



**Dipartimento di Statistica**  
**"Giuseppe Parenti"**

Dipartimento di Statistica "G. Parenti" – Viale Morgagni 59 – 50134 Firenze - [www.ds.unifi.it](http://www.ds.unifi.it)

W O R K I N G P A P E R 2 0 0 6 / 1 1 1

Organizing Administrative Data  
for Statistical Purposes:  
a Case Study

Lucia Buzzigoli, Cristina Martelli



Università degli Studi  
di Firenze

# ORGANIZING ADMINISTRATIVE DATA FOR STATISTICAL PURPOSES: A CASE STUDY

Lucia Buzzigoli and Cristina Martelli  
*Department of Statistics “G. Parenti” - University of Florence, Italy*  
*e-mail: lucia@ds.unifi.it*

**Abstract:** In Italy the public administration is actually undergoing a very deep improvement and innovation process: in this new context local bodies are required by law to build statistical information systems oriented to monitoring, auditing, management control and government decision support.

The paper presents the working project that the Department of Statistics “G. Parenti” of the University of Florence is carrying out together with the Municipality of Florence, aiming at the organization of an efficient informative system inside the administrative structure. The project is the first experience of this kind in Italy and it opens interesting areas of development for the use of statistics in policy and management settings.

**Keywords:** Public administration reform, decision making process, statistical information system.

## 1. Introduction

Over the last few years, governments all over Europe have been engaged in projects of administrative improvement. In Italy, too, a number of laws from the 90s onwards have radically changed the relationship between the political sphere and the administrative one and have produced significant changes in the public management style. Statistical methods are gaining a new role inside the reformed public administration. In particular, the law provides for the use of statistical information systems in local management control together with a set of indicators which are considered of great importance for decision making support.

This paper stems from a working project involving the Department of Statistics “G. Parenti” and the Municipality of Florence. It aims to describe how administrative data can be organised in statistical information systems and the new role of statistics in this framework.

The content of the paper can be summarised as follows.

The next paragraph (§2) synthetically describes the long and difficult process of reform that states the foundation of a new model of public management in Italy. Our

attention is devoted to the administration at the level of local government and to the new system of checks and controls to evaluate the quality and quantity of public services.

The role of statistical information systems and the extent to which they can help to formalise, standardise and structure the massive amount of data held in administrative offices is presented in §3.

§4 and §5 give a brief presentation of the working project with the description of its various steps, while §6 and §7 present the results of the metadata repository survey.

Finally, the paper ends with some concluding remarks (§8).

## **2. The reform of Public Administration in Italy**

All over the world governments are undertaking a substantial process of administrative and managerial improvement (Pollit and Bouckaert, 2000). This process has also been labelled as “global governance reform” to stress the fact that, despite some national differences, all governments have to face the same challenges and the same paradigm of change (Lynn and Stein, 2000).

Italy has not been immune to the international trend: in particular, the modernisation process under way in our country is connected to the European one (Willemsse R. and F. Fleurke, 2005), which is going towards the definition of standards for institutional change.

In Italy the main purpose of the reform was to overcome the complex and rigid system of forms and procedures, norms and dispositions, derogations and exceptions, which regulated the civil service in the past. This transformation involved all the levels of public administration, but in our case we are interested in the effect of these new practices on local government (Ladu, 2000).

The decennial process of reform began with the Parliamentary Act no.142/90, which, at local level, differentiated the roles and the functions of political organs (which plan the overall strategy) from those of the administrative ones (which have the responsibility to carry out the goals determined by the political level). The successive legislative decree no.23/93 (modified by no.80/99) established that each department in the local administration has an amount of resources – and a corresponding budget - to gain the predetermined objectives and that the department have the consequent responsibility of gaining those objectives.

Finally, two laws of 1997 (Bassanini laws) introduced the need of computerised systems inside administrations and the principles of good management and customer satisfaction (i.e. adjustment of services to the needs of the customers) as fundamentals of a modern government.

In synthesis, it is clear that the objective of legislation was to move from a highly bureaucratised society whose main principle was that of 'legality' to a new model of public management where the method of evaluation is the result of the administrative activity, that must prove how well public services are provided.

Moreover, the rapidly merging information technologies (from Internet to database management and information systems reengineering), which have introduced new means of data storing, managing and transferring, have improved the access of citizens to information, transformed the way the organizations are structured, and intensified the process of modernization of the new public administration model.

### **3. Statistical information systems: their role in Public Administration**

Local governments cannot work without a proper informational basis (Martelli, 2002): the legislator has faced this peculiar aspect of modern governance, providing for a set of internal checks and controls oriented to system supervising and to delivered services evaluation. The law explicitly refers to a unified statistical information system (S.I.S) that contains the quantitative information to be used for management, control, strategic control and evaluation (law no.59/1997; Dipartimento della Funzione Pubblica, 2001).

Every decision supporting context, in fact, must rely on a statistical information system designed to answer to informational requirements. In recent time the distinction between informational and operational systems has been specified in data warehousing perspective (Kimball, 1996; Inmon, 1995 e 1996) and many tools and products have been proposed. Usually the concept of datawarehouse is not wide enough for local administration statistical information systems: the extension of their pertinence, the role of rules, laws and bureaucracy itself makes the issue more complicated than in firms.

Following the OECD<sup>1</sup> definition, a statistical information system is “an information system oriented towards the collection, storage, transformation and

---

<sup>1</sup> <http://stats.oecd.org/glossary>

distribution of statistical information”; despite the apparent simplicity of the definition, a statistical information system has some significant peculiarities that make its production particularly problematic: the observed entities are frequently very complex with significant interrelationships, and the role of time is usually more critical than in operational systems. Time series are always requested and very often archival data are not updated (in the sense that the new data substitutes the old one) but preserved for posterity. Retrieval of data is often “volumetric” , that is, based on two or more conjunctive range searches.

A S.I.S projecting activity always starts from the required indicators (Buzzigoli, 2002) and from the models that must be estimated; afterwards the necessary data must be produced or detected in the administrative data bases and mapped into the new statistical information system. Whenever informational gaps are detected new ad hoc sources must be developed. Therefore two are the fundamental goals that must be performed when building a local administrations oriented statistical information system: a clear image of the required indicators and a very accurate knowledge of data sources, with a particular concern to administrative data sets.

### ***3.1 The metadata repository***

The role of administrative data sets is essential for realizing Statistical Information Systems for Public Administrations: local administrations have, in fact, the availability of rich amount of data directly deriving from their daily activities. These data originate from the accomplishing of administrative processes and their content describes (even if often in a non systematic way), these processes in every single step. The lack between these kind of data and the statistical one is very wide (Statistics Denmark, 2000): diversity in coding strategies, usage linked bias, selection, and, in general, all the problematic aspects due to the fact that statistics production is only a “side effect” of their administrative informative role. This aspect, in particular, means that measuring procedure is out of the statistician’s control and that administrative concepts are not necessarily harmonized with the statistical ones. Even so, despite these serious inconveniences, these data are strategic and, for the detail level of their descriptive potentialities, they are not easily replaceable by other statistical sources: in particular, from a statistical perspective, administrative data take advantage of specific and important qualities, like the dimension and broad coverage of the collected

information (for example, the population register) and, its low economic cost (because it derives directly from administrative activity). This latter aspect is particularly true in the prospective of administrative systems able to generate harmonized and systematic operational archives, in which the phase of data transformation from the administrative to the statistical context could be positively low.

Administrative data sets have nevertheless another critical characteristic: very often they are very well known only by their users while the “system”(intended as the institution in a wide sense) ignores their existence and their informative potentialities: in terms of meta-information, administrative data sets are provided (and not always) only with some description oriented to their operative maintenance and not to their statistical exploitation. In particular, it is usually very difficult to get answers to questions concerning (i) validity and relevance ( can administrative data be used for estimating the concepts that are sought for in statistics? Are these data relevant?); (ii) reliability (do the data faithfully reflect reality?); (iii) precision (are data recorded with a degree of precision suited to the needs of the statistic?); (iv) temporal dimension and comparability over time (changes in legislation can alter data definition or content); (v) timeliness of information ; (vi) scalability issues (must the results be aggregated in some sense?); (vii) internal comparability.

For all these reasons a public administration oriented to give statistical value to administrative memories must have a good metadata repository: it will be precious, for instance, whenever arises the need of monitoring government decision consequences, because it will help to find out all the archives related with the topic of interest, along with their technical features.

The role of metadata is essential in facilitating sharing, querying and understanding the content of statistical information over the lifetime of data (UNECE Secretariat, 2000). An important issue in this respect is metadata quality (i.e. the degrees to which metadata serve their purpose). Strategies for the management, control and nurturing of metadata through metadata collection, production, storage and dissemination processes must necessarily be properly designed. The standardization of methods will be of great help and the dissemination of “best practices” could contribute to better data quality.

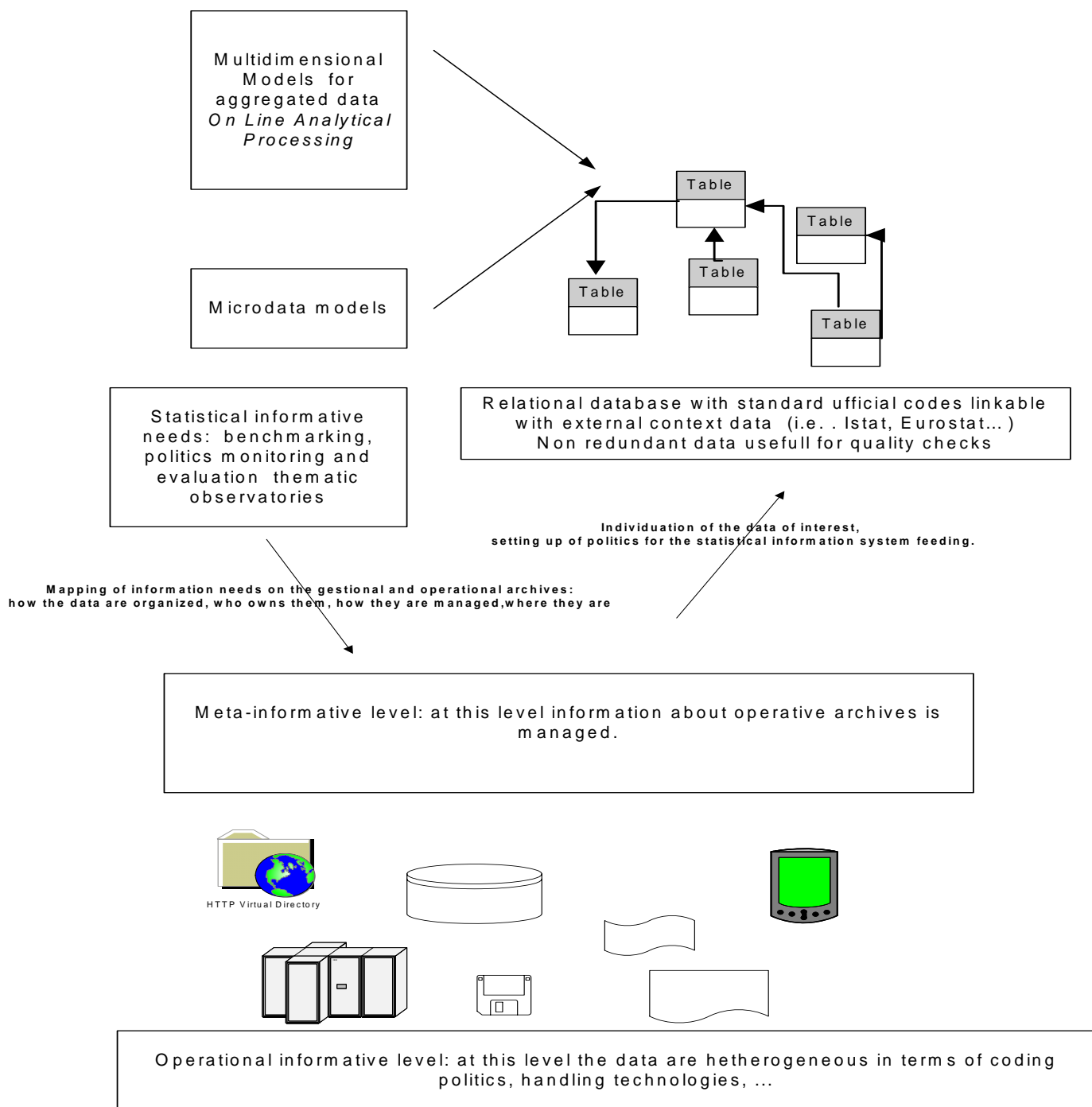
Figure 1 shows the role of metadata repository respect the operational and statistical information systems.

The figure is organized around two different levels: at the bottom the operational and administrative informational level is represented. In a Public Administration this information system is usually very heterogeneous in terms of supports, coding politics, organizational levels and data possessory title. Due to this information fragmentation it is usually quite difficult to find data of some interest respect to specific statistical needs.

The figure shows how the metadata repository is functional to individuate and to focus administrative data with statistical potentialities.

At the upper part of the figure the statistical informational level is represented. This layer is thought to be very homogeneous: the operational data have in fact been restored, systematized and harmonized. The result is a statistical relational data base, that can be used both in terms of microdata and of aggregated, multidimensional data.

The schema respects the already mentioned fundamental methodological issue in statistical information system production: informative and operational systems have been kept apart, because the difference between the two contexts in terms of users, technology and requirements suggests not to overlap these two functionalities.

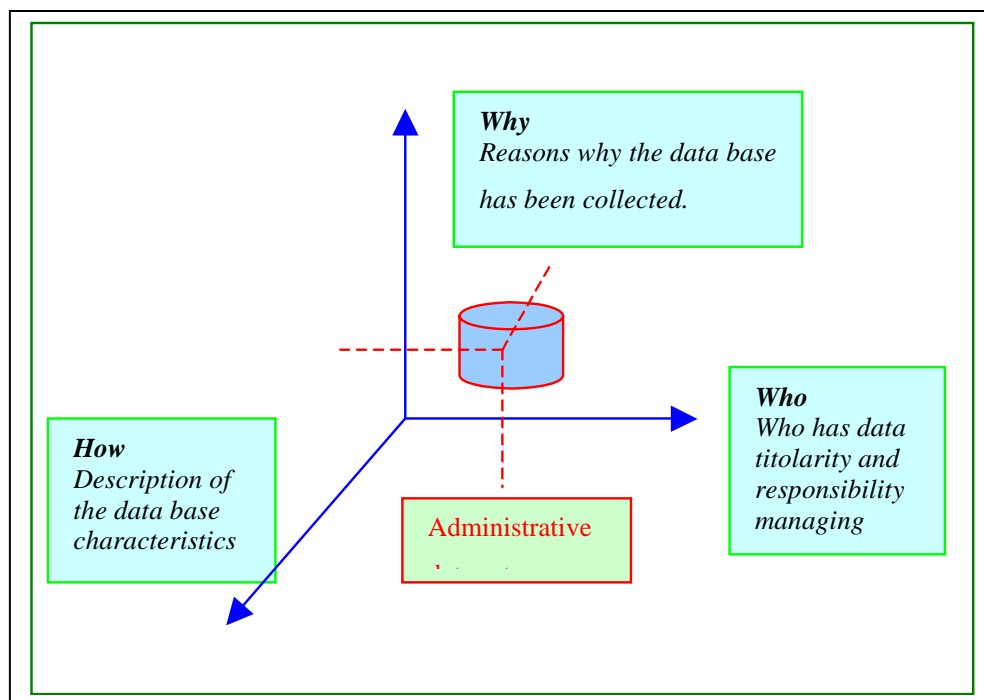


**Figure 1 The role of metadata repository between operational and statistical information systems**



### 3.2 The metadata repository survey

After having discussed the role and the importance of the metadata repository in the frame of a Public Administration Statistical Information System, it is important now to discuss the modalities that have been adopted for the Municipality of Florence metadata repository: it is important to note that, in this specific situation, administrative archives were occasionally supplied with some meta-information whose detail level was however not enough to develop a good metadata repository. In consequence of that, a survey was planned in order to describe every single archive in terms of three main axes (Figure 2). It must be stressed that such a survey was not interested in data collection, but only in metadata gathering. In other words, in this first step of S.I.S construction, there was not the interest in having the availability of administrative data but only in drawing a sort of “map” of them



**Figure 2 main descriptive axis in metadata repository**

The conceptual model of the metadata survey is made by several actors: in particular in Figure 2 is shown the interdependent role of organization, laws and storing technology to draw an archive description able to inform on situations like: “Where can I find data to study the consequences of the application of a certain rule/law?” or “Can I link together two or more administrative data sets (on the basis of the adoption of same

standard codes)?” Is very frequent, in fact, to discover potential information wasted for lack of homogeneity in coding politics that don’t allow any linkage, or for the typology of storing technology that in practice forbid the statistical usage of the data.

Through the “who” axe it is possible to set a very useful linkage between the metadata database and the local administration organizational structure. When the Municipality of Florence metadata survey has been designed, a particular attention has been paid to information linking between administrative archives and organizational data.

At the moment the possibility of adding a fourth axe is under investigation: the idea is to enrich the archive description with some links to the protocol structure of the documents recorded in the archive. One of the most serious limiting repository characteristic is actually represented by the fact that the metadata structure doesn’t allow to reconstruct in which sequence the archives are involved in an administrative procedure: this information is obviously strategic in every occasion in which the objective is to organize statistical data oriented to administrative processes monitoring.

### ***3.3 Role of the relational model in the metadata repository***

The Municipality of Florence metadata repository is based on a relational architecture. This choice was not foregone.

A statistical information system, in fact, is not oriented to managing and, in this sense, the non redundancy issue, typical of the relational approach and usually compulsory in operational data bases, can be partially released. Many authors (Kimball, 1996) underline that in informational data bases the relational approach could be sometimes too difficult to adopt, and that a multidimensional data modelling achieves a more friendly approach to data querying and retrieving.

It is well known that the four hardest thing for a data base to do are: 1) join tables, 2) aggregate data, 3) sort data, 4) scan large volume of data. By the way these are the four most often requested operations in a decision support system. In order to get around these four operations some authors suggest that the data base:1) use a denormalized or dimensional model ( to avoid the joins) 2) use summary tables (avoid aggregation) 3) store data in sorted order (avoid sorting) 4) rely heavily on indexes (avoid scanning large volumes of base table data). All of these operations require to the

data base designer to “know” what queries will be asked. Once you “know” all the questions, you can create a dimensional model that will provide high performance for these queries.

In the case of PA metadata repositories this is not the case. The repository, in fact, describes a strongly evolutionary reality. The pace at which rules, organizational context, classifications change can be, in fact dramatically high, and, in some sense unpredictable.

The biggest problem with dimensional models is that they are designed on the base of factors that are already well known: they are not adaptive to new relations that may be discovered or imposed and in fact discourage the acceptance in the domain of the informative system of such relationships.

#### **4. The working project with the Municipality of Florence**

The work project of the Department of Statistics and the Municipality of Florence is based on the preceding reflections and has two distinct objectives: first of all, it represents a significant effort to provide a coherent description of the current informative system of the Municipality; secondly, it will be used to identify the critical areas of the system, where statistical methodology can give a substantial aid in better tracking administrative activities and providing policy makers and other users with data that describe their activities and use of services.

The project work is organised in three steps, which are not independent, but strictly related to each other.

The first, long and complex step aims to describe the Municipality organization itself in order to identify the relationships among the various levels of political and administrative structures, to describe the decision process and to determine, as a consequence, the informational needs. This component could be labelled as the “informal human system” of the general information system (Land and McGregor, 2002), and is generally organised on a hierarchical basis. This is helpful to draw the diagram of departments and offices of the Municipality.

The second step concerns the survey of the data which are readily available inside the various departments to see how these data are organised in order to produce information and to form the basis of the decision making system (Martelli, 2002). This

sort of ‘data census’ involves the survey of all the informative sources and is made by means of an electronic questionnaire, which reports an analytical description of each (not necessarily electronic) file in terms of metadata: the output goes directly into a database, which will be studied together with the diagram of the organization obtained in the first step.

In summary, the first two phases should form a clear report on the current situation of the informative system in the Municipality of Florence, which is a fundamental framework to identify the critical areas and the forward targets.

The third step is the logical conclusion of all this work and could be labelled as “analysis and action”: starting from a critical analysis of all the collected data and of the organization we will try to spot weaknesses and threats, strengths and opportunities of the information system. Particular attention will be given to set of indicators used by the various departments and by Management Control Area, in order to find whether they do add value to the information system or not; this analysis will probably cause the refinement of existing indicators, the removal of the unnecessary ones and the introduction of new indicators that are important in the effective use of local monitoring systems, in order to define a coherent framework of statistical measures. Naturally, this phase can be performed only in strict coordination with Municipality departments.

## **5. The structuring of the informative basement: work phases**

From a statistical point of view, the most interesting step is the second one, that is the description of the system memory. The survey was recently completed and the results we achieved with the analysis are opening some interesting research paths.

The general lines described in the preceding sections inspired the working plan organization. Inside the administration there are huge informative resources that are still not completely exploited, containing data that are generated by the ordinary activity of the body. At present this information cannot be interpreted, but for specific cases, as a system: for this reason the accurate survey of all the file types that are present on the various departments, never done before, can be a good starting point for a correct utilization of them.

Particular attention was paid to the availability of metadata, to the check of definitions and classifications homogeneity, to the existence of relational databases, etc., also in view of the necessary quality control procedures.

The work plan is organised in three phases (see fig.5.1): the first phase is in continuous development, the second phase is already accomplished, the third one is now in progress.

### ***5.1. Phase 1: predisposition of work instruments and of human resources***

During this phase we prepared all the work instruments (conceptual, documental, technical and human resources) that can be of help to organize the survey and to carry it out: we projected the work phases, the conceptual modelling of the basement, we prepared the survey questionnaire, we choose an electronic calender to collect metadata on the timing of survey activities, and we drafted the survey manual.

In this stage of the project it was soon evident that the project represented for the municipality an important cultural innovation: for a successful result the statistical culture must be spread inside the municipal administrative organization.

The first step towards this aim is represented by the appointment of one statistical referee inside each department of the administration. This reveals the intention to give to statistics an institutional role inside the municipality and to state its transversal role in the structure.

The new figure of the statistical referee is becoming more and more frequent in Italian public administration; as long as we know an official definition is not available; we could try to define this professional role as a human resource who is able to represent a stable referee for all the problems concerning the use of administrative data for statistical purposes. As a consequence, statistical referees have an essential role not only in our project, but also in the general plan of revision of the usage of statistics in public administration.

In particular, in our case they have to build a sort of coordinated team of consultants able to form and maintain the archives, to innovate the administrative management and to cooperate in implementing the statistical information system and make its use easier.

Each statistical referee has the duty to contribute to the statistical normalization of the administrative data of his department and to maintain the informative flow which updates the general database, gaining an active role in the survey.

Special training courses for the referees have been organised that had the objective to make them sensitive to their new duties and had two different aims: propagating the statistical culture among the referees and giving them a common language for the successive implementation of an integrated work plan for the various departments.

As far as methodological issues are concerned, it is necessary to spend some words on the technical and conceptual instruments that have been arranged for the survey.

Obviously, the most important instrument is the questionnaire: its realization and informatization was rather burdensome since the data to survey are numerous and complex.

It resembles similar models used for the survey of the informative basements of the regional and provincial administrations (Region of Tuscany and Province of Florence) but, compared with them, it has been modified in order to adjust it to the specific administrative context of the municipality.

As we will see later, other modifications have been made during the survey in the test departments, in consequence of the various problems that have been met with: for example, the problem of “storage” files (containing “old” administrative data ), the problem of large collections of paper resolutions and decrees, the problem of the files stemming from the use of administrative software, etc..

Other problems derived from the implementation of the questionnaire on Intranet and from the creation of the database.

The resulting questionnaire is very analytical and it asks numerous information for each file: identification elements (title, spatial and temporal references, etc.), organizational references (office, referee for the documentation), content information (variables, classifications, etc.), realization, updating and managing modalities.

The survey doesn't aim at producing a simple description and enumeration of existing files; its purpose is the creation of a “database of databases”, comparing files that have been created and managed in different administrative structures by means of

homogenous parameters; this database should become a dynamic documentation tool which could be easily consulted and updated.

### ***5.2. Phase 2: instruments and procedures testing***

At the beginning, due to the complexity of the project the survey was carried out in only three administrative departments ('test departments'), in order to test instruments and procedures.

The test phase had several purposes:

- to sensitize the political and administrative decision makers with respect to the work project;
- to evaluate the response of the decision makers and, in general, of municipal officers to the survey and to judge their level of cooperation;
- to organize the survey from an operational point of view;
- to monitor the work done by interviewers and statistical referees;
- to test the questionnaire;
- to prepare a quality control protocol that will be periodically run on the database;
- to verify the informative potentialities of the file with the first statistical analysis.

This phase has been carried out in the following administrative departments: Culture, General Secretariat and Economic Development. These three departments are very different with respect to size, structure and competence: the Economic Development department carries on regulatory activities with many external subjects; the General Secretariat deals with the management of the municipality system; finally, the Culture handles activities producing final results and presents very special file types (libraries, historical archive, etc.).

This experimental survey involved the departments' referees and nine external interviewers, all graduate or last year under-graduate.

A series of meetings have been organized to involve the administrative officers into the survey: the sequence of meetings started from the higher hierarchical levels (executives and managers) and gradually continued in successive ones. The final step was the actual surveying activity, that is the input of data in the electronic questionnaire,

an activity carried out by the interviewers together with the referees and the file manager.

This procedure was supervised with the help of an electronic calendar, that can be easily updated and examined via web by the interviewers and the project team. With this software it is possible to continuously monitor the survey procedure and to create a database containing information on the temporal distribution of meetings and survey occasions, on their length, the workload of interviewers and offices. These quantitative data has been of great help for the organization of the final work plan of the survey.

To coordinate the work of the interviewers and to define rigorous and common rules of conduct periodical meetings have been organised. The final result of these brainstormings is the interviewer's manual, a handbook which contains work procedures to be followed during the survey, and that can be very helpful to solve not-standard situations.

During and after the test phase (see the next paragraph) the work instruments were often revised: corrections were needed to eliminate the errors pointed out by the quality control protocols and to fit them to circumstances that were not expected.

This continuous evolutionary updating is the characteristic of the work.

Moreover, preliminary statistical analysis were made on the informative basement database.

On the basis of the results of quality control protocols and of statistical analysis both the questionnaire and the survey procedure were constantly updated and revised.

This phase was presented in detail in Buzzigoli, Innocenti and Martelli (2005).

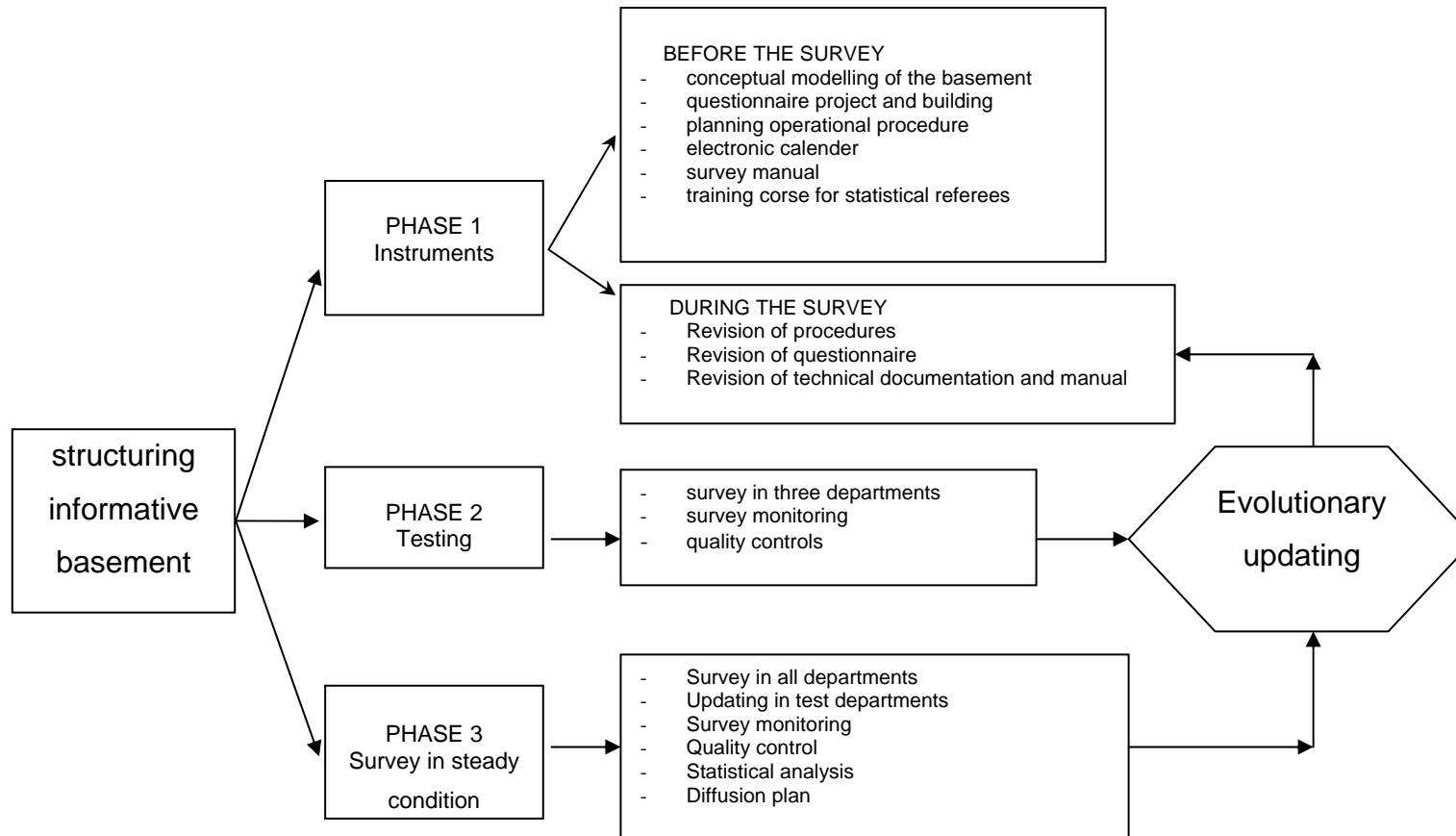
### ***5.3. Phase 3: survey in steady condition***

The results of the phase test have produced the standard sequence of procedures needed for carrying out a correct survey in all the other departments.

At present the survey has been conducted in almost all the departments, while the updating of the informative basement in the test departments has not been carried out yet.



**Figure 3 Work phases**



## **6. Some results**

In order to present the results obtained during this work project we distinguish three different levels of analysis: organizational, operational and methodological. For each of them we are going to point out the difficulties that have been faced and the solutions that have been proposed.

### ***6.1 Organizational level***

During the test phase we evaluated the response of both the political and administrative decision makers and of the personnel of the various departments, the participation and involvement of statistical referees, the interaction between referees and interviewers, etc..

The political decision makers had some hesitation due to the necessity of coordinating the project with other initiatives, to an insufficient understanding of the project's utility and also to the fear of results measurement. Moreover, the project could be interpreted as an interference of the bureaucratic bodies in their decision process and not as a contribution.

The administrative decision makers often considered data as a private property (even if the survey focussed only on metadata and not an data) and they sometimes showed an evident suspicion towards any attempt of activity measurement (the data survey was interpreted as an inspection tool).

The role of statistical referees was fundamental because they represent the connection between the work group and the administration and they make the survey easier: they have to present the survey, to explain its meaning and its importance, to involve the respondents in order to get more exhaustive and accurate information and to overcome possible beginning suspicions.

The work of the interviewers has been fundamental in the start of the survey, because the questionnaire is rather complex and they had to test it with care and attention. A group of interviewers, who gained enough experience in the test phase, have supported the statistical referees also in the sequel of the survey.

Further organizational criticalities have to be tackled with appropriate countermeasures: some of the professional resources could fail during the project; it

could be impossible/difficult to observe the scheduled deadlines; it could be necessary to support the project with marketing initiatives; the project could overlap with other initiatives regarding strategic planning and control.

### ***6.2. Operational level***

The survey was rather complex and it took a long time: the test phase was carried out from May 2003 to March 2004; the total survey lasted a couple of years. The analysis made on the database generated by the electronic calendar were useful to monitor the survey, to quantify the amount of work that has been done in this phase and to organize the sequel of the survey.

On average the duration of the interviews is rather long: two hour and a half at the General Secretariat and almost three hour and a half at the Economic Development; this fact proves that the survey activity is rather complex and demanding, although the first interviews surely suffer from the inexperience of interviewers and referees.

Another important and helpful instrument in the test phase is the interviewers' manual, which is the result of the experience of the interviewers and of the difficulties they faced during their work. The final decision about the specific rules to adopt in particular circumstances was made during the periodical meetings of the work group.

### ***6.3. Methodological level***

The first expected result of the test phase was the evaluation of the questionnaire; in particular, the main purpose was to check the adequacy of the surveying tool.

Even if the first version of the questionnaire already contained different paths for different kind of files (e.g. paper and electronic files), the empirical check suggested to broaden the number of alternatives. The data census must include both "current" files (currently used in the offices) and "closed" files (still available for the offices, but not updated any more) and, finally, the "historical" files: all these file types have different characteristics that must be described with different criteria inside the questionnaire.

In this sense the questionnaire revealed good properties for taking "a photograph" of a single archive in a static dimension, while it was not able to point out the link among different archives and to describe the temporal evolution of the archives in a dynamic dimension.

This problem will be of particular importance when the basement will be used for decision purposes (it could be necessary to trace the information flows inside an administrative procedure) or for statistical purposes (it could be interesting to build time series of administrative data) or, the updating of the database will be carried out.

The next step is focussed on the database: first of all it has been subjected to an accurate quality control. The control protocol exploits all the potentialities of the database with a set of queries aimed at locating errors and anomalies.

These queries are applied periodically to constantly monitor the database formation.

## 7. Some statistical analysis

The survey activity produced a huge database containing numerous information on the administrative files. The first statistical analysis were aimed to quantify the dimension of the evolving basement and to examine the characteristics of the files.

At the same time, some accurate quality controls were performed on the database.

In this paragraph we show some first results: we present some general tables in order to point out the main characteristics of the files we surveyed and the most important critical issues we encountered.

When reading the tables please note that with the same questionnaire can be surveyed several files (for instance, an Excel file which contains several sheets): most of the tables refer to questionnaires, and not to the files.

In all, 2778 questionnaires, 5062 files, 32561 variables containing approximately 146,022,784 records were surveyed.

Most of the files (Table 1: 63.3%) are on paper and are generated by internal procedures, that is during the usual administrative activity of the offices (Table 2: 93.7%).

**Table 1- Number of questionnaires by file type**

<b>File type</b>	<b>#</b>	<b>%</b>
paper	1758	63.3
electronic	979	35.2
software	36	1.3
other	5	0.2
<b>Total</b>	<b>2778</b>	<b>100.0</b>

**Table 2 - Number of questionnaires by origin.**

<b>Origin</b>	<b>#</b>	<b>%</b>
internal procedure	2604	93.7
external source	103	3.7
survey	19	0.7
elab./appl. of models.	42	1.5
missing	10	0.4
<b>Total</b>	<b>2778</b>	<b>100.0</b>

Moreover, files are usually updated only occasionally (see Table 3) and usually refer to the whole municipality (Table 4).

**Table 3 - Number of questionnaires by updating frequency**

<b>Updating frequency</b>	<b>#</b>	<b>%</b>
yearly	115	4.1
half-yearly	12	0.4
quarterly	15	0.5
monthly	97	3.5
weekly	20	0.7
occasionally	1948	70.1
other	561	20.2
m.v.	10	0.4
<b>Total</b>	<b>2778</b>	<b>100.0</b>

**Table 4 - Number of questionnaires by territorial reference**

<b>Territorial reference</b>	<b>#</b>	<b>%</b>
municipality of Florence	1895	68.2
district	305	11.0
other	578	20.8
<b>Total</b>	<b>2778</b>	<b>100.0</b>

As far as the use of the single files is concerned, files are seldomly accessed from outside the office where they are created and managed (Table 5), they are often oriented at managerial purposes (Table 6), it is unusual to find different levels of users (only for 41 files) and file access is hardly ever formalized by laws or regulations (only in 79 cases). Moreover, only in 282 situations there are specific diffusion channels.

**Table 5 - Number of questionnaires by access (multiple response)**

<b>Access</b>	<b>#</b>
operational referee	2742
others in the office	2254
others outside the office	58

**Table 6 - Number of questionnaires by purpose (multiple response)**

<b>Purpose</b>	<b>#</b>
documentation/study	601
managerial	2689
statistics	176
decision support	525

As far as the electronic files (together with software procedures) are concerned, almost all of them are under Windows operative system (about the 95%) and are created by the office itself (table 7).

**Table 7 – Electronic files and software procedures: number of questionnaires by realization**

<b>Realization</b>	<b>#</b>	<b>%</b>
same office	958	86.5
Information Systems Department	38	3.4
external firm	54	4.9
other	57	5.1
<b>Total</b>	<b>1107</b>	<b>100.0</b>

Most of electronic files are spreadsheets (Table 8: 65% on average) and, in particular, are produced with Excel (Table 9: 41.1% on average). In the various offices electronic spreadsheets are often made by the employees on their own initiative to carry on more efficiently their current activity and, more often, to fill in the gaps of electronic procedures that, although very complex and analytical, do not meet all the informational needs of the administrative unit. 94.2% of the files are stored in PC connected with intranet (Table 10).

**Table 8 - Electronic files and software procedures: number of questionnaires by data organization**

<b>Data organization</b>	<b>#</b>	<b>%</b>
DBMS	171	15.4
Spreadsheet	721	65.1
Sequential ASCII file	2	0.2
Data set	8	0.7
Image	19	1.7
Other	186	16.8
<b>Total</b>	<b>1107</b>	<b>100.0</b>

**Table 9 - Electronic files and software procedures: number of questionnaires by software**

<b>Software</b>	<b>#</b>	<b>%</b>
DB2	1	0.1
Oracle	21	1.9
Informix	1	0.1
Dbase III	12	1.1
Access	121	10.9
Excel	455	41.1
Word	296	26.7
Other	200	18.1
<b>Total</b>	<b>1107</b>	<b>100.0</b>

**Table 10 - Electronic files and software procedures: number of questionnaires by hardware**

<b>Hardware</b>	<b>#</b>	<b>%</b>
server	40	3.6
PC on intranet	1043	94.2
PC not on intranet	24	2.2
<b>Total</b>	<b>1107</b>	<b>100.0</b>

From Table 9 we can also note the low diffusion of database management softwares (in particular, only 93 questionnaires belong to a database management system) and the relevance of the item ‘other’, which seems much higher than expected.

The analysis of the 32519 variables shows that more than a half (18380) are text variables and that only a few (1066) make use of a classification; in this case the classification only occasionally refer to standardized rules (256 use a classification created inside the municipality and 258 variables use ‘ad hoc’ classification). Other interesting results can be summarised as follows: there are only 2 georeferentiated files and only 11 georeferentiation variables, although we expected to find many of them in particular contexts (such as the Economic Development Department); there are checks on the admissibility of variable values in only 206 cases and checks on the compatibility of variables in only 180 cases.

Also the most widespread names of variables are not standardised: for instance, for the register number three different names are used in the files (“protocol”, “number of protocol”, “protocol name”); many more names are used for the identification of citizens (“name”, “surname”, “name and surname”, etc.).

## 8. Concluding remarks

The recent introduction of management principles and techniques in the Public Administration environment has resulted in a coherent and significant use of statistical methods in the administrative context. This innovation elements however risk to remain at a formal level if they are not matched with terms and methods standardization and with the adoption of a concrete systemic prospective.

In this sense only the analysis of existent administrative realities with their real data comparability problems, collection effort burdens, etc., can give an accurate picture of what can/must be done to translate the legal regulations in effective government instruments.

In this paper a methodology for projecting and developing a statistical information system in support to local government administrations has been presented. A statistical information system is based on an exact comprehension of the processes to evaluate: the adoption of formal conceptual modelling languages may represent a solution to the task. A particular attention has been dedicated to the problem of building the statistical data sources necessary to evaluate the indicators requested by the users. Administrative data sets and archives are, in this sense, an important opportunity provided that an exact description of their characteristics and their usage limits is drawn. A metadata survey is the first step towards this objective: particular consideration must be paid to the problem of the conceptualisation of the survey itself, in order to get a full linkage with the public administration's organizational system.

## References

- Buzzigoli L. (2002), "The new role of statistics in local administration" Paper presented at the Conference Quantitative Methods in Economics, Nitra, Slovakia, December 5-7.
- Buzzigoli L., R. Innocenti e C. Martelli (2005), "La statistica a supporto della governance a livello locale: il caso del Comune di Firenze", *Statistica Applicata*, vol.1/2005.
- Dipartimento della Funzione Pubblica (2001), *Manuale operativo per il Controllo di Gestione*, Analisi e strumenti per l'innovazione. I manuali, Rubbettino Editore, Catanzaro.
- Gibert P. et M. Andrault (1984), "Contrôler la gestion ou évaluer les politiques?", *Politiques et Management Public*, n.2.
- Gormley W.T. and D.L. Weimer (1999), *Organizational Report Cards*, Cambridge, Harvard University Press, 1999.
- Inmon W.H. (1995), "Multidimensional Databases and Datawarehouse", *Data Management Review*, Jan..
- Inmon W.H. (1996), *Building the Data Warehouse*, Wiley e Sons.



- Kimball R. (1996), *The Data warehouse Toolkit*, Wiley & Sons.
- Ladu G. (2000), “The Reform of the Public Administration in Italy”, Paper presented at the Fourth International Research Symposium on Public Management (IRSPM IV), 10-11 April 2000, Erasmus University, Rotterdam, The Netherlands.
- Land F.F. and M.K. McGregor (2002), “Effective Use of Internal Information”, Working Paper 119, Department of Information Systems, London School of Economics and Political Science.
- Le Moigne J.L. (1997), “Progettazione della complessità e complessità della progettazione”, in Bocchi G.-Ceruti M. (a cura di), *La sfida della complessità*, Feltrinelli, Milano.
- Lynn L.E.Jr. and Stein S. Jr. (2000), “Globalization and Administrative reform: What is Happening in Theory?”, Paper presented at the Fourth International Research Symposium on Public Management (IRSPM IV), 10-11 April 2000, Erasmus University, Rotterdam, The Netherlands.
- Martelli C. (2002), “Statistical Information Systems for Local Government”, Paper presented at the Conference Quantitative Methods in Economics, Nitra, Slovakia, December 5-7, 2002.
- Martini A. e G. Cais (1999), Controllo (di gestione) e valutazione (delle politiche): un (ennesimo ma non ultimo) tentativo di sistemazione concettuale, II Congresso AIV, Napoli.
- National Center for Public Productivity (1997), *A Brief Guide for Performance Measurement in Local Government*, Rutgers University, Campus at Newark.
- Pollitt C. and G. Bouckaert (2000), *Public Management Reform*, Oxford University Press.
- Stame N., “Evaluation in Italy: experience and prospects”, *Evaluation*, Vol.4, n°1, 1998
- Statistics Denmark (2000), “The Use of Administrative Sources for Statistics and International Comparability”, Statistical Commission and Economic Commission for Europe Conference of European Statisticians, 48th Plenary Session, Paris, June 13-15, 2000.
- UNECE Secretariat (2000), “Report of the September 1999 Work Session on Statistical Metadata”, Statistical Commission and Economic Commission for Europe Conference of European Statisticians, 48th Plenary Session, Paris, June 13-15, 2000.
- Willemsse R. and F. Fleurke (2005), “Local decision-making and european rules and policy”, Paper to Research Group II, European policies- national, regional and administration forms, EGPA Conference, Bern, Switzerland, 31 August– 3 September 2005.

Copyright © 2006

Lucia Buzzigoli, Cristina Martelli