GUIDE-LINES FOR THE DEVELOPMENT AND MAINTENANCE OF A GEOINFORMATION UTILITY IN A DISTRIBUTED ENVIRONMENT

C.M. Paresi, M.M. Radwan

International Institute for Aerospace Survey and Earth Sciences (ITC)
P.O. Box 6, 7500 AA Enschede, The Netherlands
Tel. +31-53-4874339, Fax. +31-53-4874335
E-mail: PARESI@ITC.NL, RADWAN@ITC.NL

Commission IV, Working Group 6

KEYWORDS: Database, Development, Exchange, Information, Integration, Model, Systems, User.

ABSTRACT

Extreme high costs in data acquisition, and increased user demands for sophisticated applications, and thus for transparent access to information, regardless of the hosts on which it resides, combined with advances in databases systems and communications technology, are rapidly changing information management from a traditional centralized perspective to a distributed one. This requires inter-operability among heterogeneous hosts, operating systems, data sources and data structures. The tools and services for supporting the development and management of applications in such an environment must change consequently. Currently "island solutions" are connected by ad-hoc approaches.

A research project has been started at the ITC, aiming at the development of guide-lines for the development and maintenance of a Geoinformation Utility enabling integration, at different levels of decision making (horizontally: across different thematic databases; vertically: from local to national levels), of distributed databases, in a federated system perspective; the ultimate goal being to enhance the availability of relevant, up-to-date, and integrated land information, timely and at an affordable price, to support decision making processes related to country's sustainable development.

This paper reports on progress of the above mentioned research. It includes in particular findings related to:
- the development of a planning framework for the development and maintenance of a Geoinformation Utility based on an analysis of the institutional, technical and economical critical success factors for such a development;
- the development of Geoinformation Utility mechanisms, in a federated environment;
- an outline of a generic approach to support the cost-effective development of such an Utility.

1. BACKGROUND

The way ahead for the information society might be summarized by the following statements [M. Bangemann, 1994]: *"Throughout the world, information and communication technologies are generating a new industrial revolution already as significant and far-reaching as those in the past."*, and *"This revolution adds huge new capacities to human intelligence and ... changes the way we work together and the way we live together."*

Indeed, advances in communications technology, development of powerful desktop workstations, extreme high costs of data acquisition, and increased user demands for sophisticated applications, and thus for transparent and cost-effective access to existing information, regardless of the hosts on which it resides, are rapidly changing information management from a traditional centralized perspective to a distributed one. This requires inter-operability among heterogeneous hosts, operating systems, data sources and data structures. The tools and services for supporting the development and management of applications in such an environment must change consequently.

Therefore the need for the development of a methodology consisting of models, methods and techniques that support the comprehensive planning, construction, integration and maintenance of a Geoinformation Utility, in a federated environment.

2. GEOINFORMATION UTILITY

Information is a commodity; it has value, can be bought and sold, and value can be added in various processing steps; it is even a very special commodity, as it is re-usable; but it is only tradeable if it is known, wanted and available. Geoinformation, like any other commodity should be provided by an utility: a Geoinformation Utility.

A Geoinformation Utility is needed to improve access, sharing, integration and use of Geoinformation, to support decision making at different levels (horizontally: across different thematic databases; vertically: from local to national levels). Indeed, as already stated, extreme high costs of data acquisition (60 to 80% of the project's costs), and extensive development of sophisticated applications, increases user demands for cost-effective access to existing information. Also, wide and easy access to

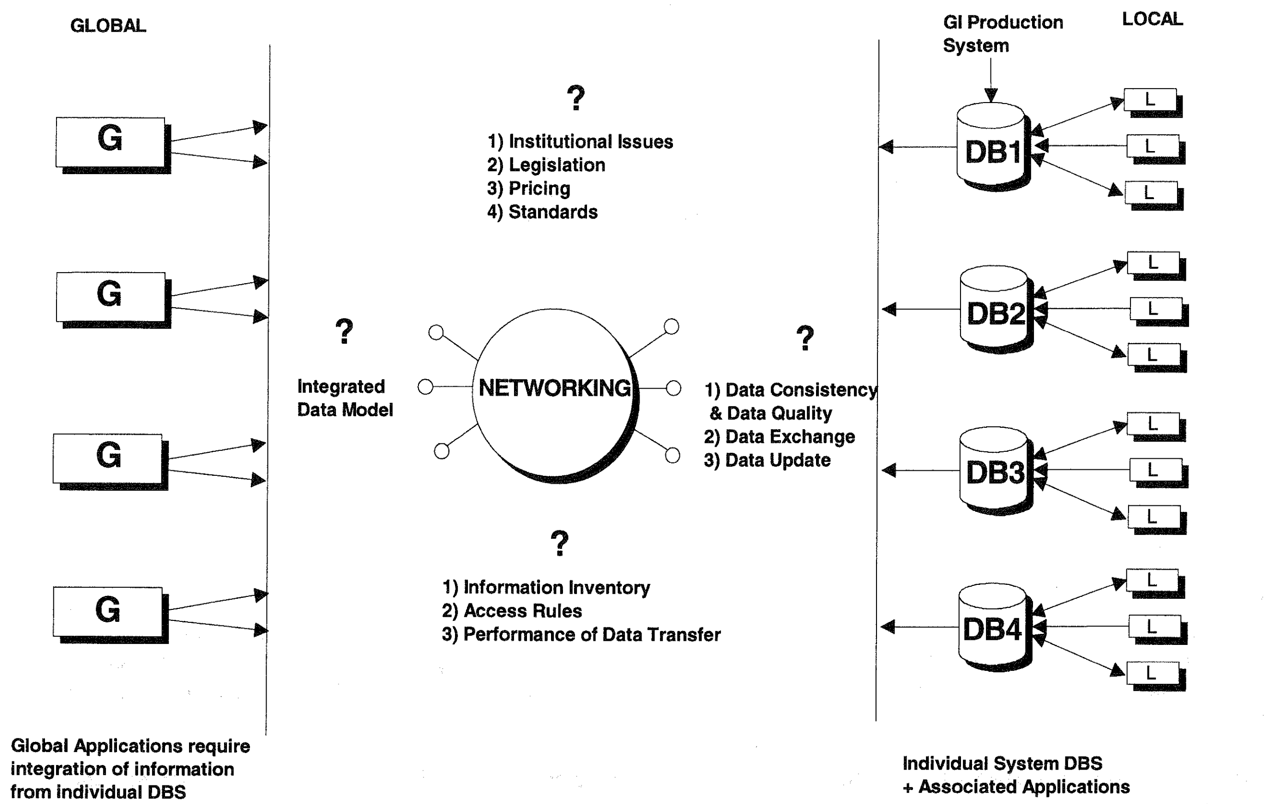


Figure 1: A conceptual model for a Geoinformation Utility

information, requests the development of a platform to formulate, agree upon, and enforce information policies and legislation.

A Geoinformation Utility can be described as a set of institutional, technical and economical arrangements, to support the availability of relevant, up to date, and integrated geoinformation, timely and at an affordable cost; to support decision making processes related to country's sustainable development. A high level conceptual model of such a Geoinformation Utility is given in Fig.1 [C. Paresi, M. Radwan, 1995].

3. TRENDS

Co-ordinate actions, at top political level, related to the development of National/Regional/Local Geoinformation Utilities in North America (eg. FGDC in the USA, Delta X and CANOGIS projects in Canada), Europe (E.U. DGXIII report on "GI2000, Towards a European Geographic Information Infrastructure [EGII]"; creation and actions of EUROGI; work of CEN on normalization), Australia (establishment of a National Resource Information Centre), and Africa (AFRICAGIS, African Information Highway) to name but a few, are proves of the relevance of those developments for the planning and management of the country's scare resources.

From a technological perspective, progress in dealing with aspects of heterogeneity in a distributed environment; advances in data exchange standards in the context of heterogenous database design; and advances in data communication (highway/super-highway), are proves of the trends toward operationalization of inter-organizational information systems.

4. A RESEARCH AGENDA

In many countries, there is a nation-wide pressure to create a national Geoinformation Utility, that makes geoinformation, hosted in geographically distributed GIS's in the country, accessible by the GIS user community. The objectives are to minimize duplication in information gathering, and maximize data sharing; furthermore, to

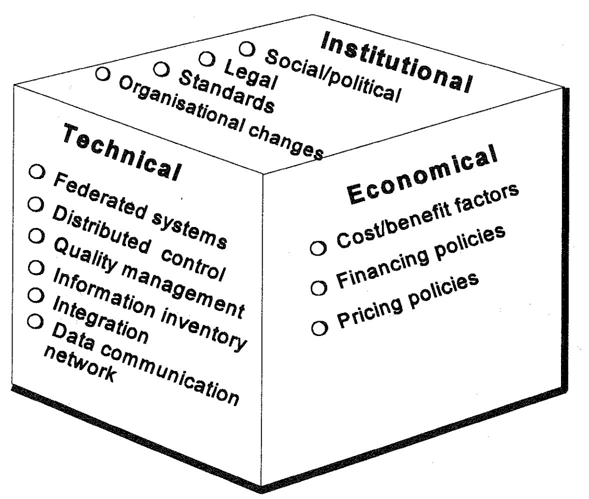


Figure 2: Integration of institutional, technical ai economical aspects of Geoinformatics

facilitate access to geoinformation in such a heterogenous environment, and to provide the mechanisms to ensure that standards for data production and exchange are available and maintained. The system should also support integration, at different levels of decision making (horizontally: across different thematic databases; vertically from local to national levels), of distributed databases in a federated system perspective. At the ITC, research in this field focuses on the integration of institutional, technical and economical aspects of geoinformatics, to realise the gradual implementation of a Geoinformation Utility in different countries. This integration is represented in Fig.2.

5. RESEARCH PROGRESSES

5.1 Definition of a planning and control framework

The development of an Information Utility in a well-coordinated and integrated manner requires the setting of a comprehensive and stable planning framework in time perspective. The planning framework should seek to translate the long-term Geoinformation Utility development objectives into strategic specific courses of action in which each step provides a tangible results that are relevant to the needs and of a high priority. However, each step will also be subject to the general policy direction that is required to control the overall development process [S. Juma, 1994; A. Bassolé, 1995].

The planning framework should define organizational activities within the first stage of developing the geoinformation infrastructure. It should set up the policy framework and institutionalises the executive planning and control. It further should spell out what activities should be performed and what should be achieved in this specific period of time. It should finally serve as a framework for the preparation of detailed operational planning, execution and control of the tasks to be carried out. The policy framework should cover specific issues concerning the various types of data to be collected, data communication and quality standards, marketing, pricing and sales of products, and fee structure, data access, use and distribution. Other policy matters include, data ownership, data integrity, security and control.

A National Advisory Board is proposed to direct and control the development process. Under this top level Committee are the various task forces responsible for dealing with each individual subject or groups thereof expressed in the policy framework. These are Task Forces on standards, access and use policy, organisation structure and changes, marketing, costs and revenue structure, and on system integrity.

5.2 Technical framework

5.2.1 A model for a Geoinformation Utility in a federated perspective

The coordinated sharing and interchange of computerized information can be controlled among autonomous databases by a Federated Database System (FDBS) which is a collection of independent and possibly heterogeneous database systems united into a loosely coupled federation, in order to share and exchange information. The basic elements of the federated architecture are individual components (geoinformation systems) that wish to share and exchange information. A component may be viewed as an autonomous database. The components must maintain as much autonomy as possible; however, the components must be able to achieve a reasonable degree of information sharing. The model illustrated in Fig.3 is in principle technology-independent. Each participant's database contains a specific land-related information plus other types of information, useful for their specific applications; also each database contains a land-related information which is of interest for other participants and in some cases this information is vital for their decision making process. Furthermore, some applications on national level (global applications) require integrated information from the participating (different) databases [A. Al-Ansari, 1994], [M. Radwan, Y. Bishr, 1995], [Y. Bishr, 1996].

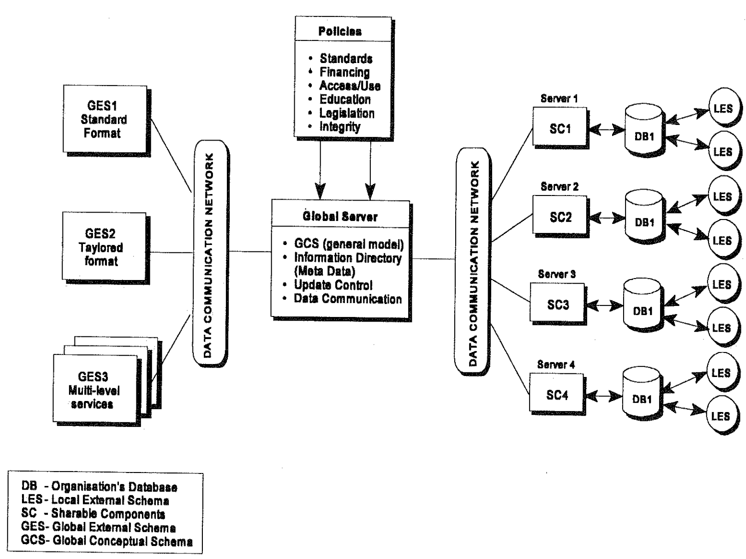
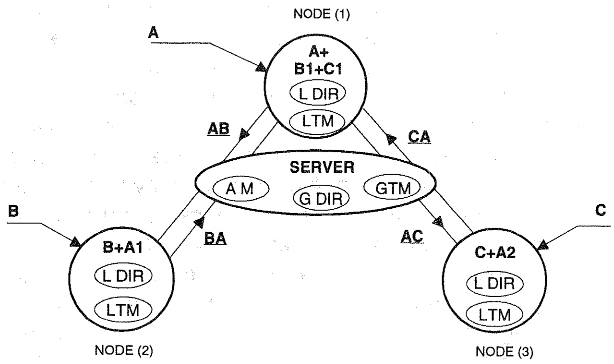


Figure 3: A proposal for a model of Information Utility

5.2.2 Update control in a distributed environment

The client/server is an advance concept in the database field, which is maximizing the degree of data sharing and system performance between different users. That concept can be used in the proposed model. There are two main objectives for that model: to propagate the update from the original copy to the imported one, and to verify the update from authorised node, and then propagate to other nodes. The proposed server has for main functions to support different views for different users (i.e. the abstract A1 from A for node 2); to control the updates in the net (ie. to propagate any update which took place at any primary node to the abstracted copies in the net; to include criteria [specifications] to evaluate the updates that might take place at other [non-specialised] nodes; to make use of the updates made at other nodes [after being approved] into the master data set A as well as the other copies of A in the net). Controlling the updates between different nodes can be applied by a set of rules [I. El-Sharaki, 1994].



L.DIR : Local directory	LTM : Local transaction manager
G.DIR : Global directory	GTM : Global transaction manager
A.M. : Application Manager	

Figure 4: A model for update control in a distributed environment

5.2.3 Analysis of well established Information Systems Development Methodologies

The development of a Geoinformation Utility shows properties that must be carefully analyzed and designed in order to meet the requirements of the customers; in terms of customers, one should differentiate between Local Users, around the information node in the network, and Global (Federal) Users requiring integration of information from various nodes. The development of such an utility is to be guided by System Development Methodologies that can best analyze the related complex problems (eg. distribution of data and processes, communication aspects and fault tolerance), structure them and make them easier to deal with.

Four well established classes of System Development Methodologies (Soft, Structured, Socio-Technical, and Object Oriented) have been examined in an attempt to identify an optimum methodology for the development and maintenance of a Geoinformation Utility based on Federated Database Systems [P. Addai, 1995]. Critical success factors for this development and maintenance have been identified, together with the activities to achieve them. The assessment of the methodologies has been based on how well they support those activities.

The result is that none of the well established methodologies fully support the development and maintenance of a Geoinformation Utility; they are too much targeted to the development of centralized information systems. An optimum mix of methodologies is recommended. Considering the four generic phases of the development of information systems, ie. System Strategy and Planning, System Analysis, System Design (global and detailed), and System Realisation (including implementation, operation and evaluation), a mix of Soft and Structured methodologies is recommended for the Strategy and Planning phase, and a choice between, or a combination of Structured (project phasing approach) and Object-Oriented (product approach and horizontal/vertical integration) methodologies for the remaining phases of the system development.

5.2.4 Other on-going research projects

* ***Quality management:*** Quality of a product or service is referred as meeting the requirements and/or expectations of the client, and is of strategic importance for an organization. In a distributed environment, information is generated by heterogenous databases and processes following different quality policies and systems. The quality management of the Geoinformation Utility is complicated, and its integrity can be affected, by those heterogenous quality management processes. Total Quality Management (TQM) standards and practices in industry are analyzed, the need of geoinformation production organizations for TQM is assessed, and an attempt is made to develop TQM models, as well as implementation and maintenance strategies as an assurance for Geoinformation Utility integrity [M. Sarpoulaki, 1994; D. Musiega, 1996].

* ***Prototype for federating heterogeneous GISs to support decision making:*** The main objective of the research is to develop a spatial decision support system prototype for federating heterogeneous GISs in order to share data and knowledge in a distributed environment, in the framework of developing a multi-decision support system for environmental decision making [M. Radwan, Y. Bishr, E. de Espinoza, T. Mabote, 1996].

* ***Geoinformation infrastructure for data and services sharing:*** The objective of the research is to set guidelines and to develop a prototype for accessing database contents from different application nodes; the system aims at federating heterogeneous systems where the heterogeneity could be semantic, syntactic and/or schematic. It includes the design of, and access to meta data and the evaluation of wether it should be decentralised or not; the assessment of the different alternatives of networking, and outline criteria for selecting the most proper networking protocol that supports the geoinformation infrastructure requirements (including performance analysis of alternatives); and the study of alternatives for the design of the federation's database schema (i.e. standard schema vs. tools for on-line design of database schema, as required by the client) [Y. Bishr, 1996].

5.5 Economical framework:

Beside the institutional and technical issues considered earlier, and in connection with these, justifying the development of a Geoinformation Utility is the economic component of the process. This is particularly important given the extent of the problem space in terms of status and number of organizations involved, range and spread of potential users, domains susceptible to be affected by its use, and geographic zone of influence.

Decision making is in economic thinking a matter of costs and benefits, which assumes that in decision making we are guided by the perceived costs and benefits of our actions. However, research on the topic reveals that in practice it is rather difficult to apply the common technique of Cost-Benefit Analysis (CBA) to justify geoinformation projects, comparatively to other economic activity areas like industry, agriculture, or trading. In addition, the development of a Geoinformation Utility induces opportunity cost for the society, and association of public or social-oriented objectives with commercial objectives.

A model combining the prediction and aspiration concepts for providing an economic shape to the problem of developing a Geoinformation Utility has been developed. The method is an extension of the CBA that takes into account the long-range characteristic of the Geoinformation Utility investment and provides, in addition to the content of the CBA asset, a way of integrating the eventual post horizon growth of the system. A decision model is provided using two decision rules based respectively on the present worth of the project, taking into account the cash flows during the post-horizon time, and the minimum growth rate of the return in the post-horizon time [A. Bassolé, 1995]. This proposal is made mainly because the problem to solve is a long-range investment problem, that in addition is to be considered as a public investment, despite the private sector involvement. Indeed, the main goal of the Utility is not purely commercial; it is also social in the sense that it aims at generating in the long run optimal use of geoinformation by avoiding redundant data collection throughout the country, thus leading to savings in both the public and the private sector, and also at reducing uncertainty in decision making. Its services, at maturity should reflect the character of public utility the same way telephone or electricity services do. As such, it is better treated as a long-range public investment.

6. CONCLUSION

The reported research efforts and implementation of the results in countries like Argentina, Burkina Faso and Colombia, to name but a few, have demonstrated that there is no single solution for the development and maintenance of a Geoinformation Utility, will it be at a national, regional or local decision making level. However, a generic approach to the problem can be considered as shortly outlined hereafter.

Country's related socio-economic constraints are deterministic; combined with locally determined information needs and requirements (from a business management perspective), they permit to determine the Critical Success Factors for the development and maintenance of Geoinformation Utility in the specific context; the challenges and the need for change. A subsequent analysis of the existing information supply (based upon eg. a S.W.O.T. analysis), of the information requirements, and of the Critical Success Factors in terms of activities to be undertaken, will lead to the definition of a planning framework, including institutional, technical, and economical strategies, and the definition of priorities for development. This will in turn lead to a workplan for the development of a Geoinformation Utility [M. Radwan, 1995].

7. REFERENCES

P.A. Addai, 1995; Selection of an optimum methodology for the development and maintenance of a Geoinformation Utility; ITC M.Sc. thesis.

A. Al-Ansari, 1994; Guidelines to establish a National Land Information System for Bahrain; ITC M.Sc. thesis.

M. Bangemann, 1994; Europe and the global information society; recommendations to the European Council.

A. Bassolé, 1995; Towards the development of a National Geoinformation Infrastructure in Burkina Faso: the "Réseau d'information urbaine" as a strategic starting point; ITC M.Sc. thesis.

Y. Bishr, 1996; A hierarchical spatial canonical data model; towards federating heterogeneous GISs; XVIII ISPRS Congress, WG III/IV.

I.F. El-Sharaky, 1994; The concurrent update for parcel based Land Information in distributed environment; ITC M.Sc. thesis.

E. de J. Espinoza, 1995; Development of a federated database system in a distributed environment for environmental decision making at different scales; ITC M.Sc. thesis.

S.Y. Juma, 1994; Incorporating institutional issues in the development of a Geographical Information System; ITC M.Sc. thesis.

D. Musiega, 1996; Quality management for a geoinformation production organization; ITC M.Sc. thesis (under preparation)

C.M. Paresi, M.M. Radwan, 1995; Development and maintenance of a Geoinformation Utility in a distributed environment; an analysis of influencing factors; AFRICAGIS 1995, Abidjan, Ivory Coast.

M.M. Radwan, 1995; Information System Design; Internal report on a consulting mission to IGAC, Bogota, Colombia, 1995;

M.M. Radwan, Y. Bishr, 1995; Towards a multi scale Decision Support System for watershed management; AFRICAGIS 1995, Abidjan, Ivory Coast.

M.M. Radwan, Y. Bishr, E. de J. Espinoza, T. Mabote, 1996; Federating heterogeneous databases in a multi-level Decision support System for watershed management; a client/server approach; XVIII ISPRS Congress, WG IV/6.

M. Sarpoulaki, 1994; Quality assurance system for the geoinformation production programme of NCC of Iran; ITC M.Sc. thesis.