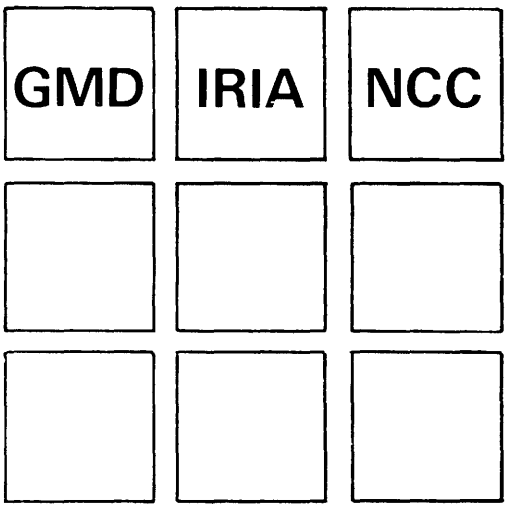


STUDY ON

DATA SECURITY AND CONFIDENTIALITY

FINAL REPORT

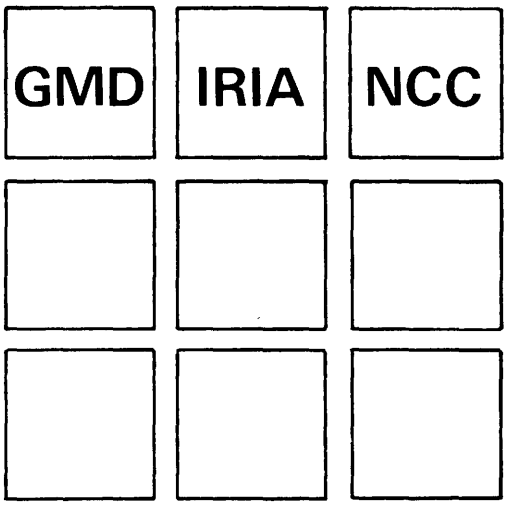
to the Commission for the European Communities

Volume 5 of 6

Section 5: Technical aspects of the right of access

by F Bancilhon
L Joinet (counsellor)

JANUARY 1980



STUDY ON

DATA SECURITY AND CONFIDENTIALITY

FINAL REPORT

to the Commission for the European Communities

Volume 5 of 6

Section 5: Technical aspects of the right of access

by F Bancilhon
L Joinet (counsellor)

JANUARY 1980

Contents of all volumes

Contents of section 5

5 Technical aspects of the right of access

5.0 Introduction

All national laws and draft laws, and international organisations' draft directives, seeking to protect individuals' privacy from abuse by computers, contain, in general, two types of provision: those seeking to ensure what one might call the 'right of secrecy', ie to prevent too great a disclosure of personal information about an individual; the others would come under the heading 'right to know', seeking on the contrary to ensure that the individual has free access to information referring to him which is stored by other people.

The recent realization of the possible dangers of computing, and the various laws, draft laws and international directives which have resulted, date only from the end of the 1960s. At this time, computing was 20 years old, and had developed independent of such legal concerns. However, one technical aspect was already highly developed, and was to be directly usable in implementing of this legislation: the area of protection and security. In fact, technical progress in protection and security directly contributed to a better implementation of the right of secrecy, and that is why the scientific and technical community has long ago answered the legal efforts by work in these precise

areas. At the same time, one can state that the scientific or technical counterpart of the right to know has not been developed.

In this study, we therefore intend to explore on a technical level, what could be the implications of this right, which some have said represents the only real novelty in these laws.

First let us briefly consider what the right to know consists of. There are four essential elements:

(1) the right to be aware of the existence of the files: this is the right of the public to know or to get to know of the existence of all files containing information on physical persons, whether held by the state or the private sector;

(2) the right of an individual to be informed of the existence of information referring to him in a given file. This right is distinct from the first one in that it is an individual right and not a public one. Also it is perfectly conceivable for one of these rights to exist independently of the other. For instance, we shall see that the 'subject notification' method satisfies the second type of right, but not the first type. Conversely, it is conceivable that the 'publicly available list' method might satisfy the first type of right, but not the second.

(3) the right of an individual to know the contents of the information which refers to him in a given system. This is obviously the essential part of the right to know, and it motivates the main part of our study. The right certainly presupposes the existence of the right of the second type, but not the first.

(4) the right of an individual to demand correction of information referring to him which is shown to be false. This right is not strictly part of the right to know, but is such a natural extension of the right of access that we judged that this study would be incomplete without it. We shall also study, as a supplement to this right, the right of propagation, which consists of requiring (as some laws have provided) the sending of corrections to everybody to whom the information has been sent.

For each of these four rights, we shall therefore consider the different methods of application and their consequences, and the technical problems which these rights give rise to. We shall also particularly bear in mind that computing is an area in constant and rapid development, and that to base our study on the present state of the part could lead to it being rapidly outdated. So we shall try, starting from recent developments, to extract the trends for the years to come and to put the problems in these perspectives.

5.1 Technical context and long-term trends

The aim of this study is to analyse the collection of technical problems linked to the right of access produced by the laws on computing and privacy. Rather than study this problem at the present moment, ie taking an instantaneous view of the state of computing methods and techniques, it seems necessary, in a field which is well known for its very rapid and deep development, to understand what types of developments will occur, so that the conclusions we might draw are not invalidated by new technological advances.

To understand these developments, we study section by section recent developments in computing. It is not a complete study, but simply a case of extracting the elements which relate to our specific problem: the right of access.

To present these developments, we divide computing into three. Computing involves (1) storing information, (2) processisng it and (3) communicating or circulating it. On the other hand, adopting the classic hardware/software division, we shall study each of these sectors in turn.

5.1.1 Hardware development

In telecommunicatins, the greatest technological impact is indisuptably due to the appearance and gradual spreading of data communications networks.

The large networks. By these one means the networks which cover a geographical distance of at least hundreds of kilometers. It is commonplace to say that the recent technological progress in this field, and particularly the appearance of the packet-switching technique, have contributed and will contribute to an accelerated development of networks in diverse sectors of application: specialised networks (available to a one class of user), private networks (internal to a company), and public networks (accessible to anybody). The long term effects of this development will be:

(i) standardisation of interfaces. Because the networks connect mixed hardware systems, the interfaces must be standardised. This trend will be more marked in the public networks;

(ii) standardisation of file formats. When networks link two systems which use different software, two methods can be applied: conversion of the local files to the same type as the files on the site to which they must be communicated (and great effort is invested in the problems of conversion), or standardisation of files, which would enable any file to be used on any site;

(iii) an increase of data circulation, and a consequent increase in the amount of data stored; partly because more circulation leads naturally to more storage (each site naturally stores the data it

receives), and partly because all data which goes through the network has necessarily been put into a transmittable form, therefore into a storable form.

Local networks. By local network, we mean one extending over a fairly limited georgraphical area (a few kilometers). These networks use two types of relatively simple connection (a loop or a star shape); they are usually intended to interconnect different centres of the same company. Their effect will be to increase data circulation within the organisation. At present, in an organisation (administration or business) of a certain size, the information is kept at relatively partitioned sites, without anybody having a global idea of the system or its possible uses. This information is therefore much less used then it could be. For instance, in one university, each department (lectures, registration, expenses, accommodation, medical services, etc) holds and manages a file of students. These files are not connected, ie updates (changes of address for instance) are not transmitted from one file to another. Therefore the accuracy of the data is poor, and each department has access to only a small amount of information. The appearance of a local network in such an environment will improve the quality of the data, and also will increase amount of data accessible by the administration as a whole, and increase its complexity. More generally, one

can say that the effect of spreading the local networks will be a dual one: in highly decentralised and relatively partitioned structures, there will be a tendency towards integration. Conversely, in highly centralised structures, there will be a tendency towards distribution, that is, the network will act as a distributer, putting information within reach of those concerned. In both cases, the information will circulate more, will be more reliable, and also more complex (by sharing of miscellaneous date.)

Mass storage. All the infomation discussed in this study is at present kept in so called mass storage. These stores are most frequently magnetic tape or magnet disc. Data which is regularly processed, accessed or transmitted must be stored on disc (at least when it is processed): only discs have acceptable access speeds. The technical feasibility of this storage depends on the characteristics of the equipment on the market, essentially: capacity (how many characters can be stored on one disc unit?), speed (how long does it take to gain access to an item of data on disc when one knows its address?), throughput (how many characters per second can be read?), and price (price of purchase or hire, cost of maintenance, of air-conditioning, etc.). It is certain that the quantitative development of data storage depends on how developments in these characteristics come onto the market. Thus it is appropriate to examine, for each type of mass storage, what variations can come.

Conventional discs. For low-capacity (floppy) discs, as much as for high-capacity ones (hundreds of millions of characters), without revolutionary changes, one can observe a constant improvement in performance at constant price (roughly double the capacity and throughput every 18 months, access time remaining constant). This regular decrease in the cost of storing information must help to speed up the creation of files.

New technologies. Can the appearance on the market of the 'new technologies' (bubble, CCD, RAM) revolutionise information storage methods? There is some doubt about bubble memories, which, already six years old, should replace conventional discs ... in two years time. Progress of integrated circuits will bring the first RAM discs (semiconductor memories), which are clearly promised a great future. Also we must mention the appearance of the first so-called 'intelligent' discs, ie those with a computing capability, although this is not really new technology. General use of such equipment would lower the cost not only of storage but also of processing.

Optical discs. Particular mention should be made of the optical disc, which, in our opinion, can represent the great innovation of the early 1980s. It is simply a digital form of the video disc already used for storing television images. One writes to the disc by perforating the film with a laser beam (this writing is irreversible,

ss that one ca write on it only once. This may seem to be a major limitation, but in fact the ridiculously low cost of the medium means that one only needs to find new methods of management to meet this new technical problem). The precise technical specification is still secret, but is of this order: the disc itself would cost FF 100, the reader about FF 10,000. To this, one must add that the unit operates in any environment, ie without expensive air-conditioning. The disc capacity would be of the order of a billion characters, the throughput of the order of 10 megabits per second. The drastic fall in storage cost which the introduction of such equipment onto the market will represent, will open storage possibilities in new sectors (archives, for instance). Anybody will then be able, for a minimal cost (FF 100!), to store enormous quantities of information (1 billion characters represents three years of a daily newpaper!).

5.1.2 Software

In software, the progress is neither as rapid nor as impressive as in hardware. One can consider the three following facts essential:

(i) Software science is still very imperfect: the development and maintenance costs of software are still inhibiting. The complexity of the systems software (operating systems and data base management systems) is increasing. Its reliability remains uncertain, and a great deal of maintenance is necessary.

(ii) Previously, data management systems were marked by a dual development: (1) first an integrating phase, where, for the sake of the quality of the information and its integrity, there was a tendency to integrate as much as possible all the files of an organization, and to control them together; this phase corresponds to going from fileing systems to data base systems; (2) then a distributing phase, where, because of cost, availability and the improved circulation of information, the data is distributed to the sites where it is used. This phase (which is at its beginning) corresponds to going from data bases to distributed data bases. However, the two phases have both corresponded to: (1) an increase in the availability of the information, (2) an increase in the complexity of systems (their management is increasingly difficult) and an increase in the complexity of the structure of the data (integration of miscellaneous data, decentralised management by different users);

(iii) there is still a very great delay between research (by the manufacturers) and practice (by computer users): many users are still on the filing systems, much computing is still of batch type. The system possibilities for structuring and manipulating the data (high level language) are still under-exploited, in particular the dual development described above is still very much in progress.

5.1.3 The appearance of the electronic office

Without any doubt, this is greatly expanding, and with a market which will develop strongly. It seems that the manufacturers' policy in this area is one of small steps: that is, to develop the market slowly, gradually offering products with ever-greater possibilities, without providing at a stroke the complete office equipment (which would in fact be technically possible). Thus one gradually goes from the typewriter to text processing etc. This avoids rejection because of too forecful a development, and gives companies time to make the necessary adaptations (training of staff, etc). But it also avoids any possibility of going back. What will the long term results of these changes be? It is difficult to evaluate them, but it is clear that more and more information will be put into memorisable, computerized or trasmissible form. Moreover, this concerns information which, up to now, was not computerized: archives, mail, conversation, notes, etc. Therefore it is a completely new area of information which is affected by computing.

5.1.4 Attempts at standardisation

A major restriction on the interconnection of information systems up to now has been the great variety of hardware and software. In fact, even when one has recognised that in a site A and a site B, there is, in one redundant

information, and in the other mainly complementary information, so that there would be an interest (at least from the point of view of those who store and use this information) in integrating or interconnecting the two sites, the cost of this operation is usually prohibitive. But this cost is largely derived from the incompatibility of the software, the hardware or the communication protocols.

Attempts at standardisation are thus necessarily in the context of getting rid of this type of barrier. For a number of years, this type of effort has been made in two areas. First, telecommunications: standardisation of communication interfaces, transmission protocols; then data bases: standardisation of languages, particularly data base description languages. It is certain that such attempts are exposed to political and economic problems, but such long term tasks sometimes succeed (cf COBOL). It is also certain that this applies to the developments we have discussed: storage of more data, and more complex data.

To conclude this quick survey, one can say that the most recent technical progress will have these effects on computing practices:

(1) more data will be stored, whether because the same quantity will be accessible to more people, or because new types of data will be in a suitable form for storage;

(2) this data will have a more complex structure, because it will depict a more complex reality, and because it will come from sharing several sources of different natures. This implies for instance, that the concept of a file made up of a series of recordings of the same type will be replaced by more complex structures, which are also more difficult to grasp.

Having clarified these two points, we now study one after another the four components of the right to know.

5.2 Should the public know of the existence of files?

Here we are concerned with the right of access of the first type: the right of the public to know of the existence of files. The general idea is that the public should be able to acquire a good knowledge of the level of filing. This therefore implies that each individual should be able to access the following information: a list of files, content of the files, current processing carried out, people referred to, number of files, size etc. Thus it is within the scope of this right that, for example, the press or consumer associations should be able to access this information.

One should note that this right is different from the right of access to personal information, and also from the right to know whether one is referred to by a given file. In fact, if each individual's knowledge of files were limited to those which refer to himself, then (1) each individual would have only a very limited view of the level of filing, which would help to keep some security in this area (at least to that individual); (2) the work of the press, whose job it is to spread information, and the researchers concerned with the problem, would be practically impossible because of the small amount of information to which they would have access.

To enable an individual to have access to a particular item of data, one must either impart this information to him, or tell him one or several places where he can go to find it. In the present case, the information represents a very great amount (at least a list of all files). Thus it is unrealistic to send this information to each individual (by post or another method). (It is notable that for another case, the list of telephone numbers, such a method is practiced for just as vast an amount of infomation; but in this case, sending the directory to each individual who has a telephone is justified by the intensive usage made of it, and when the use is less frequent, for instance, the directory of an area far away from the dwelling, systematic delivery is no longer used. In the case which we are

concerned with, the use would not be as intensive as that of the directory. Thus it is clear that the right of the public to know cannot be satisfied by sending the necessary information to each individual (one must note that in the countries where the right of access works by notification, this notification concerns only files containing information on the individual to be notified, and thus only applies to the type 2 right; notification is therefore not adequate for type 1).

Having established that the information should be made available to the public in one or several places, one must study the means of carrying this out. For this study, four parameters are to be considered:

(i) the number of places where one can access the information

(ii) the types of places where one can access information (is there only one, or are there different places for different sectors of activity of types of information?)

(iii) the equipment used to store information, and the methods of access to this information

(iv) the content of the information put at the public's disposal.

It is not a matter of settling this question by proposing one supposedly optimimal solution, but rather of suggesting alternatives, and analysing their implications.

Let us study each parameter:

The number of information points: the choice can vary from a single point where all the requests converge, to several points (of the order of ten) in the case of local offices, or to a large number of points (of he order of a thousand) if the information points are used by other already existing organisations (post offices or town halls for instance). The problem is clearly that of allocation of costs, and of knowing whether for a given cost it would be better to disperse one's efforts or to concentrate them in a single point.

The advantages of a single point are considerable: by concentrating the methods available, one can develop a fairly sophisticated organisation, both in the quality and quantity of available information, and in the access to this information. Also one can consider that the role of spreading the information in question might be taken up by the press and consumer associations. In this context, the single information point would be a powerful tool for their use.

The problems of the single-point structure are also to be considered: access to the public is more difficult geographically , requiring this service to be accessible by letter and by telephone requests. The risk of such a structure becoming bureaucratic also exists, and this would contradict the very idea of this right of access.

The intermediate solution (local offices) reflects the disadvantages of a single point structure. With an equal service, the ocst of this will be much higher, but computer networks might be used to distribute the information.

Finally, the extreme solution (thousands of points) gives the advantage of very easy individual access, and the self-publicity ensured by the very existence of these points (though the slight awareness of the Canadian public, which still does not fully know of the existence of the list of files in post offices, may lead us to doubt this). The main disadvantage is clearly that, in this case, only the simplest means of supplying information can be made available. Also one should mention that such a mechanism also presupposes the existence of a central organisation for collecting and spreading the information.

Types of information points: In the case where there is a multiplicity of information points, one has the right to consider whether they should all be identical, or if several types of points could be set up.

For instance, one could envisage a certain specialisation of points by area of activity, (private/public sector, or, in more detail, by ministry or branch of activity).

One would then end up with a hierarchical system with a centralised information point where one would have global and general information at one's disposal. At this point, one could obtain a global view, but in little detail, of the existing files. Then there would be points (attached to ministries for instannce) at which one could obtain information in more detail on more specific areas. The obvious advantage of such a system is the flexibility, for each specialised point could be adapted to the types of request that it received.

The content of information to be made available to the public:

One agrees generally that this content should be the list of files of existing persons. Thus the problem basically consists of knowing how much detail the description of each of these files should contain. Two parameters are basically to be considered at this level: one is the quantity of information given in each file, the other is the understandability of this information. With regard to quantity of information, we shall not consider this aspect in depth here; for a more detailed study, the reader can refer to 5.4, where we consider the essential components of an information system. Here we simply mention briefly that the description of a file should contain a description of the structure and the content of the file; additional data (journals and archives if they exist) and what processing is currently carried out on this file, with particular reference to deduced data.

It is appropriate to recall th conclusions of 5.1, where we stressed the growing complexity of the structure of stored information. It is this complexity which will make the problem of description increasingly difficult. Different degrees of detail in the description are possible of course: starting with a brief description of the essential information on the file and its main purposes, which can be summarised in a few lines, up to a complete description of the schema which describes, in detail, the contents of the file (a complete description of the schema can occupy dozens of pages, and this does not include the description of the programs).

Setting aside these problems of the fulness of content of the description, great attention must be given to the understandability of this information. Describing a collection of data and above all application programs is not an easy task, and often poses problems even for computer experts.

Finally, with regard to the quantity of information collected in this respect, we draw attention to the following fact: the basis of most data protection laws is the wish to protect the individual from possible abuse in the intensive use of filing systems. Thus it is a matter of monitoring the use which can be made by companies and the state of information acquired on individuals. Without judging the way in which these laws meet their objectives, one can however say that, if their

application led to a higher level of filing, that is to say if it contributed toward making more information on individuals available to the private sector and the state, one would have reason to question whether the objective of these laws had been effectively achieved. But, it is well known that, hitherto, individuals have been partly protected (naturally) from these threats for tehnical reasons: basically by the fact that one can still only make use of a small part of the possibilities of computers, that great confusion still rules in computer practices, that information is still unreliable, and that interconnections and possible correlations (even within one organisation) are usually not made (5.1). One could sum up by saying that the individual has been protected by technical imperfections. How long this protection will last is beyong this study, but it is certain that colllcting of information about existing files and making them available to all will tend to remove this natural barrier (by reducing the entropy of the system). In this context, the cure risks being worse than the illness. We do not aim to simply abandon this collecting of information, but the risk must be considered. Moreover, measures to monitor the use which can be made of the file of files must certainly be applied (similar for instance to the measures controlling the use of the central population register in Sweden).

The storage equipment for this information and the methods of access:

This is a very complex problem. Of course the choice is closely linked with the choices made in connection with location and with the content of the stored information. The determining factor will also be that of cost. We shall not detail here the list of possible solutions, which range from the most manual to the highly computerized. It is clear that, in view of the area concerned, and the necessity of making general or specific information, statistics or trends available to the public and press and researchers, the temptation is very great to choose highly computerized solutions. In this context, without suggesting a particular degree of computerization, we would like to insist on a point which seems to us of paramount importance: numerous questions linked with the development of computing in society have recently been raised. An essential question which has motivated the data protection laws is that of filing systems. Another question, no less important, is that of the systematic introduction of computing into everyday life. It can be considered that certain computing devices, if they were conceived with this aim in mind, would contribute to an improvement in the quality of everyday life, although recent experiences have raised doubts in some minds. Information distribution points and the 'file of files' surely provide a unique opportunity to experiment with such devices? We must hope that particular care is taken in computerising these

information centres, and that, in particular, they become a demonstration of what 'human' informing could be. There is such a unique opportunity for research in this area, that it would be much more regrettable to miss it, seeing that users of such systems are already aware of these problems.

5.3 Should the individual know of the existence of information concerning him in a file?

This is a matter of offering to each individual the knowledge (or the possibility of obtaining it) of the fact that information about him is kept in a file. This right is partly distinct from the public's right to know of the existence of files, which is a right concerning all files, but it is a question here of an individual right concerning the list of files in which the individual features. Of course, it is also distinct from the right of access to the content of this same information, to which it is to some extent a preamble.

The ultimate aim is therefore that each individual has or may acquire the list of all the files containing information which refers to him; this must of course be obtainable within an acceptable cost and time delay. Thus, for instance, the solution consisting of giving the individual the list of all people responsible for the files to contact, would not be acceptable: even if access were free, the necessary time would be prohibitive.

How does one satisfy such a demand? As we have already mentioned in 5.2, there are two essential means of informing an individual of a given fact: one can on one's own initiative impart the information to him, or one can tell him where to get it. The methods to satisfy the right of the second type are classified in two categories: the notification method and the 'centre of information' method.

5.3.1 Notification methods

These consist of an obligation on the person responsible for a file to inform the individual of the existence of data referring to him in the file.

First one must determine when to inform him. If it is solely a matter of informing him of the existence of data, then it is sufficient to notify him when he is first recorded. A certain time lapse will be permissible, for the notifications will probably be batch-processed (unless it is incorporated into the methods of recording new subjects in the data base). The time lapse would then be that of a batch cycle (1 or 2 months for instance). But, if the notification concerns not only the existence of data but also its content, the problem of frequency of notifications arises.

An important problem connected with the notification system is that of confidentiality. In fact, one must realise that, in a notification system, the data flow is equivalent (at least in the case where notification of the existence of data is accompanied by the content of this data) to the notorious centralized file obtained by joining all existing files. That is to say, if this data, instead of simply passing (from the person responsible for the file to the data subject) was stored at some point, that point would constitute the centralized file. Of course, the data flow is certainly much less a potential danger than a file representing the information exchanged in this flow, but, on the one hand, it represents a threat to the confidentiality of individuals, and on the other hand, one should be aware of the phenomenon, and take all the necessary precautions. One of these precautions is ensuring that the recipient of the notification is indeed the subject referred to in the file. This poses essentially two problems:

(i) The file must contain the address of the subject, and this address must be correct. One might say that most files of individuals contain their addresses. But this is not so for 100% of files; and even if most of the files do contain the address, it is because they were set up before the data protection law, when one did not have to justify the necessity of stored data.

Therefore, very often, this address is there 'in case' it is needed. It is certain that if each holder of a file had to justify the existence of the address, the practice would become less frequent. Of course, in a notification system, the system itself justifies storing the address, but this is another case where the application of a law aiming to control the storage and processing of data works towards more intensive filing, and the feasibility of such a system should be considered. Further, the fact that the address would have to be correct poses problems of the same nature, but on a higher level. In fact, the mechanisms for ensuring the rapid updating (within a few days for instance) of changed addresses are such that they also cause more intensive filing. One can cite in this respect the Swedish experience, which ensures reliable addresses by the existence of a central population file (a type of file which the public has shown a definite revulsion from in certain countries).

(ii) Mechanisms should be set up to control the reception of information. First of course in the case where notification goes with communication of the content of the information, but also in the case of simple notification of existence, for the existence of a person's name in the file is, usually, information in itself, which can, in certain cases, be very confidential (legal or police files for instance).

In this case too mechanisms of some complexity should be used, using registered mail.

With regard to notification, a particular mention should be made of 'implicit notification'. Implicit notification consists of regarding, in certain cases, the subject to be 'implicitly' aware of being mentioned in a file, either because of a personal characteristic, or because he has consciously carried out some action which has caused him to be put into the file. For instance, one can consider that a person signing for life assurance should know that the information about him is in the file of the insurance company with whom he signed the contract; in the same way a person, because he is employed by a company, can know that he features in a salary file of that company. Let us examine in more detail the basis of implicit notification. The problem is the nature of the information and the reason for its belonging to the data base. Two cases can be cited: information belonging to an information system for a specific reason which is, usually, the system's reason for existing. The two examples mentioned above correspond to this criterion. These are in fact extreme cases of 'determinist' files, corresponding to the case 'tell me who you are, I will tell you where you are filed' (one should note that this assertion also works the other way for this type of file: 'tell me where you are filed, I will tell you who you are'). At the other

end of the spectrum is a file in which the presence of information on an individual is there 'by chance': either because this information comes from another file with which some type of connection has been set up, or because of a systematic collection of information that is not motivated by a specific reason, or finally because the reason for the filing obeys a specific rule, but this rule is so complicated to grasp that it appears more realistic to speak of an upredicatable presence (this is the case of all files which contain information on people with a more or less distant link with the central subject of the file, for instance information on members of the family of a subject of the file). To sum up, the presence of information relating to an individual in a file can be either deterministic or unpredictable.

Parallel to this aspect, one can think of the two methods of presenting information: one can give the raw information, or one can give rules enabling this information to be deduced. For instance, one can say 'there is a record in your name in the salary file of the Dupont company' or 'if you are or have been employed by the Dupont company, then there is a record in your name in the salary file of this company'.

Clearly, the first method corresponds to the explicit notification method, the second to the implicit notification method. In view of the previous remarks, one can therefore make the following assertions:

(i) the implicit notification method applies only to information whose presence in the base is deterministic;

(ii) for the notified subject the result is the same, that is to say, he hold the same information (with a little mental effort!);

(iii) this is true only if the method of deduction is effectively communicated to the subject; thus one must find a method of making known all cases of filing where the implicit notification operates. This could for instance be accomplished by a small book describing all determinist files and the circumstances which lead to inclusion in these files. In the absence of such a collection of rules, implicit notification would be a catch;

(iv) the enormous advantage of implicit notification is that it respects the confidentiality of each individual better than explicit notification. In fact, raw information no longer circulates, and thus does not risk parasitic spreading, as the subject himself deduces it from information which only he knows.

5.3.2 Access methods

These are methods where the data subject does not receive information but has available to him information distribution points where he can acquire the information he wants. Let us quickly consider the feasibility of such an approach. If one wants any person who applies to such an information centre to be able to obtain a list of all the files in which he is mentioned, it is necessary for a file containing a list of all the files to be at this point, and, connected to each file, the list of data subjects (or rather the list of all the people with, for each person, a list of files in which they are mentioned, which comes to the same thing, and is only different in its structure). We believe that such a solution should be rejected without hesitation, not because of expense or technical difficulty, but because it would constitute an accummulation of data, whose dangers are obvious. If the existence of named information in the 'file of files' must be excluded for security reasons, what solutions can be conceived? There are simply the methods which enable, from the description of the file, to deduce whether or not a given individual is present in the file. One can make two comments on this:

(i) here one finds exactly the same problem as with implicit notification, that is, one provides the applicant with the rules for belonging to a file, and he can then deduce, from information which he holds on his own case, the existence in such and such a file of information concerning him. The same precautions apply therefore to this method: it is only effective for 'determinist' information;

(ii) it is clear that the mechanisms set up to satisfy the right of type 1 (the publis's right to know of the existence of files) will be used again here. Thus it would be conceivable to duplicate the management system of the file of files by an interrogation system enabling an applicant to determine, from some of his characteristics (not his name, of course), a potential list of files in which he is mentioned. Precautions should be taken to prevent the storage and archiving of applicants' questions. Also, it will be necessary to give a clear explanation of the operation of the system and to state its limits.

5.4 Should the individual be able to know the information about him in a file?

This right is the heart of the right of access. It aims to reach a state where each individual would know, or havd means of knowing what information concerning him is

stored in all systems. It should be clear that by 'have means of knowing', one understands that this right can be exercised at a reasonable cost, and by cost, we mean financial cost, time cost and cost of effort in understanding. That is, every individual should, for a moderate expense or none at all, without having to devote a great amount of time to it, and without being a genius at deciphering administrative formulae, be able to exercise his right of access.

At this stage, two preliminary observations can be made:

(1) The type 3 right of access supposes the existence of the type 1 right of access: only if any individual knows or can know, for any system, of the existence of information concerning him in that system, could he know its contents. Therefore, the implementation of the type 3 right of access presupposes that one has solved the problem of the type 2 right of access.

(2) In the same way as for the type 2 right, there are broadly speaking two methods of implementation for this right of access, the notification method and the request method. In the notification method, the person responsible for the file, on his own initiative, communicates to the data subject the contents concerning him. In the request method, the

data subject who has to make a request to the person responsible for the file, this request being answered by the contents which concern him being sent to the data subject. This is a question of basic choice on which we shall not attempt to make a decision. We content ourselves with studying here the implementation of the right in these two cases. These two methods of application are partly different and partly the same:

- the initial phase of the request is specific to the request method; in this phase, the data subject identifies the file or files, and the one person or persons responsible for the file(s), then he formulates and sends his request;

- when the person responsible for the file receives the request, he is in the same position as a person responsible for the file who notifies it on his own initiative; that is, he knows the identity of the data subject, and he must send him the information he has on him. To do this, he must solve three problems: first, he must define the information about the data subject, secondly, he must retrieve it, and finally, he must send it.

These two preliminary observations justify the following action: first we shall study the problem of setting up a request which is suited to the request method, then we shall study the method of access, dividing it into three phases: definition, retrieval and sending of the information; for these three points, it will be necessary to begin by recalling to some extent the way information is represented, structured and managed in a system.

5.4.1 Initiation of the request

To illustrate the different type of problem which varying organisations can meet in the initiation of a request, let us consider the two following examples.

First situation: the computer centre of the university of ... receives a request for access in the following form: 'I have been a student at the University of ... for two years, my student name is ... my name is ... I live at ... and I wish to know the contents of the information which you have on me in the student file and in the University Campus lodgings file'.

Second situation: The French Minister of Education receives a request for access in the form: 'My name is ..., I would like to know the information which you, or the department you are responsible for, have on me'.

By quickly analysing these two requests, one can immediately see that the first will be relatively easy to carry out, that is to say that the cost of the reply will be very low, but the second request will necessitate a very great effort, and with a result which may not be satisfactory. Let us try to see why.

Four characteristics distinguish the first request:

(i) the person who is responsible for the file is identified well, ie the person to whom the data subject addresses himself is the person who has a good knowledge of the system. On the contrary, in the second request, all that can be hoped for is the existence of an organisation capable of sending this request to more local levels;

(ii) the system and parts of the system of interest to the applicant are identified well. The data subject does not ask for everything, but specifies what interests him. On the contrary, in the second example, no indication is giien and the reply supposes that there is a complete list of files and systems dependent on the minister in question;

(iii) the data subject is identified well for retrieval, ie enough information is given to enable the data to be accessed (the fact that he is a student and his registration number, for instance). In the second request, one does not know what type of individual it is: is he recorded as a parent of a pupil, a teacher, a student ...?

(iv) the data subject is identified well, by evidence; that is, enough information has been given to assure the person responsible for the file that the request has indeed been made by the individual in question, and not by somebody else. On the contrary, in the second request, anybody at all could have wanted to obtain information on the individual in question.

To sum up, it will be relatively easy to reply to the request if the file and the person responsible for the file are identified well, and if the subject of the request is also identified well. The type of request ill depend of course partly on the arrangements for applying the law.

The law can in fact fix methods of exercising the right of access. Thus it is the law or custom which will decide how precise the descripion of the file and the data subject must be.

With regard to the precision of the file description, it is only in the case where the applicant can obtain a precise and clear description of the file which exist in a certain sector that he can be asked to refine his requests. We see here the interactions between the rights of type 1 and the type 3. The clearer, the more complete and widely spread the description of the files and their structure is, the easier it will be for the file holder to satisfy the right of access. If no description of the files is given, then the question will be very vague (and probably more numerous). Besides, solely from the point of view of cost, the general description of the files will only have to be made once (when the file is started, or on the date when the law starts to apply), whereas an expensive search through all the files will have to be made for each request.

With regard to precision in the description of the data subject, we have already seen that two types of precision are necessary: those connected with the proof of identity (5.3) and those connected with the problem of searching for information in the system. A problem arises at this level: the information necessary to facilitate the search varies from one system to another, so fixing by law the particulars to be given in all cases, would lead to the applicant being asked for more information thatn the file usually contains on him.

Therefore, in general, only the person responsible for the file can tell the applicant what particulars are necessary. Thus it is still on the level of file information that the problem rests. By describing the structure and the content of his files, the person responsible for them will be able to make his access keys known.

5.4.2 Reply to the request for notification

Assuming that the problem of the request is solved, we are now in the position where the file holder has received a request by a data subject , or must, by law, notify him on his own initiative. As we have previously explained, the reply can be divided into three phases: (1) defining the information, (2) retrieving the information, (3) sending the information. To analyse these problems, it is first necessary to understand how information is represented, structured and managed in a system.

5.4.2.1 The essential components of a system

For the problem we are concerned with, in an information system we ca distinguish basically three parts:

- basic date (that which the system aims to store)

- auxiliary data (this is additional data necessary for correct operation of the system)

- programs (these are applications and uses made of the data).

We do not claim that such a subdivision is general, but it shows all the aspects to be considered in understanding the technical problems which access to information raises.

Basic data: This represents the real world. It has a certain structure. The creation and restructuring of the base is the task of the 'data base administrator' (whether this administrator is one person or a group). It is he who, using the information communicated to him by future users, decides on the structure to be given to the base.

To gain an idea of the complexity of the structure, it is adequate to know that in certain systems, the number of persons consulted to decide on the strucute can be up to a thousand. Using this 'information about information', the administrator can make a choice of structure, that is to say that he decides on the existence of a certain number of entities (for instance, in a company there will be employees, divisions, orders, etc), each entity will be characterised by attributes (for instance, for an employee: age, address, social security number, etc.).

Then, there will be relationships between these entities (an order made by an employee, an employee belonging to a department, etc). It is important to note that there is no single best way of structuring a base:

- no single way, because one can represent certain information in different ways. For instance, the belonging of a student to a given course can be represented by a 'relationship' between the 'objects' 'course' and 'student', but also as an object which could be called 'registration', and which would have among its attributes a student's name and a course name. There are many examples of such possible alternative structures for the same 'facts' about the real world. The administrator is therefore faced with a multiple choice situation to structure his base;

- no best way, for in fact there is no single criterion in favour of one choice but a multiplicity of possible criteria:

  performance, ease of updating, of access, of formatting, ...

In order to set up this base structure, the administrator uses a declaration language, one suitable for the data base management system which the administrator uses, and part of the data base software supplied by the manufacture.

The data base structure description in the declarative language constitutes what is called the schema.

Auxiliary data. Roughly speaking, one can distinguish two kinds of auxiliary data: that which assists the operation of the system (ie which helps to increase the time of access to information), and that which helps to ensure a more accurate functioning of the system.

Data with the objective of improving performance: this is basically an index which speeds up retrieval on criteria. For instance, in a file of individuals, one can index the file by the name of the individual. Thus, when one is searching for a given person, one search in the index of names will give the address of that person's record. The same file can be indexed by each other attribute (age, profession, etc), or by several attributes at the same time. Indices increase the access speed, at cost of the space required and the time lost in updating (the indices have to be updated whenever a record is added or deleted).

Data for ensuring accurate functioning. No information system is reliable: it depends on equipment which is liable to breakdown, on data base software which contains errors, and on user programs which also contain errors.

When a software error, a hardware breakdown or a hangup of some sort occurs, work is stopped in some functions, and it is usually impossible to start again exactly in this state (especially as this state is usually 'incoherent'). Recovery mechanisms must be provided. These are usually more or less complex, depending on whether one is in a batch or conversational environment. But generally, they require logging data, which stores all the transactions carried out on the base, and duplicate data, which is a 'snapshot' of the base at a particular moment.

Physical and virtual information

Having briefly considered the essential data of a system, it is necessary to get to know the different ways information can be represented in such a system, in order to understand the problems of the right of access. Largely one can distinguish between information physically presented in a base and virtual information, which is not present but which can be extracted from the base.

Physically represented information: This is the simplest way of representing information. One simply writes the information one desires to represent in a specific record. For instance, the individual's social security number is written into a record associated

with the individual in question (this information is coded, but it is easily decodable). The result of this method is that the data is easily accessible, checkable and readable. For instance, if an individual asks for information concerning him, it is easy to read this list of data.

Virtual information: This information which is not physically present in the base but which can, at any time, be reconstructed. A certain number of obsevations can be made about this type of data: (1) the amount of this deduced information is potentially infinite. In fact, each programmer can, each time he writes a programe, deduce new information from the base; (2) it seem unrealistic to want to log this information because, firstly, the amount of it varies constantly (it seems unrealistic also to note each new application), and further because even in a not-very-large organisation, it is not always possible to log application programs; (3) regarding information connected with a person, one can distinguish between information with and without added value: information without added value is that which the person responsible for the system can deduce from the data referring to the individual. Therefore the individual himself can also deduce it. For instance, in France, knowing only a person's social security number, one can deduce his sex, the year of his birth and in which area he was born.

In the same way, from a person's taxable income and the number of dependents, one can deduce the amount of direct tax he pays. On the other hand, the value-added information is deduced by the person in charge of the file from the data referring to an individual and other information. In this case, the individual in question could not reconstruct this information himself. In this case one can say that the file knows more about the data subject than he does himself. This type of information covers very varied cases startings, with the most simple. If knowing pupils' marks, one can deduce their assessment (which they themselves do not know); knowing a company's production, one can deduce its share of the market, etc.; and the most complicated cases: profile programs, which determine, from the list of a doctor's prescriptions, whether he is 'normal', or more sophisticated programs for studying correlation which, using age, number of children and monthly electricity and gas payments, determine whether one would be a good or bade customer!

In all these cases, there is information on individuals which is not represented physically in the system, but which is available at any time to the holder of the system.

5.4.2.2 Defining information about an individual

The first problem which confronts the person responsible for a file is when he receives a request for access, or when he has to notify a data subject, is that of definition: what information referring to the individual in question is the system?

First let us set the problem in contect. A data base or a collection of files is in fact only an attempt to represent reality. Thus one can reasonably start by trying to define, in the 'real world' what actually is information referring to an individual. One can classify information concerning an individual in the following way:

(i) first, the characteristics of an individual which concern him on his own, and which he knows; eg his age, his place of birth, his salary, etc. This poses no problem a priori

(ii) a second category includes characteristics or an individual whether judgements on him by a third person, or data deduced about him which he does not necessarily know (one sees again the concept of deduced data with added value). The problem of this type of data is not a problem of definition, but a problem of conflict, as some interpretations

tend to treat this type of data as statistical secrets (for deduced data) or protect it to respect the confidence of the person who had made the judgement (for opinions on the individual).

(iii) the third category concerns the description of objects or entities which are directly linked to the individual in question. For instance, the details of his house, the loans he has obtained. There is no great problem here, since the entity in question is connected only with that individual.

(iv) finally, the fourth category concerns the description of entities which have a connection with the individual in question, but which are either from other persons, or are linked to other individuals. For instance, the company where a person works is certainly an entity related to that person, so should one consider that a description of that company is a piece of information about that person? The reply depends very much on specific situations: in the case of a private company, one is tempted to reply in the affirmative, but for a large company, one would say no. In the same way, if one says that the description of a small company is part of the information concerning the director of the company, should one include a description of the employees in the description of the company?

To sum up, we could say that there s a field of information which is clearly defined, a confused area in which only a pragmatic approach will permit a decision, and a conflicting field where the problems are not only of a legislative nature as they are connected with statistical secrecy and conflicts of confidentiality.

Now let us consider the problem of defining the information in a system. It is certain that all the 'real world' problems will arise in one form or another. The person responsible for the file who has to define this information must, as we have already shown, take into account

- basic data (files)
- auxiliary data (logs, archives)
- programs.

Information in the basic data

The administrator should start from the description of the schema (5.4.2.1) in making his decision. Then he will have in front of him the collection of entities (persons or objects) and their description, and all the relations between them. Then the problem is similar to that of definition in the real world.

One must identify the type of applicant, that is, decide by which entity he is represented, and the description of this entity that belongs to the retrived information. Next, one

has to identify which among the other entities have the relationship with the individual such that the description is part of the information retrieved. For this retrieval, only a pragmatic common sense approach will lead to an acceptable result. In the present state of system, such a step will not pose too much of a problem at the time. In many cases, the files in which one is interested have a relatively uniform structure with a record connected with each individual. However, as we have already shown, (5.1), by integration of files and the extension of data storage to now areas, the structure of information will increase in complexity. Above a certain level of complexity, it is conceivable that examination of th schema will not be adequate to define information, and then only an interactive search would be practicable.

Information in auxiliary data

Let us recall that this auxiliary data is basically made up of transaction logs (if they are kept), out-of-date copies of the base, and archives (these archives include for example, files which come from an outside source, having been used to update the base). This data contains information on individuals, and the right of access should apply to it without restriction. In fact there is a trend, among file holders, not to count such information as 'accessible'. Nothing can justify this exclusiion: either these files are used, and thus the subject should be able to exercise the right of access, or they are not used (this argument is sometimes put forward to exclude this data) and can be destroyed.

Information on application programs

Thus information essentially gives rise to two problems. First the legal problem of knowing if the information should be transmitted to the applicant because it concerns him, or whether the work carried out by the person responsible for the file to produce this data makes him the owner of this data. In this study we are restricted to technical aspects of the right of access, and it is not our place to settle this question. However, it seems that it is a vitally important question, deserving in-depth study. Let us simply mention that the current legislation does lettle to approach this problem, as it is mainly concerned with recorded data, while most of this data is not recorded but generated on request (it is certain that a more restrictive law on this type of recorded data would provoke a flight towards data which is not recorded, but can be generated by a program).

The second problem is of a more technical nature. It concerns the difficulty of getting hold of all the programs. As we have said, the list is without limits. These programs are written by a large number of people, and there is not always a complete list of operating applications, especially in large organisations.

5.4.2.3 Retrieval of information connected with an individual

Having specified the information relative to an applicant, one must extract this data from the base: this is the problem of retrieval. First let us recall briefly the structure of a filing system (or a data base). The basis is usually made up of a collection of files. A file is a collection of recorded data with the same structure. The files are distinguished in the structure by the ways in which they are accessed. A file may be sequantial, that is to say made up of an ordered set of recordings which can only be accessed one after another. A file can be indexed, ie a key (for instance the name or social security number in the case of recordings which represent individuals) enables direct access to the record corresponding to a given value of the key. A single file can be indexed by several keys. In addition, a certain number of links exist between the files which enable one to move through the base. For instance, one can have a custome file and an order file, each record on the order can have a link with the record of the corresponding customer; in the same way record of the customer can be linked to a list of orders made by the customer in question.

The first question to settle is whether retrieval should be made in batch processing or real time. This depends on two factors: the circumstances in which the retrieval is done, and the costs incurred. With regard to the circumstances of retrieval, three cases can be considered: (i0 a notification system, (ii) a request system in which one receives a request by letter, (iii) a request system in which one receives a request expecting an immediate response (ie the applicant appears in person, or telephones). In the two first cases, one has the choice between batch processing and real time, in satisfying the time limited for reply which are fixed by law (frequency of notification, or maximum response time). Within these limits, the choice depends on cost criteria. In order to evaluate the cost, one must know the computing time and the number of disc accesses necessary to satify the demands, first in the case of a single request, then in the case of a set of n requests. The cost is not usually a linear function of n, the marginal cost of an additional request tending to decrease. This retrieval cost will be evaluated taking into consideration the means of access which exist in the base (index and links) and the particulars provided by the applicant(s). Next, the frequency of the necessary accesses will be evaluated: in the case of notification, it will be fixed by the legal frequency, in the case of the request method, only measures or estimates will enable it to be evaluated.

Above a certain frequency of demands within a given period, it will be more economic to store all the requests during that time, and then to reply to them by a single search through the base. On the other hand, if the number of requests is very small, as is expected from first statistics on the exercise of the right of access in countries like Sweden or the United States, one must simply execute the retrieval program on request. It may well prove that total notification of all the individuals in the system is less expensive than a system of response to individual requests, especially if such a program can be included with other access procedures to the file. For instance, in the case of an insurance file, notification could be made at the time of the annual renewal of the policy.

Now let us come back to the case where one has to give an immediate response. It is then necessary to access each record directly (that is, one cannot peruse an entire file to find information about of person, except on very small files, as it would take too long). Thus, there must be a direct access key to the file(s) containing records relative to the applicant, and paths of access from these files to other files containing information directly or indirectly connected to the individual. Only such a system will permit a response to the applicant in real time. One should note in this respect that a

reorganisation of a system in order to gain a better right of access is improbable, partly because of the small number of requests expected, and partly because of the enormous cost of such a reorganisation (which can be reckoned in man years).

5.4.2.4 Presentation of the results

Once the information is defined, then retrieved, it will be communicated, ie one must write it in a comprehensible language, then send it or present it to the applicant. In countries like Sweden which have practised the right of access for a long time already, the way in which these results are presented has not always been entirely satisfactory. This may appear surprising: is not computing an information science, should it not have mastered communication mechanisms? But, in a particularly simple case of communication, where the 'computer experts' must transmit the data they have stored on an individual to that individual, there are serious comprehension problems for the individual who is notified.

One can find explanations of this surprising phenomenon: first, the computing community has created more communication problems than it has solved, for instance the technique of program documentation and its transmission between programmers has still not been mastered. On the other hand, computer experts have the

natural tendency of any group with specialized knowledge to keep it to themselves, in order to defend rather than share the power given by this knowledge. In this way the behaviour of the computing community has contributed towards giving the computer 'magical power', rather than taking away its mystery. Thus some people say that the user must make an effort to understand the computer, but the view should be reversed: the computer is there so that the user does not have to make an effort. To illustrate these two approaches, we given an example: to inform an individual of the content of a record, one can send him the following message: 'the content of your record is F4 17 59 14 153' with a notice (written as small as possible) explaining the codes 'an F in column 4 means etc ...'. Once can extend the process by writing the translation of the codes on the back of the record, which has the advantage of forcing the user to turn the page between each code name.

A better approach consists of using a program to print in plain text, with the necessary commentary, the decoded record. Such a program is very simple, so that developing it is not a problem, and the execution time is minimal. Of course, the moe data bases increase in complexity, the more this communication problem will become complicated. In this example we have only mentioned the case of a file with a very simple structure.

If the information is spread over several interconnected files, the explanation will be more difficult. But we think that at present the problem is not a technical one, but rather a problem associated with the view one has of what man-machine relationships should be.

5.5 Can the information be corrected?

We are now concerned with the fourth type of right of access: the right of correction which is sometimes associated with it. It is certain that this right is not strictly speaking an integral part of the right of access, but it is such a natural progression that it is difficult to separate them: in fact, a right of access at the end of which one could not exercise a right of correction if the data where questionable would be difficult to conceive. We have therefore decided to include in this report a brief study of the right of correction and the right of propagation.

The right of correction consists of allowing the individual to demand correction of the data which refers to him, and which he knows is stored in a given system.

First one can consider the question of knowing what is correct data: memorised data is in fact only the representation of a fact from the real world; let us say, therefore, that at present, data is correct if it is a

faithful representation of reality at this moment. It is essential to realise that, taken as a whole, data is not correct in a system:

- first on a 'microscopic' level, for the complex system passes through numerous updatings which give it a state of incoherence;
- for reasons linked with reliability of equipment (store, peripherals, communication lines);
- for reasons connected with the reliability of the software, first the data-base software, but basically the application programs which usually are uncontrollable;
- because of time lapse to be considered: because there is always time btween the moment when reality changes and when this change is taken into account by the system (ie change of address);
- for reasons of collection, because errors may creep in all along the chain: at the time of collection itself (filling in of a questionnaire), at the time of coding, etc.

It is therefore in the interest of the person responsible for the file to improve the quality of the data by multiple control mechanisms. From the viewpoint of the person responsible, the right of correction can be considered as one of these controls. But for this it is necessary that: the individual may have access to this data, that is, the right of access of type 3 operators in full (for instance implicit notification is not to be used).

It seems that the correction of data does not pose any technical problems in its execution. Usually it is a question of the updating of information, which is a standard procedure in most existing systems.

The right of propagation

This right consists of enabling an individual who finds an error in his data not only to correct it in the system, but also this correction is communicated to every person to whom the information has been transmitted. Let us examine the procedures necessary to satisfy this demand.

Two situations can be possible: (i) access to information in the system is not limited, that is to say that the list of persons (or systems) having access to the data base either is not limited, or is too large or too variable to be kept up to date; (ii) access to data is limited to a restricted and well-defined users or systems.

In the first case (which is not necessarily excluded by data protection laws, eg in Sweden), the list of all the people who have had access to this information must be available at each demand for correction. This presupposes that one sets up a storage mechanism which, at each access, will note the information accessed and

the applicant's identity (thus every applicant must be identifiable); the upkeep of a very detailed log will be necessary. At first glance it seems that the cost of this (storeage space), updating time, management time) is disproportionate to the advantage it gives. Thus the conclusion would be that in such a situation (which in fact should concern only non-sensitive date), the right of propatation is unworkable. In the second case, one can choose between two approaches; the first is to apply the previous method, which is feasible if the number of potential sites to which the information has been circulated is low. The second consists of storing nothing at all, and circulating the corrections to all the potential sites when they are to be corrected. The choice between these two solutions is made on two criteria;

- the first is consideration of cost; it is then necessary to evaluate the frequency of access, the frequency of correction, taking account of the number of sites and comparing the cost (computing time, disc access and transmission) for each of the policies;
- the second is the criterion of confidentiality. In fact, in the first method one must store data which might turn out to be sensitive, and in the second method there is a wide distribution of the items of data, which could pose a problem. Only a complete knowledge of the situation would permit a decision here.

5.6 Conclusion

In this study, we have examined the different components of the right of access as it is defined in data protection laws. We have shown by a study of recent technical developments in computing that the general development will be on the one hand towards a greater storage of information, on the other hand towards data whose structure will be increasingly complex. It is in this context that we have studied what technical measures need setting up to satisfy the four components of the right of access: the public's right to know of the existence of files, the right of the individual to know the existance of information concerning him in a file, the right of the individual to know the content of the information concerning him in a file, and finally the right of the individual to correct information in a file concerning him which has proved to be incorrect.

For each of these rights, we have presented the possible solutions and studied their implications. We wish in conclusion to emphasise two particular aspects.

First of all, the collection of mechanisms which must be set up to satisfy the right of access present a potential dange to privacy. In fact, these mechanisms necessitate very often storing of information which seem to contradict their purpose, which is control of the accummulation and use of data. In certain cases, like

that of the right of type 2, the accumulation of data is such that it must be forbidden: one must simply give up full application of this right. In other cases, for instance the exercise of the right of propagation, potential dangers are less evident and only a more detailed analysis would enable this to be resolved.

The second comment concerns computing devices which will be produced to satisfy the right of access: programs to reply to demands for access, a management system for the file of files, etc. It appears essential to us that particular care is taken in designing these products, and particularly that they are well adapted at the level of man-machine dialogue. This should be considered as a reseach field of the first priority.

5.7 Bibliography

F W Hondius: Emerging Data Protection in Europe. North Holland Publishing Company. New York 1975.

La Protection des données en Europe. Strasbourg 1975.

Informatique et Libertés. "Rapport Tricot" (2 vol) La Documentation Française. Paris 1975.

Report of the Committee on Privacy - Her Majesty's Stationery Office. London 1972.

Banque de données: Enterprises et Vie privée. Actes du colloque de la Faculté Notre-Dame de la Paix à Namur 25/26 Septembre 1979.

Protection des données de caractère personnel: 1968-1978: Synthèse des documents établis par l'OCDE, le Conseil de l'Europe et la CEE. Document OCDE. Janv. 79.

Submission of evidence to the Committee on Privacy (Younger Committee). British Computer Society, Privacy Committee. March 1971.

Submission of evidence to the Committee on Data Protection (Lindop Committee). British Computer Society. October 1976.

Jon Bing: A comparative outline of privacy legislation. Comparative Law Year Book 1978.

R Schomerus: The German Federal Data Protection Act. Data Regulation European and Third World Realities, conference proceedings.

P Seipel: Computing Law. Stockholm 1977.

Paul Seighart: Privacy and Computers. London 1976.