The Balance Between Data Quality and User’s Information Needs – Thinking Strategically for Rational Decision–Making

Francis CHIGIINGE, Botswana

Key words: Accuracy, Standard, Spatial data quality, Reliability

SUMMARY

The validity of a decision is a function of the reliability of the data from which it is derived. Spatial data now plays a core activity in most, if not all organisations. GIS derived results invariably impact negatively or positively on operational and planning decisions. Logically this implies that careful use of GIS-supplied information requires an accuracy assessment of all data involved. Furthermore, the accuracy and quality of data required for different applications is not usually homogeneous, implying that there is a general lack of specified standards of accuracy in GIS data. This calls for the need to determine the reliability of data, which can conveniently be done through considerations of data quality versus the user’s needs. Arriving at a useful, meaningful decision will depend on striking a convenient balance between data quality and user’s needs. This paper explores a cognitive process, the balance (or trade –off) effect that has been shown to influence good decision-making. A brief analysis of elements of data quality is given.

The user’s information needs are influenced and guided by identifiable objectives in an organisation. In the final analysis, the data transformation processes should anticipate, identify and satisfy the user’s needs. It is in pursuance of this data validation that a balance (trade-off) between data quality and user’s information needs results. Perhaps as part of system process design, we should consider what is value (in terms of data quality) to end-users. This brings in the element of ‘technical competency’ in the whole activity series of transforming data. However the main issue to be considered is the selection and optimisation of the data handling and manipulation processes. This leads to defining ‘what data quality to consider’, and this inevitably brings to the fore the role or contribution of these data transformation processes to the needs of users. There is need to define ‘what business we are in’; in a more specific way, what is our product that the end user wants to pull.

A consideration of the future state of these processes is also very important. Like in any normal life cycle, changes in aspiration levels (demands, interests, wants and from end users) will in the long run bring desirable changes to system processes. Already this is evident in that traditionally spatial data have been presented on analogue maps and manipulation has been by manual means. This resultantly constrained the volume of data that could be handled at any one time. This whole scenario changed with the advent of GIS, and more recently GPS positioning applications.

The Balance Between Data Quality and User's Information Needs – Thinking Strategically for Rational Decision–Making

Francis CHIGIINGE, Botswana

1. INTRODUCTION

Change is always important in an organisation's life cycle. This change is more often than not managed in situations of uncertainty and risk. Spatial data is fast taking centre stage on the information super highway, and the problems caused by errors, inaccuracies and imprecision in these spatial data sets are quite obvious (Kenneth E. Foote and Donald J. Huebner). The amount of data available to users is on the increase, as the information architecture is enhanced by positive changes to data access, data management and data analysis, in short, the data system. The management of this system cannot take place in unbounded premises. Decisions need to be made based on certain parameters, and these parameters are defined and shaped by the organisation's environment.

From the end user's point of view, it's no longer sufficient to define quality in terms of accuracy only. The information architect needs to seriously consider and determine user requirements. These requirements can be data set specific, but all the same embodying the elements of data quality. This inevitably calls for professional bodies and organisations to institute best practices and programmes for continuous improvement.

2. DECISION MAKING

It is quite important from the onset to separate and distinguish problem solving with decision-making. In problem solving, one attempts to seek entirely new outcomes or options whereas decision making entails selecting a choice from a set of given alternatives. Both though are processes with distinct stages. In rational decision-making, we are quite clear about what is expected and the final selection of an option provides evidence of the basis on which a decision was reached, and is thorough and systematic. It's more suitable to apply in the whole activity series of data capture, summarisation, processing, analysis and presentation. The particular strategy though, is in response to changing conditions (as dictated by the organisation's environment), product line (map, data plan, diagram, etc.), quality or service, geographic coverage (of the spatial data product or coverage in terms of sphere of influence) and the opportunity sought (is it a new spatial product development, or improvements are done to secure a certain niche in the market).

The implementation of a particular strategy will depend largely on support systems and resources at the organisation's disposal (human, financial and technical). The decision-making techniques, which are widely used in organisations, are largely based upon the rational model of planning, strategy and structure (Stoner and Freeman, 1992). Simon, 1976, challenged the rational model and argued that good decision –making is impossible without high quality information. The other dimension is that problems facing organisations are frequently not well defined. The ability to think rationally is limited by skills, intellectual ability, knowledge, values and interests (Clemen, 1990, Hamel and Prahalad, 1994, Miller, 1992).

The fact is conceded that in complex environments, the rational decision-making comes short. Despite the criticisms, it simplifies things a bit if the underlying assumptions and opinions hold. User's needs can be characterised by internal and external environments. The complexities offered by these environments necessitate high quality decision-making (Van Der Vyver Glen). Decision makers face a rapidly changing world where depending on past experiences alone is not good enough. We are faced with a turbulent world in which the traditional 'good judgement' viewpoint is weakened and the level of external interference is increasing due to globalisation. This should not be misunderstood to mean that globalisation is the causative urgent or otherwise. It is just the author's view of the effects of globalisation. The argument in some quarters is to push for consensus decision-making or to resort to adaptive decision making. While consensus decision-making fully utilises the resources of a group, it is time consuming and may in some cases present difficulties. In most instances, trade-offs have to be involved, as the various decision alternatives will not fit all concerned.

Decision makers need to be aware that differences of opinion are natural and should be expected. What guides decision-making should be what is feasible and likely to yield a satisfactory solution. Resorting to adaptive decision-making techniques breeds internal conflicts within organisations. This is because there is now a multitude of information sources and data can be easily acquired. In view of this, there may be need to have paradigm shifts of how we think about technological, social, cultural and organisational issues. This inevitably creates internal conflicts, which may skew our perception of what really needs to be done. Decision-making may in the final analysis tend to be problematic.

3. QUALITY VERSUS ACCURACY

According to ISO 9000 series, quality is defined as a characteristic that a product or service must have. Organisations must strive to develop and maintain products or services that are; reliable, usable and repairable if they are to remain competitive. Furthermore, services should be courteous, efficient and effective. All these characteristics are of great importance if products enhancements and corrective actions on any reliability issues identified by internal or user analysis are to undertaken. A product or service of high quality must successfully serve the purposes of users. Reeves and Bednar, 1994, give a more specific view, describing quality in one of four general ways; as excellence, value, conformance to specifications, or meeting or exceeding users' expectations. If these quality characteristics are inherent in a spatial dataset then there is value addition to a user manipulating that spatial dataset.

Foote and Huebner define accuracy as the degree of conformity with a standard (the true value). This is actually a component of quality and can be used in a needs assessment to find out how useful the product or service is. Accuracy should not be confused with precision, which is the degree of refinement in the performance of an operation, or the degree of perfection in instruments and methods used to obtain a result. Precision relates to quality of an operation and accuracy relates to the quality of the result. Both precision and accuracy are components of quality of a product or service. These terms have ready application and use in the sphere of spatial data issues.

4. SPATIAL DATA QUALITY ELEMENTS

The right data is that which provides the end user the sort of information he needs about the data. This information empowers the user successful execution of an application at hand. This means there is need to define a high quality (fitness for purpose/use) set of attributes to support the application or task at hand. Spatial data quality elements give information on the fitness-for-use of a spatial database by describing why, when and how accurate the data are. Furthermore, there is also the need to know data formats and types of documentation available. The main idea is to arrive at a framework for both data producers and users. This framework ensures careful, proper documentation on the part of information architects. With respect to users this mitigates the effects of possible misuse of data.

Guptill and Morrison, 1995, outline seven elements of spatial data quality, which were identified by members of the International Cartographic Association's Commission on Spatial Data Quality. The seven elements of data quality are;

- Positional accuracy- this refers to the absolute and relative accuracy of positions of geographic features.
- Attribute accuracy- this refers to the accuracy of the quantitative and qualitative information attached to each feature.
- Completeness- this is the degree to which geographic features; their attributes and their relationships are included or omitted in a data set.
- Logical consistency- this deals with relationships encoded in the data structure
- Lineage- describes the history of spatial data, including descriptions of the source material from which the data were derived, and the methods of derivation.
- Semantic accuracy- refers to the quality with which geographical objects are described in accordance with the selected model.
- Temporal information- describes the accuracy of temporal observations. It relates to accuracy of temporal effects on the spatial data quality such as quality indicators of update ness (date of last update), rate of change and temporal lapse.

These spatial data quality elements have been developed as a result of the increased spatial data production by certain governments and the private sector, mainly because there were no required quality standards. The US Geological Survey and the British Ordinance Survey have long had a mandated quality control i.e. national accuracy standards. Concern was also high for the increased use of GIS as a decision support system. Implications of using low quality data would result in spiralling litigations. Secondary data sources, propagated by the Internet, coupled with the development and growth of data translators, and transfer systems meant the possibility of getting poor quality data was real.

As a result of all concerns, metadata standards are fast taking centre stage in the development of spatial data infrastructures. Metadata or data about data describe the content, quality, condition and other characteristics of data. Metadata is a critical component in a data warehouse environment and helps users in many ways to know what data is available, how it is organised and what it means.

5. INSTITUTIONAL FRAMEWORK ANALYSIS

According to Simon, 1976, management is decision making and that the context of management is becoming increasingly uncertain and risky. Today, it is a fundamental premise of management theory that good managers require well developed decision making skills and the ability to make sound ‘judgement calls’ (Peters, 1993, Webber, 1997). Today we live in a world in which desirable and undesirable changes are continually thrust upon us. Webber, 1997, argues that many professionals today are overwhelmed by the complexities of the tasks at hand that they struggle to make informed strategic decisions.

Perhaps this has something to do with the institutional frameworks. The establishment of the National Spatial Data Infrastructures in most of Southern African countries is a welcome development, which should facilitate the maintenance of standards on data. Future research on geospatial information is likely to centre on these National Spatial Data Infrastructures and much will depend on what will have been achieved in terms of data accessibility, standards and data content validity. National Spatial Data Infrastructures envisage a system of linkages in data sharing, providing that ‘ data collected at one level is suitable for all levels’ (www.ngis.gov.bw). There should be a remarkable improvement in access to information in land administration and management for the betterment of communities they serve. These establishments are deliberate institutional set ups that consider a participatory approach to the development of metadata standards.

Within the SADC region, SDI’’s are at various stages of development, with South Africa (fully developed), Botswana, Malawi and Namibia (some components running) and the rest are yet to take off the ground (Paradzayi, C, 2004). Those developing have much to learn from past experiences, and in this regard, experiences of the National Spatial Information Framework (NSIF) of South Africa will be of utmost importance.

In the design of these projects, it is important to construct a very robust institutional framework for successful implementation and monitoring. No rule of thumb exist on what should precisely be done, serve that policy makers should consider how best to adapt to their own situations. But there is a common denominator in all this. Firstly, the effectiveness of, and respect for, the institutions dedicated to implement the SDI’s are critical to the success spatial data infrastructures. Secondly, it is essential that there be transparency and publicised procedures for spatial data infrastructure processes. Thirdly, decision making at macro and micro-management levels must be clearly divided so that the same decision is not continually revisited or revised. The other aspect is concerned with clear allocation of roles and responsibilities confined to different decision platforms.

The foregoing discussion is quite important for proper, efficient and effective management of scarce resources at the disposal of organisations. The other institutional framework characteristics are the extent to which spatial information is available to the users (Khan, Diane, Strong and Young). Accessibility is conditionality which the product or service thrives on. The design of processes and setting up of procedures should also deal with the extent to which the volume of information is appropriate for the task at hand. This ensures that users do not unnecessarily download huge amounts of information that do not serve their purposes.

The overall goal should be to develop appropriate systems that promote balanced and sustained optimisation of resources through adaptive and investigative research and technologies. For this to happen SDI's should establish adequate feedback systems for internal self-sufficiency. The idea of learning from other countries or organisations cannot be over-emphasised. A sound institutional framework is built upon a strong ability to evaluate local conditions and technological developments in other countries or organisations.

6. TRADE-OFFS (BALANCE) IN DECISION MAKING

The institutional set up of the SDI's have to recognise that data quality is a multidimensional concept. That being the case, we must rely on trade-offs (balance) among aspects of data quality. These trade-offs are a balance between what the experts envisage and what the data user desires. It is suggested to be purposefully flexible depending on the level of expertise and interest of the users (Carol, S.). Technical competency will play a major role here. Institutions with high technical dimensions will be more inclined to deliver what the user community expects.

This makes the development of any standards to be interactive and consultative. Meetings, conferences, workshops etc. should be encouraged to properly scan the entirety of the organisation's environment. The introduction of the Tribal Land Information Management Systems (TLIMS) by the Botswana Government, which is meant to integrate activities in Land administration and management, will mean survey data taken for one purpose may be applicable and used in another sphere. This will ensure information is available at local level to make decisions for local interests (Musisi, Nkambwe). The ultimate aim is to have less time in gathering information for decision-making and afford more time in analysing the data. This way, information can be used in more creative ways.

Any GIS project development must have quality designed into the elements of spatial data. The aim is to meet and surpass users' needs and provide superior value. Projects should be structured in such a way that they meet users' stated and un-stated needs or requirements.

7. WAY FORWARD

It is important to take note that mistakes in databases may be undetectable unless the user has conflicting or collaborating information available. This makes it imperative to make data accessible so that errors can be corrected as they are found, thereby helping the design of systems in future. We must realise that the community indeed has diverse needs and it is important to have end user queries and reporting tools. Having database links is a way meet diverse needs of users. Another option is the creation of file extracts, involving the use of FTP (File Transfer Protocols) or having custom programs written to access data using SQL (simple Query Language). These can range from Perl scripts, which just extract data, to full applications written in Powerbuilder, C, Java, etc.

There should be no limits to what users can do, because we are catering for the diverse set of users. Users should be able to get standard reports at the push of a button, and others in more analytical ways. Users can get pre-built queries and not expected to build reports themselves (Scott Thorne). Most people are comfortable running pre-built queries, but modifications should be able to done to suit oneself.

The needs or requirements of users can be captured in a variety of ways; direct discussion or interviews, surveys, focus groups, users' specifications, observation, warranty data, field reports etc (Kenneth Crow, 2002). Organisations should benchmark against best practices to improve product development in areas of spatial data dissemination. Crow, 1999 observes that we are in an era of new business strategies, new organisational approaches, new business processes and new enabling technologies. Forward-looking organisations should continually improve their product development processes to take advantage of the situation and so raise their business profile.

Management should aim to improve systems and processes, not forgetting that the greatest asset in an organisation is personnel. There is need to understand best practices (what is possible), assess strengths and weaknesses and carry out a gap analysis and improvement plan. It is also important to determine critical success factors to provide strategic direction (what is needed).

8. CONCLUSION

The variety of spatial data, users, and uses makes expectations difficult to manage. It was shown that organisations could do more for themselves and users of their end products if they plan for success. Whilst simplicity works for all, flexibility and diverse functionalities are necessary to strike the inevitable trade-off between system design elements and end user perspectives.

Data need to be presented in a simple form and make interactions with data as simple and straightforward as possible. An easy to understand –easy to use structure is advocated. It was pointed out that quality and usefulness of spatial data could be improved throughout the development and maintenance of a spatial data infrastructure, and that quality must be high for users to perform their tasks successfully. Spatial data infrastructures will eventually mean that data from different sources is combined. A decision on set standards should be made before that happens, ensuring that data elements used are similar.

REFERENCES

Crow, Kenneth, 2002 'customer focused development with QFD'

Glen, Van Der Vyver, 'the overconfidence effect and IT professionals'

Guptil, CG, Morrison, J. L. (1995) 'elements of spatial data quality'

Kahn, B. K., Strong, D. M. and Wang, Y. R. 'information quality benchmarks: product and service performance'

Kenneth E. Foote and Donald J. Huebner, 'error, accuracy and precision'

Paradzayi, C. (2004), 'spatial data infrastructure as a vehicle for sustainable development in Zimbabwe'

Peter Young, Adam Smith International, 'building an institutional framework for privatisation; the importance of strong institutions'

Peters, T. (1987), 'thriving on chaos'

Reeves, C. A., Bednar, D. E., 'defining quality: alternatives and implications'

Stoner, J., Freeman, R., (1992), 'management'

Thorne, S., 'a warehouse design; keeping it simple'

Webber, R. (1997), 'modern imperatives'

www.ngis.gov.bw 'national spatial data infrastructure for Botswana', accessed 16/06/2005

www.praxiom.com 'new ISO standards', accessed 15/08/2005

BIOGRAPHICAL NOTES

Francis Chigiinge is currently working as a Land Surveyor for Kweneng Land Board, Ministry of Lands and Housing, Government of Botswana. He holds a Bachelor of Science Honours in Surveying from the University of Zimbabwe (1991), and is currently undertaking MBA studies with the Midlands State University, Zimbabwe. He is registered to practice as a Professional Land Surveyor in Botswana and Zimbabwe. He is a member of the Metadata working group, Botswana National Spatial Data Infrastructure. Before coming to Botswana, he was a full time Lecturer in the Department of Surveying and Geomatics, Midlands State University, Zimbabwe from June 2003 to January 2005. He lectured modules in Principles of Geomatics and Principles of GPS Surveying. Prior to that, he was Senior Land Surveyor with City of Gweru, Zimbabwe.

CONTACTS

Mr. Francis Chigiinge
Kweneng Land Board
Private Bag 006
Molepolole
BOTSWANA
Tel. + 267 5920355
Fax + 267 5920360
Mobile + 267 71441388
Email: fchigiinge@yahoo.com