Telecommunication and GIS - opportunities and challenges

Dr. Andrew U. Frank
University of Maine and
National Center for Geographic Information and Analysis
Orono ME 04469 USA

(207)581-2147
email: FRANK@MECAN1

Abstract

This paper discusses the contrast between the rapid development of information technology, including telecommunication, and the slower pace of the use of the new technology. It differentiates the rates of change between hardware and software innovation and the slowness of organizational adaptation to new technology.
Telecommunication technology is the prerequisite for sharing of data in a Geographic Information System. The technical solutions seem to be available but we have to see what impediments have still to be overcome. A short list of potential forms of organizing the sharing of spatial data is given and specific problems analyzed. The major difficulty is with updating data shared by several organization and a conceptual model is lined out. The formal description of data quality and its assessment with respect to tasks is another major open problem identified.
The telecommunication industry can contribute to the success of GIS by providing the technical means to distribute and share spatial data. It profits from this potential by acquiring the base spatial data from other sources and thus freeing its own efforts to maintain its network and facility data.

1 Introduction

Information technology is a rapidly changing field and its innovative ideas and accomplishments will affect the design and use of future geographic information systems. Telecommunication as a technology will affect the development of GIS and the telecommunication industry will benefit from using the GIS technology.

In order to make some predictions about GIS technology within the next decade, I will start with an assessment of the development of new technology. The analysis shows that we tend to underestimate the development of new computer hardware. This seems to apply to telecommunication technology as well. On the other hand, our expectations about improved complete systems are higher than what industry can deliver.

The major challenge is the adaptation of organizations to this new technology and the identification of new ways of using it. This is particularly true for telecommunication. Telecommunication permits transfer of data and thus affects how organizations interact. This is of particular importance for GIS, as one of the major concepts in spatial information systems is to share the considerable expenses of collecting and maintaining the data - especially the spatial data - and use it for multiple purposes. It is succinctly expressed in the

Frank, Andrew U. "Telecommunication and GIS - Opportunities and Challenges." In *Networking Spatial Information Systems*, edited by P. W. Newton, P. R. Zwart and M. E. Cavil, 235-50. West Sussex, England: Wiley, 1995.

notion of 'multipurpose cadastre'. However, the technological advance alone does not yet guarantee success.

It will therefore discuss the problems created by sharing data between organizations. The interpretation of data can vary, depending on the experience of the person and the task to be executed. Variability increases with distance between the collection and the use of the data, thus data sharing increases this problem. In order to gain some understanding a taxonomy of different organizational arrangements for data sharing is outlined. We are also able to individualize some of the components of the problem, e.g. the influence of spatial concepts, data model or data quality.

The technology advances create much excitement. A careful assessment shows that the most difficult problems today are concerned with the integration of new technology, which by itself is not of value, into the organizational framework. Telecommunication resolves some old problems in this field and facilitates data sharing without time delay. But removing this problem just opens our view of the next one, namely organized sharing of data.

2 GIS and its components

Spatial information systems are used here as a general term to describe a system for the storage and management of spatial information. A spatial information system can contain a variety of spatial data and be used for a multitude of applications. While other (conventional) information systems, such as office information systems, deal with business data, the data represented in a spatial information system are closely tied to some location in space. The software system must then provide the appropriate methods to analyze and manipulate these spatial data.

Geographic information systems (GIS), land information systems (LIS), and automated mapping and facility management (AM/FM) can be seen as subgroups of spatial information systems suitable for specific spatial applications. LIS is most commonly associated with legal information about ownership of land, records of deeds, and multi-purpose cadastres, while AM/FM systems are tailored for mapping and managing the infrastructure of public utility companies, such as utility or telephone companies. GIS is a frequently used term for small-scale systems covering a larger geographic area. They are most often used for planning and resource management, while LIS and AM/FM are primarily used for administrative tasks.

A spatial information system consists of four major parts. In order to predict future developments, it is beneficial to examine their development separately:

- the institution--people, their management, organization, and methods--using the information system;
- data describing some part of reality and stored in the information system;
- programs to manage data and derive the desired information; and
- computer hardware to store, communicate, process, and represent data.

3 Technological Revolution

3.1 Hardware

New hardware is being produced at a great rate. The trends commonly observed are: CPU speed doubles every year, memory capacity--both primary and secondary--grows at a similar pace; at the same time, the physical size of a whole system shrinks. Communication speeds increase rapidly. Currently, there are neither physical nor practical indications that the climax of this development has been reached. However, the actual penetration of new technology is slower than the trade journal make it appear. Often several years pass from announcement of a new product to its widespread use.

In lieu of discussing these trends in detail, a description of the power of a personal GIS workstation, as it may become available during this decade for a price that is less than the price of a new car, is probably more illustrative:

- a CPU with 500 million instructions per second (MIPS),
- 500 Megabytes of main memory,
- 5 Gigabytes of storage space on hard disks and additional 50 Gigabytes on optical disk,
- a workstation screen with 2000 * 2000 pixels, and
- a communication device with 100 Megabits/second transfer rate connected to networks that reach a very large number of other devices

Compare these figures with our current workstations: 2-5 MIPS, 4 MB main memory, a 100 MB harddisk, 600 * 900 pixels screen, and a 10Megabits/sec Ethernet adapter.

3.2 Software

Software development is expensive and time consuming. Recognizing the growing gap between improvements in hardware and software during the last decade, the problems of the current software crisis have been widely discussed. Costs of software systems have dramatically grown, but the worst problem is that we often seem not capable of producing the software necessary at all.

What are the reasons for such poor "performance"? Compared with hardware developments, innovative accomplishments for software engineering have been rare and most of the concepts and ideas used today are quite dated:

- The programming languages currently used in industry, such as COBOL and FORTRAN, are almost 35 years old. The popular "new" languages, such as Pascal and C, have been around for 20 years and are slowly taking over.

- UNIX, the "latest" standard in operating system dates from the time when slow teletypes were the common character input device and at least its user interface is optimized for this past hardware.

- Novel ideas, such as parallel programming, are not yet ready for commercial programming, despite tremendous research efforts. The hardware is available, but the necessary software tools to exploit the innovations are missing.

3.3 Data

The economic life cycle of hardware is nowadays about 3 to 5 years and software systems are used for 7 to 15 years. Data managed and processed by such systems persists for a much longer time. In a GIS, data may be kept unchanged for several decades. For example, parcels or communication lines in a rural area change very seldomly.

At the same time we observe the enormous costs for collecting spatial data and keeping the collections up to date. Over the lifetime of a GIS we observe an approximate 1:10:100 ratio between the costs for hardware, software, and data, respectively. In order to reduce the costs of a GIS, one should concentrate on improving data collection and maintenance.

Our primary attention must be focused on the maintenance of data collection, their long-term usability. Loading data into a system for the first time may initially appear to be the key impediment; however, such a collection without updates will become quickly useless.

3.4 Institution

Introducing a new technology in an organization requires that the organization adapts to it. From our everyday experience we know that organizations are very slow to change. Introducing new technology has the potential to change some aspect of an organization. As organizations are slow to adapt to new technology, the technology is either not introduced or is introduced but not used to its full advantage (and sometimes not used at all). There are good engineering methods to construct new hardware and there is also software engineering emerging as an engineering discipline. However, there is no comparable method to guide us in the adoption and introduction of new technology in our institutions.

3.5 Assessment

It is surprising to observe that the development of new hardware creates so much excitement and receives so much attention in the public eye - at least in the technology oriented societies of the anglo-american world. This, despite the fact, that much of the new hardware is not really used (or usable) and at best serves to replicate previous manual methods without fulling realizing its potential. This is specifically true for the field of telecommunication and computer networks, the only widely used application is electronic mail, which duplicates the postal service . Less important, but still widely used are services to distribute stock prices, duplicating the old 'Reuter' system. Really new, innovative uses are extremely rare and not widespread.

4 Challenges and problems

Part of the challenge stems from the quantitative change that we have outlined above - faster and less expensive hardware every year. Another challenge is responding to the qualitative changes that become possible. Unfortunately there are often problems associated with coping with rapid change.

4.1 Rapid Change in Technology as a Problem

New equipment is placed in the market at a growing pace in order to replace "old" equipment. This shortened life cycle causes specific problems to the establishment of information systems.

• Decisions to buy a new computer are difficult in the light of new models appearing on the market every 2-3 months. In consequence, the market for used equipment is very volatile and prices drop rapidly.
• The costs of introducing a new system are high. Problems are created by the transfer of existing programs and data to the new system. Training efforts are necessary as well.
• Maintenance of a computer system has become a major economic factor, with systems becoming economically obsolete due to maintenance costs which are higher then the cost of new equipment.

Since the pace and frequency at which new technology is being introduced is so high, it becomes increasingly difficult to make technical and managerial decisions. A delicate balance between using the latest advance and creating a slightly outdated, but stable environment must be found. Clearly, the decisions for a system should be started with an analysis of the organizational goals and the data available. Then one should proceed to consider software solutions and only at the end decide on the hardware. Observing this order minimizes exposure to technology change.

4.2 Qualitative change

Change not only allows us to do the same things faster and with less cost, but it can qualitatively affect the way we do things. Probably our biggest challenge is to understand how to use this new potential to our best advantage.

In order to minimize change, in a first phase,'replication of existing services', the new system completely and exactly replaces the previously manual task. This causes least disruption in the way the organization works, but also makes only a limited use of the potential of the new technology. It carries forward all the constraints the previous technology imposed on the organization and to which it adapted over the years. Only in a second phase, 'reorganizing to exploit new technology', the new technology leads to a solution which truly exploits the potential of the new technology and liberates the workflow from the restriction of the previously used one.

It is not only the larger capacity and faster processing speed of devices that change the way computers can be used, but also their physical size and weight. The new laptop computers which weigh less than 3 kg does not replace the standard personal computer, but are used during travel and in meetings. It effectively replaces the notepad and pencil which we all use to carry around--with the addition of access to personal databases, spreadsheets, etc. One can predict that in the near future, GIS functions will be implemented on machines which can be carried into the field. How will we use them?

4.3 Adaptation of personnel

Problems with the introduction of innovations result--among other things--from the fear that the new technology will reduce work and make positions redundant, and the fear of existing personnel that they will not be able to cope with the new technology. Both are quite unfounded, but nevertheless very important psychological problems: general experience is that GIS does not lead to a reduction in the workforce and with appropriate training efforts, current personnel continue with new job descriptions. In most cases personnel are very proud of the increase in quality of their work due to the GIS system.

4.4 Change in power structure

A second set of problems results from the fact that "information is power," even within a single organization. The introduction of information technology (and telecommunication especially) is clearly changing the access to information and thus reduces the perceived power of those that have access today. There are no easy and simple solutions to these problems. We can only be extremely careful not to have the technical solution dictate organizational arrangements which require extensive organizational adaptations.

5 Influence of telecommunication technology on GIS

Communication technology is, for the GIS applications, an eminently important enabling technology. It permits sharing of data, instantly and ubiquitously. It takes away from the user the need to know and understand where data is produced, stored and distributed. It brings a lot of potential, but this potential is not necessarily easy to use as it requires adaptation in software, the organization of the data and the institutional arrangements. -

5.1 High speed communication links

If high speed data communication between offices becomes available data can be exchanged quickly. Currently one of the limiting factors determining the architecture of a GIS is the bandwidth available for communication for medium and long distance. The need for interactive graphics at the users workstation and the lack of high speed communication forced the storage of data at the workstation site.

The maintenance of a central repository and the sharing of the same data between many users has been hindered by this architecture. This has forced organizations to enforce policies and mechanism to prevent more than one user updating the same dataset. With the new high speed links using inexpensive cabling, a central repository and remote graphical displays becomes potentially feasible.

5.2 Interactive network protocols

Not only the communication speed but the total turn-around time is critical for interactive work. Current local area technology permits general interactive work but has delays that are not acceptable for operations which require immediate feedback (e.g. pointing with a mouse, menu selection etc.). In lieu of investing in improving telecommunication a study for optimal distribution of tasks between a central processor and a workstation was undertaken at MIT. The X windows [] protocol allows a distribution of tasks that has all processes critically dependent on immediate feedback run locally but still permits the bulk of the application program to sit on a central server.
This protocol has become a de facto industry standard, available from most of the workstation manufacturers today and it works well to connect equipment from different vendors. It is even available on so called X-terminals, which are a new, extremely smart kind of display terminal. X windows allows access to an application on a brand A machine from a workstation or terminal of brand B, independent of underlying operating system or hardware instruction set.

5.3 Packet networks permit distributed databases

If public service packet networks over wide areas become feasible, communication with multiple points becomes economically available. We can then build databases which are widely distributed and thus require multipoint communications between them. This is a known technology but not widely used in the GIS community. However, it could be used to connect databases in branch offices of large public utilities or organize regular data exchange between multiple utilities in the same area.

5.4 Assessment

Telecommunication as a technology has great potential, but is, as other advances in hardware and base technology, slow to penetrate the application fields. We can see two limitations, the lack of software to exploit it and the lack of acceptable organizational structures to use it.

• The necessary software to use the communication technology is not yet ready. It should be sufficient to remind you that only this year acceptable mail programs became available, such that users need not be concerned about gateways and other details of the low level layers. Many of the large electronic mail research networks in the U.S.A. are now connected, but the commercial mail servers are still not communicating - and this more than 15 years after the success of Arpanet. Distributed database software - researched for more than 10 years - is just now emerging commercially and not many wide area network uses of distributed databases are known. Similarly, the new GIS software on workstations, most of them likely connected to a local area network with high speed communication, does not make use of these facilities except for file transfer.

• Technology can only be used if the humans using it have the appropriate social patterns that govern the usage; such patterns were available for electronic mail or for the distribution of stock quotes. They seem however to be missing for the organization of data exchange on an ongoing basis. Within these social patterns we may count the expectation for cost of services. At least the research community got used to 'free' electronic mail (and other wide area network services) and it seems difficult to establish acceptable charge schemes.

It seems that the fundamental concepts to build the software to use the advances in communication technology are mostly available. We primarily lack the tools to harness the complexity of some of the issues and have problems to design, build and market commercial software. From a research point of view, understanding the organizational aspects is more of a challenge. We need to understand what data we communicate and how they can be used.

6 Fundamental problems of sharing of data

The technical advances of telecommunication allows sharing of data and it appears as if many of the technical problems were have been solved or are at least close to a solution. How is the GIS community going to use this? There are some fundamental problems with sharing data between organizational units. They stem from the character of information. Understanding these limitations in the first place allows us to avoid problems caused by expectations that are too optimistic and should help us to select arrangements and procedures which will avoid at least some of them.

6.1 Data and reality

Humans perceive an image of the real world they live in and interact with. They express the results of their experience in words or other symbols which communicate these perceptions to others. Data in a computerized database are just an extreme case of such symbols, extreme because they are more formalized than the symbols in natural communication between human beings and the person originating the communication and the recipient are not in close contact, sometimes even do not know each other..

The symbols communicate not a state of reality, but only the reality as perceived by that individual. Perception is affected by the experience of the person and the task to be solved. It is necessarily selective, because we cannot deal with the complexity of reality. It is exactly this flexible selectivity of the perception process which is at the base of humans effective dealings with reality . We can switch very quickly from one viewpoint to another, which may entail a completely different set of detail to be considered.

The expression of the results to others in symbolic form depends on agreed upon codification of ideas into symbols. This is a process, influenced by culture, tradition etc. It is surprising to note to which degree professional education and experience, or organizational "culture" changes ones vocabulary.

This all sounds hopelessly subjective and you will ask, why does human communication work at all, why can we successfully build and use databases? The relatively new experientialist viewpoint points out, that some of the base experiences are common to all humans. Experiences like eating, gravity etc early on in life are essentially similar to all humans, as is the physiology of the perceptive apparatus (eyes etc). Thus the early and fundamental experiences human make, and which shape later perceptions, are essentially similar. Cultural and other differences influence only later, more advanced level of understanding the world .

There are some significant problems based on an apparent difference in the construction of classification schemes in human thinking and in formal systems. In formal systems, classes are built such that they assemble all the entities which have a given set of properties (e.g. all things which bark and have four legs are dogs). Cognitive scientist indicate that humans form classes more around a prototype (e.g. the prototypical dog would be a german shepherd) and then include other entities based on their similarity with this prototype. Some people hesitate to call a small chihuahua, 'dog', but nearly everybody would include a german shepherd, that lost one leg in an accident, in the class of 'dogs'.

These problems are more important in geographic information systems than in most commercial applications. A commercial system and its organization defines its own reality; most of the concepts in a commercial system do not exist in reality but are defined by the organization. Users expect from GIS that the data represent reality faithfully, using whatever concepts they are used to or are applicable for their task.

6.2 Modelling of geometry

6.2.1 Spatial concepts

A very special problem arises with all systems that have to deal with geometry: there are several conceptual bases for perceiving space. It is not as simple as geometry seemed back in school. Euclidian geometry is just one method to structure our spatial experience, a very important one, but not the only one. Humans are very good at switching between different spatial concepts and selecting the most appropriate one for a task. As an example, if you plan to drive from your home to the university, you will use spatial concepts that deal with spatial objects in a sense close to euclidian geometry to get to your car, but you will use a network concept to navigate the highway system. Current GIS have difficulty in modeling space in more than one way. AM/FM systems make extensive use of network concepts, geographic information systems often think of space as a collection of points and cadastral systems use a topology structure.

6.2.2 Spatial data models

From the spatial concepts we use to structure our perception of reality we have then to build data models. A data model, as usually defined in the database literature is a set of conceptual tools provided to the user, the database administrators specifically, to structure the data. The geometric data model is the set of tools to model spatial aspects of reality and it differs from the spatial concept by two requirements: it must be formal and it must be implementable. A spatial data model should provide a set of operations and definitions to describe the spatial concepts used to structure the data.

A data model must also be implementable. Most concepts of space define space as infinitely divisible (between any two points there is another point) and extending in any direction to infinity. Thus a discretization is necessary to pass from the concepts to the model, a typical step in many spatial modeling efforts. Geometry on a grid - this is how euclidian geometry is essentially carried out in a computer system - is substantially different from the euclidian notions we learned in school and has not yet been formalized.

6.2.3 Geometric data structures

A geometric data structure is then a specific implementation of such a geometric data model. If a geometric data model describe what geometric concepts are available and what operations are provided to manipulate them, the geometric data structure deals with the details of how this is achieved. Typically there is more than one data structure that implements the same data model. For example, the raster data model can be implemented using full arrays, run length encoding or quadtrees. The operations these implementations offer should produce the same results, but they may be widely different with respect to performance, storage space needed etc.

6.3 Communication between organizations

The specific problem of sharing data is therefore to communicate meaning in a situation where the perceived reality differs, because the two organizations have different goals and work on different tasks. Clarifying these differences is very difficult, because they appear only in the details: most people will agree what a building is, may even agree on the difference between commercial and residential. But how do you define the classification of an old factory, that is remodeled and includes several flats and a public library?

There is obviously no correct answer, and it all depends on what you intend to do with the data .
If you share data with another organization then you have to understand that organization and its goals in addition to your own. If your use and their use of the data is very similar you will likely encounter less problems than if you differ substantially. For example it may be easier for two public utilities to share the 'land base', then sharing of data between a utility and the town building code enforcement department.

6.4 Data quality affecting sharing of data

The concepts that are applied to organizing the data are not only the categories used for classification of the data but also the data quality requirements. There are a number of aspects to be considered, which range from geometric precision, completeness of the data set, specificity of the categories and quality of assignments of entities to classes, level of updatedness etc. Organizations may agree on a data set, but their expectations on data quality, in any of its dimensions, may be very different. Increasing data quality is expensive and therefore the level of quality necessary has to be defined carefully at the outset. Unfortunately, we do not yet know a lot about how to characterize data quality. Data quality descriptions are crucial for the exchange of data between organizations and the U.S. Committee for Geographic Data Exchange Standards prescribes the inclusion of 'lineage descriptions' with the data. Ultimately we will need formalized descriptions which can be used in analytical processes.

6.5 What data are easier to share?

From this introduction of some fundamental problems, it follows that it is easier to share data that are commonly used and well defined than data which is very specific to a task. Topographic mapping as a common source for spatial information is only successful, because the set of objects mapped is reduced to object classes which are widely used and have commonly accepted definitions.

7 Characteristic forms of organization of data sharing

Sharing data means, in most cases that one organizational unit acquires a dataset from another one on a one time or permanent basis. In this section I will just characterize the different situations as they may arise e.g. in a public utility or telecommunication industry, and provide a framework to classify different arrangements, all of which are viable.

7.1 Spatial data used as a backdrop

Spatial data needs context to be understandable. A set of points on a screen do not tell much unless we see some other features on the same screen that provide a spatial context. If an organization holds some spatial data, i.e. data with a relation to locations in space, which by itself does not carry enough context to be understood, they may acquire a data set that provides this spatial context. It is easy to understand this as a 'backdrop' that is projected behind the data of interest to establish spatial context.
There is no intent in this case to construct relations between the acquired data and the own data. The two data set exist in different realms and are only graphically overlaid and visually integrated by the human operator. An new, updated 'backdrop' may be acquired later and replace the current one, without affecting the own data.

7.2 Spatial data used in order to start quickly

In order to get a quick start, an organization may decide to acquire spatial data from another organization. The organization uses this data temporarily until it has built its own spatial data set. It may use the data acquired to become part of the longterm maintenance set. There will be links established between their own data and the geometric data (this differentiates this case from the 'backdrop' use mentioned previously).
A promising method seems to use detailed remote sensing images, which provide an excellent overall spatial precision, are up to data to the date they were collected, etc. and then digitize on the screen as one needs the data. Existing data can be overlaid, lacking spatial registration can be detected visually and additional points digitized on the screen etc.

7.3 Spatial data acquired for long time usage

The most intense form of sharing of spatial data is the case where one acquires a spatial data set to integrate with one's own data and plans to receive regular updates for the acquired data from its source.
This intense form of data sharing requires agreements not only on the original delivery of data but also on the maintenance aspects. The problem is not only to get the system started but to achieve long term usability of the data.

This is a concept, popular in Australia and other Commonwealth countries, but not so much in the USA, due to a difference in the copyright for survey data. Technically the TIGER files distributed by US Geological Survey and Bureau of the Census may fill such a role, but based on a wide distribution and no formal arrangement for update distribution.

8 Types of problems encountered

8.1 Problems related to differences in spatial concepts

The integration of data requires that there is some common ground in the spatial concepts employed. A network data set (that does not contain some metric data in form of coordinate values for the nodes, e.g. the famous map for the lines of the London underground) cannot be integrated or overlaid with any map data that uses an euclidian concept. We observe that in most cases there is sufficient commonality in the spatial concepts.

8.2 Problems related to spatial position

In nearly all cases spatial position is expressed using coordinate values. In order to integrate datasets from different sources, we have to transform them to a single coordinate system. But even if the coordinates are expressed in the same projection system, there may be differences. Coordinates are determined based on the given coordinates of control points and these may differ from one agency to another. It may become necessary to locally adjust the coordinates of the data acquired to one's own data. Piecewise linear methods (conflation) have been studied and seem to work, but the manual effort to determine common points and supervise the computation is still considerable.
This problem is least important when acquired data is used as a backdrop and is most difficult if geometric data acquired and own geometric data must be merged. It is non existent, if all the geometric data is acquired.

8.3 Problems related to data quality

A subproblem of spatial integration is the assessment of the quality of the result. This is a field where research is very important, both for economic and liability reasons. Maintaining data with more precision than necessary is very expensive. Combining data with different qualities may lead to results which are not reliable and expose the organization to error and liability.

8.4 Problems related to differences in spatial data model

The data to integrate may be based on very different spatial data models. We may for example see a mostly network oriented model used by a utility and a raster model of land use, integrated by a planning agency. If the data needs to be integrated beyond a visual overlay, i.e. if we want to apply analytical processing and need to establish linkages between entities in one set and entities in the other, one has to reconcile the data models.

8.5 Problems related to entity definitions

On the level of individual data entities that describe individual real objects,in database terminology 'instances', we have to understand the methods each data set uses to define and delimit the entities. If the integration of the two data set is intense, one will need to establish linkages between data entities which describe the same real objects. This is often attempted by just comparing location of objects (using coordinates). Entities may have identifiers, which may occur in both data sets. The most important example is certainly postal addresses and post codes.

9 Update of shared data

If a dataset acquired should be used over a longer period of time, it has to be updated. The same economic pressure that leads to the acquisition of the data from another source in the first place will also indicate in most cases that the maintenance of data should be shared. The savings from data sharing are primarily savings from sharing the effort to maintain the data and avoiding the situation that several agencies all go out and measure the same new building or road individually and separately. Therefore, considering integration of data from another source entails in most cases studying the problem how later updates can be integrated.

This is a relatively new problem, which has not yet received sufficient attention. A formal model can be built in terms of database theory and this will help to clarify some notions and point out some important aspects. We deal with a particular situation of a distributed database, i.e. a database which consists of data stored in more than one location.

A few conclusions can immediately be drawn from this model:

• It is extremely important to maintain a set of stable and unique identifiers for each data entity in the shared data. The provider has to guarantee that these identifiers will not change as long as the data entity does not change and it is probably advisable, not to reuse identifiers of entities deleted. From practical experience we know that these identifiers should be artificial, i.e. constructed by the database management system and not have any significance in the real world.

• From viewing this problem in terms of a distributed database, it becomes evident that the updates must be distributed as 'instructions to change' (with the necessary details) and not as a distribution of a complete new set of entities. If a new set of entities is delivered to the acquiring site, it must painstakingly sort out the changes, i.e. identify all the data entities which were deleted, changed or added. This may be difficult at least and will not recover information about the reasons for a change, which may be important for the acquiree to adapt his data set.

• The acquirer of the data may not change the entities acquired, but has to leave them as they were received from the provider, in order to insure that the data set is an accurate copy of the dataset the provider had at the time that communication between the two partitions of the distributed database ceases. This is a reasonable demand which would also follow from the administrative principle of determined responsibility: the responsibility for maintenance of this part of the dataset is with the provider.

10 Conclusion

The advancement of technology, especially telecommunication technology, creates opportunities and challenges for GIS. The ability to exchange data rapidly with minimal delay and at low cost will greatly influence the technical design of GIS. As a primary result, sharing of spatial data between organizations will become technically possible. The telecommunication industry could benefit from such sharing, as it may gain access to spatial data held by other organizations and may reduce its efforts to collect and maintain 'general purpose' data and concentrate on its specific installation data. It may also use sharing of data within its own organization to exchange data between branch offices etc.

The rapid progress of information technology, of which telecommunication is a part, is not always used to its full potential. Some innovations, like electronic mail and particularly fax, have found extremely fast and wide distribution. Others, like distributed databases, GIS etc. have not yet realized their full potential. Observations suggest, that information technology that duplicates functionality which is currently used but of lower quality (eg. postal service) may be accepted quickly. Other innovation, that affect the way organizations work, are introduced very slowly, restricted by the pace with which organizations can adapt.

We see thus, that the limit to the introduction of new technology is less with the technology itself than with the readiness of organizations to change. Telecommunication will allow sharing of data between organizations - one of the major goals of GIS - which is for most organizations new and thus difficult to adapt to. We have discussed some possible forms for sharing spatial data and the problems they entail in some detail. Much more work, particularly in the formalization of data quality, is required.

Technology should be introduced to serve an organization to achieve its goals better; it should not be introduced for its own sake and it should not force the organization to change to accomodate technical restrictions.
Telecommunication has the technical potential to liberate the collection, maintenance and storage of spatial information from a single place and make the data accessible from any place as the need may be. Most of the low level

technical problems seem resolved and it remains to work on the organizational ones, which include, such diverse topics as fair pricing, responsibility, accessibility and privacy etc. The tools are there to build more valuable GIS - but more remains to be done.