

2013

Adapting geospatial business intelligence for regional infrastructure planning

Rohan C. Wickramasuriya Denagamage
University of Wollongong, rohan@uow.edu.au

Pascal Perez
University of Wollongong, pascal@uow.edu.au

Jun Ma
University of Wollongong, jma@uow.edu.au

Matthew J. Berryman
University of Wollongong, mberryma@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/smartpapers>



Part of the [Engineering Commons](#), and the [Physical Sciences and Mathematics Commons](#)

Recommended Citation

Wickramasuriya Denagamage, Rohan C.; Perez, Pascal; Ma, Jun; and Berryman, Matthew J., "Adapting geospatial business intelligence for regional infrastructure planning" (2013). *SMART Infrastructure Facility - Papers*. 113.

<https://ro.uow.edu.au/smartpapers/113>

Adapting geospatial business intelligence for regional infrastructure planning

Abstract

Business Intelligence (BI) has traditionally been used in organizations as a strategic tool to maximize profit. When coupled with Geographic Information Systems, however, BI can be transformed into a cutting edge decision support system for planning local and regional areas, as we demonstrate in this paper. Local and regional governments often face a major challenge in terms of developing a holistic view upon disjointedly operated utility services in their jurisdictions due to data silos. This limitation has become a serious impediment to infrastructure planning and regional adaptation to changes. Geo-BI provides tools to manage data coming from multiple and disparate sources, and visualize them through online interactive userinterfaces. The SMART Infrastructure Dashboard (SID) is an innovative Geo-BI solution that includes an open-source ETL (Extract, Transform and Load) toolkit to handle various datasets, a spatially-enabled data warehouse hosted in PostgreSQL/PostGIS and proprietary BI software for creating and administering analytical reports and dashboards. SID allows planners and policy makers to analyze the interplay between the use of infrastructure services, demographics and weather parameters across multiple spatial and temporal scales. Furthermore, SID enables planners to run various what-if scenarios related to projected consumption patterns, service vulnerability of utility networks, and transportation demand management. Future research involves enabling the analysis of networks of networks through SID to understand the propagation of cascading failures and benefits in interconnected utility networks.

Keywords

infrastructure, regional, planning, intelligence, adapting, business, geospatial

Disciplines

Engineering | Physical Sciences and Mathematics

Publication Details

Wickramasuriya Denagamage, R. C., Perez, P., Ma, J. & Berryman, M. J. (2013). Adapting geospatial business intelligence for regional infrastructure planning. In J. Piantadosi, R. S. Anderssen & J. Boland (Eds.), 20th International Congress on Modelling and Simulation (pp. 1-7). Australia: The Modelling and Simulation Society of Australia and New Zealand Inc.

Using Geospatial Business Intelligence to Support Regional Infrastructure Governance

Rohan Wickramasuriya*, Jun Ma, Matthew Berryman and Pascal Perez

SMART Infrastructure Facility, University of Wollongong, Wollongong NSW 2522,

Australia. (rohan@uow.edu.au, jma@uow.edu.au, mberryman@uow.edu.au,

pascal@uow.edu.au)

* Corresponding author

Telephone: +61 2 4239 2344

Fax: +61 2 42981489

Email: rohan@uow.edu.au

Address: SMART Infrastructure Facility, University of Wollongong, Wollongong NSW
2522, Australia.

Abstract

In many developed countries including Australia, infrastructure services at local and state levels are being provided by an increasing number of disjointed public and private agencies. There is an urgent need for an integrated view on the provision and use of these services for better governance and productivity. Developing an integrated view is challenging due to the dispersion of relevant data sets and the underlying complexity of increasingly interconnected infrastructure networks. Using a case study in New South Wales (Australia), we demonstrate how tools and processes in Geospatial Business Intelligence (Geo-BI) can be harnessed using a collective design approach to develop an integrated solution; the SMART Infrastructure Dashboard (SID). While providing a much needed planning and policy support tool for the local governance of infrastructure services, SID pushes the boundary of Geo-BI beyond its traditional use.

Keywords: Infrastructure Services, Local Governance, Geospatial Business Intelligence, Australia

1. Introduction

Provision of infrastructure services to communities is a fundamental requirement, and this has been traditionally viewed as the responsibility of governments at various levels.

However, in many developed countries, including Australia, the private sector is increasingly dominating the provision of public infrastructure services, including both transport and utility networks (Cannadi & Dollery, 2005). While this increasing private sector involvement eases the pressure on local and state governments' limited resources, it inevitably brings new challenges in terms of monitoring and regulating services provided by several disjointed organizations. These challenges are further exacerbated by the fact that modern infrastructure

networks are highly interconnected (Wang, Hong, & Chen, 2012). Hence, local and state governments urgently need an integrated view on infrastructure networks and services for better governance and planning of cities and regions.

This integrated view on infrastructure networks and services could be realized through recent advances in Information and Communications Technology (ICT) research (Lim & Lee, 2010; Mauree, 2010; Chen et al., 2011; IBM, 2012) and enhanced by a collective design approach. However, to be useful, this new generation of tools has to overcome two technical challenges. First, given the decentralized nature of infrastructure network management and service provision, these tools have to accommodate highly diverse and dispersed data sources. Second, they have to handle the underlying complexity of operations on individual networks, as well as the interconnectedness of infrastructure networks. Moreover, as with any other decision-support system, these tools need to exhibit positive usability traits such as performance, user-friendliness and intuitive visuals. One method, with which positive usability can be achieved is through collective design approaches whereby several iterations between designers, data providers and information users are employed to evolve software from proof-of-concept to an effective end-product.

We suggest that Geospatial Business Intelligence (Geo-BI) can be adapted to meet the requirements of an integrated solution for local and regional governance of infrastructure services. Business Intelligence (BI) refers to “the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance” (Gartner, 2012). Getting the right people involved in the process is an indispensable aspect (Loshin, 2003), that is implied, but often goes unnoticed in that working definition of BI. In terms of processes, a BI project involves data acquisition, data

warehousing, data analysis and mining, and reporting and presentation (Shi, Peng, & Xu, 2012). Geo-BI is an improvement on this traditional BI approach made possible by integrating Geographic Information Systems (GIS) with BI (Angelaccio, Basili, Buttarazzi, & Liguori, 2012). This integration, though technically challenging, opens up a myriad of new and exciting ways to analyse and present data. Given that the majority of data collected by organizations has a geographic reference (Franklin & Hane, 1992), Geo-BI provides the spatial perspective which was missing in traditional BI (McHugh et al., 2009). Hence, we opt for Geo-BI as a solution for integrated infrastructure governance.

The aim of this study is to develop a robust, easily accessible and user-friendly Geo-BI solution, the SMART Infrastructure Dashboard (SID), which can harness diverse and dispersed datasets to support decision making related to local governance of infrastructure services. SID aims to inform planners and policy makers about the current and past states of infrastructure systems and services, as well as their spatial and temporal interdependencies. SID also enables future planning by allowing users to run various ‘what-if’ scenarios based on user-defined parameter values. Our approach to Geo-BI marks a radical change to the way BI has been traditionally used. Both BI and Geo-BI are conventionally used to deal with data from a single organization with the aim of improving profit and performance. The use of Geo-BI to access data from multiple providers and operators in order to analyse interdependencies of infrastructure systems and services is relatively recent (Sinnott et al., 2011; Mokrech et al., 2012). Further, the use of Geo-BI in this context is not only targeted at improving the profit and performance of individual infrastructure services, but enables policy makers and stakeholders to view the response of an integrated suite of interconnected utilities and services, to different scenarios, based on a complex-systems approach. Currently, SID includes the following utilities: electricity and water distribution, as well as sewage and solid

waste collection and treatment. In this paper, we demonstrate, using a case study approach, how Geo-BI could be transformed into an integrated solution for the local governance of infrastructure services.

2. Study Area, Stakeholders and Data Types

Following Olszak and Ziemba (2007), we started our Geo-BI project with a *design* phase whereby we identified the study area, relevant stakeholders, and data requirements for the project.

The study area corresponds to the Illawarra region (New South Wales, Australia). This coastal region, located south of Sydney, is made of five Local Government Areas (LGAs): Wollongong, Shellharbour, Kiama, Shoalhaven and Wingecarribee (Figure 1). The first four LGAs occupy the coastal plain limited on the east by a forested cliff, while Wingecarribee LGA spreads across the southern tableland, west of the cliff. According to the 2011 census from the Australian Bureau of Statistics, the population of the Illawarra region is 413,216 persons, of which 46.6% live in the Wollongong LGA (Wilkinson, 2011).

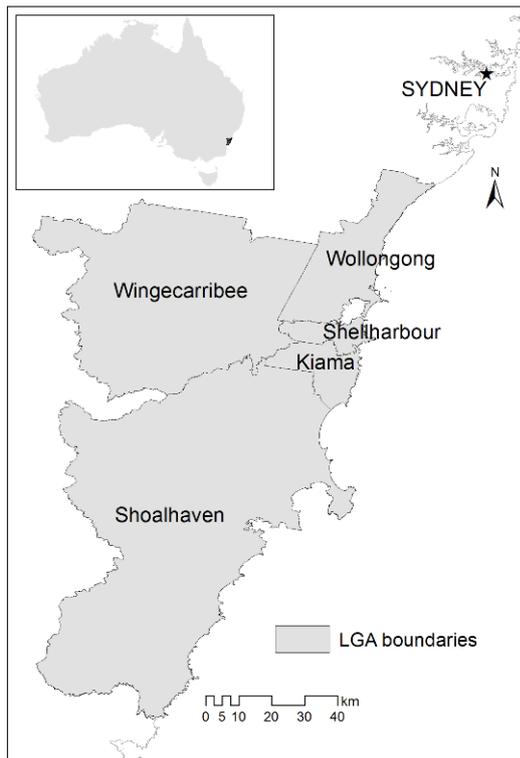


Figure 1. Study Area (Illawarra region in New South Wales, Australia)

Although the geography and topography of the Illawarra region have helped to create clear delineations for each utility network, relatively well separated from neighbouring regions, authority and management vary considerably across utilities and jurisdictions. For example, the electricity distribution network is managed by a single operator (Endeavour Energy) for the whole region while water distribution is split between a private operator (Sydney Water, servicing Wollongong, Shellharbour and Kiama LGAs) and two local agencies (Wingecarribee and Shoalhaven LGAs). Likewise, a single private operator (REMONDIS) manages solid waste collection in Wollongong and Shellharbour LGAs while the three other LGAs administer their own facilities. Hence, we identified the five LGAs and all the private utility operators as stakeholders in the SID project.

From the SID perspective, the stakeholders were both data providers and end users of the dashboard. Thus, we collected a diverse set of data from the stakeholders including geometric datasets of utility networks, service usage or consumption at various geographic levels over various time periods, water discharge at reservoirs and pumps, water quality at various points in the network, power consumption of assets such as treatment plants and pumps, waste collection routes, and quantity of waste collected. As early interactions with stakeholders showed their interest in correlating utility data with demographic and climate variables, we identified relevant databases from the Australian Bureau of Statistics (ABS) and the Bureau of Meteorology (BOM) to be added to SID (see appendix for a comprehensive list of data and providers).

3. Technical Architecture and Work Flows

There are two approaches to BI; managerial (Liautaud, 2000; Tomic & Milic, 2013) and technical (Watson et al., 2002; Farooq, 2012). The managerial approach concerns the way BI generates knowledge required to make strategic decisions using gathered data, while the technical approach is seen as the tools and methods needed to support decision making (Petrini & Pozzebon, 2009). We first describe the technical architecture and workflows involved in building the Geo-BI, and then how the planners and policy makers could use this collection of tools and methods for infrastructure governance. Figure 2 gives an overview of SID and the main workflows involved.

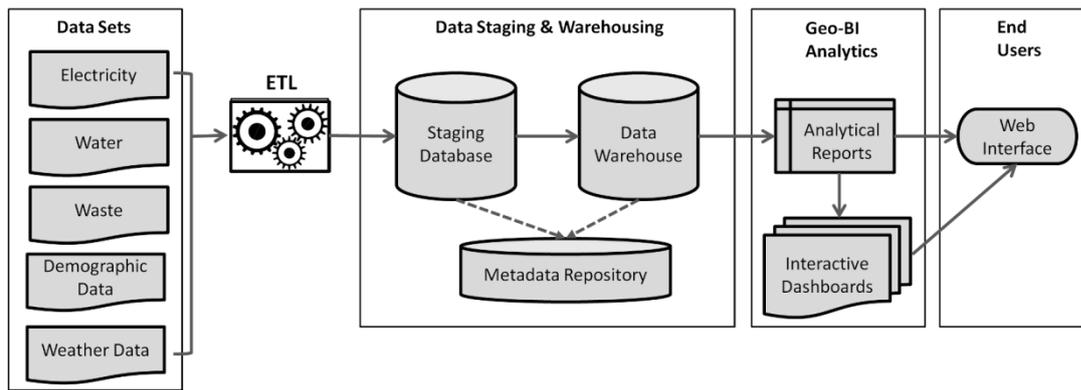


Figure 2. Technical Architecture and Workflows of SID

3.1 *Extract, Transform and Load*

SID receives data from several providers in diverse file types (e.g. Excel spreadsheets, plain text files, CSV files, ESRI shapefiles) and in heterogeneous structures (e.g. number and types of columns). Extract, Transform and Load (ETL) is the standard process adopted in migrating such data into an optimized data warehouse environment. We discuss the structure of this data warehouse later in this section.

The ability to conduct analysis at multiple spatial scales has been identified as an essential component of SID. ETL plays a crucial role in shaping data to give SID this ability.

Stakeholder consultation revealed that they are interested in using two geographic hierarchies in the analysis. One is the Australian Statistical Geography Standard (ASGS) released by ABS in 2011 (ABS, 2011). In this geographic hierarchy, the coarsest granule is a State and the finest granule is a Statistical Area Level 1 (SA1). The closest counterpart to an LGA, a typical granule in Government Administrative Hierarchy (GAH), in ASGS is the Statistical Area Level 3 (SA3). Stakeholders are keen to use a fusion between GAH and ASGS, from LGA down to SA1 through SA2, as the second geographic hierarchy. Accordingly, the two geographic hierarchies used in SID are:

- a) SA3 → SA2 → SA1
- b) LGA → SA2 → SA1

Data sets that we received for SID show obvious inconsistencies in the used geographic standard and the spatial resolution at which data is aggregated. For example, we received electricity consumption data at Census Collection District (CCD) level, which is the finest granule in the geographic standard used by ABS in its census up until 2006. On the other hand, water consumption data has been aggregated at postcode level, which is a completely different geographic standard. We explain how ETL handles this issue using ABS demographic data itself as an example.

We were interested in interpolating demographic variables, such as population and dwelling numbers, between two census years, 2006 and 2011. The most recent census data (2011) has been released at SA1 level, whereas 2006 census data are at CCD level. By definition, an SA1 is smaller than a CCD. However, an SA1 may be formed by several fractional CCDs due to boundary mismatching. To address this issue, we use a mapping between CCD and SA1 based on population. If a CCD contributes population p from its total population P to an SA1 area, then the mapping ratio from this CCD to SA1 is p/P (denoted by r below). We use this ratio to convert a demographic variable at CCD to SA1 level. Suppose S is an SA1 area which is covered by m number of CCDs (C_1, \dots, C_m) with ratios r_1, \dots, r_m and d is a demographic variable, then d_1, \dots, d_m are the values of d on those CCDs. Then the value s of demographic variable d in S is calculated as:

$$s = \sum_{i=1}^m d_i \times r_i \quad (1)$$

or

$$s = \frac{\sum_{i=1}^m d_i \times P_i \times r_i}{\sum_{i=1}^m P_i \times r_i} \quad (2)$$

where (1) is used for the total amount and (2) is used for the per capita value.

A similar approach is followed to convert utility data we received at various spatial resolutions.

3.2 *Star Schema*

The data warehouse is based on a Star Schema design. Star Schema is a widely used data model for data warehousing in various real-world applications (Kimball & Caserta, 2004). It organises data into one or more fact tables referencing any number of dimension tables. A star schema can simplify join queries and provides the capability to analyse data from multiple dimensions, thus enabling a user to perform various drill down, roll up, slice and dice operations on data. Kimball & Caserta (2004) provide an in-depth description of star schema and its variations used in data warehousing. Figure 3 illustrates two linked star schema used in SID data warehouse. These two schema are linked using common dimension tables.

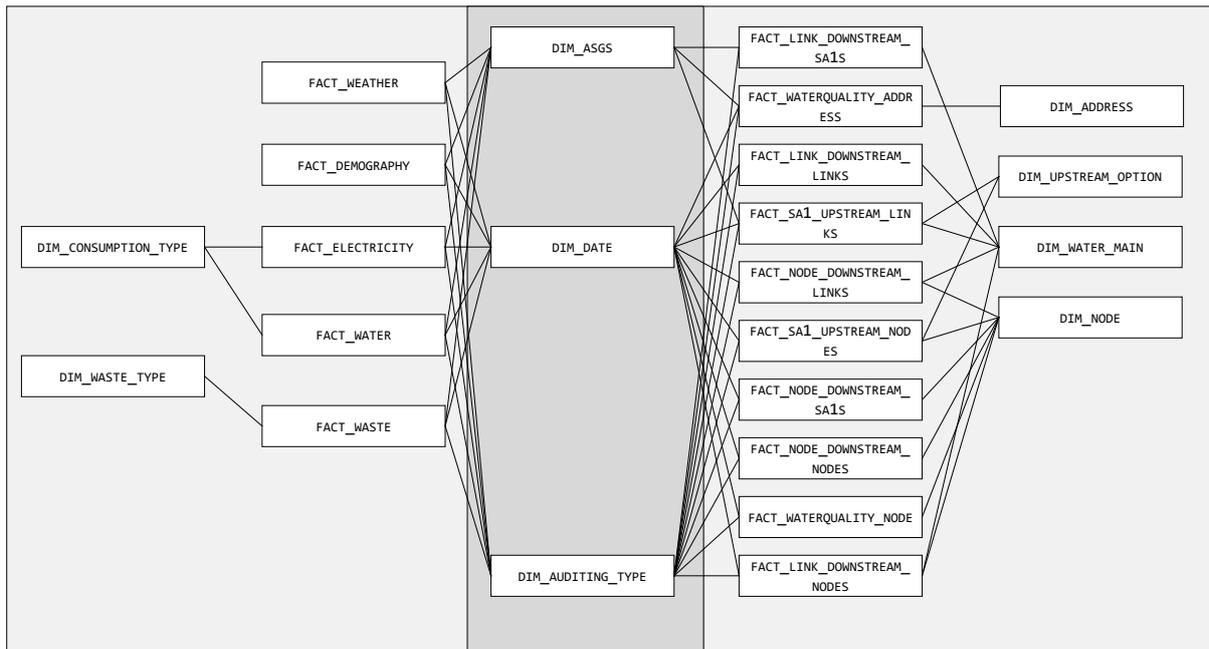


Figure 3. Overview of the star schema used in SID

The dimension table *DIM_ASGS* models the geographic hierarchy adopted by ABS in its 2011 census. Note that this dimension table contains polygon geometry fields to store spatial extents of each census area. Dimension *DIM_ASGS* supports drilling down and rolling up operations based on two geographic hierarchies described previously. Additionally, this dimension also supports postcode-based analysis, which is an important spatial disaggregation on its own right. The dimension *DIM_DATE* models the date based on the ordinary calendar and the Australian season and financial calendar. This dimension enables multi temporal analysis of utility consumption. The fact table *FACT_WEATHER* models temperature and rainfall derived from BOM daily records. Point source weather data collected from BOM were converted into surfaces using Kriging techniques prior to extracting average values at SA1 level. The fact *FACT_DEMOGRAPHY* models demographic information extracted from and generated based on ABS census results. Based on the census results released in 2006 and 2011, demographic data in non-census years has been generated using linear interpolation.

In our data warehouse, each utility networks is modelled using a separate star schema. However, these schema and the schema for utility consumption are interconnected through common dimensions. For clarity, only the star schema for water network is depicted in detail in Figure 3. This schema allows the user to perform several analyses: tracing quality of supplied water from household level all the way upstream to the treatment plant, querying the pumping cost for a unit volume of water from the source to a household, and running ‘what-if’ scenarios such as finding all affected SA1 areas given the failure of a particular pump or a water main.

3.3 Geo-Analytics and User Access

Geo-analytical tools are used to create interactive reports such as maps and charts, and interactive dashboards that assemble related reports. These reports are built on top of the optimized star schema-based data warehouse. End users access this interactive content via an online portal. The user interface is visually rich and comprises of easy-to-use controls like filters that provide keys to intricate analysis while effectively concealing the complexity of calculations and database queries from the user. Visualization is a primary focus for the user interface as highly intuitive visuals play a pivotal role in successful policy support tools (Lindquist, 2011).

4. Selection of Software

PostgreSQL on Linux (Ubuntu) virtual machines is used to provide a fast, stable and standards compliant database, with an excellent geospatial extension (PostGIS), a key

requirement for analysis and display of infrastructure data. Pentaho Data Integration was selected as a mature and open-source-licensed ETL package.

We paid a particular attention to BI suites in software selection. Some of the rigid criteria against which we assessed potential BI software included, (a) the ability to support the full range of vector data - points, lines and polygons, (b) ability to perform multi-scale spatial analysis using drill down and roll up through spatial hierarchies, (c) ability to link map-based reports with all other report types such as bar charts, radar charts and statistical bubble charts in dashboards, (d) level of complexity involved in building interactive content such as reports and dashboards, (e) visual quality, analytical breadth and interactivity of reports and dashboard, and (f) support. It is apparent that most of these criteria evaluated various ‘geospatial angles’ of the potential BI software. This is essential as any BI project with a strong geospatial flavour must select BI tools that can take full advantage of the geometry of the objects being studied (Shi et al., 2012). Yellowfin BI software scored well in majority of the areas prompting its selection for SID. GeoServer was chosen to serve security-enabled ancillary spatial layers required by map-based reports in Yellowfin.

5. Applications of SID in Infrastructure Service Governance

SID has a myriad of potential applications in the infrastructure domain. Before we move onto discussing such applications, we first briefly introduce the intuitive Graphical User Interface (GUI) of SID that lets users perform numerous analyses.

A dashboard in SID consists of one or more interactive reports of types ranging from maps, bar charts, column charts, pie charts, plots, radar diagrams to statistical bubble charts. These reports are arranged using a layout that supports one to three reports across the width and any

number of reports across the height of the dashboard. These reports can be linked together using common fields so that a filter applied on one report updates all linked reports. The simplest graphical operation that can be performed on a report is to hover over it to disclose tooltips with important information about the report. On a report that supports drill down functionality, users can select a particular dimension to disaggregate the information further. Reports can be generated to provide drill through functionality so that a selection of a particular field on a report takes the user to relevant child reports with additional information. Filters can be made available as lists, drop down items, sliders or using radio buttons. Related filters are usually grouped together, and can be nested if required. In the next sub-sections, we briefly discuss three applications that illustrate SID's potential in integrated infrastructure planning at the same time exposing further user interactions possible with SID GUI.

5.1 Relationships among Utility Usage, Socio-Demographics and Weather

Spatio-temporal pattern recognition and establishing association rules among variables are fundamental areas in spatio-temporal data mining (Andrienko, Malerba, May, & Teisseire, 2006). Knowledge generated in these two areas has practical implications for planning and policy making. For example, the knowledge that a certain group of suburbs consumes an increasingly high level of energy during winter can lead to several intervention actions. This could be an area-specific programme to retrofit energy efficient equipment or could be an investment in the network infrastructure to cope with the increasing demand. In order to generate such knowledge, analytical tools should support multi-spatial and multi-temporal analysis. Figure 4 shows an analytical dashboard developed in SID that caters for this demand.

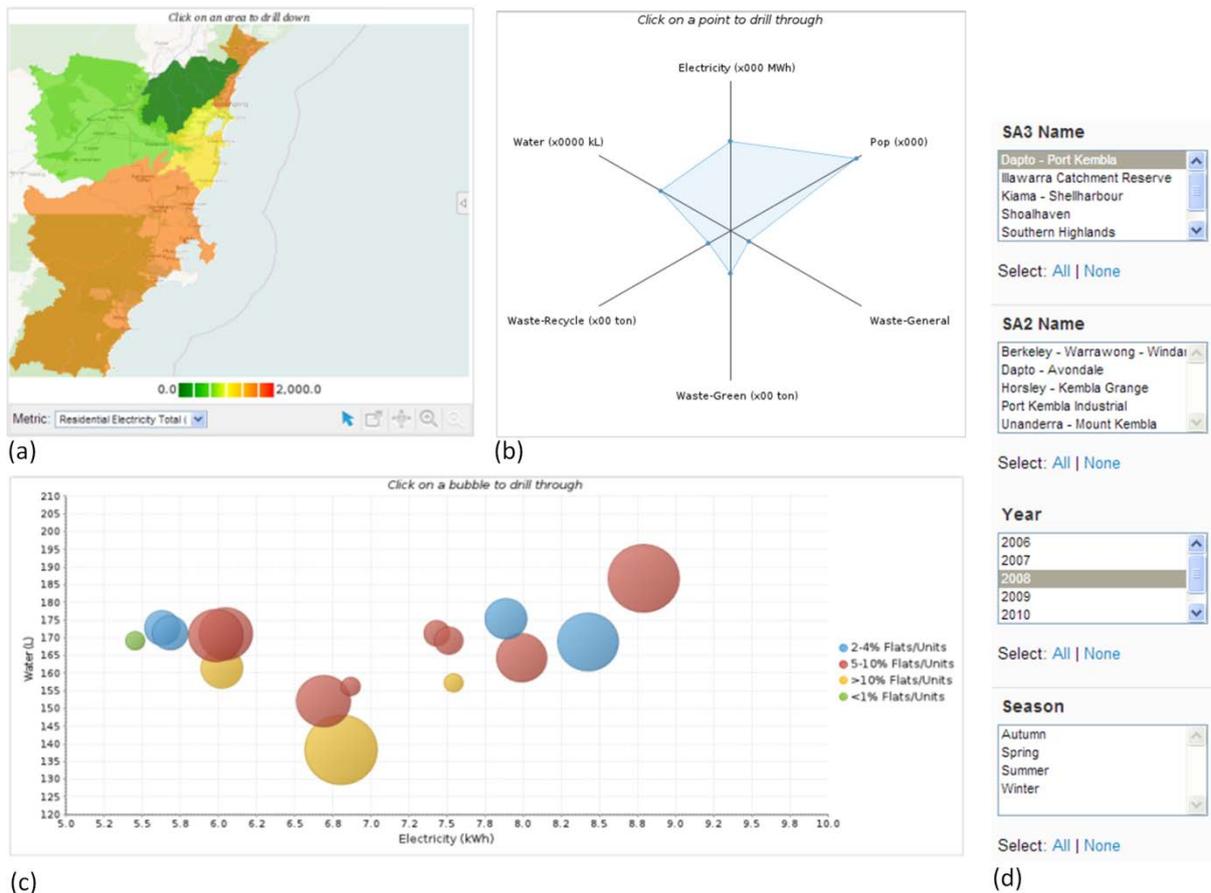


Figure 4. An interactive, geo-analytical dashboard consists of three reports and a group of filters, (a) drillable map report, (b) radar chart providing a snapshot of utility use, (c) statistical bubble chart, and (d) filter group

This dashboard consists of three interactive reports. The map report starts at SA3 level, and is drillable to SA2 and SA1 levels. The second report (radar chart) gives an overview of the utility usage for the active area on the map report. The statistical bubble chart shows the relationship between per capita water and electricity consumptions at SA2 level. A bubble in this report represents an SA2, and the radius of a bubble is proportional to the population in that SA2 while the colour of bubbles indicates the abundance of flats and units in a SA2. Filters to the right of the dashboard let users control all reports at once.

Multi scale spatial analysis is facilitated by the drillable map report, and the ability to update linked reports with the drill down or roll up on the map. Figure 5 illustrates how a user could click on a map polygon to visualize information at an increasingly finer granule. This has been made possible by the spatial hierarchy designed at the star schema-based data warehouse.

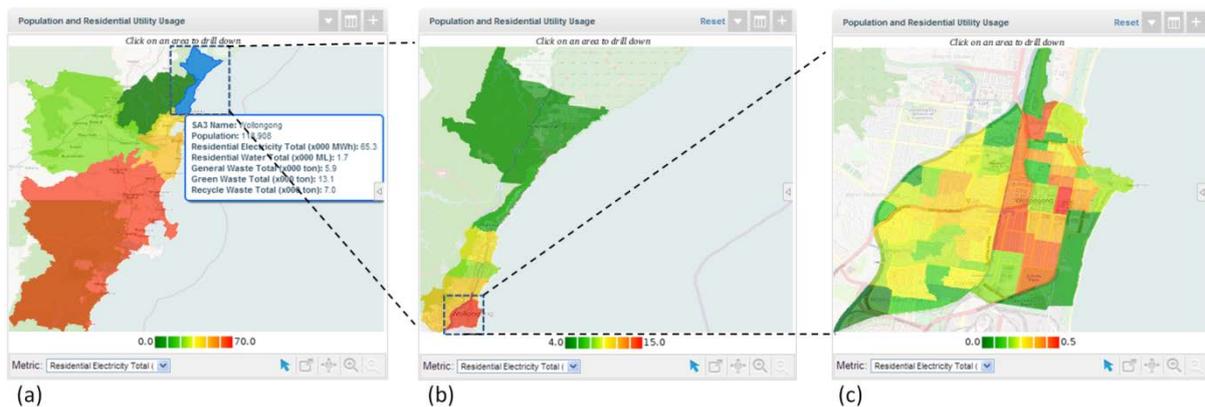


Figure 5. Multi Spatial Analysis, (a) SA3 level, (b) SA2 level, (c) SA1 level

In the map report illustrated in Figure 5, a user can choose to display indices such as population, residential electricity consumption, residential water consumption, quantity of green waste and quantity of recycle waste. This is made possible by a business view created on top of the star schema that brings together appropriate fields from the dimension and fact tables. Noting that the records in the fact tables are at the finest spatial granule (SA1), aggregate functions like sum and average are applied to these columns (e.g. residential electricity) at the report level.

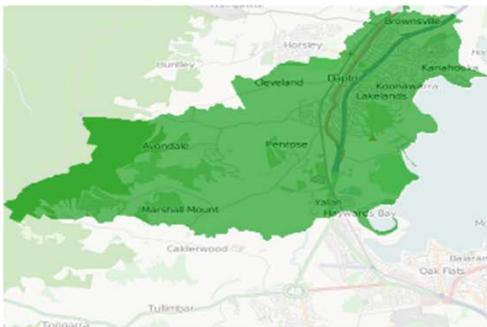
Multi temporal analysis is enabled in this dashboard through hierarchical filters. For the temporal analysis, two levels of disaggregation are possible: year and the seasons. Although SID facilitates further disaggregation in the time dimension, this application stops at season level because water consumption data are available only for seasons.

Perhaps, the most important feature of SID is its ability to cross spatial and temporal granules at will. For example, some patterns could be pertinent at SA2 level only when viewed at seasonal intervals, while some other patterns may project well at SA1 level at yearly intervals. SID has been designed to uncover such hidden patterns in data.

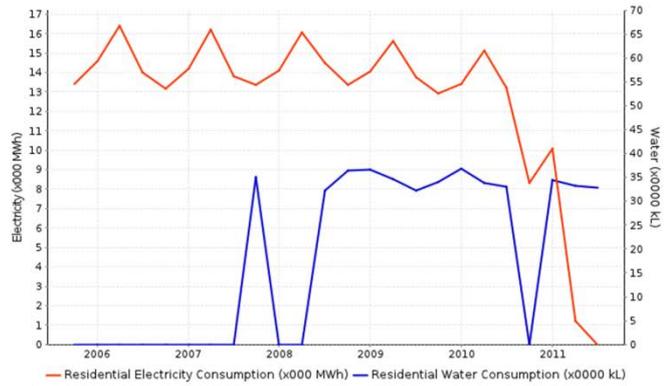
Statistical bubble chart in Figure 4 is an example for several ways by which SID uncovers interrelationships among various utilities and demographic variables at multiple temporal resolutions. While this chart enables the user to appreciate general trends, it is also an ideal tool to spot outliers. For example, the SA2 with the lowest flat and unit percentage showed moderate per capita water consumption, but consistently the lowest per capita electricity usage. SID then supports the natural curiosity of the analyst to find more about this SA2 via detailed 'drill through' reports (Figure 6).

SA2 Name	Dapto - Avondale
Population	22,099
Flats/Units Percentage	3.1
Residential Electricity Consumption (x000 MWh)	56.8
Residential Water Consumption (x0000 kL)	139.8

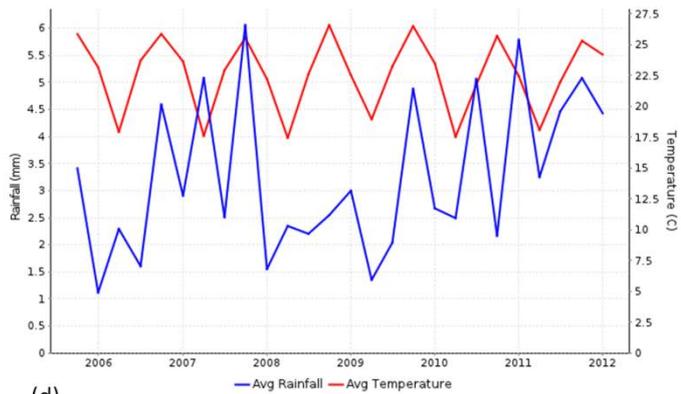
(a)



(b)



(c)



(d)

Figure 6. Child reports illustrating the relationship between utility consumption and weather at SA2 level, (a) SA2 profile for year 2009, (b) boundary of the SA2, (c) Electricity and Water time series, and (d) Weather time series

Child reports accessed via drill through (Figure 6) provides a lot of relevant information such as the boundary of the SA2, additional demographic details, and the trends of utility usage and weather parameters. These are just few examples of many possible child reports that can facilitate the formulation of hypotheses explaining why a particular area would exhibit the observed utility consumption patterns.

5.2 Service Vulnerability

Another interesting application of SID is the assessment of potential service vulnerability of a utility network. Figure 7 illustrates how a user could visualize potentially affected postcodes in case of a failure in an electricity substation.

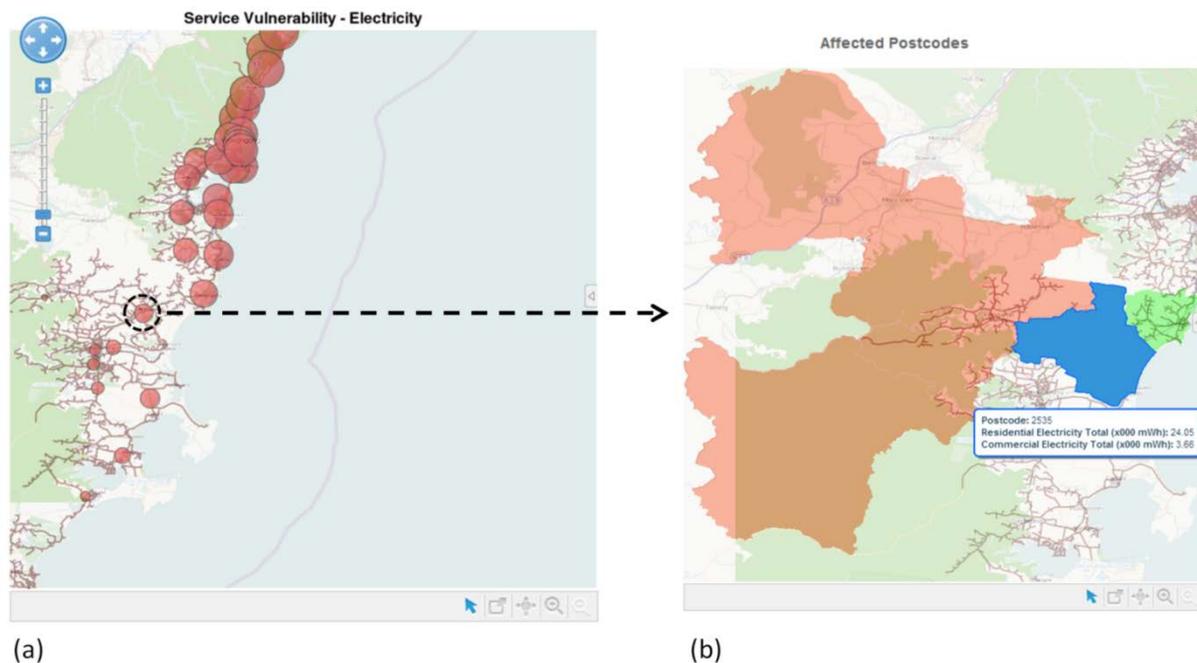


Figure 7. Potential area of impact in case of electricity substation failure, (a) Substations and network links, (b) affected postcodes during a substation failure

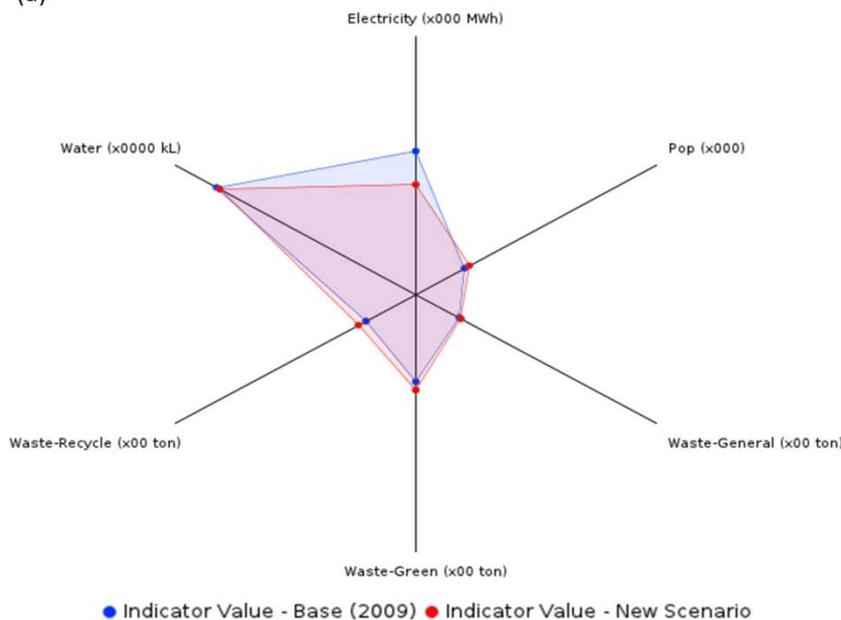
In Figure 7a, red circles represent electricity substations. A user could click on any of these substations to activate a second map showing potentially affected postcodes during a failure of that substation. It is possible to conduct this type of analysis even at the finest spatial resolution (SA1) thanks to the star schema we have designed behind the scene. However, energy distributors often employ mechanisms by which they can restore the energy supply to an area using dynamic loops and backup systems, which we have not taken into account in this analysis. Such scenarios can also be supported in SID by designing the star schema accordingly.

5.3 What-If Scenario

What-If analysis is an indispensable tool by which planners could estimate the potential impact of changes in a set of independent variables on one or more dependent variables (Golfarelli & Rizzi, 2009). Traditionally, BI solutions have been designed to analyse past data only, and they lack the ability to provide any sort of anticipation of future trends. Hence, what-if analysis is a cutting edge feature of BI that makes such tools sit on the top of the BI pyramid (Golfarelli & Rizzi, 2009). SID provides planners with ample opportunities to leverage historical data for better planning through what-if analysis. Figure 8 shows one of several integrated reports in SID that let users to perform what-if analysis.

Indicator	Indicator Value - Base (2009)	Indicator Value - New Scenario
Electricity (x000 MWh)	1,112	852
Pop (x000)	405	444
Waste-General (x00 ton)	361	377
Waste-Green (x00 ton)	678	744
Waste-Recycle (x00 ton)	416	478
Water (x0000 kL)	1,659	1,635

(a)



(b)

Parameter List ✖

Percentage Population Increase
 0 10

Electricity Use Efficiency
 0 30

Water Use Efficiency
 0 10

Recycle Waste Contamination Rate
 0 5

SA3 Name

Select: All | None

[Refresh Cached Filters](#) ➤

(c)

Figure 8. What-If Scenario Analysis, (a) Table report comparing base case and new scenario, (b) Radar chart graphically comparing two scenarios, and (c) Parameter list enabling scenario analysis

Using this integrated report, a planner can estimate the expected utility consumption for a given SA3 or SA2 under various scenarios, and compare the predicted value with the base case both tabularly and graphically. While the planner predicts the utility usage in 5 years, for example, using a certain value for the population increase, he could also see what efficiency gains would keep the expected usage within manageable limits. This report cleverly keeps the complexity of the implemented calculations behind the scene, and presents the non-savvy user with a few easy-to-use sliders and filters to perform multiple scenario analyses on utility usage.

6. Conclusion

Infrastructure services at the local and state levels are usually provided by a mix of public and private agencies. Private sector is increasingly dominant in this domain particularly in developed countries like Australia. Nonetheless, it is the responsibility of local and state governments to ensure that the public receive the best possible infrastructure services. These public agencies, therefore, urgently need an integrated view on the provision and use of infrastructure services in their legislative areas. Among several reasons for the lack of such integrated view, the difficulty to collate disparate datasets from multiple agencies and the technical challenges associated with developing such solutions due to intricate network structures and operations of interconnected infrastructure systems are at the pinnacle.

We suggest that the Geospatial Business Intelligence can be adopted to provide the much needed integrated view on infrastructure service. Geospatial aspect is crucial in this solution to take the best out of majority of the infrastructure-related data that contain a location reference. Moreover, identifying patterns and associations among various infrastructure services and usage in space-time is crucial in decision making. Needless to mention that an interactive map is a better visual when it comes to communicating location-related information. BI provides tools and methods needed to tap into diverse sets of disparate data and load them into an optimized data warehousing environment for efficient analysis and reporting. With the involvement of right people from the design phase to final usage phase, Geo-BI can be turned into a powerful tool for the governance of infrastructure services.

Using a case study for the Illawarra region of New South Wales, Australia, we demonstrated how tools and processes in Geo-BI could be harnessed to develop a user-friendly solution, which we call SMART Infrastructure Dashboard (SID), geared towards the governance of infrastructure services at the local level. This region consists of five Local Government Areas (LGAs) where infrastructure services are provided by several private agencies and local governments themselves. These private and government agencies participated in SID both as data providers and potential users.

Through a web-based portal, SID provides planners and policy makers a visually-rich interface to perform powerful spatio-temporal analyses needed to identify patterns and associations among multiple utility-related variables in space and time. Traversing through space dimension using drill down and roll up functionalities on maps and other reports is enabled by the star-schema based data warehouse. Hierarchical temporal filters ensure that the user can disintegrate the time dimension into years, seasons, months, weeks, days and

even into hours and minutes. Among many other uses of SID, the ability to carry out service vulnerability assessments is a major application area, enabled by an innovative star schema that models infrastructure networks. Moreover, SID facilitates what if scenario analysis offering a way by which planners could anticipate and plan for future trends in utility usage.

In addition to providing a proven solution for local governance of infrastructure services, our approach to BI expands its horizons beyond the traditional use in single organizational environments for strategic decision making. Despite its huge potential, Geo-BI has rarely been used outside the box. We have successfully trialled SID, our innovative Geo-BI solution, in many other non-traditional application areas including water quality monitoring in estuaries used for shrimp farming and evaluating the success of an energy efficiency programme for small and medium sized businesses in New South Wales, Australia. SID is planned to be used in the future as a solution to monitor environmental footprints of publicly managed buildings and as a solution to enhance public awareness in policy debates.

Acknowledgement

Authors would like to acknowledge Sydney Water, Endeavour Energy, REMONDIS, the Office of Environment and Heritage and the Wollongong City Council for their support. This research has been sponsored by the Australian National Data Service (ANDS) through their Application program.

References

- ABS, 2011. Australian Statistical Geography Standard (ASGS): Volume 1 - Main Structure and Greater Capital City Statistical Areas. Australian Bureau of Statistics, Canberra.
- Andrienko, G., Malerba, D., May, M., & Teisseire, M. 2006. Mining spatio-temporal data. *Journal of Intelligent Information Systems*, 27, 187-190.

- Angelaccio, M., Basili, A., Buttarazzi, B., & Liguori, W., 2012. Using geo-business intelligence to improve quality of life, IEEE 1st AESS European Conference on Satellite Telecommunications. 02 - 05 October 2012, Rome.
- Cannadi, J., & Dollery, B. 2005. An evaluation of private sector provision of public infrastructure in Australian local government. *Australian Journal of Public Administration*, 64, 112-118.
- Chen, C. S., Duan, S. X., Cai, T. & Liu, B. Y. 2011. Online 24-h solar power forecasting based on weather type classification using artificial neural network. *Solar Energy* 85: 2856-2870.
- Farooq, F. 2012. The data warehouse virtualization framework for operational business intelligence. *Expert Systems*. Article in Press.
- Franklin, C., & Hane, P. 1992. An introduction to GIS: linking maps to databases. *Database*, 15, 17-22.
- Gartner, 2012. IT Glossary. Gartner Publishing, Stamford.
- Golfarelli, M., & Rizzi, S. 2009. What-if simulation modeling in business intelligence. *International Journal of Data Warehousing and Mining*, 5, 24-43.
- IBM. 2012. City of Greater Geraldton: IBM's Smarter Cities Challenge. IBM Corporation, Armonk.
- Kimball, R., & Caserta, J. 2004. *The Data Warehouse ETL Toolkit*. Wiley Publishing, Inc.
- Liautaud, B., 2000. E-Business Intelligence: Turning Information into Knowledge into Profit. McGraw-Hill.
- Lim, A. H. L. & Lee, C. S. 2010. Processing online analytics with classification and association rule mining. *Knowledge-Based Systems* 23: 248-55.
- Lindquist, E., 2011. Grappling with Complex Policy Challenges: exploring the potential of visualization for analysis, advising and engagement.
- Loshin, D. 2003. *Business Intelligence: The Savvy Manager's Guide*. Morgan Kaufmann, San Francisco, CA.
- Mauree, V. 2010. ICT as an Enabler for Smart Water Management: ITU-T Technology Watch Report.
- McHugh, R., Roche, S. & Bedard, Y. 2009. Towards a SOLAP-based public participation GIS. *Journal of Environmental Management* 90: 2041-2054.
- Mokrech, M., Nicholls, R., Dawson, R. J. 2012. Scenarios of future built environment for coastal risk assessment of climate change using a GIS-based multicriteria analysis. *Environment and Planning B: Planning and Design*, 39 (1): 120-136.
- Olszak, C.M., & Ziemia, E. 2007. Approach to Building and Implementing Business Intelligence Systems. *Interdisciplinary Journal of Information, Knowledge, and Management*, 2, 135-148.
- Petrini, M., & Pozzebon, M. 2009. Managing sustainability with the support of business intelligence: Integrating socio-environmental indicators and organisational context. *The Journal of Strategic Information Systems*, 18, 178-191.
- Sinnott, R. O., Galang, G., Tomko, M., Stimson R. 2011. Towards an e-Infrastructure for Urban Research across Australia. IEEE e-Science Conference, Stockholm, Sweden. Access online: <http://aurin.org.au/wp-content/uploads/2011/10/AURIN-v3.pdf>
- Shi, H., Peng, C., & Xu, M. 2012. Business intelligence in construction: A review. *Advanced Materials Research*, 594-597, 3049-3057.
- Tomic, B. and Milic, T. 2013. Automated interpretation of key performance indicator values and its application in education. *Knowledge-Based Systems* 37. 250-60.

- Wang, S.L., Hong, L., & Chen, X.G. 2012. Vulnerability analysis of interdependent infrastructure systems: A methodological framework. *Physica a-Statistical Mechanics and Its Applications*, 391, 3323-3335.
- Watson, H., Goodhue, D., Wixon, B. 2002. The benefits of data warehousing: why some organizations realize exceptional payoffs. *Information and Management* 39. 491–502.
- Wilkinson, J., 2011. The Illawarra: An Economic Profile, in: N. Parliament (Ed.). Sydney.

Appendix: Summary of available data for the Illawarra Region

Provider	Type	Description	Aggregation
Sydney Water	Consumption	Water consumption by property type in kilolitres	<ul style="list-style-type: none"> Area - Wollongong LGA, Shellharbour LGA, Kiama LGA Geographic Level - Postcode Temporal Level - Quarterly, September 2008 - June 2011
	Flow	Illawarra water delivery system reservoir zone daily demand (ML/day)	<ul style="list-style-type: none"> Area - Sydney Water Reservoir Locations in the Illawarra Temporal Level - Daily, Jan 2011 - August 2012
		Water Treatment Plan - Inflow, Outflow and Reuse	<ul style="list-style-type: none"> Area - Gerringong/Gerroa Temporal Level - Monthly, Feb 2005 - Nov 2009
		Wollongong Sewerage Discharge, Litres/Second and Total Volumes	<ul style="list-style-type: none"> Area - Wollongong Temporal Level - Daily, Jan 2006 - June 2006
	Network	Illawarra Water Network Spatial Data	<ul style="list-style-type: none"> Reservoir Locations Water Mains Water Pumping Stations Water Treatment Plants
		Illawarra Sewer Network Spatial Data	<ul style="list-style-type: none"> Sewer Mains Sewer Pumping Stations Sewer Treatment Plants
	Quality	Illawarra Reservoir Water Quality	<ul style="list-style-type: none"> Area - Illawarra Reservoir Locations Temporal Level - Weekly - Jan 2006 - June 2012
		Illawarra Reticulated Water Report	<ul style="list-style-type: none"> Area - Illawarra Geographic Level - Suburbs Temporal Level - Daily/Weekly/Monthly - Jan 2006 - June 2012
		Illawarra Water Filtration Plant	<ul style="list-style-type: none"> Area - Kembla Grange Temporal Level - Weekly, Jan 2006 - June 2012
	Energy Consumption	Illawarra Plant Energy Consumption (MWh)	<ul style="list-style-type: none"> Area - Illawarra Geographic Level - Selected Sites Temporal Level: Yearly - 2006 - 2011
Endeavour Energy	Consumption	Electricity consumption	<ul style="list-style-type: none"> Area - Wollongong LGA, Shellharbour LGA, Kiama LGA, Wingecarribee LGA, Shoalhaven Geographic Level - Collection District Temporal Level - Monthly, 2007 - 2011

Provider	Type	Description	Aggregation
	Network	Illawarra Electricity Network	<ul style="list-style-type: none"> • Geographic Level – Collection District • Conductors • Distribution Stations • Switches • Zone Sub Stations
REMONDIS	Collection	Waste in Kgs collected per route. No. of Bins collected	<ul style="list-style-type: none"> • Area - Wollongong LGA, Shellharbour LGA • Geographic Level – Street • Temporal Level – Daily, 2007 - 2011
	Network	Bin collection route Spatial Data	<ul style="list-style-type: none"> • Area - Wollongong LGA, Shellharbour LGA
Bureau of Meteorology	Rainfall and Temperature	Rainfall and Temperature Data	<ul style="list-style-type: none"> • Area - BOM sites in the Illawarra
Australian Bureau of Statistics	Demographic	Census Data	<ul style="list-style-type: none"> • Temporal Level – 2006 and 2011 Census