On motivators and barriers of interorganizational GIS data sharing for address organizations in a South African SDI

Malete Daniel Sebake[1,2], Serena Coetzee[2]

[1]Spatial Data Management Unit, Council for Geoscience, Pretoria, South Africa, dsebake@geoscience.org.za

[2]Department of Computer Science, School of Information Technology, University of Pretoria, Pretoria, South Africa, scoetzee@cs.up.ac.za

Abstract

The use of spatial address data in public and private organizations has become critical in the delivery of services, such as emergency services, mail delivery and utility services, among others. In South Africa, spatial address data has been collected by different public and private organizations, with little or no cooperation on data sharing. Although there is weighty evidence about the benefits of spatial data sharing, organizations are still experiencing problems in making it happen, hence the costly duplication of efforts in data collection and storage still persists. This research uses a case study methodology to evaluate the factors that motivate or hinder organizations from sharing spatial address data among themselves, with a focus on facilitating the development of a common national address database (NAD) in South Africa. Three cases from public and private organizations, which collect and maintain spatial address data, were studied to determine the factors underlying the interorganizational GIS data sharing. This paper presents the findings of the case studies, and their implications on the interorganizational GIS data sharing initiatives as they occur in spatial data infrastructures (SDIs), particularly the South African SDI. We hope that the results of our study can guide the successful development and implementation of a single NAD in South Africa.

1. Introduction

To realize the full potential of the GIS, most researchers agreed that some type of a multi-participant establishment is required to coordinate the spatial data sharing efforts of private and/or public organizations at local, national or transnational levels, which was termed interorganizational GIS (Nedovic-Budic 1999, Nedović-Budic and Pinto 2000, Onsrud 2007). The interorganizational GIS could be seen as a precursor to the more matured, high-level spatial data infrastructure (SDI), which is the ultimate goal of interorganizational coordination and spatial data sharing. The SDI is a framework that consists of coordinated actions on policies, organizational remits, data, technologies, standards, delivery mechanisms, financial and human resources to support ready access to geographic information in the possession of both public and private organizations (Masser 2005). Even with the increased proliferation of the SDIs or GIS data sharing initiatives, there has been general inability and reluctance to share data and information among organizations (Warnecke et al. 1998).

In this research, the focus was on the interorganizational GIS data sharing applied to spatial address data as an instance of GIS data. Spatial address data refers to any spatial reference data to which an individual's information is linked, for instance utility services such as water and electricity linked to a geocoded street address (Coetzee 2008). The interest in this topic is informed by the advent of the South African address standard, SANS 1883 (SABS 2009), which offers an opportunity to the address data organizations to use common terminology and elements in order to improve interoperability and data sharing among themselves. Normally, address databases are under the jurisdiction of separate local authorities (Coetzee 2008), but there is a need for the collation of these disparate databases into a common national address database (NAD). The NAD should bring to the fore private and public entities, which have been developing address databases for different purposes into a high-level spatial data infrastructure (SDI) that aims to serve a wider audience.

This paper presents the results on the motivators and barriers influencing the interorganizational GIS data sharing within the NAD (as an SDI initiative) using a case study approach. A semi-structured questionnaire was developed from numerous literature sources including interorganizational GIS, spatial data sharing, SDIs and spatial address data (Nedovic-Budic and Pinto 2000, Masser 2005, Harvey and Tulloch 2006, Onsrud 2007, McDougall et al. 2007, Coetzee 2008). Three cases comprising of a public organization, a private organization, and a local municipality were studied, and key informants who were well-conversant with the GIS data sharing activities were identified in each of the organizations, and verbal interviews were conducted with them. For the case study method to be reliable and repeatable, the data collection protocol was put in place to ensure that the procedures used were well-documented and could be repeated with the same results when conducted again. In the data collection process, different sources of evidence, such as interviews, documentations, observations and records were used to corroborate the findings.

The new evidence about the motivators and barriers of interorganizational GIS data sharing as it applies to address data organizations provides valuable information to guide the development of a single national address database in South Africa.

2. Background

2.1 Interorganizational GIS data sharing in an SDI

The availability of GIS technology and data resulted in the ease of data exchange and integration across organizations, which in turn stimulated interorganizational partnerships in spatial data sharing. It is these interorganizational partnerships that give rise to the concept of interorganizational GIS, which is defined as '*a type of multi-participant setup … achieved through coordination between various organizations*' (Nedovic-Budic and Pinto 2000, p.456). In this multi-participant setup, public and/or private organizations at local, regional and national level are jointly involved in data acquisition and database development for a common purpose (Cash et al. 1994). Interorganizational GIS could be seen

as having a limited scope and measurable goals, as compared to the high-level, cross-cutting nature of the spatial data infrastructure (SDI) defined hereunder.

The spatial data infrastructure (SDI) is an establishment that encompass "*coordinated actions on policies, organizational remits, data, technologies, standards, delivery mechanisms, financial and human resources to support ready access to geographic information in the possession of both public and private organizations*" (Masser 2005, p.16).

From the definitions, it is apparent that interorganizational GIS operates at a lower level of the hierarchy and deals with basic coordination issues such as GIS data sharing agreements or contracts among participating organizations. On the other hand, SDI is applicable at a higher level as it addresses the framework or environment for GIS data sharing including policies, standards, or delivery mechanisms, which provide interoperability for spatial data discovery and distribution (Budhathoki and Nedovic-Budic 2007). In a common analogy, the interorganizational GIS will denote a couple of friends with manageable mutual interests while SDI pertains to a community of persons with complex relations and rules guiding their interactions and co-existence. It can be argued that the issues within the interorganizational GIS arrangements, as buildings blocks, would determine the nature of SDIs emerging from them.

2.2 Motivators and barriers for spatial data sharing

As much as there are motivators for spatial data sharing, there are similarly barriers that require attention when setting up interorganizational GIS programmes. This section presents the motivators and barriers of interorganizational GIS data sharing from numerous literature sources such as Nedovic-Budic and Pinto (2000), Masser (2005), Harvey and Tulloch (2006), and Onsrud (2007), among others. Table 3 gives a summary of these motivators and barriers in interorganizational GIS data sharing initiatives.

Table 3. Motivators and barriers for interorganizational GIS data sharing initiatives

Motivators
Sharing the cost of implementation among participating organizations; improved decision making through exchange of information
Cost-effective
Improved data quality
High returns on investment
Reduced time spent in data collection and decision making; increased data availability; More diverse or dynamic maps
Improved user satisfaction
Barriers
Institutional disincentives; historical and ideological barriers; power disparities; differing risk perceptions; technical complexity; political and institutional culture
Conflicting priorities among participating organizations; differences in GIS facilities, awareness and data handling skills; Concessions over access to information, leadership, data standards, equipment, and training
Staff turnover; lack of resources; incompatible old systems; lack of support from management
Coordination of system requirements
Lack of common data definitions, formats, and models
Differences in data quality; inadequate planning and consultation about data use; insufficient staff and technical resources
Networking costs; data confidentiality, liability and pricing

2.3 Address data in an SDI

Originally, an address was defined as an entity that describes the location where a person resides for the purposes of delivering mail, e.g. house number, street or road name (The American Heritage Dictionary of the English Language 2000). In recent times, the concept of address would apply to any reference data to which an individual's information is linked, e.g. billing, utilities (e.g. water and electricity), and credit application (Coetzee 2008).

In a survey conducted by Levoleger and Corbin (2005), it had been shown that most of the European countries have a strong foundation with regard to the historical records and address databases. The address data is collected from local level in countries such as Netherlands, Norway, Austria and the National Land and Property Gazetteer (NPLG) in the UK; while other countries have embarked on collating address data on national level, e.g. the GeoDirectory in Ireland and the AddressPoint dataset produced by Ordnance Survey in UK (Coetzee 2008). There are a number of European countries where address data is still maintained at local level and not collated into a national level, e.g. Croatia, Portugal, Germany, France and Hungary (Coetzee 2008).

In Australia, the Public Sector Mapping Agencies (PSMA) follow a semi-automated process of matching address data into the standard format of the Geocoded National Address File (G-NAF), which is distributed on a quarterly basis (Paull 2003).

In developing countries such as India and Brazil, there is lack of comprehensive address data that can be collated into an address database at national level (Coetzee 2008). According to Coetzee (2008), this paucity of address data is due to (a) large cities with slums and rural areas, which are characterized

by irregular occupation; (b) areas without street signs or individual address signs at each dwelling; (c) incomplete addressing database due to lack of quality information or the cost of creating and maintaining a detailed database in the places with irregular growth (Davis and Fonseca 2007).

Due to the necessity of address data as a prerequisite to most of the services delivered by government departments or commercial address-related service providers, the proliferation and maintenance of the address databases has been haphazard and discordant in nature. Consequently, there are different address data formats assigned by the South African Post Office, Statistics South Africa, national departments, national utilities and private companies (Coetzee et al. 2008), which can only be overcome by the use of the common South African address standard. All the affected and interested organizations need to work together to achieve this interorganizational GIS initiative, which could be valuable in providing a common national address database (NAD) for South Africa.

3. Method

In order to study the motivators and barriers in interorganizational GIS data sharing for the NAD, a multiple case study method was applied to evaluate three organizations which were collecting and maintaining the spatial address databases. The case study method was deemed appropriate to answer the 'why' and 'how' questions pertaining to the motivations and barriers for spatial data sharing among organizations. The questions to be answered in the research were the following:

1. Why will organizations share spatial address data for the development of the National Address Database (NAD)? What are the motivators?

2. Why will organizations not share spatial address data for the development of the National Address Database (NAD)? What are the barriers?

3. How will the understanding of motivators and barriers for sharing spatial address influence the existing theory on interorganizational GIS data sharing frameworks?

The organizations which were selected for the research were representative of the key role players in the address data management, which included a local municipality, a public organization and a private organization.

In the case study methodology, the key informants– who were well-conversant with the GIS data sharing activities – were identified in each of the organizations, and verbal interviews were conducted with them. For the case study method to be reliable and repeatable, the data collection protocol was put in place to ensure that the procedures used were well-documented and could be repeated with the same results when conducted again. The data collection would focus on two main sources of evidence (i.e. interviews and documentation) to ensure construct validity.

The research used a semi-structured questionnaire to enable the informants to provide information on a wide range of issues that motivate or hinder organizations from sharing spatial address data. The interview questions were constructed in such a way that they offered leeway for informants to fully

express themselves with minimal or no prejudice from the interviewer. The questionnaire was solely developed to assist the interviewer whenever the interview requires some guidance. The semi-structured questionnaire consisted of three parts as discussed hereunder.

Part 1 – Particulars of the organization: This part of the questionnaire (which was rather structured) collected the information about the organization, such as its size (i.e. number of employees), spatial data sharing equipment and resources, and whether its GIS activities are for private or public consumption.

Part 2 – Motivators for spatial address data sharing: This part of the questionnaire consists of open-ended questions, which sought to establish the motivators for spatial address data sharing among organizations including issues of cost, data quality, return on investment, effectiveness and efficiency, improved decision making and incentives.

Part 3 – Barriers for spatial address data sharing: This part of the questionnaire comprised of a list of open-ended questions pertaining to the factors that may hinder spatial address data sharing initiatives among organizations, which included the impact on revenue-generating streams of the organization, priorities, accuracy and reliability of the data, copyright, data privacy and ownership issues, staff turnover and technical resources.

The verbal interviews were recorded on the voice recorder and transcribed later for further analysis. The interview transcripts were corroborated by other sources of information, such as interorganizational agreements, website pages describing the activities of the organizations, research papers (i.e. conference proceedings and journals), internal and project reports in order to minimize biasness of the data.

4. Results and discussions

In this section, the three cases of address data organizations will be described and their respective motivators and barriers to participate in interorganizational GIS data sharing initiatives will be presented. Furthermore, the findings on the different perceptions of the three cases will be discussed to reflect on the issues which require attention prior to establishment of a multi-participant initiative.

4.1 Description of the case studies

The three cases studies were representative of key role players in the usage of spatial address data, which included a public organization (Case A), a private organization (Case B) and a local municipality (Case C).

Case A is a public organization with more than 500 employees, which is responsible for collecting and distributing census data in South Africa. The census data is used to guide the planning and policy directives of the country as a whole. Case A experienced lack of addresses for census purposes, which made its work of collecting data difficult. This public organization attempted capturing address data for its purposes, which was likely to result in a conflict with local authorities as it was encroaching into

their mandate to collect and register addresses within their jurisdictions. Case A emphasized that they were not trying to generate a national address database (i.e. a mandate of local authorities), but would like to establish a dwelling framework, which will capture the dwelling places as geocoded points or polygons with associated census attribute data. The dwelling framework was still maintained in-house by Case A to give the organization invaluable data needed for census purposes.

Case B is a private organization with more than 50 employees, which offers GIS services to its South African and international clientele. The organization specializes in location-based services which were required for geocoding, geo-marketing, mobile applications, transportation modelling, etc. In order to provide these services, the organization required a quality spatial address database as base datasets. The lack of these base datasets prompted the organization to create its own spatial address database by sourcing the address data from the local authorities in paper or digital formats. These addresses are then captured and geocoded to produce spatial address data, which is used for the development of location-based applications for their clients.

Case C is a local municipality which encompasses the metropolitan area of Johannesburg, one of the biggest cities in Africa with a population of approximately 3 million. The challenge of this local authority is to keep an updated street address register for different purposes of the municipality such as property valuations, provision of emergency services, delivery of utility services, establishment of townships and debt collection. The Land Information System (LIS) is currently the organization's sole source of property information, which is a registry of geocoded properties with their attribute data. The local authority has instituted a process whereby every property registered is simultaneously captured in the spatially-enabled LIS, thus keeping the database dynamic and up to date.

The description of the three cases indicated that address data is a common denominator in everyday operations of different organizations, despite their commercial or public inclinations. Therefore, it would make sense for these organizations to pool their resources in establishing a core spatial address database, which could be used for different applications. The focus of the next section is to identify the motivators or barriers of organizations in establishing a multi-participant initiative for spatial address data sharing.

4.2 Motivators

This section presents motivators, perceived or real, which attract the organizations to participate in an interorganizational data sharing initiative, particularly for spatial address databases.

Reduce the cost of data handling. In all the three cases, the benefit of pooling resources to capture and validate the spatial addresses was recognized as a significant motivator to establishing a common spatial address database. The use of disparate databases was seen as duplicating efforts, and a waste of time and money as the data from different sources would still need validation before it is rendered useable. Case B, a private organization, felt that it was an unfair burden for the organization to be handling spatial address data from different sources, for instance 'only metropolitan areas had useable

data, not the municipalities'. Although Case B had a doubt on the reduction in cost of handling data, the organization had a strong conviction that the cost of capturing and validating the spatial address data should be taken up by the local authorities, because it was falling within their mandate. Nevertheless, the responses indicate that there was a role for both public and private organizations (i.e. public-private partnerships) in the management of spatial address databases.

Improved data quality. Both case A, B and C identified the improved data quality as a significant motivator to establishing a spatial address data sharing initiative. The perception is that if organizations are contributing to the same spatial address database, it would foster compliance to common definitions, standards, protocols and formats, thus improving the usability of the data. The improved data quality was also attributed to well-trained personnel in terms of technical skills and GIS data standards.

Return on investment, and improved decision making and planning. According to the responses, the returns to be derived from a spatial address data initiative were manifold in terms of improving decision making and planning in both public and private organizations. By using the common spatial address database, the public organizations would unlock the potential in improving their functions, inter alia, collection of census data, collection of rates and taxes, delivery of emergency and utility services, establishment of townships/new developments and overall decision making and planning. The private organizations would benefit by focussing their energies into developing new applications (e.g. location-based technologies), instead of wasting their time and resources recapturing and validating spatial address data.

Incentivization. The public organizations (Case A and C) were not keen on any kind of incentives, because of their disposition, i.e. they were mandated to establish spatial address registers for their purposes. Only Case B, as a private organization considered financial compensation to be an appropriate incentive for recapturing and validating spatial address data, which is essentially the function of the local authorities. It is clear from the responses on the issue of incentives that the organizations did not give enough thought to the type of public-private partnership (e.g. ANZLIC-PSMA arrangement) that could be ideal in the spatial address data sharing initiative.

Other motivators. The three cases acknowledged that the participation in a common national spatial address database would create an enabling environment for organizations to use similar standards, e.g. South African national address standard (SANS 1883), thus making it possible for public and private organizations to work from the same address dataset.

4.3 Barriers

This section will present the perceptions of the three cases in terms of the barriers that obstructs or impede the sharing of spatial address data among organizations.

Negative impact on revenue-generating streams. In case of the public organizations, they were not expected to generate revenue from their efforts in developing the address database, as a result this was

not an impediment to contributing in a spatial address database initiative. Although it is not a pronounced practice, the local municipality expressed that at times they are expected to fund their own operations, and to balance the budget, they might be tempted to sell the data in their custody. On the other hand, the private organization has an inherent commercial interest, which includes selling of the value-added data and applications in their possession. This could make them a reluctant player in a spatial address database initiative where they might be expected to contribute their services for the greater good.

Priorities of the organization. For the public organization and local municipality, their priority is of collecting and registering the addresses for their own use. But, this priority would depend on whether the budget is available to maintain the SDI, i.e. 'if money is tight, it (SDI) will take the back burner'. For the private organization, the commercial priority may come first, but they alluded that it is possible to have a commercial interest, but still participate in establishing the common spatial address database.

Accuracy and reliability of spatial address data. Although the public organizations might have used the same methods of collecting and validating address data for many years, a lot of data from the custodians, such as municipalities may be suspicious in terms of accuracy and reliability. This would be due to lack of capacity building and training in GIS data handling, because 'data management is not a priority of these organizations'.

Copyright issues, data privacy and data ownership issues. For the public organizations with the mandate to distribute and share data, the copyright issues were not as pronounced as in private organizations. In public organizations, there are no licensing agreements required and data is distributed for free. But, the private organization put restrictions on its spatial address data by making the purchasers to sign licensing agreements, which stop them from sharing the data. So, issues of ownership should be considered in an SDI, to ensure that the private organizations are not alienated by SDI initiatives which advocate sharing of data at all costs. Privacy issues could be raised if personal/ private information for the person residing at that particular address is involved, but that did not to seem to be a problem in all cases as it was not difficult to distribute the address data without private data.

Lack of common data definitions, formats and models. All three cases did not show any reluctance to work within standards, as they all took part in the formation of the South African national address standard (SANS 1883). It was mentioned by Case A that it is a challenge to get the internal systems to adapt to the new SANS 1883 standard, which might take time to implement. Although there was still a concern about common data definitions, formats and models, the three cases were positive about the future in lieu of the promulgation of the national address standard.

Staff turnover and technical resources. Case A, which is a public organization, mentioned that 'attracting the right people and retaining them is a concern', as the organization experienced a high staff turnover for some time. Although both the public organization and local municipality appeared to be well-resourced, their budgets were not limitless. The high staff turnover could result in less technical

skills, thus paralysing any spatial data sharing initiative these organizations would be involved in.

Unequal commitment from organizations in an SDI. Because most of the organizations are mandated to distribute and share data, they might feel obliged to make promises to the multi-participant initiatives which they might not fulfil. Case A mentioned that 'in a public forum, organizations might make promises, but the bureaucratic structures make data sharing difficult'. There was a perception that some organizations (both public and private) use data as currency to elevate their importance, thus creating unnecessary restriction to data sharing.

Inadequate support from strategic management plans and policies. For public organizations, their strategic documents and policies would largely be supportive of building partnerships and creating an enabling environment for distributing and sharing data among organizations. For Case C, which is a local municipality, they used the Balanced Scorecard to align the targets of spatial data sharing with their strategy. So, it is important for organizations to ensure that their aspirations of contributing to the multi-participant spatial data sharing initiatives are included in their strategic documents in order to make them relevant to the organization as a whole. If it is not in the strategic document, it would risk being overlooked by decision makers.

5. Conclusions

The aim of this study was to assess the motivators and barriers for establishing an interorganizational GIS data sharing initiative among organizations which capture and maintain spatial address databases in South Africa. The research used the case study method – with a semi-structured interview – to investigate the motivators and barriers in three address data organizations. Although the findings from this sample of address data organizations could not allow us to make generalizations about their behaviour, it gave us the window of opportunity to identify significant issues that motivate or impede these organizations to participate in the interorganizational GIS data sharing initiatives or a wider SDI. The results of the three cases indicate that in establishing any interorganizational GIS data sharing initiative or an SDI, understanding the motivators and barriers which make the organizations to participate or not to participate, should be considered as the initial step in the process. In taking this study further, a net could be cast to survey a sizeable number of address data organizations in order to evaluate the patterns that emerge from their complex relationships in an SDI initiative and to validate the findings of this research.

6. Acknowledgements

References

Bhudhathoki, N.R. and Nedovic-Budic, Z. 2007. Expanding the spatial data infrastructure knowledge base. In Research and Theory in Advancing Spatial Data Infrastructure Concepts, ed. H. Onsrud, 7-31. ESRI Press, Redlands, California.

Cash, J.I.Jr., Eccles, R.G., Nohria, N., and Nolan, R.L. 1994. IT between organisations: interorganizational systems. In *Building the information-age organization: Structure, Control, and information technogies*, eds. J.I. Cash Jr, R.G. Eccles, Nohria, N. and R.L. Nolan. R.D. Irwin, Homewood, 338-396.

Coetzee, S. 2008. *An analysis of a data grid approach for spatial data infrastructures*. Unpublished PhD thesis. Department of Computer Science. Faculty of Engineering, Built Environment and Information Technology. University of Pretoria.

Coetzee, S., Cooper, A.K., Lind, M., McCart Wells, M., Yurman, S.W., Wells, E., Griffiths, N. and Nicholson, M.J.L. 2008. Towards an international address standard. GSDI-10 Conference, Trinidad and Tobago, 25-29 February 2008.

Davis, C. and Fonseca, F. 2007. Assessing the certainty of locations produced by an address geocoding system. *GeoInformatica*, 11(1), 103-129.

Harvey, F. and Tulloch, D. 2006. Local government data sharing: Evaluating the foundations of spatial data infrastructures. *International Journal of Geographic Information Science*, 20(7), 743-768.

Levoleger, K., Corbin, C.(Eds.). 2005. Survey of European National Addressing as of May 2005, v3, EUROGI, Ref. AWP 2005, viewed 29 January 2011, <http://eurogi.org/egin/Agenda.5_EUROGI_Address_Survey_REPORT.pdf>.

Masser, I. 2005. *GIS Worlds: Creating spatial data infrastructures*. ESRI Press, Redlands, Carlifornia.

McDougall, K., Rajabifard, A. and Williamson, I. 2007. A mixed method approach for evaluating spatial data sharing partnerships for SDI development. In *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, ed. H. Onsrud, 55-73. ESRI Press, Redlands, California.

Nedovic-Budic, Z. and Pinto, J.K. 1999. Understanding interorganizational GIS activities: a conceptual framework. Journal of Urban and Regional Information Systems Association, 11(1), 53-64.

Nedovic-Budic, Z. and Pinto, J.K., 2000. Information sharing in an interorganizational GIS environment. *Environment and Planning B: Planning and Design*, 27(3), 455-474.

Onsrud, H. (Ed.). 2007. *Research and Theory in Advancing Spatial Data Infrastructure Concepts*. ESRI Press, Redlands, California.

Paull, D. 2003. *A Geocoded National Address File for Australia: The GNAF what, why, who and when*. PSMA. PSMA Australia.

South African Bureau of Standards (SABS). 2009. South African National Standard (SANS) 1883, viewed 25 January 2011, < https://www.sabs.co.za/webstore/user/search.php?q=SANS+1883>.

The American Heritage Dictionary of the English Language, 4th ed. 2000. Updated in 2009. Houghton Mifflin Company, viewed 29 January 2011, < http://www.thefreedictionary.com>.

Warnecke, L., Beattie, J., Cheryl, K., Lyday, W. and French, S. 1998. *Geographical Information Technology in Cities and Counties: A Nationwide Assessment*. American Forests, Washington DC.