LIS/GIS TRENDS THROUGH THREE EXPERIENCES IN EAST EUROPE AND CENTRAL ASIA

By G.Carrai(*), R.Favilli(**), L.Morandini(***)

1 SUMMARY

The authors analyse the evolution of GIS technology over the time, through the development of Cadastre and Land Registration Systems. Over three years of consultancy for EU Technical Assistance in Kazakhstan, Albania and Bulgaria, the authors experienced three different approaches to LIS/GIS technology for cadastre and registry:

- an hybrid CAD - Relational Data Base Management System;
- a full topologic GIS with a Relational Data Base Management System;
- an universal database solution.

The analysis surveys strong and weak points of different approaches to cope with a common theme, the Cadastre and Land Registry in three countries with their own differences in data sources, organisation of cadastre and registry, funding and human resources. The paper gives the opportunity to compare different international cases and to survey the latest oncoming Information Technology tools either for geometric and semantic data to strengthen the integrity of data in a friendly user environment. A preview of GeoInformation standards and normalisation is also provided with the aims of globalisation of information and interoperability in the public administration.

2 MAIN ISSUES OF A GIS FOR CADASTRAL AND REGISTRY APPLICATIONS

Before entering into close details, it ought to be mentioned that a substantial difference between Cadastre and Registry lays with the different functions of the two of them, as Land Registry is concerned with exact identification of the real estate and its legal rights while Land Cadastre is a data base containing an inventory of real estates for the purpose of establishing various determinations; i.e. on the base of Land Cadastre it could be possible to determine the values of land for market and for taxation purposes while Legal Rights are attached to the land parcel but might not be its primary objective.

Legal rights are expressed as: no lender would advance money “against land” unless assured that a) the boundaries of the land are clearly defined; b) the borrower (or guarantor) has legal ownership rights to the land, and; c) that the land is not mortgaged to a previous lender having priority rights.

Cadastre is not “probatory” while entries of rights can be found only in the Registry.

Entry in the Registry can be made on transaction or on Court decision or by law; entry in Cadastre is an administrative procedure.

(*) Managing Director of SVALTEC – Florence Italy - E-mail: svaltec@ats.it
(**) Managing Director of Mappamondo Informatica – Florence Italy - E-mail: R.Favilli@fol.it
(***) GIS Consultant – Terni Italy – E-mail: l.moran@ita.flashnet.it

The more the Registry is kept close to the Cadastre the more legal validity is given to the Cadastral database enhancing the system to become a "positive" one based onto two sub-systems: the Cadastral sub-system and the Registry sub-system.
The Cadastre and Registry systems are at the same time one of the easiest model and one of most complex to be designed in Information Technology; conceptually, a cadastral applications is mainly modelled onto some aspects which can be defined as "internal", thus, i) a subset of a simple geometric layer (cadastral parcels) and , ii) textual information related to parcels, but it is also made of many different "external" components which all together form the final system. The interaction of said " external" components with the "internal" aspects could lead to different solutions either for the hardware and software architecture and for the interoperability between the Cadastre and Registry.
In an other article the authors are giving a broad analysis of Spatial and Quality aspects; here it could be helpful to add that Semantic and Temporal aspects contribute to form the logical and formal consistency of the cadastral database. All these aspects could be easily managed with the nowadays available Information Technology.

On the contrary, many others are the components which, because of their extreme variability, are affecting a system:

Land legislation
Basically, in the area of TACIS and Phare projects, the legislation on land and real estate properties is in its evolving phase, passing from a state ownership of the previous economic regime to a market economy; this is reflected in the lack of a clear legislative design able to state the content and procedures of legal system.
The approach to the traditional dualism between the Registry and Cadastral component of the Land Registration System seems also to be one of the main issue to be solved since some Governments are oriented toward a central system while others are oriented toward a distinct activity of the Cadastre and the Registry.

Land Tenure situation
For the time being, the term "privatisation" has brought changes in the social aspects more than economic benefits; the agrarian reform is not started yet almost everywhere and it is quite improper to speak of land tenure while it could be appropriate for real estate in urban areas.

Land & Real Estate market and valuation
In most of Tacis and Phare areas, the registration procedures and applications are quite simple and are carried out at municipal level. The valuation of a property is based on a "quality" criteria which means for rural properties land fertility and water supply while for urban properties it is mainly determined by the availability and cost of public infrastructures and services. The valuation of market price is almost everywhere on its experimental phase. Bank system in issuing mortgages is based, for instance, on capitalising of income, but also in this case there is no fixed procedures and assessment criteria.

Land & Real estate taxation
The taxation system stresses the need of a flow of information mainly concerning: size, quality, land-use and name and category of physical and juridical persons from the cadastral and registry system.

National Cadastral programme

The overall trend of Phare and Tacis countries is to develop a national programme on cadastral systems; the context of this assumption is that of a management of all data bases, maps systems to be part of the same (national) information system. In this long term project, the juridical and descriptive data on land parcels are the first priority; the sense is that of a cadastral book for inventory of the real situation comprehending base maps, surveying of plots, juridical data of ownership and land use.

Surveying and mapping

The spatial and the quality aspects are directly related to the sources of geometrical data; in countries where the surveying was maintained (Tacis), the quality of data is reasonable accurate but is referred to huge land plots (Sovkhoz, Kolkhoz) and is poor in urban areas; in other countries (Phare) many projects of field surveying were financed by Phare to reallocate the land to the former owners with a consequent huge gather of data with heterogeneous and too often not controlled results.

Main problems are lack of control procedures, no standard formats of data, too large use of local software products; the result of such approach is that software was mainly used for data production without any forecast on the future use of data in the maintenance phase while on the contrary the world market trend is leaning towards a standardisation of format and an enhancement of long experienced standard GIS software.

Staff

Staff is probably the most important ring of the chain; staff must be well trained, but, after training, which by the way is always one of the main component of EU programmes, staff must be motivated by incentives to build up a high specialised technical nucleus inside the public administration.

Availability of trained staff, computer literacy of staff, a long term plan of civil servants employment is one of the factors mostly affecting a matrix of risk in implementing systems; this is true at a greater extension if IT component is to be introduced.

Land Administration

Currently, the State retains ownership of a great deal of land and other real estate (buildings, irrigation infrastructure, roads, etc.), and has a broad range of administrative responsibilities to ensure appropriate use. The government agencies that had the authority for real estate management under the former system still exist in one form or another and retain much of their earlier mandate; to these agencies have been added several new structures.

Furthermore, it will still be involved in many aspects of land administration concerning private as well as public land.

Existing responsibilities of the State concerned with real estate include:

- Design of privatization and market reform programs
- Initial allocations of land and other real estate to users and owners
- Legal and regulatory framework for real estate markets
- Registration of interests in land and other real estate
- Surveying and mapping
- Real estate taxation
- Dispute resolution

- Land use planning and zoning
- Administration of leases and appropriate uses of state-owned land

Of these responsibilities, the first two are temporary, and will disappear when the transition is complete. The others are not; Government will continue to provide these services even after land is privately owned. The design of the Real Estate Registration System must ensure that the necessary interaction with other government agencies involved in the real estate sector is rapid, transparent, and efficient.

Information Technology

In most of the EU projects there is an excessive expectation for the IT in general; it is not well understood that an automation system is only the last and terminal phase of a long process of analysis, design, implementation. It is quite common that the analysis and design are jumped and implementation is started before any sound study and pilot phase. The overall criteria is mostly that of "waterfall" approach which might be risky even for well experienced organisation and systems (which is probably not the case of most of western countries indeed). It might be wiser to approach the problem via a "spiral" method based on:

analysis, prototype, cost/benefit assessment;

design, prototype, cost/benefit assessment;

implementation, prototype, cost/benefit assessment.

It could be sorted out that most of the very expensive hardware and software is oversized for the pilot projects, and that no computerised system is able to perform what even manually cannot work because of the lack of organisation or legislation.

3 HYBRID APPROACH IN KAZAKHSTAN: COMBINING CAD, GIS AND RELATIONAL DBMS

3.1 Kazakhstan 1994: Organisation, funding, human resources, data sources

The project " Establishment of a Provisional Land Information System and a Pilot Operation" was born in the frame of the wider project "Institutional support in the area of policy formulation and agricultural statistics" started in 1993, accordingly to the fact that Land Reform deals not only with rural areas but indeed with the entire national territory, the environment, territorial development, transportation, technological networks, town planning, Cadaster, etc. The primary objective was therefore to produce a General Guideline to outline the criteria of GIS/LIS design and of the organisation aimed; the second objective was co-ordinating the requirements of all Ministries to avoid duplication sharing data and enhancing the interoperability. *Specific project objective* was to provide technical advice and management support for the realisation of the Provisional Land Reform Information System based on the principle of rapid implementation (about three years) by making the best use of the existing mapping documentation and Cadaster data: accordingly to the Guideline, a pilot operation on a territory of a district of about 20,000 inhabitants was carried out.

Since one of the aims of the Land Information System was to support a Land Reform, then a Land Reform must have: a) an efficient, simple and well organised Cadastre system; b) a Cadaster Agency open to the public, able to deal quickly with Land and Real Estate records; c) a unified system managed by a Single Agency to classify and inventory data; d) a mapping system;

e) a provision of a clear, certain and safe Tenure Register to strengthen the confidence of owners and possessors of their rights; and f) a system to record transactions and rights to encourage long-term investment, mortgaging, leasing, sales and purchases to increase credit for agriculture and real estate development.

The project was funded by TACIS programme on behalf of Ministry of Agriculture; the technical counterpart was the State Committee for Land Relations and Land Tenure (Goskomzem) through the GosNpcZem which was the Agency appointed for issuing the State Akt. The project lasted two years during which many changes had occurred in the general framework the most relevant of which was that the Government, after a political election, decided to move the responsibility of registration under the Ministry of Justice to strengthen the legal security of possession.

Since the Kazak Government applied for a 40 Mil dollars loan from World Bank, the project was also a prototype test for the WB mission and some experts took part to the WB missions.

The skill of the staff was fairly enough for office automation but without any literacy about modern mapping and computer technologies applied to GIS; funded by a bilateral donors agreement, some very expensive workstations(Sun SparcStation, Unix OS) and software was purchased but, because of the lack of training, were not utilised.

Concerning the source of data, the geometric data of parcels and the accuracy of the work carried out by different agencies in charge for cadastral mapping was adequate to the system. The pilot project brought into evidence that, although by the means of not of the last generation instruments, the geometrical data were fulfilling the average standard requirements and the staff was skilled both in field works and in map making: these facts were considered as one of the pillars of the system.

All the edges of the parcels boundaries have been surveyed except those along natural features (i.e. rivers); for each edge the names of both of the owners of the two parcels were provided. A test carried out directly on the field books (attached to the parcel ledger) have brought into evidence that almost 75% of data were acceptable according to a minimum standard deviation less that 2 meters.

This level of accuracy was not possible to achieve through capturing data directly from maps; a comparative test carried out on raster images of maps leaded to a standard deviation of 20 meters due to the map making procedures; this led to the conclusion that, wherever possible, a direct entering of field data into the system was less time consuming, and with a favourable rate cost/benefit if compared with the availability of data and accuracy achievable.

Errors were coming mainly from :

- the lack of reconstruction of existing borders in cases where parcels are derived from existing;
- boundaries mismatching of adjoining parcel and between rural and urban areas.

Paradoxically, the preventive detection of errors brought a strong element in favour of the adoption of the digital base instead of the raster because an early analysis of data could prevent disputes in plot allocation once the transaction is ongoing; critical areas could be studied in advance, a map could be produced for public comments to be received and examined within a short period; this could short the temporary state (initial-, processing-, provisional-) before final confirmation of a parcel modification.

3.2 The need for an inexpensive solution.

The objectives of the project were those of a support to a wider agricultural policy formulation programme and in the beginning the computerisation of the cadastral component in the pilot areas was merely focused onto a geometric and a statistical approach; with this aim, the fund for the Land Information System was budgetary limited to 200,000 ECU comprehensive of equipment and a 12 man/months Technical Assistance. As mentioned above, the pilot project assumed over the time an increasing importance because either the Kazak Government and World Bank were converging their attention on the emerging issue of the Registration of Land and an Integrated Cadastral System.

Convinced that is not the power of computers but indeed the organisation and the criteria of analyses/design to make a system working, the **challenge was therefore to maximise the design phase and minimise the investment** in hardware and software.

The hardware was limited to 1 server PC Pentium, 1 PC Pentium, 1 PC 486, 1 A3 Plotter, 1 A4 scanner, 1 Drawing flex digitizer, 1 Ethernet HUB,1 Printer.

The Software was limited to Windows NT 3.51 Server; Windows WorkGroup 3.11; Relational Data Base: Oracle 7 /Windows; Windows Access

For Surveying data control and field book calculation: MOVE3 (Holland)

For cartographic editing and vector/raster management: GCarto (Italy)

For automatic vectorialization of raster images: GCarto (Italy)

For programming: Visual Basic and Visual C, Microsoft.

For topological construction: ArcCad ESRI

GIS software: ArcView ESRI

3.3 How to combine a low-cost topologic editing, a relational storage, and a data viewer.

As somebody said, "one of the biggest technical challenge we have to face nowadays is putting order in the chaos of already collected data".

The primary goal in approaching the database was therefore *not* to build up a new geographic information system, since several of them can be already found on the market; the purpose was rather to store information in a valuable form that might be exploited by an existing GIS or by new, ad-hoc applications.

A geographic system relies on geometry and on related information: there are products that use a database for both of them or that employ an hybrid structure, i. e. operating system files for the geometry plus a database for the attributes. This is, for instance, the case of ArcInfo (ArcView). In putting up the database one could have adopted the second approach, but since the clear trend in the market is to store in the database the geometry as well, it was decided to avoid any file system.

Some graphic systems tend to underestimate the importance of information directly related to the points. This is typical in the CAD approach, and too often a point is simply seen as a couple of coordinates that have a sheer geometrical meaning. In cadastral practice, a point owns a wealth of features and quite often a real story that can help in decision making.

No loss in information should be admitted in building up a database. A database ought to give back *at least* all the details that were originally available or that are going to be gathered

All these considerations led to use a single database to store both the geometry and the "external" data, exploiting all the facilities that modern Data Base Management Systems currently offer.

The choice of the Relational Database fell on Oracle Corporation because of the leadership in both marketing and technical terms among the commercial relational databases

As it is also well integrated in the Windows environment, the choice fell on ORACLE 7.1 Server on a Windows NT 3.5.1 Server, which is a robust operating system that already runs on PC's and RISC based workstations, allowing for a good scalarity as load grows

While developing new ORACLE specific applications with standard tools such as Forms, Report and Graphic, a connection to popular cadastral analysis products was sought. Among those available on PC with Windows, the choice fell on ArcView by ESRI, as it was widespread and relatively cheap.
Geometric data, if their source was that of digitising a map, were exported in DXF format and the topology inherent in our database was reshaped in ArcCAD. An ArcView procedure was written in the ESRI proprietary programming language Avenue to print out tabular and graphical reports for typical queries.
A point was uniquely identified by an integer key, derived from its coordinates. Several performance issues suggest this approach; for instance, joins between tables may not be efficient when using real numbers or the display can be faster than anticipated.
A field indicated the origin of the measured point; currently only three different determinations were considered:

- the point was recalculated by an adjustment program:
- the point's coordinates were typed in as recorded in the ledger
- the point was digitised from a map.

Since the determination is basically an enumeration, its range could have been easily enlarged to cover, amongst others, GPS and trigonometric points. The source of the determination can warn the applicationer that a same point was measured twice within a distance range (for instance, typed and digitised).
Since the *datum* was changed many times over the time and different ministries used different reference systems, transformation algorithms were implemented to store coordinates in a unique system indicating for each pair the original reference. The date of insertion in the DB was also recorded. This table was kept mainly for administrative purposes to track down the original sources where doubt might occur.
An arc is one of the basic entities in the adopted model for topology. This definition is consistent with the practice at GosNPCzem, since this arc will be marked in two different parcels showing the two (different) owners. In other words, the database can track down the owners' names of two confining parcels, i. e. parcels that share an edge.
In the project the edges were manually constructed following the indications given in the ledgers and sticking to the rule that in a closed network a point may belong to only one edge, unless it is a node. Nodes in this meaning have their own representation in the ledgers.
A polygon was defined by a sequence of one or more edges.
A relation connects the arcs to the parcels, again avoiding any redundancy, and since the referential integrity provided by ORACLE7 was employed, a consistent database was set about.
A database is **consistent** if there is no field dependent on a field that does not exist. Thus, if the parcel with ID =7 is formed by the sequence or arcs 90, 23, 56 it is sure that the arcs 90, 23, 56 exist and have already been stored in the database; in technical terms, there is no foreign key without a corresponding primary key.
In practice a few inconsistencies were found, such as overlapping parcels or islands, i.e. slivers that did not belong to any parcel. These are cases in which a different determination of the points can help deciding which one is the more reliable. Problems arisen with channels or borders that sometimes were simply sketched on the map. Software for adjustment was used, but only the recorded coordinates were stored in any case.
Concerning the textual attributes a good deal of facts was already available at GosNPCzem in a FoxPro database (files in DBF format) in cyrillic. These DBF tables were kept in their original format for the following reasons:

- DBF files can straightforwardly be translated into ORACLE tables, if demanded.
- A PC version of ArcView was used to bring about paper reports. This software works with DBF tables and can access ORACLE, but in so doing, it simply copies the information needed into local DBF files.
- The size of the pilot project was rather small and the data can be handled by a small DBMS as well.

During the progress of the project, as mentioned above, the Government decided to separate the legal component and put under the umbrella of the Ministry of Justice. The project was therefore reshaped to simulate a distributed system with an arrangement as follows:

- the coordinates and the geography of a parcel were stored at the GosNPCzem in an independent database **A**.
- The data concerning the owner (a person or a company or whoever) were kept in a database **B**.
- The legal rights were registered and handled in a database **C**.

The concept of three different databases means only that logical operations that are to be performed are different. It could be three physical databases A, B, C on different computers connected with each other, or all the information could be stored in a unique database A+B+C. Its up to the DBA (Data Base Administrator) to set the configuration and the protections according to the legal and administrative requirements.

Conclusion

The system, although based on a entry level hardware, is still running and more and more implementations have been carried out in the last period by local technical staff trained during the pilot phase. In the first two months of activity, 200 parcels have been entered in the system; during the following six months, due to the increased skill of staff and an improvement in applications' development, the same office was able to complete the Rayon and other decentralised offices have been equipped and trained.

The decision to move the registration under the Ministry of Justice has slowed down by the fact the ongoing process; the resources have been completely drawn to sustain a new issue for the Ministry of Justice so that the cadastral component has been delayed. Being the cadastral component of fundamental importance for the agrarian reform, it was decided to use other funds to follow up the pilot project. It is easy to foresee that with two different donors and two different approaches the final result will be negative; it is a further prove that technology is far ahead politicians ‘ designs.

4 FULL TOPOLOGY GIS AND RELATIONAL DBMS: THE CASE OF BULGARIA

4.1 Bulgaria 1997: Organisation, funding, human resources, data sources

The European Union and the Government of Bulgaria in 1995 have recognized the need to help land reform and the establishment of a land market with a suitable IT infrastructure. Therefore, a PHARE project was started to provide funding and consultancy services for setting some experimental cadastral offices in a *Pilot area*. After the completion of this *pilot project*, lessons learned and knowledge acquired would form the basis of the *nation-wide implementation*, which will build the infrastructure of the National Cadastre of the Republic of Bulgaria.

The funding of the project was articulated in three distinct phases:

1. Data acquisition
2. Purchase of hardware and base software (operating systems, DBMS, GIS)
3. Application software

The second phase was awarded well before the beginning of the third phase, hence, given the fast evolution of IT, the hardware part is now rated as too costly and too cumbersome to manage (the chosen hardware was composed by UNIX workstations and servers), but at the time, well before Windows NT had proved itself worthy of enterprise-wide systems, this choice made perfect sense.

In Bulgaria, as in Italy and many other countries throughout the world, cadastral services are given to the public by two distinct organizations, which belong to two different ministries: Cadastral Offices under the MRDPW, Registry Offices and Notaries (soon to be private notaries) under the MJ.

This organization lends itself to some peculiar arrangements in terms of work-flows and IT infrastructure.

From the organizational point of view some suitable exchanges of documents between the two offices are to be provided in order to issue cadastral services e.g.: to sell portion of one's parcel, the seller should first go to the cadastral Office which splits the parcel and issue a new UPI to the newly created parcel, then Cadastral Office will send the Registry the new geometry (after the splitting), and, finally, the citizen can sell a part of his/her property. After the selling Registry has to signal the Cadastral Office of the changing of ownership, backing it by a reference to the legal document of transaction.

In the making of a computerized system for the automation of cadastral chores, some issues are raised by this partitioning of tasks between two different organizations. The system has to be built around two distinct data bases, which could be, in the future, became two different instances of a single database (by means of replication). For now, considering the low-quality of telecommunication links in some part of the Bulgaria, a conservative approach was taken, hence the two separate data bases linked by data export/import over magnetic supports.

Human resources for Cadastral and Registry offices are composed mainly by middle-aged officers with little computer knowledge, so a comprehensive training program was planned. The training program issues range from operating system (UNIX), to GIS software (Arc/Info), to application software (ICS and TIRES).

Data sources ranges from: surveying using traditional methods, GPS surveying, digitizing from paper maps, digitizing from raster maps, photogrammetric stereoplotting.

Therefore, no wonder that some mistakes have been recognized; moreover, two distinct ministries took care of data capture, hence, again, no wonder that urban and rural data don't match and that neighboring settlements overlap each other.

4.2 GIS and Relational DBMS: a perfect marriage

Managing of geometric data, which is the core of GIS technology, was (and still, mostly is) done by storing graphic files on mass storage, giving an Identification number (ID) to each geometric feature and linking every geometric feature to a record containing its alphanumeric data (attributes). Nothing can be more different from the Relational Data Model, which doesn't use IDs and uses composition of attributes (known as *keys*) for identifying single records (which, strictly speaking, are called *tuples* and not *records* in order to differentiate the relational model from other ways of managing data).

The Relational data model which is the model chosen by almost every DBMS (including: ORACLE, INFORMIX, DB2, SYBASE, SQL SERVER) was conceived for dealing with alphanumeric data, and simple relationships between single pieces of data. Hence it was slow and ill-suited to deal with graphic data, which should be retrieved in large quantity to be displayed on a screen and are connected by not-so-simple relationships.
Cadastral applications comprise the managing of data which are both geometrical ad alphanumeric, moreover, a strong emphasis is put over reliability and integrity of data (especially legal data), hence four alternatives are given:

1. To use an entirely GIS-based data model, which would be the fastest way to manage geometric and alphanumeric data, given GIS use of pointers to records (also knows as *reticular data model*) ; but, also, the unsafe way, because of missing data integrity mechanisms.
2. To use an entirely relational data model, which is quite fashionable, and, given the progress in recent Relational DBMS, not so slow compared to GIS; but, also this model has its drawbacks: little knowledge exists about the use of this technology of geographic applications, it is costly, users have to build GIS functionality by themselves.
3. To use an Object-Oriented DBMS (which could be seen as an evolution of reticular data model, but powered by concepts of object orientation). This solution is, in principle, the most apt to geographic applications, but, again this model has its drawbacks: little knowledge exists about the use of this technology, it is costly (the most costly), and, as in the totally relational approach, users have to build GIS functionality by themselves.
4. To use an hybrid approach: GIS data model for geometric data coupled with a Relational DBMS for managing alphanumeric data. This is the best of two worlds, but require good communications between the two halves of the system.

Luckily, but not entirely by chance, modern GIS software packages have very good links with Relational DBMS. In the case at hand a combination of Arc/Info GIS software and INFORMIX Relational DBMS was chosen , this pair get on well each other: INFORMIX table’s attributes can be manipulated from within the Arc/Info programming language, Arc/Info alphanumeric table can be converted from/to INFORMIX tables, Arc/Info can open cursors on INFORMIX tables, Arc/Info can issue SQL instructions to INFORMIX and get the results.
Therefore, alphanumeric data relevant to the cadastre are managed by INFORMIX, so ensuring data integrity, while geometric data and manipulation is done using the facilities provided by Arc/Info.

4.3 Towards a multipurpose cadastre

The PHARE-sponsored cadastral offices in the pilot area will be provided with application software for the basic cadastral chores; although, during the study phase, a need for others functions has arisen. The aim is to keep the system as simple as possible, in order to have it ready as fast as possible and to start getting feedback from users; this aim notwithstanding a scalable solution, open to improvements and additional function is devised.
The use a of topologic data model for geometric data combined with a flexible GIS software make it easy to add, in the future, some new functionality which will transform a *legal/fiscal cadastre* into a truly *multipurpose* one; e.g. adding utility lines for generating interdict areas, adding vegetation data for the management of forestry resources, and the like.
A multipurpose cadastre can be used by many government agencies for planning of land use and for building public works like roads and railways; moreover, a multipurpose cadastre can be profitable by selling its geometric data to private firms to use them as a base for construction projects or real estates evaluation.

5 UNIVERSAL DBMS FOR ALBANIA

5.1 Albania 1997: Organisation, funding, human resources, data sources

According to recent data, Albania, has a population of 3.3 million and a total land area of 28,750 Sq. Km., covered for a 36% by forests, 23% by not agricultural lands, and the rest distributed over a gently sloping land, suitable for crop farming.

Despite of this favourable natural situation, the development of agriculture, is seriously handicapped by many different constraints and pitfalls. It is indubitable that radical changes from planned economy to marked economy has also an impact in natural resources management not only because of privatisation, but indeed because of strong fragmentation of land properties, not scientifically planned cropping, low use of mechanisation etc.. A certain arrangement of land, both in use and in property, is easily predictable when the taxation system will be implemented.

A Phare Land Use Policy (LUP) Project has been funded to define guidelines in management of resources and decision making for a development of agricultural; however, considering the present condition of Albanian information sources, the first aim of PMU/LUP is to retrieve all the existing but heterogeneous data (soil, properties, infrastructures, meteorology, irrigation, elevation etc.) and to store them in a computerised system.

At the same time the ambitious privatisation programme, which involved almost the total of agricultural land and many urban properties - 3 million new immovable properties have been estimated - made necessary the creation of a new immovable property registration system (IPRS) to provide owners with secure rights to their property. In 1993 USAID assisted the Albania Government to prepare a Land Market Action Plan. After consultation with USAID, World Bank and EU PHARE the responsibility for the IPRS and projects related to the future developments of the land and real estate markets was vested in a Phare Project for the IPRS. Phare programme is assisting in reviewing the procedures of surveying, mapping and quality control while USAID inputs have been directed towards the Land Registration. The link between geometric base and registration data , the computerisation of Titling System and the construction of a dynamic Cadastral and Registration System are the main practical aims of PMU/IPRS. Right now this means to establish in the Cadastral Offices a technical unit able to improve the production of cadastral maps by the means of sub-contracting activities to the private sector and at the same time to ensure an adequate level of accuracy and reliability of geometric data through well defined, standard control procedures.

Concurrently with this activity, the IPRS/PMU has another component dealing with map production, represented by an already working unit of photogrammetry, hosted in the Military Topographic Institute and equipped with high technology analytic stereoplotters.

Scarce possibilities of control, *laissez-faire* policy, recent situation of anarchy and disorders already caused serious, sometimes irreversible damages to large forests Albanian patrimony. For these reasons another EU PHARE project was financed, focused on woods census and defence.

5.2 Universal database

By now the three PHARE projects- IPRS, Land Use Policy, Forestry- are collaborating to the realisation of an integrated GIS . The system will be physically installed in the offices of the Ministry of Agriculture in Laprake where the IPRS, the Soil Institute and the Forestry, the main users of the system, have their own offices. This three departments have in common the parcel based map on the top of which all applications and thematic derivations have to be developed. So both the soil Institute and Forestry are to be considered as an interesting example of cadastral services external users (future predictable users could be the city plan offices, water and gas networks etc.). At the same time the system can be seen as a first nucleus of one more ambitious project of universal (spatial and alphanumeric) data-base, based on an integrated data server with the following main responsibilities:

- retrieving and distributing geographic and alphanumeric information to clients in a spatial/temporal schema
- granting to clients different levels of access for different sectors of data base: from simple alphanumeric consultation, to integrated data editing and update
- managing locks and concurrence privileges: every access to information is negotiated depending to privileges and priorities
- emancipating clients from any physical organization of data (i.e. sheets, administrative boundaries) using indexes for both alphanumeric and geographic data
- allowing variable-scale spatial elaboration (from small scale for wide phenomena geographic analysis to large scale for administrative certifications)
- emancipating clients from software dependencies using standard data connectivity protocols
- allowing integrated spatial and attributes analysis
- managing clients “proprietary partitions” of data, to study temporal evolution
- mailing information in a suitable format to other Government agencies .

5.3 The system architecture

At present PMU is concluding the phase of analysis and requirements issue of such system. CEN standards have been adopted to outline the main aspects of the conceptual model of data and functions. Different classes of geographic features and alphanumeric data, their attributes and their functions (methods) were described with appropriate diagrams. Due to different user characteristics any geographic feature have been depicted at the resolution according to the granularity of the investigation, and assigned to a feature class of which at least four important aspects are known: geometric, semantic/thematic, qualitative and temporal. For each of said class other more specialised classes have be defined following a top-down structure till the disaggregation of a class in “all-parties” relations.

For example the data model of a Feature has been modelled as shown in the following table:

Class:	Feature
Attributes:	user identifier, reference to thematic class, reference to geometric class, reference to quality, temporal aspects, etc.
Methods:	create, delete, modify, up-date, etc.
Constraints:	on creation require identifier and reference to... before to delete, trigger action in table ....
Domains:	visible at .... scale; etc..
Assumptions:	

The geometric class has been modelled has shown below:

Class:	Geometric
Attributes:	identifier, reference to a feature, etc.
Methods:	create, delete, modify, up-date, etc.
Constraints:	on creation require identifier and reference to...

	before to delete, trigger action in table ....
Domains:	visible at .... scale; etc..
Assumptions:	

Such class has been successively specialised in several classes like in the following example:

Class:	Node
Parent:	Geometric
Attributes:	X, Y, Z, etc..
Methods:	get, modify, display 2D 3D, etc.
Constraints:	require 3D coordinates when acquired, reference to...

The following sample table is one of the final result of such classification where it was possible to recognize different kind of data, their type and their users:

Name	Type	Users
Parcels	Polygon	IPRS - SOIL - FORESTRY
Buildings	Polygon	IPRS
Toponyms	Annotation	IPRS-SOIL-FORESTRY
Topographic lines	Arc	IPRS-SOIL-FORESTRY
Administrative boundaries	Polygon	IPRS-SOIL-FORESTRY
Geodetic points	Point	IPRS-SOIL-FORESTRY
Themes	Polygon	SOIL-FORESTRY
Remote Sensing	Polygon	SOIL-FORESTRY
Sampling	Point	SOIL-FORESTRY
Sampling	Semantic	SOIL-FORESTRY
Registry	Semantic	IPRS-SOIL-FORESTRY
Statistics	Polygon	SOIL-FORESTRY
Statistics	Chart	IPRS-SOIL-FORESTRY
Statistics	Report	IPRS-SOIL-FORESTRY

On the side of methods the functions the system performs have been grouped into four classes:

- Importing/Exporting of data
- Inquiry functions
- Geometric Data editing
- Tabular data editing
- Auxiliary functions

Import/Export of data functions are the ones that allow users to import/export data among different applications by the means of standard data transfer format
Inquiry functions include providing with tools to retrieve information both in graphical and semantic aspects; for the Database this is basically restricted to SQL but middleware are accepted at client level. The temporal extension of the query language is requested because of

the variability of attributes in the time. To this function is demanded also the production of maps, sketches and portion of DataBase either for public and for Government Agencies.
Geometric Data editing functions include all the manipulations of geometry, spatial correlation, up-dating, changing and displaying different map scales, etc..
Tabular Data editing functions include the manipulation of semantic data into tables which are relevant to Cadastral and Registry, Soil, Forestry; it comprehends creating, up-dating, removing data with the modality and restrictions described in the following paragraphs.
Auxiliary functions include operations like: backup and restore of data base, access control, logging and reporting of transactions, … .

Also different potential users have been integrated in the model classified as:

- Query user (QU)
- Semantic edit user (SU)
- Geometric edit user (GU)
- System Manager (SM)

The first class (*QU*) is due to perform activities in which no editing of data is involved, like plotting sketches and printing reports. The second and third user classes (S*U and GU*) are allowed to do changes in data using ONLY product's functions. The fourth user class *(SM*) is allowed to use base-software-level functions in order to edit data: base-software is intended to be GIS command line interface commands or RDBMS command line interface commands.
The difference between a *SU* user and an *GU* one is that they edit different data, geometrical data and relate attributes for *GU edit*, while *SU edit* user edits only data which relate to tabular data. Classes of users allowed to use classes of functions have been established as summarized in the following table:

Users ⇒ Functions ⇓	QU Query user	SU Semantic user	GU Geographic user	SM System Manager
Import/Export of data				✓
Inquiry functions	✓	✓	✓	✓
Geometric data editing			✓	✓
Tabular data editing		✓		✓
Auxiliary functions				✓

Software products to implement the first nucleus of the system have not be chosen yet. However the system design establishes a client/server architecture. A Relational DataBase Management System is placed in the Server unit together with GIS application. A LAN will connect the Server with the at least 4 client sites. One client site is the Mapping Production Unit where devices for data capture and data plotting will be placed. The three clients/users (the Immovable Property Registration System, the Soil Institute and the Forestry Project) will be equipped with Geographic database client software ,suitable to view/elaborate thematic maps, legend, overlay themes, raster imageries, aggregate/disaggregate information, tagging features, database querying, creating both temporal and permanent tables, manipulate proprietary attributes. External software will be interfaced by means of data import/export, using standard geographic and alphanumeric formats (in the beginning especially DXF, DBF). File sharing between clients and server is provided by Network File Sharing protocols built on top of TCP/IP networking protocol.

6 DATA-SHARING THROUGH THE ADOPTION OF STANDARDS

6.1 Standards in data manipulation

When we mention data transfer, it comes immediately to refer to converting geodata from one usually proprietary data format to another. For too many years discussions were held about the need of a standard approach, but unfortunately either the scientific and the commercial world were too busy to discuss to convert into practice the thousand of words spent to design conceptual and sometime abstract criteria. The final result is that there is a certain diffidence toward the geoinformation because of:

- it is well aware that the most expensive phase of any geoinformation is data collection and data must be of good quality before check-in procedures;
- after a period of indiscriminate and uncontrolled grown of data production, many countries start to discipline the cartographic activities mainly for budget reasons;
- those illuminated agencies that commissioned data collection for their databases realised that data are quickly outdated while in the same time other agencies collected the same set of data but, because of uncommunicability, data were not available as well;
- to underpin decision makers and planner, it is not enough to have only one set of information although very rich in attributes but it is preferable to have all the range of available data to compare realities and to simulate an action on some data and see the reaction on other data;
- in the Internet age, geoinformation cannot be the only science not ready for the globalisation of information .

It could not be sufficient a “de facto “ standard because of two main reasons:

1. from the commercial point of view, a single-vendor standard for distributing geoprocessing would have limited in the long run the user choice and therefore restricted the market;
2. the interoperability is nowadays a strong requirement in the cost/benefit analysis of computerisation of the public administration and public administration cannot accept any monopolistic market system.

There was also a trend to develop a proprietary format, especially in the early phase of geoprocessing development. After such negative experiences in USA and West Europe, it is quite frequent to notice that in many Phare and Tacis countries there are many attempts to develop and sustain “local” solutions, despite of the advises of international experts on the matter.

It is indubitable that, before the Berlin wall fall, there was a sort of “nationalism” to defend its own frontiers; it is as much true that wherever salaries of high professional software developers are very low, the development of application is much more convenient than to purchase in the market at Western prices. But it has to be pointed out that those countries have no knowledge of “up-dating” since for so many years they conceived the reality as the one, immutable, planned at the central level. In a market economy changes are very fast and geoprocessing must be reshaped and tailored with a frequency that doesn’t allow to spend resources in the basic software; it comes out that in the long run the research and development of basic software will be on the shoulders of GIS vendors while only a small part of development will be utilised for the specific user customisation.

In the last few years "de jure" standards are growing in importance in the Information Community and many countries have their experts in scientific commissions; the standards created in a formal "jured" process, are restricted to the International Organisation for Standards Technical Committee 211 (ISO TC/211) and OpenGIS Consortium. TC/211 is working at a level of abstraction above OGC, so to an important degree the two efforts complement each other, and both are necessary. ISO is to deliver a specification near the end of 1998. By then OGC will have implemented multiple engineering-level specifications, but they won't be nearly as broad as ISO's work. ISO's work isn't likely to result in immediate implementation-level specifications, so OGC effort is to implement specifications to fit into the ISO framework as implementation profiles.
From early papers circulating it is clear that ISO TC/211 is much more close to data and metadata while OGC is more focused on object technology.
The TC/211 bases its activities on the definition of a reference model, profiles of users, terminology, conceptual geoinformation schema, spatial, temporal subschemas, rules for application schema, quality, positioning services, geodetic reference systems, encoding, spatial operators, services. To describe data and made then shareable, TC/211 enforces the use of EXPRESS language as the description method according to ISO/DIS 10303 – information automation systems.
The OGC bases its activities on the definition of a general and common set of basic geographic data types that can be used to build interfaces between dissimilar geoprocessing systems; the results will be a software schema in which data and processing functions are packaged into small, discrete, interoperable modules, offering advantages such as portability and easy maintainability.
Strictly related to spatial data, topology is the most common structure in many data standards: DIGEST, NTF, ISO-CEN TC/211, to mention some, utilise topology because of the possibility to "navigate" to reconstruct the geometry of a feature independently from software application, scale and positional factors.

6.2 Standards in GIS software

During the first phase of GIS diffusion in Europe many public and private companies preferred to commit their systems to small, effective, local software-houses instead of choosing the already existing accredited, commercial packages, mostly coming from USA. There were several reasons for such a choice: artisanal software was in general of good quality; it was very often developed or adapted to users' specific requests; the prize was lower and, above all, assistance service was absolutely better.
It is easy to understand that during the last ten years the situation has completely overturned: commercial, industrialized packages quality increased quickly: graphic user interfaces became more and more friendly, relational data-bases increased performances and possibilities, client/server architectures allowed more modular and economical solutions, new software object technology allowed to integrate such products with special routines developed on customers' necessities; some of commercial packages data formats become standards for information interchange, utilized all over the world; big companies opened their branches for technical assistance nearly in every country. On the other side new software engineering standards of analysis, development and documentation need very high investments, generally not affordable by little companies, which in fact started since the beginning of '90 to focus their activities in customization and assistance.

At present the choice of GIS software is reasonably made between a certain number of well known, wide diffused, experienced packages, eventually privileging those software which supports standards for external integration and embedding.

6.3 Standards in system development

The world economy is becoming more and more interconnected, goods and services can be produced in countries where it is inexpensive to produce them and can be sold in countries where they are expensive to buy: therefore, developing countries are getting a bigger share of world markets. This increasing competition from developing countries is forcing industrialized countries to raise quality of goods and services produced, hence selling them at an higher price (because of higher quality) than developing countries products.

Quality management is an established way of raising the quality of industrial goods; introduced by Japanese firms during the '70, it is increasingly common in western countries for many industrial sectors, and it is becoming used even for production of non-industrial products, like financial services or telecommunications networks management.

Quality management is codified by some *International Standard Organization* publications, which are collectively called *ISO 9000 guidelines*; these publications describe some sound management practices which may help a firm to enhance the quality of its products. Some guides describe the application of quality management principles in specific sectors of economy.

Some organizations, connected to the ISO, certify firms which follows ISO 9000 guidelines. This public certification ensure customers that the firm has an internal control system which monitors all activities which may impact the final quality of firm's product: hence customers can easily tell high-quality products from low-quality ones and make their choice

The activity of system development, can be viewed as an industrial process, and to raise the quality of the final product (the system being built), ISO 9000 guidelines may be applied. To apply such guidelines means leveraging on sound management principles, which have been used by successful firms around the world and analyzed by scientists and researchers of major universities. For the ISO 9000 guidelines to be applied to system development, some specific best-practices guides should be considered; one important source for software development are the *Institute of Electrical and Electronics Engineers* guides and standards. These publications describe sound engineering principles in software engineering for every system development phase.

A GIS system is composed by many components, like:

* Hardware and communications
* Base software, like GIS software and DBMS software
* Application software
* Geographic data
* Organization (people and activities)

Some of these components can be bought off-the-shelf, like hardware and communications, or base software, some of them may be purchased by third parties, like geographic data, but organization and application software must be specific to the GIS system being built: therefore they should be developed ad-hoc.

Let's focus on the application software part of a GIS system: how this activity can be organized in such a way to provide high quality output ? By following ISO 9000 guidelines, of course.

As was previously mentioned, the software development is analyzed by the IEEE publications, which covers all phases of development:

* Requirements

* Design
* Coding
* Test

The requirements phase is, probably, the most important phase, it is the one in which what the system has to do is described. Failing to capture the real needs of users, brings to a system which is useless, no matter how skilled and smart the developers are. The requirements are defined in a document called *Software Requirements Specifications*, this document should be written following practices recommended in (3), therefore every requirement has to be: uniquely identifiable, unambiguous, verifiable. Which means that every requirement should be given a unique code (like RF003); that every requirement is stated using a clear language and every term must have a unique meaning; that every requirement can be tested in a finite time.

For instance:

The system has a function for checking the integrity of data

Is: not uniquely identifiable, ambiguous, not testable,

While:

DA13 System enforces referential integrity of Data Base for every function regarding editing of data

DA14 Integrity control in DA 13 includes: Uniqueness of primary keys' values

DA15 Integrity control in DA 13 includes: Presence of valid primary keys' values

DA16 Integrity control in DA 13 includes: Ensuring of minimum cardinality constraints in relationships

DA17 Integrity control in DA 13 includes: Ensuring of maximum cardinality constraints in relationships

Are: uniquely identifiable, not ambiguous, testable.

Every requirement has to be verified during test phase, if it is not testable it should be detailed in more depth or thrown away, something which is not verifiable does not belong to a computerized system.

Design phase output is described in (5) and must be carried out using widely-accepted methodology, like Entity-relationship diagrams or Unified Model Language, or Data-Flow Diagrams; this is to maximize understanding of design and to leverage on proven methodology.

Coding phase is to be carried out using sound programming practices like modularity and readability. These practices may be different using different programming languages.

The testing phase is covered by another standard (please, refer to (6)) and it is the right place in which system is verified, features tested, usually, include: compliance to requirements, performance, robustness of software against incorrect data, usability.

Another aspect of system development which is covered by both ISO and IEEE is *Verification & Validation*. This term comprise all activities which ensure the control of development, like reviews, anomalies tracking and reporting activities. The guidelines for this task are described in (4). V&V activities are recorded and carried out in a structured manner, which means that:

* Meetings have to be recorded in meeting minutes
* Actions should be uniquely identified and followed until their completion, completion of actions has to be agreed, in written form, by every party involved
* Reviews are planned to evaluate intermediate products of development. Reviews, usually, last several days and are summarized by a meeting

*	Anomalies (also know as "bugs", in software slang) are uniquely identified, recorded, and followed until they are solved. Solving of anomalies has to be agreed, in written form, by every party involved

The compliance with mentioned standards improves the overall quality of GIS systems and make them easier to maintain, therefore means fewer defects and less expensive maintenance, bringing to less money spent in the long run.

Bibliography

(1) *ISO 9000*, International Standards Organization, publishing data
(2) *ISO 9000-3*, International Standards Organization, publishing data
(3) *IEEE Recommended Practice for Software Requirements Specifications*, Institute of Electrical and Electronics Engineers, Std 830-1983, IEEE 1983.
(4) *IEEE Standard for Software Verification an Validation Plans*, Institute of Electrical and Electronics Engineers, Std 1012-1986, IEEE 1986.
(5) *IEEE Recommended Practice for Software Design Descriptions*, Institute of Electrical and Electronics Engineers, Std 1016-1987, IEEE 1986.
(6) *IEEE Recommended Practice for Software Test Documentation*, Institute of Electrical and Electronics Engineers, Std 829-1983, IEEE 1983.
(7) *LKI -Cadastral LIS The Netherlands*- B. van Osch, Dutch Kad –ITC, 1996
(8) *Land Information in Eastern Europe*, S. Hartley – GIM,1997
(9) *European Standard of GeoInformation (European Committee for Standardisation) Document prEN12160*, CEN-Brussels,1995
(10) *The Digital Geographic Information Exchange Standard (DIGEST)*, DGIWG,1992
(11) *A GIS Approach to Land-Use Change Dynamics Detection*, C.P. Lo, R. Shipman - PE&RS, 1990
(12) *Cost effective data for regional GIS bases*, L.H. Spradley - ISPRS, 1994
(13) *Feature level topology for Arc/Info using the Arc/Info-Oracle interface*, J. van Smaalen, M. Molenaar - ISPRS,1994
(14) *Design of a National Topographic database: an Object oriented approach*,B. Shahrabi – ISPRS,1994
(15) *Framework report on database management systems* - ANSI/X3/SPARC, 1978
(16) *Proceedings of the international workshop on object-oriented database systems*, K.Dittrich - ANSI-AFIPS, 1978
(17) *The OpenGIS Guide*,K. Buehler, L. McKee – OGIS, 1996
(18) *ISO/TC211 work programme* - ISOTC211, 1995
(19) *Automation of the Cadastre in El Salvador*, A. Tjalma - The Dutch Kadastre, 1995
(20) *The Thailand Land Titling Project* P. Angus-Leppan INT. GEOG. INFOR.SYS, 1989
(21) *GIS versus CAD versus DBMS:What are the differences?*, D.Cowen - PE&RS, 1994
(22) *A Perspective on GIS Technology in the Nineties*, A.Frank,M.Egenhofer - PE&RS, 1991
(23) *The Indonesian Land Administration Project*, C.Grant – GIM, 1997
(24) *First Generation OpenGIS Components*, J. Roodzand – GIM, 1996
(25) *Towards Integrated Land Registration*, G.Remetey – GIM, 1996
(26) *Views on Iteroperability and Spatial Data Infrastructures in Europe –GSDI and ESDI* P.Burrough – ASITA, 1997
(27) *Quality in GIS*, L.A. Koen – ASITA, 1997

Towards a European Policy Framework for Geographic Information DGXIIIe Luxembourg, 1997