4.8, the drafters wished to make clear that the legitimacy of a survey's purpose did not in itself absolve the controller from the duty to seek the data subject's consent. Any exception to that rule can only be justified by major public interest, as where statistical information is needed to contain epidemics, combat the evil of drug taking, investigate the scale and pattern of sexual assaults on minors or develop aid to

social groups in difficulty. Such examples, to which many more might be added, relate to matters which affect society's essential interests and in which the state has responsibilities. In such cases the guarantees on protection of sensitive data must be adapted to the objective information needs arising from the public interest. To obtain representative, up-to-date and useful information about the population's health, for example, data on individuals' illnesses and medical or hospital care received may be collected from doctors, health workers or hospitals without patients' consent. However, such practices call for stricter protection measures and tighter data security as well as for special provisions of domestic law.

86. Principle 6.3 deals with consent where data on legally incapable persons are collected and processed for statistical purposes.

a. In principle, such consent, as described in Principle 6.2, must

be given by the person legally responsible (the parents, guardian or other person, authority or body with legal responsibility for the person concerned). However, where domestic law provides that particular categories of legally incapable person may act on their own behalf, they may themselves give or withhold their consent. The drafters nevertheless felt that, in such special cases, the independent supervisory authorities (Principle 18) should take great care that the requirements for conducting the survey were strictly in accordance with the fairness principle and in particular met the requirement to seek valid consent.

b. In line with Principle 5.5, which encourages giving information to legally incapable people if they are able to understand it, the drafters took the view that, unless domestic law did not permit this, such persons should be allowed to express their desire to take part or not take part. This does not involve consent in the strict sense (which ultimately is a matter for the person legally responsible), but is intended to give the legally incapable a say in weighing up and deciding whether to give consent. In the case of minors, it may also provide a valuable lesson about data collection for statistical purposes and give them a foretaste of exercising their rights.

87. Principle 6.4 is intended to preclude all possibility of imposing arbitrary penalties on anyone refusing to provide personal data. It makes clear that there must not be any penalties for refusal unless domestic law provides for them. The term "penalty" encompasses not only criminal or administrative penalties but any disadvantage which might be incurred by anyone refusing to take part in a survey. Further, although it appears in the section on consent (which exists only where participation in a survey is voluntary), the requirement applies also, and much more directly, to the types of compulsory collection dealt with in Articles 4.3 and 4.5. (see paragraphs 71.a and 73.

7. Rights of access and rectification

88. Principle 7.1 draws attention to a guarantee in the convention: everyone must be able to obtain access to personal data concerning him/her and must also be able to have the controller rectify the data if they are wrong or inaccurate. The purpose of this provision is in particular to allow people to ensure that individual measures or decisions concerning them are not taken on the basis of incorrect or inaccurate data held in files and which could be released to third parties by a controller.

89. There is no such risk where personal data are collected and processed exclusively for statistical purposes (see paragraph 65). In addition, under Principle 3.1, the controller may legitimately modify data collected for statistical purposes and the purpose of the right of

access (to keep data accurate) is meaningless in that context (see paragraph 65.c). In the drafters' experience, too, where statistical processing has merged the individual data item with a mass of data it would require a disproportionate amount of effort for the controller to identify the item and make it available to the person concerned. In practice, assisting exercise of the right of access to data would make it necessary to delay separating out the identification data, and after the identification data had been separated out the controller would have to be able to preserve both the identification data and the other personal data for quite a long time. Clearly such measures would be contrary to Principles 3.3, 8, 10 and 11.1 of the recommendation. The drafters therefore took the view that, regard being had to the convention's objectives, it was preferable here to entertain a restriction on the rights of access and rectification as dealt with in Principle 7.2. Principle 7.2 is similar in wording to the convention, Article 9 (3) of which allows restriction of the rights of access and rectification when there is obviously no risk of an infringement of the data subject's privacy. This is the case where personal data are collected and processed exclusively for statistical purposes (within the meaning of Principles 1, 4.1 and 4.2) and where respect for privacy is guaranteed (within the meaning of Principles 3.1, 3.2 and 3.3). The drafters therefore agreed that, in so far as domestic law permits, the rights of access and rectification might be restricted where personal data had been collected exclusively for statistical purposes and where special measures entirely prevented identification of data subjects by a third party, whether from individual data or statistical results. In other words, the drafters took the view that the rights in question might be restricted proportionately to the degree of protection which could be provided by implementing the recommendation's provisions as a whole.

8. Anonymity

90. Principle 8.1 makes various clarifications concerning the anonymity of data which Principle 3.3 requires as a direct means of preventing invasion of privacy (see paragraph 67). Principle 8.1 states that identification data must be separated from personal data (that is, data informative for statistical purposes) as soon as collection, checking and matching are complete (see paragraphs 20 and 21). However, there is provision for two types of exception:

a. the controller may take steps to enable anonymised data to be used later in identifiable form for statistical purposes. That requires that it be possible for the identification data to be kept in such a way that they can later be reunited with the anonymised personal data. In such cases the data, though anonymous for the time being, can be regarded as potentially identifiable. Under Principle 8.1, this potential identifiability must be controlled by separate storage of identification data, a matter dealt with in Principle 10.1 ;

b. the drafters likewise contemplated the possibility that, for purposes other than checking and matching, it might be necessary to carry out statistical-processing operations before data had been made totally anonymous. In such cases it is necessary to apply the reinforced security measures laid down in Principle 15.2 ;

c. lastly, the drafters wished to make clear that Principle 8.1 was to be interpreted as allowing a member state that provided appropriate safeguards to keep data in non-anonymous form for archive purposes.

9. Primary collection of personal data for statistical purposes

91. Principle 9.1 lays down a rule of fairness in all collection of personal data for statistical purposes. It provides that collection be conducted with total openness, and in this connection the drafters wished to draw attention, in particular, to the requirements that data subjects

be kept fully informed and have freedom of reply. The nature and scope of these two requirements are the subject of sections 5 and 6 of the recommendation (see paragraphs 77 to 87).

92. Principle 9.2 recapitulates on the various types of primary collection which are allowed on the conditions set out in Principle 4.3: collection from the data subject (direct primary collection), collection from a member of the household (indirect primary collection), and collection from someone else who is not a member of the household

or from legal entities such as companies or public authorities (see paragraph 71). It is specified that this third type of direct collection is permissible only if there is provision for it in domestic law, which must provide appropriate safeguards. However, the drafters considered that this requirement need not apply where "there [was] clearly no risk of infringement of the rights and fundamental freedoms of the data subjects". That may apply to data collection from persons close to the data subject but who are not members of the household (for example, a teacher). It also applies to surveys of shifting populations which draw

on some specialist familiar with the populations concerned, such as a social worker. Principle 4.3 states that this is permissible only in so far as it does not pose any risk of invading the data subjects' privacy (see paragraph 71.b).

93. Principle 9.3 deals with direct collection without questioning - that is, by means of observation of acts of data subjects, such as entering a museum or travelling over a particular section of the road network by motor car.

a. Needless to say, and in accordance with the purpose principle, what is involved here is not investigation of individual behaviour but solely direct counting of items of behaviour in order to measure mass phenomena (the number of museum visitors, or the volume and type of motor traffic on a section of the road network).

b. By definition, of course, such collection of data involves neither informing nor obtaining the consent of the persons observed; in the drafters' view, however, in so far as the collection was for statistical purposes it was unnecessary here to lay down specific protection measures in that counting of items of behaviour was carried out in public and in that, in accordance with the anonymity rule, the people concerned were not identified or identifiable. Thus, whether performed manually or automatically, such collection is not subject to any particular requirement in so far as it involves neither collection of identification data nor linkage to data of that type.

c. The position is different when direct collection without questioning is performed together with gathering of identification data. In such cases there must be explicit provision in domestic law for simultaneous collection of personal data and identification data and domestic law must also lay down appropriate safeguards.

d. Lastly, the drafters took the view that, in so far as domestic law permitted it and laid down appropriate safeguards, data collected by direct observation could be linked to identification data, in particular for the construction of samples.

94. Principle 9.4 deals with processing of data on non-respondents - that is, persons who refuse to take part in a survey.

a. Even where the controller has not been given the requested personal data, he may have, or obtain from the person concerned, three types of information:

- he may have personal information in the database created for the purposes of the survey. Such information will normally consist of identification data or partial identification data (for example, an address or telephone number, and therefore the place of residence); but it may also consist of personal data from a file and concern variables (sex, age, socio-professional status, etc.) relevant to drawing up a representative sample;

- it may also be that the controller plans to record the reasons for refusal to take part in a survey and group them into various categories;

- lastly it may be that, when interviewed, some people refuse to take part in the survey but nonetheless agree to provide information of value to the representativeness of the survey (for example, information about the size and composition of the household).

b. The drafters took the view that, subject to the fairness principle

and the legitimacy of the collection purposes, the controller could use such information, but solely to ensure the representativeness of the survey (structure of the sample, correction of estimates, analysis of non-response, etc.).

c. It should be made clear that, in accordance with Principles 5 and 6, where a person has been informed appropriately and has agreed to provide some of the personal data requested or an item of information relevant to part of the survey, partial consent has clearly been given

and the collected data may therefore be processed for the statistical purposes of the survey.

95. Principle 9.5 points out that the controller must take special care in the selection of staff and in organising the work of interviewers or anyone else with possible knowledge of the data collected. The drafters took the view that, although there was no need to make any stipulations here regarding selection and management of staff, it was nonetheless essential to point out the importance of human and organisational considerations in ensuring that the survey's purposes were observed, data remained confidential and privacy was protected. Here, the controller should take steps to avert any slip-ups which might result from his/her staff's ignorance or carelessness. To that end the drafters felt that survey staff should be informed of the survey's methods and procedures, supervised by competent people, and made aware of the point of protection and security measures.

96. Principle 9.6 requires that the controller take appropriate measures to allow those questioned to satisfy themselves that the person collecting the data has proper authority to do so: the fairness rule requires the principle that they be able to check that the interviewer is bona fide, on what authority he/she is acting, and the institutional status of the controller (market-research firm, research institute, public agency, etc.). Where the survey is conducted by telephone, the controller should first inform the persons concerned in writing. The drafters believe that promoting this approach is in the interests both of the target population and of statistics agencies: both suffer when impostors assume false institutional identities and claim respectable statistical purposes in order to collect personal data for disreputable ends.

10. Identification data

97. Principle 10.1 provides: "When identification data are collected and processed for statistical purposes, they shall be separated and conserved separately from other personal data, unless it is clearly unreasonable or impracticable to do so". This formulation warrants one or two observations and clarifications:

a. Firstly, as defined in Principle 1, identification data are data, such as a name or address, that allow direct identification of the data subject. This was qualified to the effect that where data allow identification

of the data subject only if extremely complicated, lengthy and costly procedures are resorted to, the data are not identification data within the meaning of the recommendation (see paragraphs 52.d and 53.a).

b. In accordance with the proportionality principle, Principle 4.7 provides that only those identification data may be collected and processed which are necessary for legitimate collection purposes (see paragraph 75.a).

c. Where their collection is necessary, Principle 3.3 requires that they be separated from personal data as soon as these no longer need to be in identifiable form (see paragraph 67). Principle 8.1 states that this separation should generally take place as soon as collection, checking and matching procedures are complete (see paragraph 90).

d. Principle 10.1 deals more specifically with keeping identification data for subsequent processing. In such cases they must be kept separately from other personal data. The separation of the two types of data may be physical or logical, the main thing being to make sure that no connection can be established between them, unless that proves clearly unreasonable or impracticable.

e. The drafters were anxious to make it clear that Principle 10.1 must be interpreted in strict accordance with the above definition of identification data. In particular they were mindful that some member states have developed their national statistics systems on the basis of secondary collection (see paragraph 72). In those countries, statistical results are generated by matching or combining data from various administrative files, by means of a common identifier, the personal identity number (PIN), which necessarily remains attached to the personal data. The drafters took the view that there was no need to insist on separation or separate storage of PINs as long as they did not give direct access to data subjects' identity and as long as technical, organisational and supervisory measures safeguarded data subjects' privacy (see paragraph 24).

98. Principle 10.2 states that identification data may be used to create files of addresses for statistical purposes, but only on one of the following conditions:

a. domestic law permits it;

b. the data subjects have been informed that their data will be used to create files of addresses, and have not raised any objection;

c. the data have merely been taken from a publicly available file (telephone directories, professional or commercial registers, etc.).

11. Conservation of data

99. Principle 11.1 deals with the conservation and destruction of personal data collected for statistical purposes. It requires that data be destroyed or erased when they are no longer necessary for those purposes. It is specified that data are "necessary for statistical purposes" when required for the collection, checking and matching of data, to ensure the representativeness of a survey, or to repeat the survey later. In addition the principle states that personal data may be kept in two circumstances: either where they have been anonymised in

accordance with Principles 3.3, 8.1 and 10.1 or where domestic law explicitly provides for their conservation for archive purposes subject to the appropriate safeguards (see paragraph 90.c).

12. Communication

100. Principle 12 lays down the conditions on which personal data originally collected and processed for a particular statistical purpose may be communicated to a third party for further processing for other statistical purposes. What is involved here may be illustrated by three examples. Firstly, take the case of data on smoking initially collected by a private or public body for statistical purposes; such data may then be passed on to insurance companies interested in carrying out further statistical processing and analysis so as to ascertain the risk factor for various age groups. Second example: readers of a press group's newspapers and magazines agree to provide various personal data for statistical purposes (identifying readership characteristics, analysis of regional market penetration, etc.); such data may then be communicated to an advertising agency for further statistical processing (to obtain profiles of groups targeted by advertising campaigns, for instance). Third example: a public body collects and processes personal data for statistical purposes for a survey of travel and preferred modes of transport; the data may later be communicated to another body for further processing for purely statistical purposes, whether in connection with scientific research on personal mobility or for use in public transport planning. These examples have one thing in common: during initial collection, the data subjects were informed and consented to processing of their data for a specific statistical purpose, and although the specificity of a statistical purpose is always a relative matter (see paragraph 75.b) the purpose of communicating data to a third party is generally processing for statistical purposes other than those for which the data were initially collected. The drafters therefore took the view that a special regulatory framework was needed to cover such communication of personal data. The framework has six components:

a. the type of communication involved is communication of data whose collection, from the outset, was for statistical purposes and whose communication to a third party is still for exclusively statistical purposes. The lawfulness of such communication thus derives directly from Principles 4.1 and 4.2, dealing with the purposes of statistical surveys and compatibility of purposes: since Principle 4.2 acknowledges that data gathered for non-statistical purposes may be used for statistical purposes, such use is permissible in the case of data gathered for statistical purposes. This component is the subject of Principle 12.1 ;

b. principle 4.4 lays down two alternative lawfulness requirements applying both to secondary collection and to communication for statistical purposes of data collected for other statistical purposes: such communication must be necessary for the performance of a task carried out in the public interest or for the purposes of the legitimate interests pursued by the recipient of the data "except where such interests

are overridden by the rights and fundamental freedoms of the data subject". Although the drafters did not consider it necessary to reiterate these requirements in section 12, they still apply in full and the explanations concerning them should therefore be consulted (see paragraph 72) ;

c. in accordance with the proportionality rule in Principle 4.7, all communication is subject to a restrictive requirement concerning its purpose: the processing which the data recipient intends to perform must pursue a specific objective and may be carried out only for a strictly limited time. This component is covered by Principle 12.2;

d. communication must meet the lawfulness requirements laid down in Principle 4.3, under which the data subject's consent may not have to be obtained ;

e. as communication is a form of processing, the rules which apply include the provisions of Principles 3.1, 3.2 and 3.3, concerning respect for privacy. Although the drafters did not consider it necessary to enlarge on those rules in this section, they felt it appropriate to point out that:

- in any communication of data the data recipient must observe the strictest confidentiality;

- as a rule, data should not be communicated along with identification data. However, provided that domestic law lays down appropriate safeguards, data may be communicated in identifiable form when that is necessary both to a public-interest assignment and legitimate purposes of communication (for example, a longitudinal study; see paragraph 67.b);

f. lastly, where data subjects have not been informed of communication of their data to a third party for further processing, the drafters took the view that, under the proportionality principle,

- communication of personal data requires additional safeguards (see Principle 12.3);

- communication of sensitive data must be subject to specific requirements (see Principle 12.4).

101. Principle 12.1 establishes a direct connection between the purpose for which personal data are collected and the lawfulness of communicating such data to a third party. What we have here, therefore, is an application of Principle 4.1: the holder of the data not only must not use them for non-statistical purposes but also must not allow any third party to use them for such improper purposes. In other words, this principle prohibits the disclosure for other purposes of data collected for statistical purposes. However, data disclosure exclusively for statistical purposes is allowed by Principle 12.1.

102. Principle 12.2 states that data processed for statistical purposes may be communicated to third parties for specified statistical purposes of limited duration.

a. In the drafters' view there was no justification for allowing such third parties to have unspecified or unlimited use of the data. In accordance with the proportionality principle, communication must be confined to data necessary for the intended statistical processing (Principle 4.7). Sets of data or intermediate results, where they allow the data subjects to be identified, may be communicated on the same conditions.

b. The drafters acknowledged that, within the public statistics sector, communication of personal data to institutions such as national statistics institutes, statistics departments of ministries or the Statistical Office of the European Communities satisfies, by definition, the requirement that communication be for a specific statistical purpose.

c. Their finding was also that the provisions of this principle applied, among other cases, to communication of data to third parties with a view to processing for statistical purposes in the context of scientific research, subject to the following observations. As explained in the introduction to the explanatory memorandum, communication and use of personal data "for research purposes" must be limited strictly to processing operations on data of a statistical nature, aimed at characterising collective phenomena, and comprise therefore no personalised feedback to the individuals concerned (see paragraph 14). In this respect, the drafters were anxious to reiterate at this point that communication "for research purposes" may only take place in so far as the purposes of processing and the results aimed at by the recipient come

within the definitions of purposes and statistical results as set out in Principle 1. The drafters pointed out that communication of personal data must be confined to processing of that type and should envisage only results of that type. Any communication for scientific research on individuals

or which could result in decisions or measures concerning a particular individual was precluded.

103. Under Principle 12.3, in accordance with the principle of proportionality, communication of personal data to third parties for processing for statistical purposes requires additional safeguards involving a contractual agreement between the provider and the recipient of the data. Before communication, the controller must therefore lay down the conditions governing communication and use of the data.

a. He must stipulate that the data communicated are solely for the use of the recipient, who must undertake not to communicate the data to any other user except, of course, with the controller's consent.

b He must satisfy himself that the recipient of the data will scrupulously protect the data and more particularly keep them secure and will take care that personal data cannot be identified through any results which are published.

c. To make sure the recipient of the data is fully aware of the precautions that need to be taken and promises to apply them, there should be a written document specifying them. However, the drafters acknowledged that the text of a law or regulation authorising the communication of personal data or of sets of such data could be regarded as equivalent to a "written document setting out the rights and obligations of the parties".

d. They also took the view that the above provisions applied equally to communication for the recipient's use and declared purposes and to communication to enable the recipient to carry out statistical processing for the controller.

104. Principle 12.4 requires that any communication of sensitive data be subject to even stricter precautions. It is permissible on two alternative conditions: either there must be provision for it in law and, in accordance with Article 6 of the convention, the law must lay down legal and material safeguards to prevent any misappropriation; or the data subject (or his/her legal representative) must have given explicit consent. In the latter case, moreover, the drafters allowed that, even with such consent, the law might prohibit such communication (for example, on grounds of public policy).

13. Transborder data flows

105. Principle 13.1 deals with transborder communication of personal data for statistical purposes. People's increasing mobility, the global economy and the need for a thorough, up-to-date grasp of international phenomena have generated increased information demand. Here, progress in information technology and telecommunications goes hand in hand with increased statistical activity on the part of companies and multinational groups and greater statistical co-operation not only between countries but also between countries and supranational institutions, particularly in Europe. In our "information culture", transborder flows of statistical data are therefore taking on ever more importance: speedy, straightforward communication of such data is a necessity no less for the public sector than for the private sector or the scientific community. But the need for availability of data must not take precedence over safeguarding the individual's privacy. The safeguards and protection measures laid down in the

recommendation must therefore apply to any transborder communication of personal data for statistical purposes.

106. Principle 13.2 states that data protection measures in the country providing the data and protection measures in the country receiving the data must be "at least equivalent". In this connection Principle 13.2 concerns states which have ratified the convention ("parties to the convention") and which, to be able to do so, first adopted data protection legislation whose provisions were in line with the convention's basic principles. It acknowledges that legal safeguards in those countries are at least equivalent and that there is therefore no obvious reason to restrict transborder flows of data for statistical purposes between individuals, companies or public or private institutions in those countries. In so far as it complies with the principles in the recommendation, communication of personal data collected, processed or used for statistical purposes (this includes statistical results allowing indirect identification of the data subject) should not be restricted by special requirements concerning protection of privacy.

107. Principle 13.3 likewise allows transborder communication of personal data for statistical purposes between individuals, firms or institutions in the territories of parties to the convention and individuals, firms or institutions in the territories of states which, though they have not ratified the convention, have adopted legislation conforming to the principles in the convention since, in such cases, the national data protection measures are equivalent.

108. Principle 13.4 deals with transborder communication of personal data for statistical purposes to states whose protection arrangements do not conform to the principles of the convention or recommendation. In such cases there is no equivalence of protection. The drafters accordingly took the view that unless the domestic law of the sending state provided otherwise, there must be no such communication except on one of the two following conditions:

- a. either special arrangements must have been made, in particular under a contractual agreement, to bring protection and security measures into line with the convention's principles and the recommendation's provisions;
- b. or data subjects, having been informed of the possibility of their data's being communicated to a state whose legal safeguards for data protection are not equivalent to those in their own countries, must have expressly given their consent to such communication.

14. Statistical results

109. Principle 14.1 deals with statistical results. As stated in the introduction, such results ought not to require special protection in so far as they do not say anything specific about any of the persons whose data have been used (see paragraph 9). On occasion, however, it may be possible, from statistical results, to trace back to personal data by means of analysis and cross-checking and then to make a connection between the data and particular individuals (see paragraphs 27.d and 58.a). On account of this risk of indirect identification, Principle 2.1 brought statistical results within the recommendation's scope (see paragraph 61). Principle 14.1 requires that precautions be taken right up to result publication to guard against indirect identification of the data subject. Here, the drafters wished to make the following clarifications:

- a. in accordance with the definition, in Principle 1, of the "identifiable" individual, measures to prevent indirect identification may be regarded as adequate once indirect identification of data subjects would require unreasonable time and effort;
- b. the requirement that results be systematically checked to ensure their opacity applies in all cases, whether the results have been obtained from data collected for statistical purposes or from data taken from files assembled for other purposes;
- c. it was accepted that, in some surveys, it might not be possible to make sure that no data subject was identifiable from results (for example, local results might state that an occupation was only carried on by one person, who then became readily identifiable). In such cases, results might be published if the data subjects had given their consent or if the data were already common knowledge or their publication was clearly not damaging to the data subject's privacy;
- d. lastly the drafters noted the co-operation between some national statistics institutes and the scientific community on development of automated methods of "statistical disclosure control". That work is aimed at systematically applying the very measures for prevention of indirect identification which Principle 14.1 advocates and is accordingly the type of technical development which Principles 15.4 and 17.1 encourage.

15. Security of personal data

110. Principle 15.1 deals with the technical and organisational steps which must be taken to ensure data security. The drafters were well aware that to provide full protection of personal data it was not enough to lay down legal rules and that the controller had to take actual steps to prevent unlawful access to or use of data, whether accidental or ill-intentioned.

- a. The security measures which this principle requires relate, among other things, to the powers and authority of services and staff. They also include measures concerning access to installations and documents as well as movement of data-support media. Lastly they encompass procedures, logic keys, processing programmes, and encryption or scrambling of data (see paragraph 26).
- b. In the drafters' view, the duty to ensure security of data applies just as much to compulsorily collected personal data as to data collected on a voluntary basis.
- c. While the requirement to provide data security applies in all cases, the choice and strictness of actual measures may vary according to the type of data it is desired to protect: elementary security measures should be taken to protect data which have been made totally anonymous, whereas tighter security measures will be required in the case of data from which data subjects may be identified (Principles 15.2 and 15.3).
- d. Security is the responsibility no less of the controller than of his staff or sub-contractors or anyone to whom personal data are communicated for statistical purposes as provided for in Principle 11 of the recommendation. The controller has a duty to inform those employed to collect and process data of their security responsibilities. Such persons must be required to give a formal undertaking to comply with security measures.

111. Principle 15.2 requires that the security measures laid down in 15.1 be made stricter when personal data are kept in a form

which allows identification of the data subject. The controller must set up organisational filters and barriers as well as employ appropriate technology (in particular computer technology) to prevent all unauthorised access to such data.

112. Principle 15.3 likewise clarifies the provisions of 15.1 where personal data have been separated from identification data but there is still a risk of the data subject's being identified. In such cases, physical and logical access to data should be prevented, or use made of protection technology such as encryption or scrambling (see paragraph 26).

113. Principle 15.4 acknowledges that protection of personal data can only be effective with use of techniques, methods and procedures designed, refined and applied by statisticians themselves. It accordingly encourages firms and statistics agencies to press on with development of such techniques and methods.

16. Codes of ethics

114. Principle 16.1 encourages statisticians and controllers (enterprises and agencies, whether private or public) to draw up and make public codes of ethics in accordance with the principles in the recommendation. The drafters took the view that, in so far as the regulatory framework of the recommendation derived not only from the principles in the convention but also from principles and practices approved by the international community of statisticians, it first and foremost required the unforced assent of the profession. Experience has shown that effectiveness of data protection depends to an appreciable extent on the profession's acknowledgment and implementation of rules of sound practice. Thus self-regulation is an indispensable complement to legal provisions (see paragraphs 30 to 38).

a. The drafters took the view that codes of ethics should not be a matter solely for international institutions or national-level professional bodies. To ensure that the principles both of the recommendation and professional ethics are known and subscribed to by all involved in collecting and processing data for statistical purposes, each enterprise or agency engaged in such collection and processing needs to have a code of its own which is both consistent with the basic principles and rooted in the realities of day-to-day work.

b. In particular, in addition to codes of ethics, there needs to be clear information as to the controller's name, status and authority, which persons and institutions have access to personal data, and the specific protection, confidentiality and security measures which the firm or agency applies.

17. Technical development, co-operation and assistance

115. Principle 17.1 points out the need for international co-operation as regards technical knowledge, data-processing technology and automated data protection methods. The drafters noted that development of methods and technology, as encouraged in Principle 15.4, required highly qualified people and large financial resources. They accordingly felt that the recommendation must call on the appropriate government bodies to co-operate closely and set up international co-operation programmes. They specified that the aim of such international co-operation should be to promote transfer of knowledge and provide effective technical assistance to statistics agencies in less well-off countries.

18. Supervisory authorities

116. Under Principle 18.1 member states shall give one or more independent authorities responsibility for ensuring the application of the provisions of domestic law giving effect to

the principles laid down in this recommendation. This explanatory memorandum several times refers to the need for an opinion from such independent authorities in given circumstances. The drafters took the view that applying the principles of the recommendation might sometimes require special assessment of the implications of a particular data collection or processing operation from the standpoint of privacy or the controller's legitimate interests.

- a. They took the view that such an authority should not confine itself to supervisory functions but should also provide guidance and advice, both to the data subject and the controller.
- b. They also felt that where domestic law permitted and where personal data of a particular type required especially strict protection, the independent authority or authorities should be able to have direct security responsibilities, such as custody and control of logic keys allowing personal data to be united with identification data, encryption of certain data, scrambling of data, or direct supervision of communication of data to third parties.
- c. Lastly they took the view that relations between the independent authorities and data controllers should be based on genuine knowledge, on both sides, of the basic principles of data protection and the essential features of statistical work. Such relations should be governed not by systematic mutual suspicion but by an acknowledgment of each side's role and should be based on a constructive search for solutions that satisfied both the requirements of the recommendation and the legitimate interest in collecting and processing data for statistical purposes.

Footnotes

1. At present there are forty member states: Albania, Andorra, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Moldova, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, San Marino, Slovakia, Slovenia, Spain, Sweden, Switzerland, "the former Yugoslav Republic of Macedonia", Turkey, Ukraine and United Kingdom.
2. Hereinafter referred to as "the convention".
3. Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Slovenia, Spain, Sweden and United Kingdom.
4. Cyprus, Hungary, Romania and Turkey.
5. In 1989 the Council of Europe appointed its own Data Protection Commissioner who supervises personal data protection in the organisation's Secretariat.
6.
 1. Recommendation No. R (81) 1 on regulations for automated medical data banks (23 January 1981).
 2. Recommendation No. R (83) 10 on the protection of personal data used for scientific research and statistics (23 September 1983).
 3. Recommendation No. R (85) 20 on the protection of personal data used for the purposes of direct marketing (25 October 1985).

4. Recommendation No. (86) 1 on the protection of personal data used for social security purposes (23 January 1986).
 5. Recommendation No. R (87) 15 regulating the use of personal data in the police sector (17 September 1987).
 6. Recommendation No. R (89) 2 on the protection of personal data used for employment purposes (18 January 1989).
 7. Recommendation No. R (90) 19 on the protection of personal data used for payment and other related operations (13 September 1990).
 8. Recommendation No. R (91) 10 on the communication to third parties of personal data held by public bodies (9 September 1991).
 9. Recommendation No. R (95) 4 on the protection of personal data in the area of telecommunication services with particular reference to telephone services (7 February 1995).
 10. Recommendation No. R (97) 5 on the protection of medical data (13 February 1997).
7. New technologies: a challenge to privacy protection? (1989); Data protection and the media (1990); The introduction and use of personal identification numbers: the data protection issues (1991).
 8. International Statistical Institute, "Declaration on professional ethics", in *International Statistical Review*, 54/2 (1986) 227-242.
 9. Council Regulation (Euratom, EEC) No. 1588/90 of 11 June 1990 on the transmission of data subject to statistical confidentiality to the Statistical Office of the European Communities, OJL 151/1.
 10. Directive 95/46/EC of the European Parliament and of the Council of the European Union on the protection of individuals with regard to the processing of personal data and on the free movement of such data.
 11. Decision C(47) of the United Nations Economic Commission for Europe, adopted at its 47th Session on 15 April 1992 (Doc. EEC/UN E/1992/32).
 12. ECHR judgment, *Cantoni v. France*, 15 November 1996, No.45/1995/551/637, paragraph 29. Accessibility is secured where the relevant legal provisions have been published or drawn to the attention of those concerned. Foreseeability is a question of a statute's precision and clarity. However, in two areas the foreseeability of the law's effects is to be narrowly construed in view of the inherent risk of arbitrariness:
 - where a discretionary power is vested in an administrative authority;
 - where use is made of secret surveillance measures (on telephone tapping, see the *Klass* judgment 1978, the *Malone* judgment 1984 and the *Kruslin* judgment 1990; on keeping and consulting a security register, see *Leander* judgment, 26 March 1987, Series A, No.116, paragraph 51).

Here the law itself must adequately specify the conditions governing and extent of powers of interference so as to ensure the minimum degree of protection required by the rule of law in a democratic society.

13. *Sunday Times* judgment of 26 April 1979, Series A, No. 30, paragraph 46.
14. Kruslin judgment.
15. ECHR Barthold judgment of 25 March 1985, Series A, No.176-A, paragraph 29.
16. See paragraph 39 of the explanatory report to the convention for the Protection of Individuals with regard to Automatic Processing of Personal Data.
17. Cantoni v. France, judgment of 15 November 1996, No.45/1995/5512/637, paragraph 29.
18. S. Trechsel, "The right to liberty and security of the person - art. 5 of the ECHR in the Strasbourg case-law", *Human Rights Law Journal*, volume 1, Nos.1-4, Engel 1980, p.102.
19. The qualitative requirements must be assessed in the light of the specific circumstances, regard being had to the seriousness of the interference, its nature (the nature of the processing, the type of data processed, the purpose of the processing, the number of people concerned), the field and circumstances in which the interference occurs), the field and circumstances (organisational arrangements, the information system used), and the number and status of those who will be using the data. The basic legal provisions have to be sufficiently precise for the person concerned to be able to adapt his/her behaviour and foresee the consequences of his/her behaviour with a degree of certainty (see *Sunday Times* judgment, paragraph 49.)
20. The drafters of the recommendation considered that the concept of "**individual data**" covers personal data and anonymous data, that is any information relating to a given individual, whether identified, identifiable or not. However, the recommendation does not use the term "individual data", and that definition has consequently not been included.
21. At the drafters' request this interpretation of Article 5 of the convention was approved by the consultative committee of representatives of parties to the convention at its 11th meeting, 7 to 10 November 1995.