

SMART Infrastructure Dashboard: A Fusion between Business Intelligence and Geographic Information Systems

Rohan Wickramasuriya^{a*}

Jun Ma^a

Vivek Somashekar^a

Pascal Perez^a

Matthew Berryman^a

Abstract: Business Intelligence (BI) has popularly been adopted as a process that enables easy access, analysis and visualization of information through specialized set of tools for informed decision making. Two most noticeable characteristics of traditional BI is that it (a) is largely used in single-organization environments and (b) uses predominantly aspatial data. We believe that BI has applications beyond single-organization environments, but it very much requires integration of geospatial capabilities given the increasing availability of large volumes of spatial data and a growing interest to see things spatial. The SMART Infrastructure Dashboard (SID), our innovative solution that fuses BI and Geographic Information Systems (GIS), fills this significant gap. In this study, we demonstrate how SID can be used to perform spatio-temporal analysis and visualization of diverse sets of data to uncover complex interrelationships among utility usage, demographics and weather patterns at local and regional scale.

Key words: Business intelligence; Geographic information systems; Infrastructure

I. Introduction

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’¹. In terms of processes, a BI project involves data acquisition, data warehousing, data analysis and mining, and reporting and presentation². Geospatial BI (Geo-BI) is an improvement upon traditional BI made possible by integrating Geographic Information Systems (GIS) with BI³. 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 reference⁴, Geo-BI provides the spatial perspective which was missing in traditional BI.

^a SMART Infrastructure Facility, University of Wollongong, Wollongong, NSW 2522, Australia,

*Corresponding author: Rohan Wickramasuriya rohan@uow.edu.au

<http://dx.doi.org/10.14453/isngi2013.proc.49>

Though BI and Geo-BI are traditionally used in single organization environments primarily as a decision support tool that helps to achieve increased profits, we suggest that Geo-BI in particular has applications beyond this traditional use. In this study, we demonstrate how Geo-BI can be tailored to provide an integrated view on infrastructure networks and services for better governance and planning of cities and regions.

The Provision of infrastructure services to communities is a fundamental requirement, and this has been customarily 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 such services⁵. 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 interconnected⁶. Hence, local and state governments urgently need tools that can provide an integrated view on infrastructure networks and services. Such tools should overcome several technical challenges. First, given the diversity of infrastructure networks and the decentralized nature of their management, they should be able to harness diverse and dispersed data sources. Second, they have to handle the underlying complexity of operations on individual networks, as well as the interconnectedness of networks. Moreover, as any other decision-support system, these tools need to exhibit positive usability traits such as performance, user-friendliness and intuitive visuals.

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. The current version of SID includes the following utilities: electricity and water distribution, as well as sewage and solid waste collection and treatment. SID has been principally designed for the Illawarra region in New South Wales, Australia.

II. The Illawarra Region

The Illawarra, a 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 census⁷, the population of the Illawarra region stood at 413,216 persons, 46.6% of which lived in Wollongong LGA only⁸.

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.

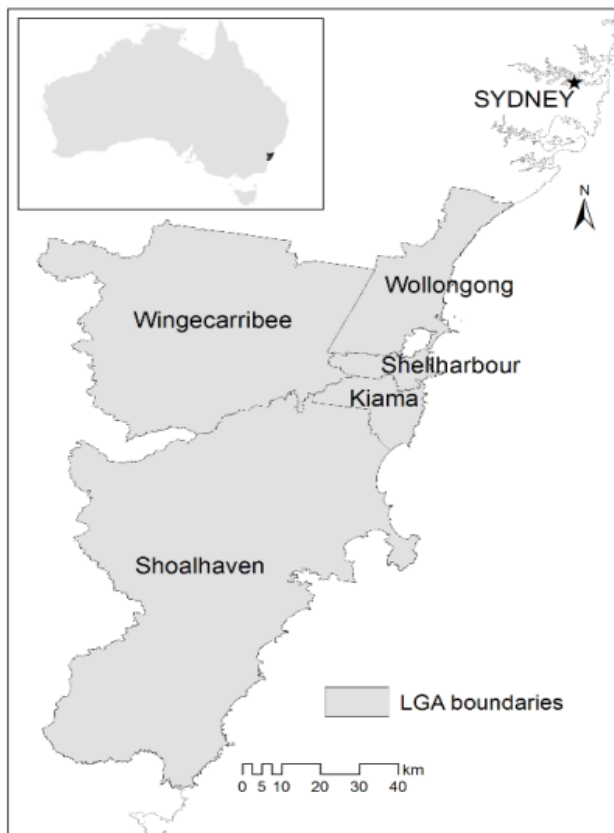


Figure 1. The Illawarra regions, New South Wales, Australia

III. Smart Infrastructure Dashboard

A. Stakeholder and Data Types

We identified the five LGAs and aforementioned private operators as stakeholders in the SID project. From SID's perspective, these stakeholders were both data providers and users. We collected a diverse set of data from them 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 incorporated into SID.

B. Technical Architecture and Work Flows

Figure 2 gives an overview of SID's technical architecture and main workflows involved. SID receives data 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). We use

the standard Extract, Transform and Load (ETL) process⁹ to migrate these data into an optimized data warehouse environment.

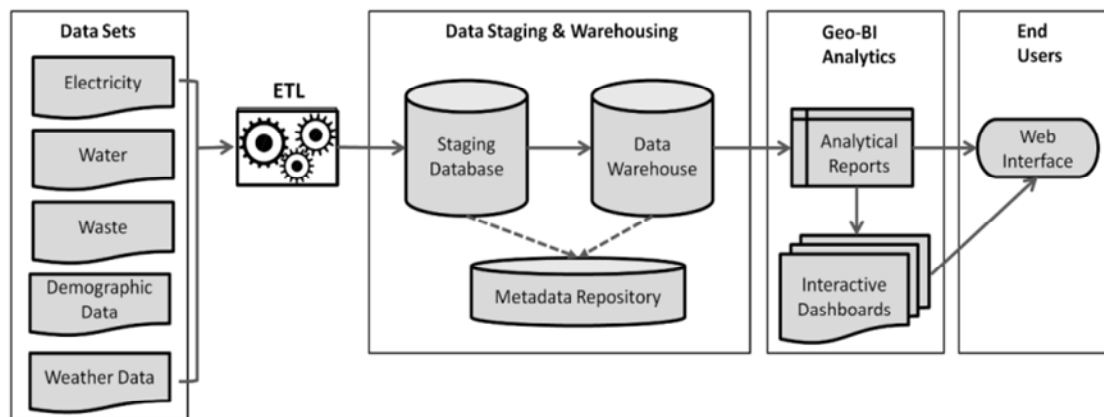


Figure 2. Technical architecture and workflows of SID

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. Currently, two geographic hierarchies are supported in SID. The first is based on the Australian Statistical Geography Standard (ASGS) released by ABS recently⁷. This hierarchy starts at Statistical Area Level 3 (SA3), and drills down to SA1 through SA2. The second geographic hierarchy is a two-step one where an LGA drills down to postcodes.

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 applications⁹. It organises data into one or more fact tables referencing any number of dimension tables. A star schema can simplify join queries and provide capability to analyse data from multiple dimensions, thus enabling a user to perform various drill down, roll up, slice and dice operations on data. In our data warehouse, each utility networks is modelled using a separate star schema. However, these network schemas and the schema for utility consumption are interconnected through common dimensions.

Geo-analytical tools (proprietary Geo-BI software *Yellowfin* in SID's case) 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 tools¹⁰.

C. Applications

SID has a myriad of potential applications in the infrastructure domain that can provide planners and policy makers with ingredients needed to ensure a better service provision to communities. We briefly discuss two such applications next.

- 1) *Relationships among Utility Usage, Socio-Demographics and Weather*

Figure 3 shows an analytical dashboard developed in SID that enables exploration of relationships among utility usage, demographic variables and weather parameters. This dashboard consists of three interactive reports. The map report starts at LGA level, and is drillable to postcode level. Moreover, the other two reports can be controlled by this map report. A user can select one out of several metrics, for example total residential electricity consumption, to be displayed on the map. 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 postcode level. A bubble in this report represents a postcode, and the radius of a bubble is proportional to the population in that postcode while the colour of bubbles indicates the abundance of flats and units in a postcode. Filters to the right of the dashboard let users control all reports at once.

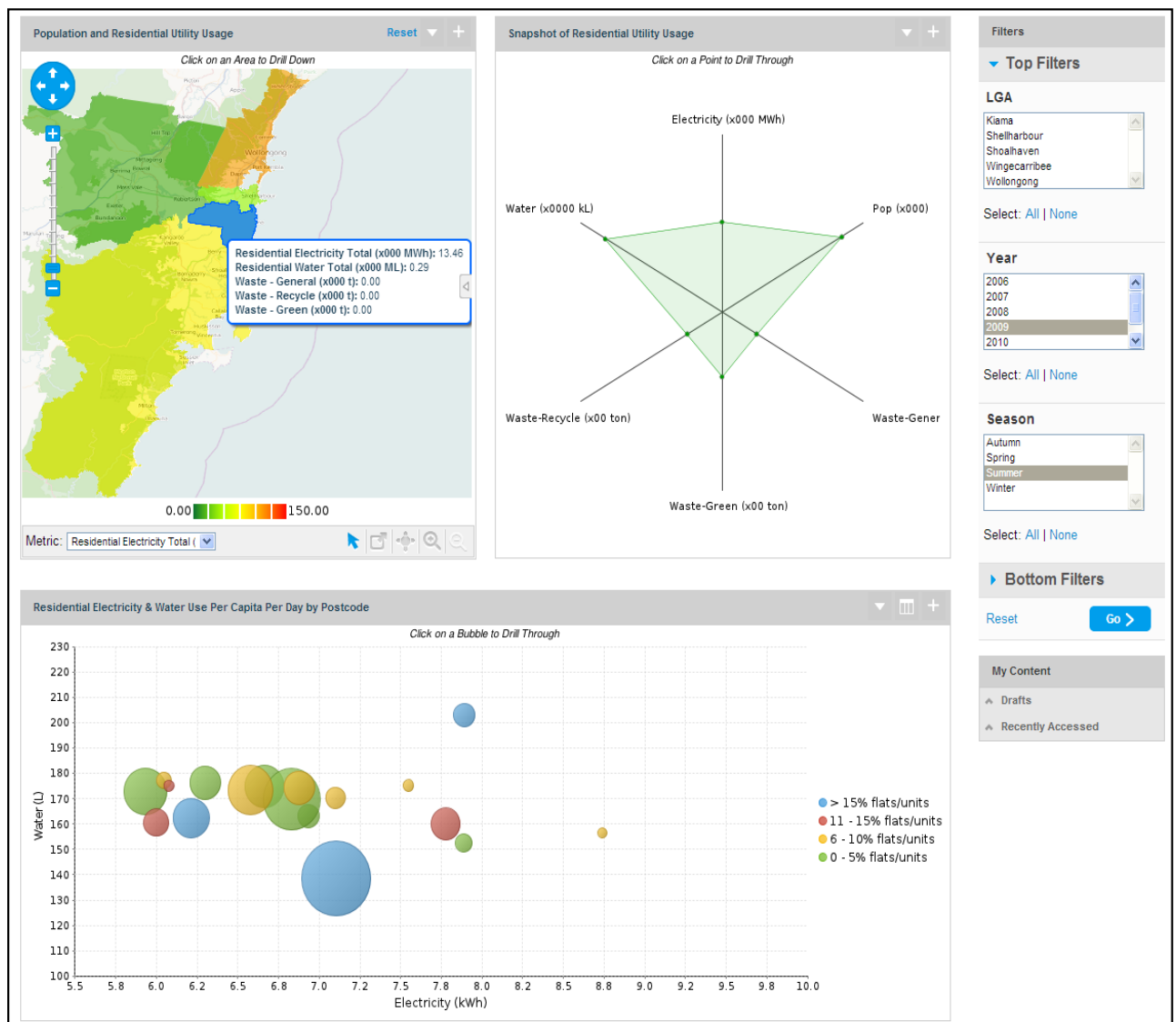


Figure 3. An interactive, geo-analytical dashboard used in SID

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. Multi temporal analysis is enabled in this dashboard through a hierarchical filter. For the temporal analysis, two levels of disaggregation are possible: year and the seasons. Although SID facilitates further disaggregation in the time dimension through month, weeks, days up to minutes, this particular application limits this disaggregation to season level only. The main reason for this was the water consumption data for which the finest available temporal granule was season.

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 postcode level only when viewed at seasonal intervals, while some other patterns may project well at LGA level at yearly intervals. SID has been designed to uncover such hidden patterns in data.

Statistical bubble chart in Figure 3 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.

D. 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 variables¹¹. 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 modern and cutting edge feature of BI that makes such tools sit on the top of the BI pyramid¹¹. SID as a Geo-BI solution provides planners with ample opportunities to leverage historical data for better planning through what if analysis. Figure 4 shows one of several integrated reports in SID that let users to carry out what if analysis.

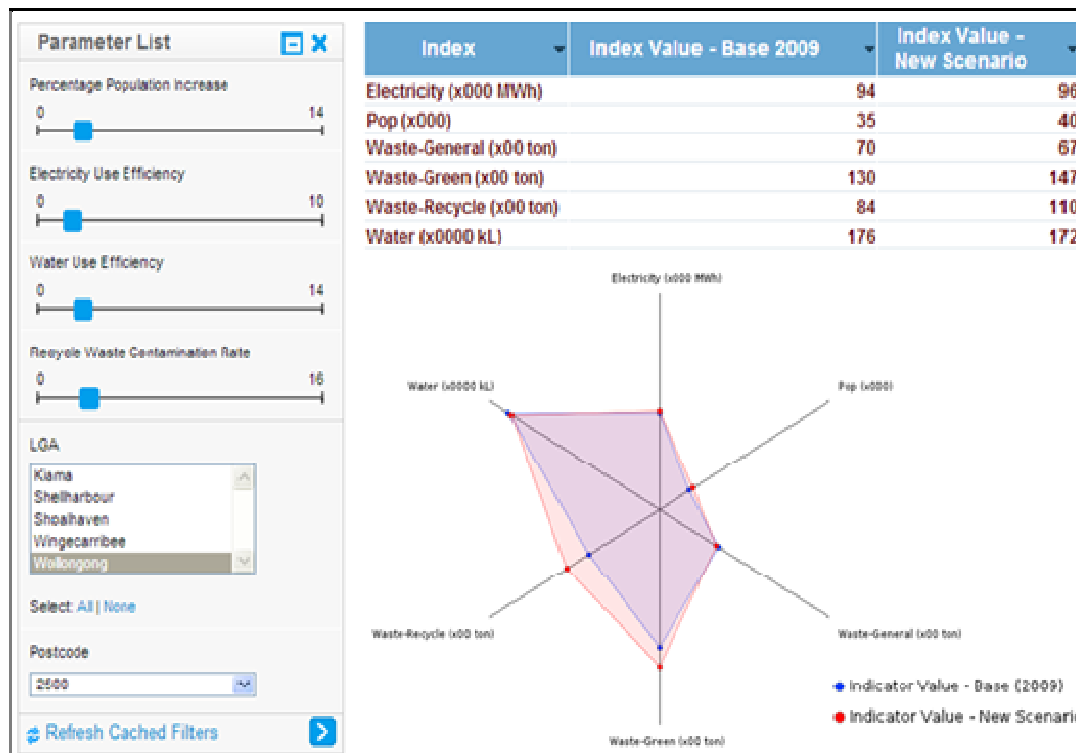


Figure 4. What-if scenario analysis

Using this integrated report, a planner can estimate the expected utility use for a given LGA or postcode under various scenarios, and compare the predicted value with the base case (table and graph). 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 keeps 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.

IV. Discussion and Conclusions

Despite the traditional use of Geospatial Business Intelligence (Geo-BI) as a tool to help increase profits within a single organization, we believe that it possesses a largely untapped potential that warrants its use as a generic decision support system in planning and policy domains. In this study, we demonstrate that the Geo-BI can be adopted to provide the much needed integrated view on infrastructure service at local and regional level.

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 termed the SMART Infrastructure Dashboard (SID), geared towards the governance of infrastructure services at the local and regional levels. 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. Moreover, SID facilitates what if scenario analysis offering a way by which planners could anticipate and plan for future trends in utility usage. Future research will involve extending SID's capabilities to investigate the propagation of cascading failures in interconnected networks.

Acknowledgement

Authors would like to acknowledge Sydney Water, Endeavour Energy, REMONDIS and the Southern Councils Group. This research has been supported by the Australian National Data Service (ANDS) through their Application program.

References

- ¹Gartner, *IT Glossary* [online glossary], Gartner Publishing, 2013, URL:
<http://www.gartner.com/it-glossary/>

- ²Shi, H., Peng, C., and Xu, M., “Business Intelligence in Construction: A Review”, *Advanced Materials Research*, Vols. 594-597, 2012, pp. 3049-3057.
<http://dx.doi.org/10.4028/www.scientific.net/AMR.594-597.3049>
- ³Angelaccio, M., Basili, A., Buttarazzi, B., and Liguori, W., “Using Geo-Business Intelligence to Improve Quality of Life”, *IEEE 1st AESS European Conference on Satellite Telecommunications*. Rome, 2012.
- ⁴Franklin, C., and Hane, P., “An Introduction to GIS: Linking Maps to Databases”, *Database*, Vol. 15, 1992, pp. 17-22.
- ⁵Cannadi, J., and Dollery, B., “An Evaluation of Private Sector Provision of Public Infrastructure in Australian Local Government”, *Australian Journal of Public Administration*, Vol. 64, No. 2, 2005, pp. 112-118.
<http://dx.doi.org/10.1111/j.1467-8500.2005.00457.x>
- ⁶Wang, S.L., Hong, L., and Chen, X.G., “Vulnerability Analysis of Interdependent Infrastructure Systems: A Methodological Framework”, *Physica a-Statistical Mechanics and Its Applications*, Vol. 391, 2012, pp. 3323-3335.
- ⁷ABS, *Australian Statistical Geography Standard (ASGS): Volume 1 - Main Structure and Greater Capital City Statistical Areas*, Australian Bureau of Statistics, Canberra, 2011.
- ⁸Wilkinson, J., “The Illawarra: An Economic Profile”, *e-brief 18/2011* [electronic document], NSW Parliamentary Library Research Service, Sydney, 2011, URL:
[http://www.parliament.nsw.gov.au/prod/parlment/publications.nsf/key/TheIllawarra:AnEconomicProfile/\\$File/Illawarra+Region+An+Economic+Profile+GG2.pdf](http://www.parliament.nsw.gov.au/prod/parlment/publications.nsf/key/TheIllawarra:AnEconomicProfile/$File/Illawarra+Region+An+Economic+Profile+GG2.pdf)
- ⁹Kimball, R., and Caserta, J., *The Data Warehouse ETL Toolkit*, Wiley Publishing, Inc. 2004.
- ¹⁰Lindquist, E., “Grappling with Complex Policy Challenges: Exploring the Potential of Visualization for Analysis, Advising and Engagement” [online document], *HC Coombs Policy Forum*, 2011, URL:
https://www.anzsog.edu.au/media/upload/publication/108_Lindquist-Discussion-paper-from-roundtable-2.pdf
- ¹¹Golfarelli, M., and Rizzi, S., “What-if simulation Modelling in Business Intelligence”, *International Journal of Data Warehousing and Mining*, Vol. 5, 2009, pp. 24-43.
<http://dx.doi.org/10.4018/jdwm.2009080702>