

A LOW CARBON LIVING SPATIAL DATA HUB



**RP 3003 - A Low Carbon Living Spatial Data Hub: Establishing the case for the
development of an integrated spatial data server and analysis capacity for the CRCLCL -
FINAL REPORT**

A. Taeihagh and B. Randolph

City Futures Research Centre, University of New South Wales

with options modelling by Jared Berghold and Penny Cooke, Intersect

November 2013

City Futures Research Centre, Faculty of Built Environment,

University of New South Wales, Sydney 2052, NSW, Australia

Telephone: +61 (2) 9385 7777 / +61 (2) 9385 5117 Fax: +61 (2) 9385 5935

Email: cityfutures.contact@unsw.edu.au Website: www.cityfutures.net.au



An Australian Government Initiative



This study was undertaken by UNSW, Intersect, Swinburne University and CSIRO on behalf of the CRC for Low Carbon Living.

November 2013

ISBN:

@ CRC for Low Carbon Living

CRC FOR LOW CARBON LIVING LTD

ROOM 202-207, LEVEL 2

TYREE ENERGY TECHNOLOGIES BUILDING

UNSW SYDNEY NSW 2052 AUSTRALIA

National Library of Australia Cataloguing-in-Publication data:

CRC participants involved:

Participant 1: UNSW

Participant 2: Swinburne University

Participant 3: CSIRO

Participant 4: Intersect

CRC participants consulted:

University of South Australia

University of Melbourne

Curtin University

Acknowledgements

We would like to acknowledge the contributions of many of the members involved with CRC for Low Carbon Living. In particular program/project leaders who supported our efforts in better understanding their projects and involvements in CRC:

Peter Newton, Swinburne University

Stephen Wright, CSIRO

Andrew Higgins, CSIRO

Chris Ryan, University of Melbourne

Deo Prasad, UNSW – CRC LCL

Janette Hartz-Karp, Curtin University

Michael A P Taylor, University of South Australia

Nicholas Holyoak, University of South Australia

Paul Osmond, UNSW

Steffen Lehmann, University of South Australia

Terry Burke, Swinburne University

Tomi Winfree, Swinburne University

In addition to those listed above, we would like to thank Maria Schwensen and Paul Hopkins (CRC) for providing us access to detailed information in regards to the projects in programs 2 and 3, Andrew Tice (UNSW) for his insights in regards to geospatial information and systems, and Shahla Ghobadi (Intersect) for her support in the development of the technical options for the Data Hub.

Executive Summary

This scoping project sits across two of the three program areas of the CRC. It aims to assess the scope for developing an integrated building level and cadastrally organised relational database and analysis capacity to link the full range of geo-coded data which will be necessary to operationalize the Living Laboratory concept as well as the broader objectives of the CRC LCL. It also reviews the requirements generated by the need to undertake a range of survey and spatial modelling activities at a range of spatial scales to both support specific research projects as well as to form a significant research tool in its own right.

The analysis of the CRC LCL funded projects and discussions with the program and project leaders makes it clear that a large number of the current funded projects use spatial data, are interested in accessing spatial information and will generate outputs at different spatial scales. Furthermore, in future there will be an increase in the number of such projects as CRC LCL matures and moves from the initial pilot and scoping studies phase to more specific place-based research across Australia.

The key outcome of this scoping study is the specification options presented below for an integrated spatial data facility – the Low Carbon Living Data Hub – that will provide a central resource for all research undertaken under the CRC, stimulate further research that would use and add to the Data Hubs dataset holdings as well as links the empirical outcomes into other major eResearch initiatives currently under development in Australia

Four possible options considered for the LCL Data Hub are:

1. Basic Data Hub – provides storage infrastructure and support
2. AURIN-Enabled Data Hub – adds an analysis capability in the form of integration with the AURIN portal

3. Extended AURIN-Enabled Data Hub – extends the analysis capability of the LCL Data Hub
4. Enhanced Spatial Analysis Hub – provides a fully customised, advanced analysis capability for the LCL Data Hub.

The report concludes by recommending Option 3 (Extended AURIN-Enabled Data Hub) – that provides the data storage infrastructure and support, enables analysis via the AURIN portal, adds an automated data feed from the originating sources for datasets, and provides analysis tools and additional functionality more closely aligned with the needs of the LCL research cohort.

Keywords: Geospatial Data, Data Integration, Data analysis, Data Hub, Low Carbon Living, Living Laboratories, Urban Analysis

Table of Contents

Acknowledgements.....	5
Executive Summary.....	6
1 Introduction.....	10
2 Background Information.....	12
2.1 AURIN.....	14
2.2 Data Integration.....	15
2.2.1 Precinct Information Modelling (PIM).....	17
3 Project Methodology.....	17
3.1 Time Frame.....	18
3.2 Initial Scope and Project Refinement.....	18
4 Findings from Data Analysis and Discussions with Project Leaders.....	22
4.1 Identification and Analysis of Program and Project Partnerships.....	22
4.2 User Needs Assessment.....	24
4.3 Living Laboratories.....	31
4.4 A Proposed Data Structure.....	34
5 Proposed Specifications for the Data Hub.....	40
5.1 Defining the Data Hub.....	41
5.2 Options Overview and Rationale.....	41
5.3 Cost Estimates Overview.....	42

5.4	Data Hub Options	43
5.4.1	Option 1 – Basic Data Hub	43
5.4.2	Option 2 – AURIN-Enabled Data Hub	49
5.4.3	Option 3 – Extended AURIN-Enabled Data Hub	55
5.4.4	Option 4 – Enhanced Spatial Analysis Hub	60
5.5	Trade-off Analysis.....	62
6	Conclusions and Recommendations	65
	References.....	68
	Appendices.....	69
	Appendix A1 List of organisations and respective datasets in AURIN portal (AURIN 2013 – retrieved in October 2013).....	69
	Appendix A2 List of CRC output and utilization milestones that the development and use of the Data Hub can directly and/or indirectly support	75
	Appendix A3 List of Latest CRC Commissioned Projects.....	76
	Appendix A4 List of Potential Future CRC LCL Projects.....	77

**RP 3003 - A Low Carbon Living Spatial Data Hub: Establishing the case for the
development of an integrated spatial data server and analysis capacity for the CRCLCL -
FINAL REPORT**

1 Introduction

CRC Low Carbon Living aims to provide a range of social, technological and policy tools to facilitate the adoption of cost effective low carbon products and services to help reduce greenhouse gas emissions and avoid negative consequences of climate change. Most of the low cost carbon saving opportunities are known to be associated with the built environment (CRC LCL, 2013) and the CRC LCL is assembling a necessary critical mass and a diverse range of stakeholders to provide alternative infrastructure and community engagement solutions and demonstrate pathways to carbon reduction in Australia. Specifically the CRC Low Carbon Living seeks (CRC LCL, 2013):

- A reduction in built environment carbon emissions by 10MTCO₂-e/yr by 2020
- A projected economic benefit of over \$684 million
- Adoption of new government policies and industry building models that contribute to achieving the government's 2050 greenhouse gas emissions reduction target in the built environment
- Development of young researchers with skills in the low carbon built environment.

CRC LCL (2013) aims to achieve these goals through the following programs:

Program 1: Integrated Building Systems: development of building-integrated multipurpose solar products, low-carbon-lifecycle building construction components/ materials, and integrated design, energy rating and reduction methodologies

Program 2: Low Carbon Precincts: by understanding the interdependence of energy, water, waste, transport systems and buildings and reducing the carbon footprint of precinct infrastructure through the development of better tools and planning techniques

Program 3: Engaged Communities: understanding and influencing consumer behaviour and decision-making to reduce the carbon intensity of modern lifestyles.

With CRC's ambitious targets and goals it is evident that researchers and project teams will require, accumulate, and use extensive amounts of data. Therefore the development of a central Data Hub to store, manage, share and use data with all of the partners involved in CRC becomes an important consideration. The development of a Data Hub will make it possible to link the empirical outcomes of various research initiatives in a spatially integrated manner. Therefore, adopting a systematic and rigorous approach for understanding the user needs in terms of data sources, data requirements, and Data Hub features is necessary to ensure the accuracy of this scoping study.

The quality and accuracy of the information gathered has a direct impact on the design and specification of alternatives generated for the development of the Data Hub and ultimately the recommended approach for costings, location, staffing requirements of the Data Hub, and its successful implementation.

Furthermore, adopting the systematic user needs assessment ensures correct alignment of the initial vision for the Data Hub formed by the project funders and the final specifications and detailed architectures proposed by technical experts. A poor understanding of the user needs is considered as one of the most important causes of failure in computer systems (Davis, 1993). Faulk (1997) points out that the relative cost for correcting systems errors escalates rapidly in progressively advanced stages of the systems development life cycle.

The background information is discussed in Section 2. Section 3 describes the project methodology, time frame, and approach in identification of key project leaders and objectives. Section 4 illustrates the findings from data analysis and discussions with stakeholders. The proposed specifications with their costing, pros and cons are presented in Section 5. Recommendations and conclusions are described in Sections 6 and 7.

2 Background Information

In order to better take advantage of the research outputs from CRC funded projects and further stimulate spatial research the CRC will rapidly require the capacity to provide and integrate a wide range of data across the variety of spatial scales. It's likely core datasets available will include data on:

Building Scale

- Location, structure, adaptation and performance, including embodied energy;
- Energy rating such as Green Star, Nabers, BASIX, etc.;
- Residential Use

- Households (demographics, behaviours and attitudes relevant to energy consumption, including psychographical data drawn from qualitative and quantitative interview and survey research);
- Individual household energy and water consumption and waste generation data drawn from utility companies and local authorities;
- Household transport usage and mode share, including data on vehicle ownership and emissions profile;
- Non-residential Use
 - Business energy and water consumption and waste generation data drawn from utility companies and local authorities;
 - Business transport usage and mode share, including data on vehicle ownership and emissions profile;
 - Business and employment structure including energy use and customer/client profile and behaviours;

Precinct and City Scales

- ABS Census data
- Building Stock information
- Information on the distributive networks for key infrastructure including transport, energy grids and ICT;
- Community-based renewable energy projects
- Urban development patterns
- Pedestrian information
- CO₂ Emission data

- 3D data on strata and higher density residential and commercial/industrial building typologies;
- Geo-spatial data climatic characteristics such as urban heat island effects, rainfall, wind patterns, temperature and vegetation cover.
- Development approval data for new build, demolitions and alterations;
- Property valuation and rental characteristics.

2.1 AURIN

The aforementioned datasets are vital for CRC LCL and initiatives such as Australian Urban Research Information Network (AURIN) already provide access to some of these data through their portal.

AURIN's mission is to provide an infrastructure to facilitate access to a distributed network of aggregated spatially referenced datasets and information services, facilitate discovery, navigation, modelling and simulation, interpretation and communication of data by many stakeholders in order to increase understanding of patterns of urban development and to form and model urban growth for a sustainable future of Australian cities (AURIN, 2013).

A significant range of organizations already contribute to AURIN as of October 2013 (current datasets provided by these organizations are listed in Appendix A1) and it is anticipated that AURIN's data holdings will be significantly enhanced in the coming years as AURIN extends its capacity and scale across Australian jurisdictions.

Moreover, by linking the CRC *Low Carbon Living Data Hub* into the development of the AURIN and related eResearch initiatives and datasets, research focussing on properties, households and businesses can benchmarked against similar communities in comparable

spatially defined neighbourhoods to allow the inclusion of “control” examples to be built into the research across the program areas.

2.2 Data Integration

As the population in urban areas is increasing and our socio-technical systems are becoming more complex, better integration of technological, environmental and social data becomes more important for increasing our understanding of our urban systems and for improving the functionality and design of our cities. Moreover, maintaining, updating and sharing large datasets can be costly and time consuming.

All of the relevant spatial datasets required by CRC LCL researchers will have a key locational attributes that will need integrating into a central geo-spatial relational data set. This means that a spatially integrated data set will need to be organised at the land use cadastre level through a geo-codable address or building lot number. These building and household datasets will need to be integrated with available additional data such as ABS Census data at small area geographies, especially using the new Census geographies, that will provide critical socio-demographic profiling data. But there is also a key need to integrate the property household and neighbourhood level data that the CRC research will itself generate during its life cycle.

For example, if an educational project in the *Engaged Communities Program* needs to assess how the intervention has played out at the household level, then addresses of targeted households can be linked to their energy usage derived from utility data and building characteristics (including synthetic prices or rental values) derived from property level data. Similarly, households selected to implement new technologies in the *Integrated Building Systems Program* can be linked to property level data on building type, structure and orientation and

energy consumption data from their energy utility, as well as any behavioural data gathered from household surveys.

In addition to this, an integrated geospatial dataset based on the cadastre level will provide a data rich sampling frame for conducting local or national household or business surveys or for deriving addresses for the recruitment of panels or focus groups for behavioural psychographical research which will be central to a range of research to be undertaken by the CRC.

There is a clear need, therefore, for the CRC to establish a centrally managed and accessible property level dataset that integrates the wide range of property, business and household data that will be generated throughout the lifetime of the Centre.

This will need to link to other datasets that are likely to be made available through the AURIN initiative at local and national level, as well as other yet to be identified spatially identifiable datasets. This facility will form a central resource for the entire CRC and will grow into a nationally, and indeed, internationally significant dataset, providing the opportunity for a wide range of researchers to access and add to the data and use the analysis capacities.

In this way, over the lifetime the CRC, the Low Carbon Living Data Hub will develop into an unrivalled source of information on the development, diffusion, and adoption of low carbon technologies, behaviours and outcomes in the wider Australian community. As such, this will be a world leading development and will also provide a potentially long lasting legacy for the CRC well into the future.

2.2.1 Precinct Information Modelling (PIM)

Geographic Information Systems (GIS) and Building Information Modelling (BIM) along with socio-demographic and socio-economic data are an important part of urban analysis and modelling. GIS handles large scale geospatial data and BIM is used for modelling individual buildings. An opportunity exists to bridge BIM and GIS systems conceptually and technically (Song et al., 2010; Zhang et al., 2009; Newton et al., 2013). Newton et al. (2013) aim to bridge the GIS and BIM systems through new built environment digital modelling technologies termed PIM and identify the following opportunities for further development of the PIM concept:

- PIM object library to provide a mechanism for locating and linking to reference data;
- A precinct viewer based on the PIM object library format that be accessed from any precinct assessment tool;
- A precinct object ontological dictionary to facilitate accessing and sharing of reference data;
- Establishment of shared PIM databases for Living Laboratories.

3 Project Methodology

The development of CRC LCL Data Hub will need careful specification, costing, and understanding of the trades between different options. Therefore, development of the technical specifications for the Data Hub have been conducted in close collaboration with *Intersect*, the NSW University funded eResearch support company (<http://www.intersect.org.au/>).

3.1 Time Frame

This scoping study has been conducted in four phases:

- Initial scoping and project refinement 1 – 30 May 2013
- Discussions with key partners and agencies 1 – 31 June 2013
- Development of alternative Data Hub specifications July - Aug 2013
- Delivery of final scoping report with recommendations 30 October 2013

3.2 Initial Scope and Project Refinement

Following the commencement of the scoping study, in May-June 2013, the conceptual framework for conducting the scoping study was refined based on the preliminary analysis of the projects/reports and discussions with CEO of CRC LCL (Prof. Deo Prasad) and leaders from Program 3 (Dr. Stephen White) and Program 2 (Prof. Peter Newton). This was due to strong spatial and conceptual relation between Precincts and Living Laboratories identified during preliminary data analysis and discussions. Figure 1 illustrates the conceptual framework used for conducting the scoping study.

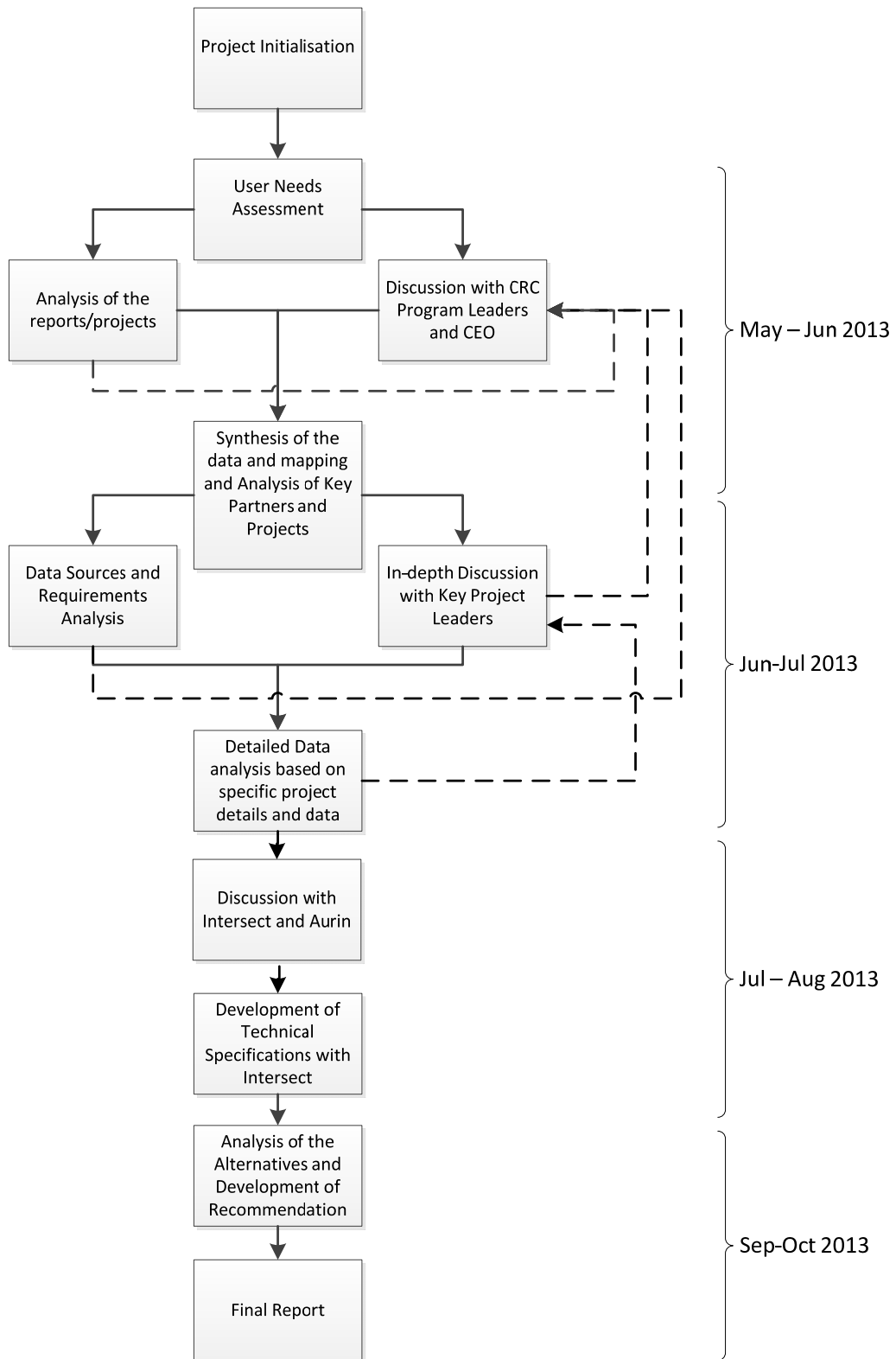


Figure 1 Conceptual framework of the DATA HUB Scoping Study

The components of this study after initial scope and project refinement were:

- Analysis of CRC reports and projects
- Discussions with CRC LCL Leadership to understand CRC-wide and program specific visions, goals and challenges.
- Synthesis of the Data from project descriptions, and mapping and analysis of key projects and partners.
- Understanding Data Sources and Requirements and discussion with key project leaders to understand project goals, data sources and requirements.
- Analysis of project level data and establishment of a range of base datasets required for a cadastrally-based spatial database that would provide key integrated data sets for the Living Laboratories.
- Discussions with Intersect and AURIN to explore compatibility and possibility of integration of the Data Hub with national level initiatives.
- Development and analysis of technical requirements for LCL Data Hub and establishment of technical architectures and specification alternatives with Intersect
- Analysis of the generated alternatives and development of the scoping study's recommendation to CRC for implementation of the Data Hub.
- Write up of the Final Report.

As indicated in Figure 1, successful implementation of this scoping study required iteration and consultation with Program Leaders which was carried out in the project's lifespan. Furthermore, in specific instances similar approach were taken in development of the data sources and requirements

with specific project leaders on case-by-case basis. Figure 2 depicts the workflow used for user needs assessment.

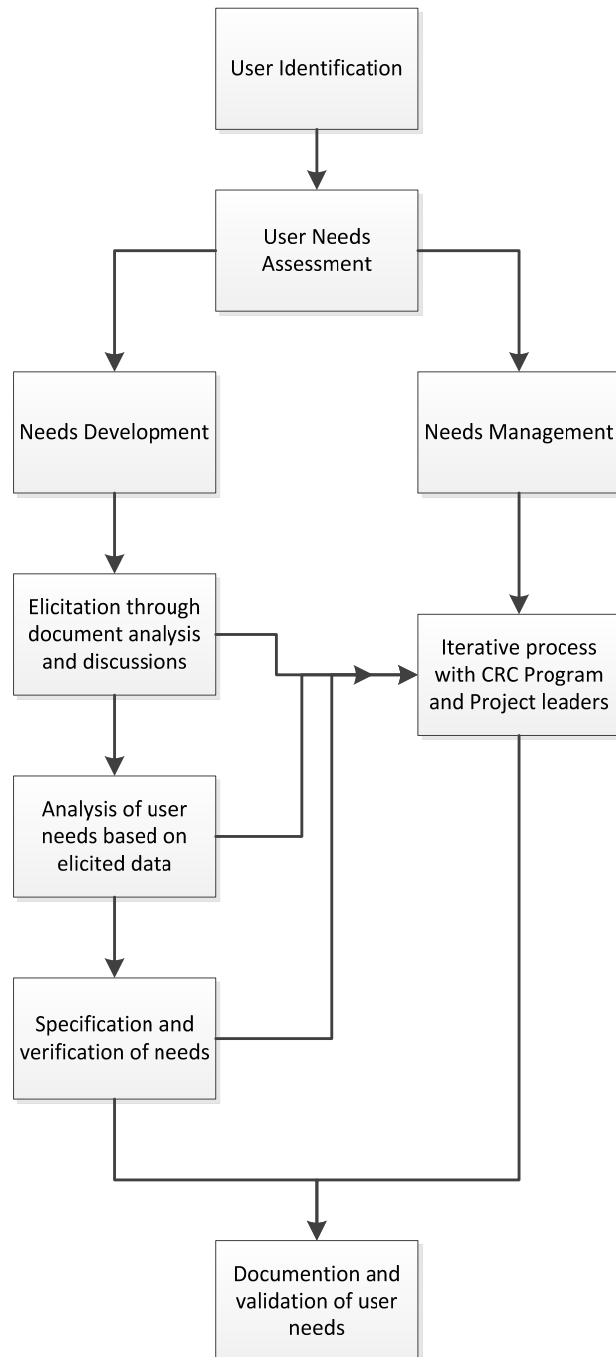


Figure 2 User Needs Assessment Workflow

4 Findings from Data Analysis and Discussions with Project Leaders

4.1 Identification and Analysis of Program and Project Partnerships

Following the initial project refinement phase, identification of the key project partners in the first two rounds of Programs 2 and 3 was carried out based on the documents provided by CRC LCL and through discussions with CRC LCL leadership¹. Figure 3 illustrates a 2-mode network of projects and partners in Programs 2 and 3 in June 2013. In Figure 3 the blue nodes indicate project partners and red nodes indicate the projects. The size of the nodes indicates the number of connections it has (e.g. a larger blue node indicates that researcher is involved with a larger pool of projects and similarly the larger the size of the red node indicates a larger number of researchers are indicated as project partners based on the documents provided). The turquoise link between a partner and a project indicates that researcher is the project leader of that project and the red links indicate that the researcher is a partner on the project. Mapping out this information was carried out to understand the involvement and interaction of researcher in various projects and identifying the key players in the first two rounds of funded projects in programs 2 and 3 of CRC LCL.

¹ Note that access to Program 1 project documentation was not possible. An initial assessment of Program 1 summaries suggests that at this stage, projects in this area will not generate outcomes with explicit spatial attributes. This can be revisited in the light of subsequent funding rounds.

A LOW CARBON LIVING SPATIAL DATA HUB

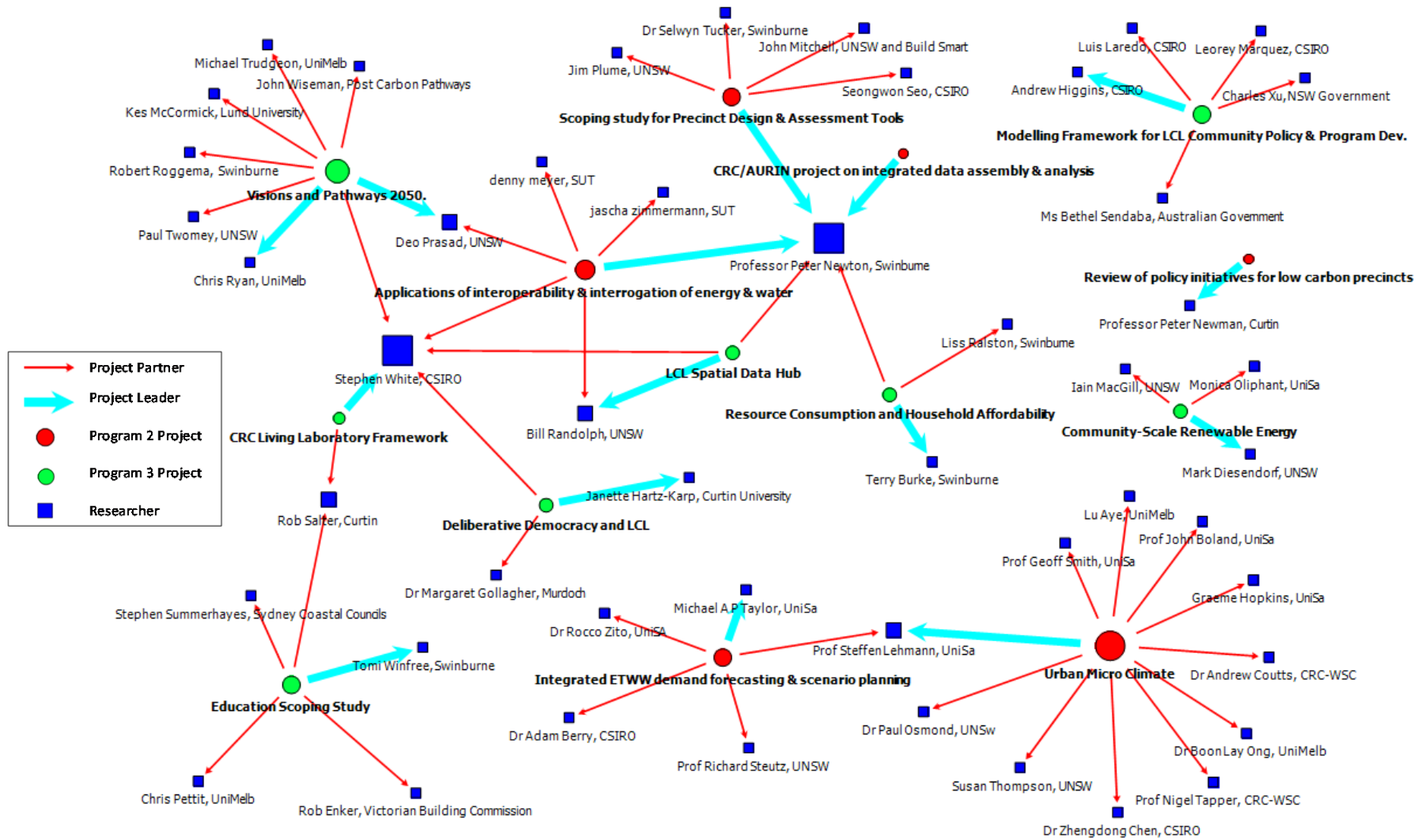


Figure 3 2-Mode Network of Projects and Partners in Program 2 and 3 (June 2013)

A LOW CARBON LIVING SPATIAL DATA HUB

4.2 User Needs Assessment

Following the identification of project partners, detailed document analysis based on the project briefs and proposals was carried out and all of the project leaders in Programs 2 and 3 were contacted in June and July 2013. The aim of the communication with project partners was to understand their needs and establishing the requirements for the development of an integrated building level and cadastrally organised Spatial Data Hub for CRC LCL. Moreover, whether each project had spatial data inputs/outputs was explicitly established and discussions were held to understand whether key researchers anticipated potential future projects during the life time of CRC that might have spatial components. Table 1 provides an overview of the contacted project leader/participants in Programs 2 and 3.

A LOW CARBON LIVING SPATIAL DATA HUB

Table 1 Overview of the contacted project leader/participants

Project	Program Leader	Contact	Mode(s) of Contact		
			Email	Phone	Meeting
Resource Consumption and Household Affordability	Terry Burke, Swinburne University	tburke@swin.edu.au	X	X	
A Modelling Framework for Low Carbon Living Community Policy and Program Development	Andrew Higgins, CSIRO	andrew.higgins@csiro.au	X	X	
Deliberative Democracy and Low Carbon Living	Janette Hartz-Karp, Curtin University	j.hartz-karp@curtin.edu.au	X	X	
CRC Living Laboratory Framework	Stephen White, CSIRO	stephen.d.white@csiro.au	X	X	X
Education Scoping Study	Tomi Winfree, Swinburne University	twinfree@swin.edu.au	X	X	
Opportunities and challenges for the development and implementation of community-scale renewable energy projects	Mark Disendorf, UNSW	m.diesendorf@unsw.edu.au		X	X
Visions and Pathways 2050	Chris Ryan, University of Melbourne	cryan@unimelb.edu.au	X	X	
	Deo Prasad, UNSW	d.prasad@unsw.edu.au	X		X
Integrated ETWW demand forecasting and scenario planning for precincts	Michael A P Taylor, University of South Australia	michael.taylor@unisa.edu.au	X	X	
	Nicholas Holyoak, University of South Australia	nicholas.holyoak@unisa.edu.au	X	X	
Urban Micro Climate	Prof Steffen Lehmann, University of South Australia	steffen.lehmann@unisa.edu.au	X		
	Dr Paul Osmond	p.osmond@unsw.edu.au			X
Review of national and international policy initiatives and programs for low carbon precincts	Professor Peter Newman, Curtin University	p.newman@curtin.edu.au	X		
Scoping study for Precinct Design and Assessment Tools	Professor Peter Newton, Swinburne University	pnewton@swin.edu.au	X	X	
Innovative applications of interoperability and interrogation of energy and water					
Joint CRC/AURIN project on integrated data assembly and analysis					

A LOW CARBON LIVING SPATIAL DATA HUB

The summaries of conversations with project leaders/partners in Phase 1 and 2 of Programs 2 and 3 are presented below:

RP 2002 Integrated Energy, Transport, Water and Waste (ETWW) demand forecasting and scenario planning for precincts, Michael A P Taylor and Nicholas Holyoak, University of South Australia.

The major research initiative in Program 2 is concerned with PIM and is proposed as the underpinning data platform for all of the Program 2 research. Major aspect of PIM research in program 2 is on identification, definition and specification of precinct-based objects and their representation.

RP 2002 has strong usage of spatial data in both inputs and outputs. This project focuses on the analysis of energy, travel, water and waste as domains and their interactions in precinct demand forecasting. Newton et al. (2013) suggest that this project linked to potential Living Laboratories can act as a test-bed for checking the forecasting approaches developed and can be used for comparison of the existing precinct assessment tools.

The primary input data are aggregated (SA1 or Mesh Block level) ABS census data on population and housing. Data on employment, education and commercial/retail activities will also be used at precinct level. Within the precincts the data would include information in regards to layout of buildings, facilities, their usage and their spatial connections. Due to its nature this project has interest in infrastructure networks and their components. This includes, but is not limited to, data on location, types and services of:

- transport networks (e.g. locations of bus and tram/LRT stops and railway stations in the vicinity of a precinct, access arrangements and car parks).

- energy supply networks,
- water supply networks,
- waste removal networks, and
- telecommunications networks (for future travel demand forecasting involving travel substitution).

Further formal determination of the common and domain specific datasets will be carried out at the later stages of the project

RP 2003 Review of national and international policy initiatives and programs for low carbon precincts, Peter Newman, Curtin University

Data is currently part of the CRC SI and it is possible that arrangements can be made for these data to be incorporated into the CRC LCL at a later date.

RP 2005 Urban Micro Climate, Steffen Lehmann, University of South Australia, Paul Osmond, UNSW.

This project focuses on the development of a cross-disciplinary, multi-scale understanding of Australian cities' micro climates and will compare the built environment in three major Australian cities - Sydney, Melbourne and Adelaide. It will focus on the interplay between urban form, density, surfaces and ambient temperature and Urban Heat Island effect.

The project will focus on environmental management, decision making and urban policy and planning issues. It will operate at building, precinct and city scales and it is certain that spatial data will be used as input and output at different scales. The exact detail of the data is not clear yet as the project is in early stages and the details will be worked out in future workshops among project participants.

RP 2001 Scoping study for Precinct Design and Assessment Tools, RP 2004 Innovative applications of interoperability and interrogation of energy and water, RP 2007 Joint CRC/AURIN project on integrated data assembly and analysis, Peter Newton, Swinburne University.

Prof. Newton is currently involved with PIM and the development of systems to assist in precinct data assessment, visualisation and modelling with the CRC for Low Carbon Living. Furthermore, Prof. Newton is the program leader of the Water and Energy lens in AURIN.

Program 2 in general has a very strong spatial focus in terms of data inputs and outputs as the focus of the program is on the PIM concept and spatial. RP 2001 Scoping study highlights the spatial data required for PIM and outlines the features of and requirements for PIM (Newton et al, 2013). Prof. Newton's vision for precinct based Living Laboratories envisages longitudinal analysis of communities with precincts and view's precinct based Living Laboratories and the Data Hub as a legacy for CRC LCL.

RP 3001 Resource Consumption and Household Affordability, Terry Burke, Swinburne University.

The current project does not have any spatial data inputs. The outcome of the research will have relevancy at city scale. Currently there are no anticipated future projects that have spatial components.

RP3002 A Modelling Framework for Low Carbon Living Community Policy and Program Development, Andrew Higgins, CSIRO.

This project focuses on forecasting the uptake of retrofit packages in commercial buildings under a range of government policies. The Agent-based model simulates adoption of retrofit with a typology of representative buildings and scales up to the office building stock across the Local Government Area (LGA) and across the state of NSW. The spatial granularity of the data is LGA and the commercial building stock data is provided by Geosciences Australia. Other data sets that may be used in the future include, for example, the actual uptakes of each policy program, such as a list of buildings that successfully submitted to the Green Building Fund. The data currently used is mostly contained in spreadsheets, rather than being geocoded.

RP 3004 Deliberative Democracy and Low Carbon Living, Janette Hartz-Karp, Curtin University

Email and Phone conversation. No relevant spatial data to date.

RP 3005 CRC Living Laboratory Framework, Stephen White, CSIRO

This project focuses on the establishment of the CRC LCL's Living Laboratories. Living Laboratories' goal is to test theories and innovative low carbon technologies/solutions in communities, and observing and measuring their impact. As such the project does not have spatial components, but since Living Laboratories will be established in communities dispersed in different neighbourhoods/precincts across Australia there will be a strong spatial component to these Laboratories and a suite of datasets and indicators need to be developed to capture common and unique aspects of the Living Laboratories that can often be influenced by the location of the communities under study (See Section 4.3 for more detail).

RP 3006 Education Scoping Study, Tomi Winfree, Swinburne University.

Email and Phone conversation. No relevant spatial data to date.

RP3007: Opportunities and challenges for the development and implementation of community-scale renewable energy projects - CRC for Low Carbon Living, A/Prof Mark Disendorf, UNSW.

This project currently focuses on how community-based renewable energy projects can be developed in Australia with focus on institutional structures, governance processes and sources of finance. It will inform and provide information guidance for different stakeholders (communities, utility companies, local government, etc.) to develop renewable energy projects.

It is natural to expect that once deliverables are ready the next step for developing community-based renewable energy projects is to identify suitable locations across Australia for development of such projects through spatial analysis and simulation using the outcomes of this project and other relevant datasets.

RP 3008 Visions and Pathways 2050, Chris Ryan, University of Melbourne

This project focuses on exploring visions, scenarios and policy pathways for a low carbon built environment. Initially the project will focus on engagement strategies to gather stakeholder views of desirable future possibilities for Australian cities. Then scenarios and transition pathways will be developed for overcoming barriers and better managing the complexities of the built environment.

It is anticipated that in later stages of this project once the visions have been developed, in order to establish baselines and develop scenarios different types of datasets relating to

material and energy usage, technology, socio-economic and environmental aspects will be used. Furthermore, in future this project will allow spatial considerations for development of policy measures and governance structures for achieving these visions and conducting high level spatial analysis will become possible.

4.3 Living Laboratories

White and Salter (2013) define a ‘Living Laboratory’ as “an organisational arrangement, where the impact of introducing a change process or a new product/service (intervention), can be monitored and observed in a real world community with diverse stakeholders.”

The CRC LCL is committed to creating a minimum of nine living laboratories. We envisage that the researchers of these Living Laboratories will be conducting experiments/monitoring changes and will generate various kinds of data streams. These data streams can be used for populating databases, enhancing the Data Hub, maintaining longitudinal data. Thus the CRC LCL Data Hub could provide a central storage and analysis capabilities for all of the Living Laboratories. It will allow researchers of these Living Laboratories and other members of CRC to better understand the experiments being conducted, collaborate, built upon previous research and conduct long-term longitudinal studies in the areas that have been designated as Living Laboratories.

Furthermore, by concentrating the projects in a limited number of specific geographical locations defined by CRC researchers (typical suburbs and neighbourhoods at precinct level) and using the data address level cadastre it becomes possible to link to household and building scale data (socio-demographic, energy, water, transportation, waste etc.) and integrate them with precinct level data (e.g. ABS SA1 Data) and infrastructure information through an integrated

database. This provides the possibility to use the ABS SA1 data to form some of the required information or perhaps use special information commission from ABS (at cost) to define certain precinct locations and amalgamate the census data to address level data into those precinct areas.

This approach allows creating a repository of datasets, studies, interviews for future assessments (e.g. using survey for households or businesses) to test out different scenarios for education, material and energy consumption, transportation, adaptation and adoption of new technologies, climate change and resilience (e.g. urban heat island), or demographics and economic outcomes. It will enable examining whether particular targets can be reached in with a given set of conditions and measure the impact of different policies/strategies taken by government or businesses. If a well thought structure is utilized for capturing the data the value of this information intensifies over time as more data is accumulated.

Moreover, if the locations are chosen to be sufficiently different based on socio-demographic parameters (income, age, population density etc.) or other defining factors (once the Living Laboratories are running and the data has been stored in the Data Hub) it would be possible to generalize upwards and be able to deduce information at higher geographical scales. By taking advanced spatial modelling approaches, using census data, monitoring the Living Laboratories and conducting interviews and surveys and deploying the projects in the same space the possibility might arise to identify interesting cross connections and perform longitudinal studies that would otherwise not be possible.

A complementary approach, proposed by White and Salter (2013), would be to accept all communities that are interested in becoming part of the Living Laboratories scheme in the CRC in order to increase the diversity and number of the laboratories given the limited resources of CRC. As this approach will result in a diverse range of dispersed Living Laboratories in which

CRC will have a limited control, an alternative approach would need to be taken. Aside from the specific datasets and indicators that are likely and relevant for specific projects and locations, perhaps CRC LCL could demand a core set of datasets and indicators from these Living Laboratories as a condition for acceptance to CRC to make it possible to conduct a uniform set of base assessments. This can also facilitate the comparison of the experiments and observations of these diverse Living Laboratories with the designated and predefined neighbourhood/precinct Living Laboratories to test whether outcomes are replicable in different geographical spaces, contexts, and timescales.

Establishing the Data Hub will allow the CRC to have a high impact and tangible effect on research on low carbon living after CRC LCL's funding expires. This legacy is crucial as the progress towards low carbon outcomes will certainly not be resolved within the next six years. Having the Data Hub as a base of long term longitudinal projects that could continue beyond this time will be an exceptional contribution from CRC to Australian built environment research community along with providing world class urban data.

To best capture value from emerging Living Laboratories would, in our view, entail establishing a CRC LCL Data Hub as soon as possible. At one level it would support the co-development of PIM and Living Laboratory concepts and having dedicated personnel to monitor the outputs of the projects and work with researchers to add information to the Data Hub would add significant value from the initiative. Similarly, White and Salter (2013) suggest a Living Laboratories Co-ordinator as part of the Living Laboratories team which can take the role specified above. This position could either be added to the CRC's core office team or located in a partner organisation with existing capabilities in this area.

4.4 A Proposed Data Structure

Table 2 highlights our findings in regards to the current and future data usage/needs of projects funded in rounds 1 and 2 of Programs 2 and 3, based on the analysis of project documents and discussions with project partners. It is evident that a large number of the current funded projects use spatial data, are interested in accessing or are likely to generate spatial information that would be recorded at a variety of spatial scales. As indicated in Table 2, many of the current projects have potential for follow up projects that are likely to have increase spatial data use and outputs. This is especially true if the Data Hub becomes increasingly populated by CRC data and links to wider spatial data eResearch infrastructures are developed further.

Furthermore, in future there will be an increase in the number of projects with spatial components as CRC LCL matures and moves from initial pilot and scoping studies phase to more specific place-based research in Australia. By way of a preliminary estimate of such future use, we have appended an illustrative assessment of the project approved by the CRC subsequent to the research on this project being developed in Programs 2 and 3 (Appendix A.3 and A.4).

Figure 4 highlights a number of databases and data types identified through analysis of project documents and discussions with project partners that were identified for this project. They are categorised as building, precinct or city scale depending on their nature and granularity. The Table therefore summarises the likely range of data that the Data Hub might be expected to acquire in the first iterations of project outcomes over the next two years.

A LOW CARBON LIVING SPATIAL DATA HUB

Table 2 Spatial Data in Funded Projects in Programs 2 and 3

Project	Current Spatial Data			Future Potential of Follow up Projects		
	Building	Precinct	City	Building	Precinct	City
Resource Consumption and Household Affordability , Terry Burke, Swinburne University			X		?	X
A Modelling Framework for Low Carbon Living Community Policy and Program Development , Andrew Higgins, CSIRO	X	X	X	X	X	X
Low Carbon Living Spatial Data Hub , Bill Randolph, UNSW				X	X	X
Deliberative Democracy and Low Carbon Living , Janette Hartz-Karp, Curtin University				?		
CRC Living Laboratory Framework , Stephen White, CSIRO				?	?	?
Education Scoping Study , Tomi Winfree, Swinburne University				?		
Opportunities and challenges for the development and implementation of community-scale renewable energy projects , Mark Disendorf, UNSW					X	?
Visions and Pathways 2050 , Chris Ryan, University of Melbourne, Deo Prasad, UNSW			X		?	X
Integrated ETWW demand forecasting and scenario planning for precincts , Michael A P Taylor and Nicholas Holyoak University of South Australia	X	X		X	X	Likely
Urban Micro Climate , Steffen Lehmann, University of South Australia, Paul Osmond, UNSW	X	X	X	X	X	X
Review of national and international policy initiatives and programs for low carbon precincts , Peter Newman, Curtin University					?	?
Scoping study for Precinct Design and Assessment Tools; Innovative applications of interoperability and interrogation of energy and water; Joint CRC/AURIN project on integrated data assembly and analysis , Peter Newton, Swinburne University	X	X		X	X	Likely

A LOW CARBON LIVING SPATIAL DATA HUB

Building Scale	Precinct Scale	City Scale
BEEC register, green star, basix	ABS Census Collection District - Demographics and housing Stock	ABS Household Expenditure Surveys - Demographics
NABERS/API/RP Data, Property Council of Australia	Building Stock - LGA or ABS Census Collection District	Key infrastructure (transport, energy grid, ICT)
Ausgrid metering data	Estimation of level of emission reduction	Effect of behaviour change of household on demands
PCA market reports / Geosciences	Housing data for Sydney	Integrated energy, travel, water, waste estimation
IPD Green cities and APD building better returns	Temperature	Co ₂ emission at household level
Socio demographic and socio psychographic data	Rainfall	Urban density
Owners and tenants info	Humidity	Land use pattern
Commercial building stock and usage	Sunlight	City scale scenarios
Waste and material flow (per capitq, recycle/reuse rate, recovery rate, consumption patterns, supply chain and disposal)	Forecast if different stock of technologies (pc, appliance) at locations with high granularity	Regional datasets (Ergon, Twonsville, James Cook, qcif, lbm smart cities...)
Household/business energy consumption	Wind patterns	
3d building data	Evaporation and comfort	
Property valuation and characteristics	Aggregated data to ABS Mesh level	
Economic variables	Community-based renewable projects	
Albedo and facades	Relevant data holdings at precinct level	
Water demand	Relevant CRC databases (LCI, GIS, WSC)	
Water efficiency rates	Relevant Aurin databases to be identified	
Water recycling rate	Population growth estimation	
Electricity generation	Urban development patterns	
EV data	Industry sectors	
PV installation data from ORER and HW	Urban heat island effect	
Gas data	Public space	
Consumer info	Urban greenery	
Thermal photography	Landscape feature	
Roof space	User behaviour	
Building occupancy day/night	Supply nodes	
Household/business transport usage (total trips, destinations, forms of transport, timing, routes)	Effects of various factors on micro climate (land use, landscape features, greenery, effect of materials and building envelopes)	
Tax breaks for green buildings scoping study and RIS and buildings baseline report		
Building location, structure, adaptation, and performance		
Surface material		

Figure 4 Potential input/outputs data at different scales

A LOW CARBON LIVING SPATIAL DATA HUB

As Figure 4 indicates a large number of datasets are or have the potential to be used during the lifetime of the CRC LCL. A major hurdle faced by the researchers is the challenge of acquiring detailed embodied energy and utilities data due to issues surrounding privacy and commercial activity. Moreover if/when data is gathered, integration of the data from different datasets has various challenges relating to differences and difficulties in matching the scales, quality and time horizons used in collection of the data. Aside from challenges for simulation and modelling activities, in the context of CRC LCL such issues will be highlighted in the development and use of precinct assessment tools and comparison of different living laboratories.

Accurately illustrating the interactions of the various datasets that can be potentially used in the Data Hub for the purposes of this paper is not possible. Instead, Figure 5 depicts the anticipated indicators that can be used and their interactions at different spatial scales (each indicator uses one or more datasets and might rely on data from a number of other indicators). For instance, the development of environmental indicators at precinct scale might rely on demographic information derived from dwelling level information as well as on information such as temperature, duration of sunlight, wind patterns, rainfall and humidity which are often only available at precinct/census tract scale, combined with usage data for transportation, energy, water and water (which might be available at household/building scale but are more likely to be only accessible at a higher spatial scale) as well as high level information in regards to climate and heat island effects at city scale.

Figure 5 is not intended to be exhaustive, but is intended to provide an indicative schema of the range of key datasets that a CRCLCL Data Hub will need to include as the CRC develops. Critically, this range of data will build over time to an unparalleled resource which CRC

researchers can access for further secondary research. A key ability will be to draw down data for use in individual research projects and to return new data back to the hub. This kind of data structure will become increasingly important if, for example, additional program areas are incorporated into the CRC, such as urban infrastructure.

A LOW CARBON LIVING SPATIAL DATA HUB

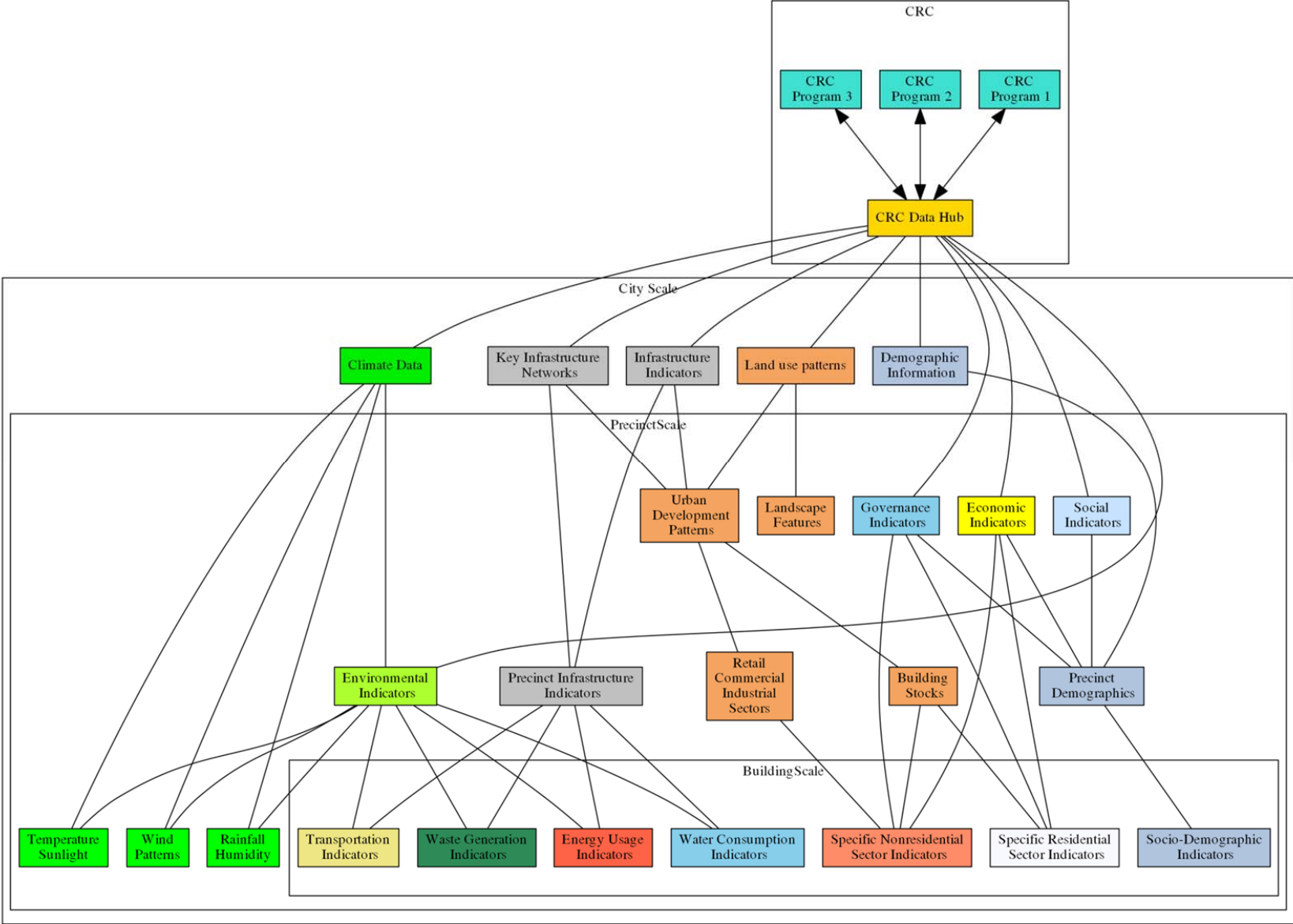


Figure 5 Indicators at different spatial scales (each requiring specific datasets (not shown) and data inputs from other indicators)

5 Proposed Specifications for the Data Hub

The findings of this scoping study indicate that the CRC Data Hub needs to:

- Increase accessibility and act as the main repository of data for CRC;
- Host the different types and scales of data used for urban studies;
- Provide secure access rights to CRC members and other select organizations in case-by-case basis;
- Provide a large storage space to host as the spatial data such as satellite and aerial imagery, spatial surveys, life cycle inventory databases, and core indicators and data for assessment in living laboratories (e.g. real-time monitoring of energy use) and guarantee the integrity of the data;
- Accommodate a relatively large number of users, facilitate collaboration, and be accessible to non-technical experts;
- Connect to large federated or distributed portals and datasets such as AURIN;
- Support the integration of BIM and GIS systems through PIM;
- Be able to accommodate development of visualization tool such a precinct viewer or decision support systems that will be used for precinct assessment and scenario analysis such as eco-efficiency assessment of distributed energy generation technology options in different geographical locations (Newton *et al*, 2013).

Based on the findings from the user need assessment, four alternative specifications were considered for CRC LCL Data Hub. The following section sets out the options analysis undertaken by Intersect that provide a business case to support consideration of the most appropriate approach to the development of the LCL Data Hub.

5.1 Defining the Data Hub

The Data Hub proposed will provide a centrally managed data storage facility for all datasets generated as part of CRC LCL activities.

It is anticipated that a large number of projects associated with the LCL will generate datasets suitable for storage with the Data Hub. These datasets are envisaged to total 100 terabytes in size. The number of users of the Data Hub is estimated to be 500 in total number; however, cost estimates were also provided also cover 50 and 5,000 users.

The Data Hub proposed has been specified with enough storage capacity to cater for all datasets generated over the CRC's six-year time frame, with options to expand that capacity should the need arise in the future. Nothing about the way in which the Data Hub has been specified will preclude certain types of datasets from being added to it. The analysis features of the Data Hub have been specified to provide sufficient performance capability to handle 50, 500 and 5,000 users.

The Data Hub is specified to be a long-term solution for the management of LCL data. Additional funding will see the Data Hub extend 6-10 years beyond the initial six-year time frame of the CRC.

5.2 Options Overview and Rationale

Four possible options are provided for the LCL Data Hub. The four options are:

5. Basic Data Hub – provides storage infrastructure and support
6. AURIN-Enabled Data Hub – adds an analysis capability in the form of integration with the AURIN portal
7. Extended AURIN-Enabled Data Hub – extends the analysis capability of the LCL Data Hub

8. Enhanced Spatial Analysis Hub – provides a fully customised, advanced analysis capability for the LCL Data Hub.

Within each option, one or more features are defined. In some cases, all features are obligatory for an option to offer a viable solution. In other cases, some features are optional. Each option builds upon the previous one and in some cases, includes features from one or more previous options. Details of the possible combination of options and features are provided in the sections below. The reason for specifying an array of options and features is to allow for the optimal combination of cost and functionality level to be achieved (see Trade-off Analysis section for more details).

5.3 Cost Estimates Overview

The costs stated below are estimates only. While every effort has been made to take into account all potential costs incurred by the establishment and continuing operation of the Data Hub, and to estimate them accurately, the cost estimates are made in the context of several unknown factors, including a lack of detailed user requirements, unknown future costs for computer hardware, unknown future costs for the use of eResearch infrastructure, possible changes to software licence costs, and other issues that may come to light in the future. Uncertainty also exists around the total number of users, total size of the datasets to be hosted by the Data Hub, and the analysis features required by researchers. Assumptions have been made about these requirements, as outlined in other parts of this document, however, as the details become clearer, the cost estimates presented will need to be revised.

Given the combination of a relatively long project time of 6 years and long horizon of 6-10 years, it is difficult to anticipate the change to computer hardware costs. Whereas storage infrastructure costs will almost definitely reduce over time, product end-of-life time

frames will lead to the requirement to replace hardware approximately every 3 to 5 years and data centre hosting costs are likely to increase.

All storage related cost estimates are calculated assuming a base of 100 terabytes. Within reasonable limits, it is possible to calculate additional storage costs by using the per-terabyte estimates provided.

Some features of the Data Hub refer to the potential use of existing eResearch infrastructure as alternatives to self-setup options. eResearch infrastructure mentioned are the NSW RDSI node for storage and the NeCTAR Research Cloud for compute. Both projects are yet to establish a cost model to sustain the infrastructure beyond the end of 2014. In both cases, cost estimates are based on comparable commercial services.

Note that cost estimates for personnel are based on the Intersect membership cost-recovery rate of \$600 (ex GST) per day per person. This cost is projected by assuming a 5% per annum increase. One FTE is calculated as 220 days in a year, which takes into account weekends, public holidays, annual leave and sick leave. No buffer has been built into the estimates for personnel; rather, we have provided them as an estimate of what we expect the project to cost, based on our experience of similar efforts.

Note that all costs have been estimated ex GST, unless indicated otherwise.

5.4 Data Hub Options

5.4.1 Option 1 – Basic Data Hub

The Basic Data Hub option represents the minimum infrastructure and support required to provide a long-term storage solution for the CRC LCL project data (see Figure 6). It assumes a total storage capacity of 100 terabytes. There are options to expand this capacity, and cost estimates are provided for this.

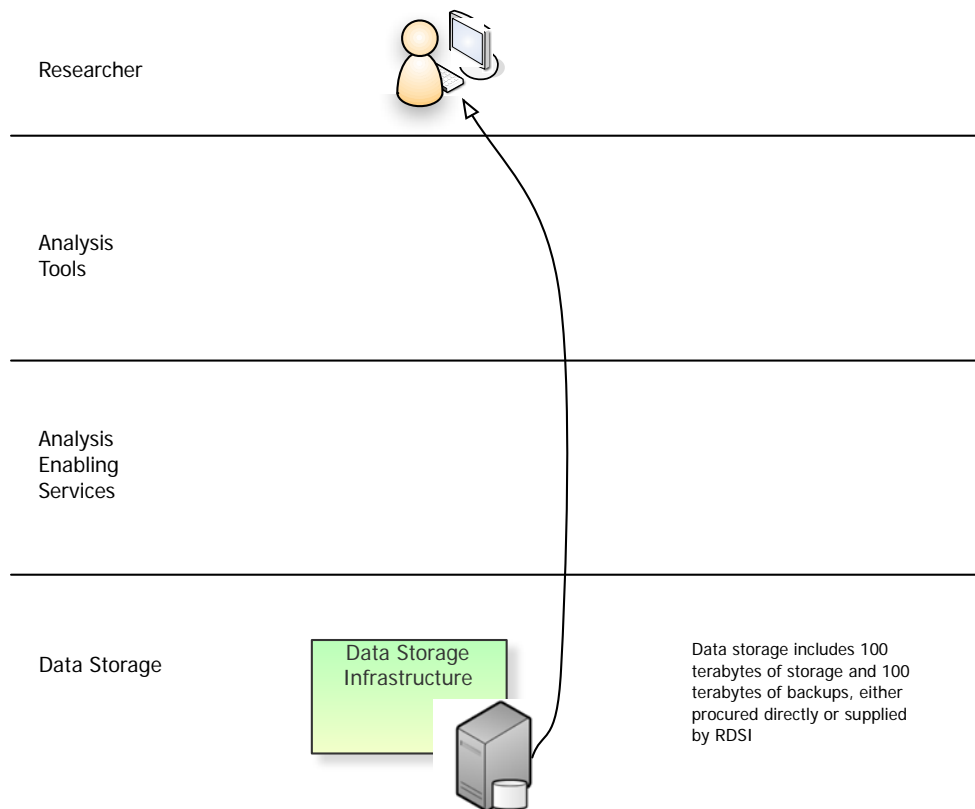


Figure 6 Option 1: Basic Data Hub

This option provides raw access to the data, but will not offer any services required to access the data via external geospatial analysis platforms such as ArcGIS or the AURIN portal.

Strengths:

- The minimum possible cost to deliver a long-term storage solution for the CRC LCL.

Weaknesses:

- Access to the data will be primitive and not well suited to allowing analysis of the data.
- It will be difficult to compare data in the Data Hub with broader national datasets.
- Does not include an analysis capability, relying instead on analysis tools external to the Data Hub.

Total Cost:

Depending on the various alternatives chosen (as detailed below), the total cost of this option ranges from approximately \$365,000 to \$635,000. All of the features listed below are obligatory for this option to be a viable solution.

5.4.1.1 Feature 1.1 - Data Storage

Sufficient data storage is required to house datasets for all CRC LCL projects that have data to contribute to the Data Hub. It is anticipated that a significant number of CRC LCL projects will have datasets that are relevant to the Data Hub, although the number of datasets contributed by each project will vary considerably. Some projects will contribute just 1-2 datasets, whereas others may contribute more than 10 datasets. The size of each dataset is also anticipated to vary considerably. The types of datasets that will be hosted by the Data Hub and their indicative size ranges, per dataset, are as follows:

- Survey and other tabular datasets – 10 to 100 megabytes
- Spatial data and time-series data – 100 megabytes to 100 gigabytes
- Satellite imagery, aerial photography and Lidar data – 100 gigabytes to 5 terabytes

To illustrate the sizes of some of the larger datasets, one example is Ausgrid's phasor dataset, which is in excess of 2 terabytes. The cost estimates for this feature are based on the assumption that 100 terabytes of storage will meet the needs of the Data Hub. Per terabyte costs are provided below to allow approximate extrapolation of costs for additional storage.

The various alternatives for housing the CRC LCL datasets are:

1. Self-managed storage system hosted at UNSW (approx. \$150,000)
2. Self-managed storage system hosted at a commercial data centre (approx. \$390,000)

3. Utilise RDSI storage infrastructure (approx. \$330,000)

Note: Each cost estimate in the list above is for the six-year period.

The first option above assumes that by hosting the storage infrastructure at UNSW, no cost for any of the services typically provided by a commercial data centre (i.e. space, installation, electricity, network usage, and basic system monitoring) will be incurred by the project. This has yet to be confirmed with UNSW Central IT, however, at the time of writing, this is our understanding based on discussions with the UNSW Faculty of the Built Environment Computing Unit.

Data hosting the infrastructure at a commercial data centre (included in the cost for alternative 2, above) is estimated at \$240,000 over the six-year time frame. This assumes all required equipment consumes 5 kW of power. Note that the power consumption will increase as the disk capacity increases in a roughly linear fashion (e.g. by adding an additional 100 terabytes of storage, the power consumption will roughly double). This will increase the costs of data hosting in a similar fashion; however, the exact increase in cost will depend on the cost model of the chosen data centre.

Alternatives 1 and 2 above (i.e. self-managed storage) include costs to procure disk storage, backup infrastructure (sufficient for a single copy of all data hosted by the Data Hub) and supporting equipment (e.g. network switch, server rack, cabling, etc.). The cost for this infrastructure is approximately \$670 per terabyte over a six-year period. This cost takes into account the requirement to procure replacement hardware a second time during the six-year project. This is based on the typical end-of-life time frame for IT infrastructure being 3 to 5 years.

Also included in the cost estimates for alternatives 1 and 2 is support and hardware maintenance. This usually takes the form of a maintenance service contact with the original

supplier of the hardware. Depending on the level of service required, this is likely to cost 20% of the original purchase cost of the hardware per year.

The third option leverages eResearch data storage infrastructure in the form of RDSI (Research Data Storage Infrastructure). Provided the CRC LCL datasets qualify, storing data with RDSI will not incur any costs to the Data Hub until after the end of 2014. However, as the cost-recovery model for RDSI post-2014 has yet to be determined, estimates have been used for this proposal to generate the figure of \$330,000 quoted above.

Cost:

Based on the three alternatives shown above, the cost estimate for data storage for the Data Hub ranges from approximately \$150,000 to \$390,000 for the six-year time frame.

5.4.1.2 Feature 1.2 - Data Preparation

To maximise the benefits that can be drawn from a dataset added to the Data Hub, a certain amount of preparation will be required prior to its inclusion. The nature and amount of preparation will vary but, in general, the following will be required:

- Data Cleaning – This will include fixing typographical errors, inconsistencies and invalid data; removing incorrect, incomplete or irrelevant records; removing duplications.
- Convert Data to Common Spatial Levels and Formats – It is a requirement of the Data Hub that datasets are integrated spatially, organised ideally at the property/building level or land use cadastre level. This will include geocoding of some datasets. This will enable PIM (Precinct Information Modelling), BIM (Building Information Modelling) and GIS analysis as required by users of the Data Hub. At the same time, it will be beneficial to reduce the disparity in data formats across the datasets by converting them to an agreed set of common formats. Done once in the preparation of datasets prior to inclusion in the Data

Hub, this will reduce the time taken by individual researchers having to convert the file formats themselves. For some files, such as survey data, tabular data and time-series, this may involve ingesting the data into a database.

- Establish Data Preparation Workflows – Some datasets included in the Data Hub will be periodically updated. In the case of those datasets, it will be more efficient to establish a repeatable data cleaning and conversion workflow that can be performed across the dataset each time a new version is received.

Cost:

Where possible, data cleaning and conversion activities should fall back to the original data custodian. Some administration work will be required in gathering datasets, rejecting ‘dirty’ data and co-ordinating format conversions. Even so, the Data Hub will be required to perform some data preparation activities and, in particular, will need to establish data preparation workflows.

Based on data preparation activities being required for 50-100 datasets, it is estimated that one FTE will be required to spend 3-6 months on the three tasks outlined above – approximately \$35,000-\$65,000.

Note that this estimate only covers the effort required to prepare the initial set of data for the Data Hub. Ongoing data preparation is included as part of the next feature, Feature 1.3 - Operations.

5.4.1.3 Feature 1.3 - Operations

Operations include any tasks that support the day-to-day running of the Data Hub.

This will include:

- Continuing dataset preparation and migration – as CRC projects contribute new datasets, technical staff will need to add these new datasets. This includes all activities outlined above in Feature 1.2 – Data Preparation.
- Monitoring, evaluation and reporting – gathering usage statistics, producing annual reports for the LCL (for usage patterns for the Data Hub) and making changes to the Data Hub to maximise its benefit to users.
- Business continuity activities – ensuring that the Data Hub is able to maintain a consistent service and is in a position to recover from threats to the service, either planned (e.g. upgrades, backup, restore) or unplanned (e.g. hardware failure, human error, security attack).

Cost:

It is anticipated that technical staff will be required to undertake the operational tasks listed above. Wherever possible, automated systems and software should be used (e.g. for monitoring systems, statistics gathering and reporting). As an estimate, one person, one day per week is required for the six-year time frame – approximately \$180,000.

5.4.2 Option 2 – AURIN-Enabled Data Hub

The AURIN-Enabled Data Hub option provides the data storage infrastructure and support as outlined in Option 1 and extends it with the infrastructure and services required to make the datasets available for analysis via the AURIN portal (see Figure 7). It includes the provisioning of a GIS server, which will host the chosen GIS server application software. It also includes establishing a GIS database and enabling the appropriate service (e.g. Web Feature Service) to provide a data feed to the AURIN portal.

Strengths:

- Introduces an analysis capability to the Data Hub, in the form of the AURIN portal.
- Provides researchers with access to the compute power offered by the AURIN portal.
- Provides the capability to process LCL data alongside other national datasets.
- Enhances the discoverability of data hosted by the Data Hub.
- Provides a familiar interface to researchers used to the AURIN interface and its tools.

Weaknesses:

- Analysis capability is limited to that offered by the AURIN portal.
- More expensive option than the Basic Data Hub.

Total Costs:

Depending on the various alternatives chosen (as detailed below), the total cost of this option ranges from approximately \$565,000 to \$815,000. Note that this range includes the cost of all three features from Option 1, and is based on the number of users being 500 (for the cost breakdown for 50 and 5,000 users, see details below). For this option to be a viable solution, both features listed below are obligatory.

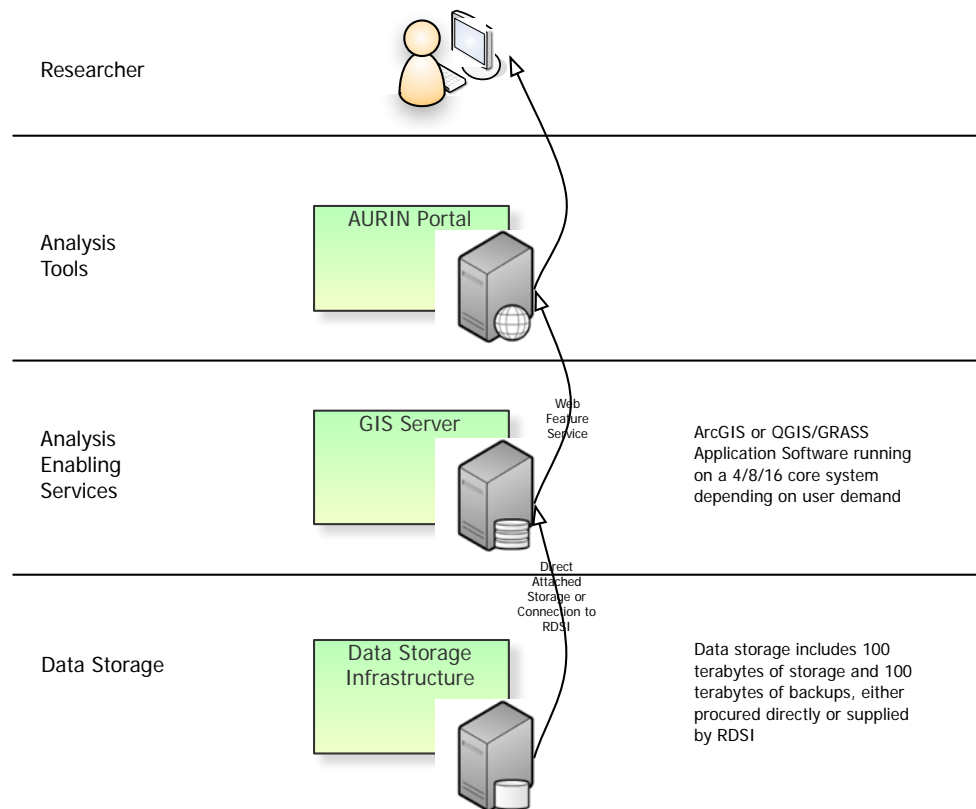


Figure 7 Option 2: AURIN-Enabled Data Hub

5.4.2.1 Feature 2.1 - GIS Server

In order to integrate the Data Hub with the AURIN portal, it is required that a GIS server be established. This encompasses an application server and accompanying software, which together respond to requests made by users of AURIN to access data hosted by the Data Hub.

There are many factors that will affect the level of performance required of the GIS server hardware, however, the most significant factor will be the number of simultaneous users and the types of analysis those users are performing. As a rough guide, a single 8-core machine utilising the Intel Xeon E5-2690 chipset and running at 2.90 GHz will support approximately 200 simultaneous users of ArcGIS Server². It is assumed that by provisioning a 16-core system to support 400 simultaneous users, this will be sufficient for a total users base of 5,000 users.

In terms of provisioning the GIS server hardware, there are two alternatives. The first option is to procure a system for the project; the second is to leverage the existing NeCTAR Research Cloud eResearch infrastructure.

The first of these two alternatives assumes that the infrastructure will be co-hosted with the data storage infrastructure as detailed above (see Feature 1.1 - Data Storage). This means that no additional data hosting costs need to be taken into account. The cost to procure a 16-core system with 256 GB RAM, which will support a total user base of 5,000, is approximately \$40,000 over the six-year time frame. Systems offering less performance required for fewer users will cost approximately \$20,000 for 500 users (8-core system) and \$10,000 for 50 users (4-core system) over the six years.

The second of the two alternatives above leverages the NeCTAR Research Cloud to host the ArcGIS/GIS application software. The NeCTAR Research Cloud allows researchers to provision virtual machines with varying performance specifications, to aid their research. A GIS server could be setup as either a single compute node on the NeCTAR Research Cloud, or as a cluster of compute nodes, depending on the flexibility of the chosen GIS software. Utilising the NeCTAR Research Cloud will not incur any costs to the Data Hub until after the end of 2014. The cost-recovery model post-2014 is yet to be determined, however, estimates have been used that place the costs at \$25,000, \$10,000 and \$5,000 for 5,000, 500 and 50 users respectively, over the six-year time frame.

The ArcGIS Server software is available in three editions – Basic, Standard and Advanced, with additional optional features.³ Some further analysis is required to establish the appropriate edition of ArcGIS Server for the Data Hub, however for the needs of the Data Hub's users as they are currently understood, the Standard edition is recommended. Note that

² [wiki.gis.com Capacity Planning Tool 32nd Edition \(Spring 2013\)](http://wiki.gis.com/wiki/index.php/Capacity_Planning_Tool_32nd_Edition), http://wiki.gis.com/wiki/index.php/Capacity_Planning_Tool_32nd_Edition

³ ArcGIS 10.2 for Server Functionality Matrix White Paper (Aug 2013),

<http://www.esri.com/software/arcgis/arcgisserver/features/-/media/Files/Pdfs/library/brochures/pdfs/arcgis-server-functionality-matrix.pdf>

the cost of the ArcGIS Server licence depends on the number of cores in the system the software is running on.

ArcGIS Server Standard edition is estimated to cost approximately \$120,000 for a 16-core system. For an 8-core system the cost would be approximately \$60,000 and for 4 cores, \$30,000. If it is assumed that the University of New South Wales has an unlimited ArcGIS site licence that can be used for this project, there should be no cost for the ArcGIS Server licences.

QGIS (Quantum GIS) and GRASS (Geographic Resource Analysis Support System) offer viable, free alternatives to ArcGIS. Both are freely available open source GIS software solutions that can be used independently or together (e.g. with QGIS acting as a front-end to GRASS). As with a large number of open source software packages, there may be additional costs incurred in their use, in the form of increased set up time, maintenance effort and time to learn. Note that the supporting hardware required for an alternative open source solution is assumed to be the same as that required for ArcGIS.

Both ArcGIS and QGIS can be extended with additional components and plug-ins, however ArcGIS is potentially a more mature product with a greater feature set, meaning that it is more likely to meet the needs of researchers without the need for custom enhancements.

Cost:

- GIS Server Hardware – approximately \$5,000 to \$40,000, depending on the alternatives, as outlined above.
- ArcGIS/QGIS/GRASS Server Software Licencing – approximately \$0 to \$120,000, depending on the alternatives, as outlined above.

5.4.2.2 Feature 2.2 - Integration with AURIN

The AURIN portal is a four-year \$20 million Super Science project that will provide infrastructure for built environment and urban researchers so that they can access datasets and

other information services via a distributed network. AURIN is based on urban research disciplines' needs and provides a well-developed way for researchers to access and use a range of relevant datasets, with the functionality to combine with the researchers own data sets. The user accesses the AURIN portal services via their web browser.

For the purposes of integrating LCL with the AURIN portal, the Data Hub will be considered a sub-project of the AURIN project. The AURIN team has specific requirements for the integration of data from AURIN sub-projects. The AURIN preferred method for establishing the data feed is that of the Open Geospatial Consortium (OGC) - Web Feature Service Interface Standard (WFS). The data feed functionality includes being able to use, get or query features based on spatial and non-spatial constraints, create, delete and/ or update feature instances.

In order to integrate the LCL Data Hub with AURIN, the following activities will need to take place:

- Provide the requisite documentation to AURIN.
- Establish a GIS database (e.g. MS SQL for ArcGIS or PostGIS for QGIS/GRASS) and ingest data into the GIS database.
- Convert datasets to the appropriate format, where required.
- Convert existing or add metadata to the datasets in the format compatible with the AURIN portal.
- Establish the data feed to AURIN using a WFS (Web Feature Service), or compatible service.
- Maintain the data feed to AURIN.

Cost:

The estimated cost of integrating the Data Hub with AURIN is approximately \$160,000. This is based on the requirement for 1.2 FTEs over a 12-month period.

5.4.3 Option 3 – Extended AURIN-Enabled Data Hub

The Extended AURIN-Enabled Data Hub option provides the data storage infrastructure and support as outlined in Option 1, enables analysis via the AURIN portal as outlined in Option 2, adds an automated data feed from the originating sources for datasets, and provides analysis tools and additional functionality more closely aligned with the needs of the LCL research cohort (see Figure 8). This option includes three independent features that can all be considered optional: GIS analysis software, extensions to AURIN/ArcGIS and an automated data feed.

Strengths:

- Builds on the analysis capability established by Option 2 features and shares all of its strengths.
- Provides access to mature analysis products from the ArcGIS framework.
- Provides additional functionality that meets the specific needs of the LCL research cohort, including secure access to data and improved sharing capabilities.
- The automated data feed makes receiving updated datasets more efficient.

Weaknesses:

- Significant additional costs to develop custom features.
- Any risks associated with undertaking software development work.

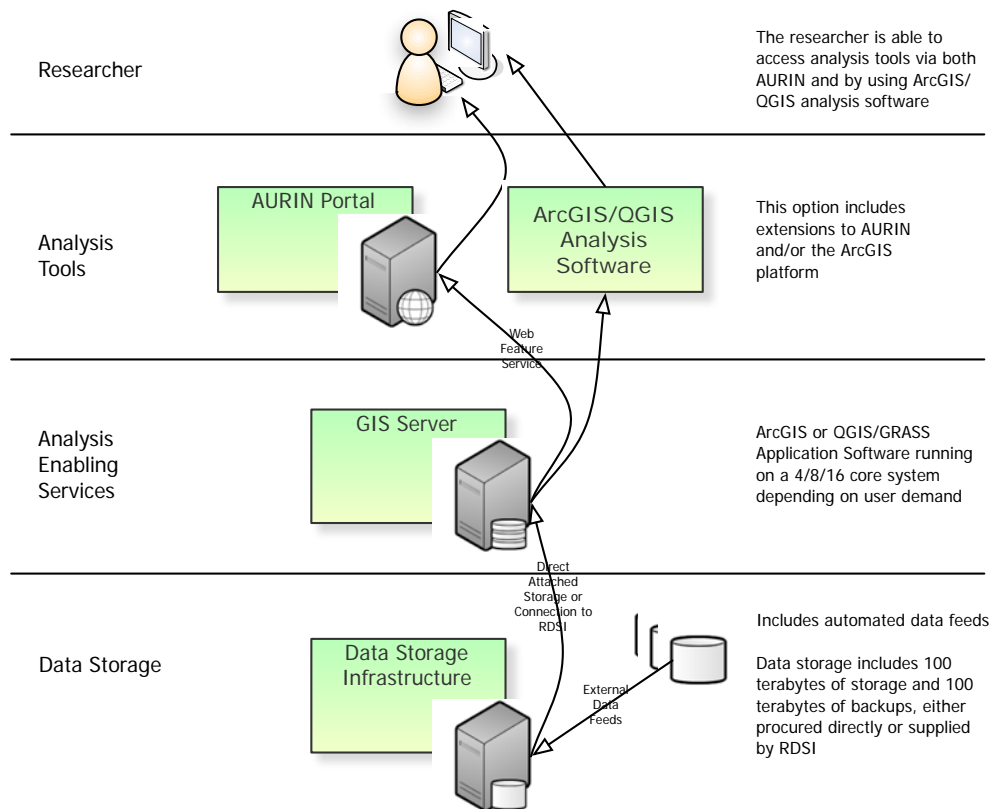


Figure 8 Option 3: Extended AURIN-Enabled Data Hub (shows all optional features enabled)

Total Costs:

Depending on the various alternatives chosen (as detailed below), the total cost of this option ranges from approximately \$750,000 to \$1,910,000. Note that this range includes the cost of all three features from Option 1 and both features from Option 2. It is based on the total number of users of the Data Hub being 500 (for the cost breakdown for 50 and 5,000 users, please see details below).

None of the features included in this option are obligatory. So, the cost will reduce as follows, depending on the alternatives chosen:

- Only Feature 3.1 - ArcGIS Analysis Software - \$535,000 to \$1,565,000
- Only Feature 3.2 - Extra-AURIN Features - \$665,000 to \$1,075,000
- Only Feature 3.3 - Automated Data Feeds - \$620,000 to \$900,000

5.4.3.1 Feature 3.1 - GIS Analysis Software

The ArcGIS platform provides its users with a wide range of tools for analysing geographic data, maps and other data. The tools focus on providing users with remote access to data, a wide variety of analysis techniques and a means to share findings. An assessment of the most appropriate tools for the Data Hub user cohort should be undertaken before selecting which tools to purchase. Potentially, the users themselves could purchase tools, rather than funding this via the Data Hub.

As an indication of potential cost, the ArcGIS Desktop Advanced + all extensions package licence can be purchased for \$250 per user per year (this assumes that licences can be obtained via Intersect, which qualifies under the Esri Non-Profit Organisation Program). The full set of features of this package can be viewed on the Esri website⁴.

The estimated cost for the ArcGIS analysis software is \$7,500,000 for 5,000 users over six years, \$750,000 for 500 users over six years or \$75,000 for 50 users over six years. These costs are calculated as a factor of the number of licences and number of years required. They do not take into account any volume licencing or other discounts (these can be obtained through negotiation with Esri, or via CAUDIT). Similarly, these costs assume researchers are not covered by their institution's site licence. As with the GIS Server (see Feature 2.1 - GIS Server), there are open source alternatives, which will offer viable, free alternatives to ArcGIS analysis software. The disadvantage of using an open source alternative, such as QGIS, is that some users may be alienated if they are familiar with products from the ArcGIS suite. Note that it should be possible to connect QGIS analysis software to an ArcGIS server, and vice-a-versa, however, this could add an overhead to the management of the Data Hub if incompatibility issues are encountered.

Cost:

- ArcGIS/QGIS Analysis Software Licencing – approximately \$0 to \$7,500,000, depending on the alternatives, as outlined above. Note that it is highly unlikely the full figure of \$7,500,000 would be encountered, as explained above. In practice, we can assume a more likely level of use would be 500 users, with an associated estimated cost of \$750,000.

5.4.3.2 Feature 3.2 - Extended Functionality

It may be the case that the AURIN portal does not provide the full range of analysis functionality required by users of the CRC LCL data. While the AURIN portal is currently funded until mid-2015 and may secure further funding, there are no guarantees about what analysis features will be added to the existing set. The Data Hub project provides an opportunity to expand the range of analysis features offered by AURIN. This can be achieved by undertaking one or more of the following:

1. Leverage the analysis tools offered as part of the ArcGIS platform (see Feature 3.1, above)
2. Contribute improvements to the AURIN source code repository
3. Develop AURIN plug-in components using the OMS3 (Object Modelling System) framework
4. Develop software for the Data Hub, that enhances the functionality offered by AURIN
5. Develop custom components or add-ins for ArcGIS.

The estimated cost of each feature varies considerably depending on the amount of additional functionality required. We currently know that, depending on what the AURIN portal can provide in the future, the Data Hub will require at least the following:

- Secure access for commercially sensitive data
- Ability to easily share research outcomes with other users
- Ability to import data, manipulate it and tag it, making the results available to other users.

Users of the Data Hub may have additional needs beyond this list, so a detailed analysis of user requirements is required before establishing a more accurate cost estimate. There may also be variations in the efficiency of the software development effort, depending on the approach taken. As an example, ArcGIS is a mature, complex software system, so developing custom components for ArcGIS may involve additional effort up front to understand the system.

Cost:

- Two developers working for 6-12 months to enhance the AURIN analysis functionality – approximately \$130,000 - \$260,000

5.4.3.3 Feature 3.3 - Automated Data Feeds

This is an optional feature, which involves the establishment of data feeds that automatically transfer data from the originating source to the Data Hub. The advantage of implementing this feature is that it makes it easier for the Data Hub to receive frequent updates to the data and reduces the effort required to setup and maintain data preparation workflows. Note that for each data feed that is established, there will almost certainly be work required at the originating end. It is also assumed that it will not be technically feasible to setup automated data feeds from every originating source. It is likely that not all data custodians will have the required infrastructure and knowhow to support this.

Cost:

The estimate only includes the effort required by the Data Hub and does not include the effort required at the originating source in establishing the data feeds. The estimate is based on 0.2 FTE over a period of 2 years to establish the automated data feeds and 1 day every 4 weeks required to maintain the data feeds – approximately \$85,000 over the six years.

5.4.4 Option 4 – Enhanced Spatial Analysis Hub

The Enhanced Spatial Analysis Hub option provides a comprehensive data storage and advanced analysis system, customised for the specific cohort of researchers that will make use of data from the LCL. This option includes all of the features outlined in Option 1, optionally the GIS Server feature from Option 2 and optionally the GIS Analysis Software and Automated Data Feeds feature from Option 3 (see Figure 9).

Strengths:

- Provides the flexibility to customise the solution for the specific cohort of users.
- Reduces the risk that features required as part of AURIN are not built by the time the AURIN project ends, mid-2015.
- Eliminates the waiting time for acceptance by AURIN for inclusion as a sub-project.
- No requirement to extend the functionality of ArcGIS, as additional features can be built into the custom analysis portal.

Weaknesses:

- Expensive alternative to reusing the already existing AURIN portal.

- Removes ability to compare LCL datasets with broader national datasets. The CRC would have to renegotiate access to datasets already included in the AURIN portal.
- Involves developing a large, complex software system, and therefore inherits any risks that stem from that, including: failure to deliver on time, changes in scope resulting in loss of important features, etc.

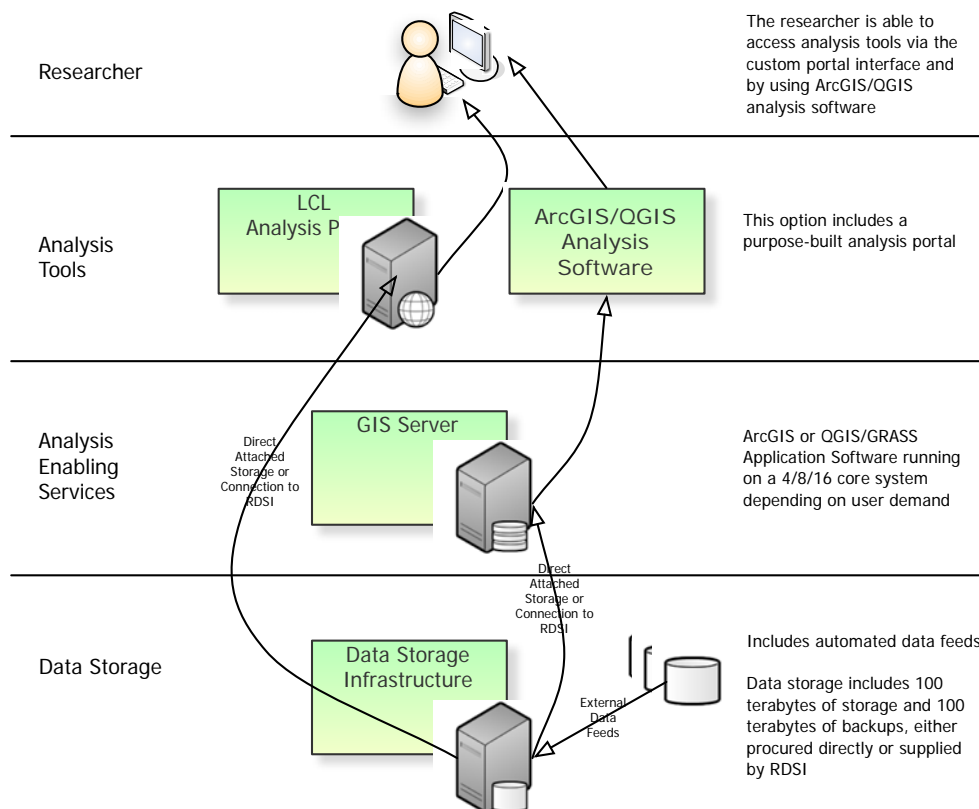


Figure 9 Option 4: Enhanced Spatial Analysis Hub (shows all optional features enabled)

Total Costs:

The total cost of this option ranges from approximately \$2,100,000 to \$3,900,000. Note that this range includes the cost of all three features from Option 1, the GIS Server feature from Option 2 and both the GIS Analysis Software and Automated Data Feeds feature from Option 3. The GIS Server and GIS Analysis Software costs are calculated assuming 500 users, so additional costs may be involved if increasing the number of users beyond this

number. See the individual feature sections for more detail. By removing all of the features from Option 2 and 3, the cost range for this option can be reduced to \$2,000,000 to \$3,000,000.

5.4.4.1 Feature 4.1 - LCL Analysis Portal

The LCL Analysis Portal provides analysis capability similar to the AURIN portal, but customised to the needs of researchers that will use the datasets housed by the Data Hub. This will provide the flexibility to develop a platform that meets the specific needs of the users of the Data Hub. Examples of some of the functionality that could be afforded by this level of control over the analysis platform include:

- Fast access to data due to a fine degree of control over connectivity to the underlying storage infrastructure
- Advanced combinatorial analysis based on ability to customise for specific datasets that make up the Data Hub

The estimated cost of this feature will depend on the final requirements. A detailed analysis of user requirements is required before establishing a more accurate cost estimate.

Cost:

Six to nine engineering staff (including Software Developers, QA, UI and Project Manager) working for 24 months to develop an AURIN-like system – approximately \$1,600,000 - \$2,400,000.

5.5 Trade-off Analysis

The options that have been presented differ considerably in relative cost and level of functionality being offered. To properly assess the optimal combination of cost versus level of

functionality, the requirements of the Data Hub need to be considered. Users will require the following functionality from the Data Hub:

- Access to a wide range of data across the variety of spatial scales at which research will be conducted
- Spatially integrated datasets, organised at the property/building level or land use cadastre level
- Ability to analyse spatially integrated datasets
- Ability to link to broader national datasets, and include these in analyses
- Ability to analyse results using filters based on broader national datasets

While all four options presented are capable of meeting all of these requirements, they will vary in the level of ease and efficiency with which a researcher is able to achieve their research outcomes. The first option, the Basic Data Hub, offers researchers the least in the way of analysis capability. While datasets will be spatially integrated, there will be no easy way to link or analyse datasets with broader national ones. The second option enables the Data Hub for use with the AURIN portal. This adds the analysis capabilities offered by the AURIN portal, introducing an efficient way to analyse LCL datasets alongside those datasets already offered via the AURIN portal. The third option builds on this by offering extensions to the functionality offered by AURIN, additional GIS analysis tools and automated data feeds from the various custodians of LCL data. Depending on the individual features chosen for Option 3, the estimated cost is approximately 1.5 to 2 times more expensive than the least expensive option, Option 1. Finally, the fourth option provides users of the Data Hub with an advanced analysis platform, capable of being completely customised for the particular group of researchers that will analyse LCL data. It is also the most expensive option of the four, at over three times the cost of the next most expensive option. Table 3 provides a summary of

the costs associated with the highest and lowest estimates for the various Options discussed above. The detailed costing underlying these summary figures are available if required.

These are indicative only, and do not include the costs of high level project management and oversight associated with implementing the proposals. They also assume that there are no costs associated with on-going access to UNSW infrastructure, software licences and IT support. These costs estimates would therefore need to be further reviewed should a decision to proceed be taken.

Table 3 Summary of Data hub Options

Totals	Minimum Cost	Maximum Cost
Option 1	\$365,000.00	\$635,000.00
Option 2	\$535,000.00	\$815,000.00
Option 3 (Full)	\$750,000.00	\$1,910,000.00
Option 3 (Just 3.1)	\$535,000.00	\$1,565,000.00
Option 3 (Just 3.2)	\$665,000.00	\$1,075,000.00
Option 3 (Just 3.3)	\$620,000.00	\$900,000.00
Option 4 (Full)	\$2,100,000.00	\$3,900,000.00
Option 4 (Partial)	\$2,000,000.00	\$3,000,000.00

6 Conclusions and Recommendations

This project has the potential to form the data and analysis ‘heart’ of the CRC in terms of supporting a wide range of research across the Living Laboratories and the three CRC Program areas and being able to link the empirical outcomes of these various research initiatives in a spatially integrated manner.

The key output of the report has been to set out a recommendation for establishing the LCL Data Hub with costings, location and staffing requirements. Based on the results presented in this report we recommend using Option 3 for the establishment of the Data Hub with the added benefit of access to comprehensive set of publicly accessible databases and analysis capabilities through the AURIN Portal that can be used by the CRC LCL experts. While Option 3 is comparable to that offered by Option 4, it is significantly cheaper. Option 3 therefore offers the most optimal combination of cost versus level of functionality. The risk with Option 3 is that it does not offer the level of flexibility and freedom of customisation that is offered by Option 4. In comparison, Option 4 introduces risks associated with building a large, complex software system and will likely result in a repeat of efforts carried out as part of the AURIN project to secure access to a broad range of national datasets. Options 1 and 2 are not considered to have a sufficient level of functionality to provide a suitable basis for the objectives of the Data Hub.

To reduce the final cost of the Data Hub, it is recommended that eResearch infrastructure, including RDSI and the NeCTAR Research Cloud be leveraged wherever possible (the cost trade-offs for these alternatives are outlined in the sections above). If no cost will be incurred to the project by hosting the storage infrastructure at UNSW, then this is the most cost-effective alternative to RDSI storage. Utilising institution site licences where

possible will significantly reduce software licencing costs. A summary of cost breakdown of Option 3 which was discussed earlier is presented in Table 4.

Adoption of this option and establishment of the LCL Data Hub itself would provide a major stimulus to further research projects that would both use and also add to the datasets it will hold. In addition, development of the Data Hub will have a direct and tangible impact on the delivery of the majority of the CRC Milestones as it would provide a platform storing, analysing, and sharing of information and development of modelling approaches, tools and support systems by different CRC partners (Appendix A2 highlights the range of output and utilization milestones that the development and use of the Data Hub is likely to directly and/or indirectly support).

A LOW CARBON LIVING SPATIAL DATA HUB

Table 4 Cost Breakdown of Option 3

Feature	Setup with Minimum Cost	MIN	Setup with Maximum Cost	MAX
Feature 1.1 - Data Storage	Storage Infrastructure (BECU Hosts) + Support Staff	\$147,200.00	Storage Infrastructure (Data Centre Hosts) + Support Staff	\$387,200.00
Feature 1.2 - Data Preparation	Data Preparation (min.)	\$33,000.00	Data Preparation (max.)	\$66,000.00
Feature 1.3 - Operations	Operations	\$179,570.50	Operations	\$179,570.50
Feature 2.1 - GIS Server	NeCTAR Research Cloud (Alternative to ArcGIS Application Server) (500 users)	\$12,000.00	Total Self-Managed Application Server Costs (500 users)	\$18,181.82
Feature 2.2 - Integration with AURIN	Integration with AURIN	\$158,400.00	Integration with AURIN	\$158,400.00
Feature 3.1 - GIS Analysis Software	QGIS/GRASS	\$0.00	ArcGIS Desktop Advanced + all extensions (500 users) Non-Profit Licence	\$750,000.00
Feature 3.2 - Extended Functionality	Development of AURIN-extra functionality (min.)	\$132,000.00	Development of AURIN-extra functionality (max.)	\$264,000.00
Feature 3.3 - Automated Data Feeds	Automated data feeds	\$85,482.62	Automated data feeds	\$85,482.62
Total (Rounded Figures)		\$750,000		\$1,910,000

References

AURIN (2013). Australian Urban Research Infrastructure Network (AURIN)
<http://aurin.org.au> Retrieved August 2013

CRC LCL (2013). Cooperative Research Centre for Low Carbon Living,
<http://www.lowcarbonlivingcrc.com.au>, Retrieved July 2013

Davis, A.M. (1993). Software Requirements: Objectives, Functions and States,
Englewood Cliffs, NJ: Prentice-Hall.

Faulk, S.R. (1997). "Software Requirements: A Tutorial", in Software Requirements
Engineering, 2nd Ed. by Thayer, R. and Dorfman, M. (Eds.), Los Alamitos, CA: IEEE
Computer Society Press.

Newton, P., Marchant, D., Mitchell, J., Plume J., Seo S., and Roggema R. (2013).
Performance assessment of urban precinct design: a scoping study, CRC LCL Final Report.

Song, Y., J. Bogdahn, et al. (2010). "Integrating BIM with Urban Spatial
Applications: A VEPS Perspective." ed. Underwood, J. & Isikdag, U., Handbook of Research
on Building Information Modelling and Construction Informatics: Concepts and
Technologies. IGI Global.

White S., Salter R. (2013). CRC Living Laboratories Framework, CRC LCL Final
Report.

Zhang, X., Arayici, Y., Wu, S., Abbott, C., and Aouad, G. (2009). Integrating BIM
and GIS for large-scale facilities asset management: a critical review.

Appendices

Appendix A1 List of organisations and respective datasets in AURIN portal (AURIN 2013 – retrieved in October 2013)

Centre of Full Employment and Equity (CofFEE)

Australia By CofFEE Functional Economic Region

Australia By Female CofFEE Functional Economic Region

Australia By Labour Force Region

Australia By Less Skilled CofFEE Functional Economic Region

Australia By Male CofFEE Functional Economic Region

Australia By Skilled CofFEE Functional Economic Region

Australia By Statistical Local Area

Australia By Trades CofFEE Functional Economic Region

City of Melbourne (CoM)

City of Melbourne Census of Land Use and Employment
(CLUE) Space Use -2008

City of Melbourne Census of Land Use and Employment
CLUE Predominant Space Use -2010

City of Melbourne CLUE Capacities 2010

City of Melbourne CLUE Capacity Measures -2008

City of Melbourne CLUE Employment By Industry -2008

City of Melbourne CLUE Employment By Industry 2010

City of Melbourne CLUE Space Use -2010

Department of Human Services (DHS)

Department of Human Services (DHS) 3 Bedroom House
Affordable Lettings by Local Government Area (June 2012)

Department of Human Services (DHS) Quarterly Median Rents
4 Bedroom House by LGA June 2012

DHS Affordable Lettings by LGA Jun-12 1 bedroom

DHS Affordable Lettings by LGA Jun-12 2 bedroom

DHS Affordable Lettings by LGA Jun-12 4 bedroom

DHS Affordable Lettings by LGA Jun-12 all bedroom types

DHS Quarterly Median Rents by LGA Jun-12 1 Bedroom Flat
 DHS Quarterly Median Rents by LGA Jun-12 2 Bedroom Flat
 DHS Quarterly Median Rents by LGA Jun-12 2 Bedroom
 House

DHS Quarterly Median Rents by LGA Jun-12 3 Bedroom
 House

Department of Planning and Community Development (DPCD)

Urban Development Program (UDP) – Broadhectare Estates -
 2010
 Urban Development Program (UDP) – Major Infill Sites 2009
 (regional only)
 Urban Development Program (UDP) – Broadhectare
 Residential Land -2010
 Urban Development Program (UDP) – Industrial Land -2010
 Urban Development Program (UDP) – Industrial Nodes -2010

Urban Development Program (UDP) – Industrial Proposed
 Areas -2010
 Urban Development Program (UDP) – Major Residential
 Redevelopment Sites -2010
 Urban Development Program (UDP) – Minor Infill Supply
 2009 (regional supply only)
 Urban Development Program (UDP) – Urban Growth
 Boundary (Aug 2010)

Department of Environment and Primary Industries (Formerly DSE)

Aggregation of Victorian Government Regional Departmental
 Boundaries (DSE) – Vicmap Admin
 FOI – Index Centroid – Vicmap Features of Interest
 FOI – Polygon – Vicmap Features of Interest
 Local Government Area Boundaries (Property) (polygon) –
 Vicmap Admin

Locality Boundaries (Property) (polygon) – Vicmap Admin
 Parish Boundaries (polygon) – Vicmap Admin
 Planning scheme overlay – Vicmap Planning
 Planning scheme Urban Growth Area – Vicmap Planning
 Planning scheme Urban Growth Boundary – Vicmap Planning
 Planning scheme zones – Vicmap Planning

Property Map Polygons – Vicmap Property

Property View – Vicmap Property

PTV Bus Routes Metro

PTV Bus Routes Regional DPS

PTV Train Corridor Centreline

PTV Tram Routes

Rail Network – Vicmap Transport

Statewide road network, line – 1:250,000 to 1 :5 million.

Vicmap Lite

VICMap Address

Victorian Government Regional Departmental Boundaries –

Vicmap Admin

Victorian Lower House Electoral Boundaries 2001 – Vicmap Admin

Victorian Upper House Electoral Boundaries 2005 – Vicmap Admin

WARD Boundaries 2008 (polygon) – Vicmap Admin

Australian Bureau of Statistics, via Landgate

Employment Statistics by Census District (2006 Census)
(ABS-080)

Employment Statistics by LGA (2006 Census) (ABS-079)

Employment Statistics by Suburb (2006 Census) (ABS-078)

FESA Product (STE) (Australia) (ABS-030)

Household by Census District (2006 Census) (ABS-071)

Household Income by Census District (2006 Census) (ABS-068)

Household Income by LGA (2006 Census) (ABS-067))

Households by LGA (2006 Census) (ABS-070)

Households by Suburb (2006 Census) (ABS-069)

Language by Census District (2006 Census) (ABS-074)

Language by LGA (2006 Census) (ABS-073)

Language by Suburb (2006 Census) (ABS-072)

Population by Census District (2006 Census) (ABS-077)

Population by LGA (2006 Census) (ABS-076)

Population by Suburb (2006 Census) (ABS-075)

PROPERTY Product (SSC) (Western Australia) (ABS-044)

Public Health Information Development Unit (PHIDU)

Agedistribution Aboriginalfemales

Agedistribution Aboriginalmales

Agedistribution Aboriginalpersons

Agedistribution Females

Agedistribution Males

Agedistribution persons

Amenable treatablemortality

Avoidablemortality

Birthplace nonEnglishspeakingresidents

Childcare unpaid

Childhealth

Chronicdisease syntheticprediction

Communitystrength

Compositeindicators syntheticprediction

Disability

Earlychildhooddevelopment

AustralianEarlyDevelopmentIndex

Education

Families

Healthriskfactors syntheticprediction

Housing Transport

Incomesupport

Indigenouststatus

Internetaccessathome

Labourforce

LearningorEarning

MBSservices

Medianageatdeath

Mothersandbabies

NonEnglishspeakingcountriesofbirth

Prematuremortality byselectedcause

Prematuremortality bysex

Preventablemortality

Privatehealthinsurance

Psychologicaldistress syntheticprediction

Residentialagedcareplaces

Screening

Selfassessedhealth syntheticprediction

Summarymeasureofdisadvantage

Totalfertilityrate

A LOW CARBON LIVING SPATIAL DATA HUB

Public Sector Mapping Agency (PSMA)

PSMA Street Network of Victoria

University of Queensland - UQeResearch

Australian Electoral Booth Catchments

Australian Electoral Booth Variables

Demographic Variables by LGA for Australia

Demographic Variables by SLA for Australia

Demographic variables by UCLs for Australia

Location Quotient (LQ) by LGA for Australia

Location Quotient (LQ) by SLA for Australia

Location Quotient (LQ) by UCL for Australia

Shift share variables by LGA for Australia

Socio-economic variables by LGA for Australia

Socio-economic variables by SLA for Australia

Socio-economic variables by UCLs for Australia

Vichealth

7 day \$ spend at a licensed premises (of those purchasing)

7 day \$ spend on packaged liquor (of those purchasing)

Adequate work-life balance

Shire of Melton

Shire of Melton Foot Paths -2012

Shire of Melton Public Street Lights -2012

The Bureau of Infrastructure, Transport and Regional

Economics (BITRE) variable by SLA for Australia

The Bureau of Infrastructure, Transport and Regional

Economics (BITRE) variable by Working Zone (WZ)

The National Centre for Social and Economic Modelling

(NATSEM) child social exclusion index by SLAs for Australia.

The National Centre for Social and Economic Modelling

(NATSEM) estimates of housing stress and estimates of

poverty variables by SLAs for Australia

The National Centre for Social and Economic Modelling

(NATSEM) indicators of child well-being variables by SLAs

for Australia

Attended arts activities or events (in the last 3 months)

Daily soft drink consumption

Inadequate sleep (Internet access at home

Lack time for friends/family

Long commute (2 hours per day)

Made or created art or crafts (in the last 3 months)

Participation in citizen engagement (in the last year)

Perceptions of safety- walking alone during day

Perceptions of safety- walking alone during night

Purchased alcohol in the last 7 days

Sedentary behaviour (sitting hours per day)

Shares a meal with family (5 days per week)

Social networking used to organise time with friends/family

Subjective wellbeing (range 0-100)

Time pressure

Visit to green space (once per week)

Visit to green space (once per week)

Volunteering (once per month)

Volunteering (once per month)

Valuer General Property Data - Victoria

North West Melbourne Region Local Government Areas -2010

A LOW CARBON LIVING SPATIAL DATA HUB

Appendix A2 List of CRC output and utilization milestones that the development and use of the Data Hub can directly and/or indirectly support

- R 1.3 Integrated design, showcase, ratings and standards
- R 2.1 Digital information platform for informed precinct design
- R 2.2 Integrated assessment of eco-efficiency during precinct design
- R 2.3 Precinct-level demand forecasting for distributed infrastructure networks
- R 2.4 Health and productivity co-benefits
- R 3.1 Transition scenarios and affordability
- R 3.2 Drivers and barriers to community engagement
- R 3.3 Living laboratories
- U 1.1 Commercialisation of building integrated solar technologies
- U 2.1 Adoption of PIM across the precinct design sector
- U 2.2 Automated precinct assessment tool
- U 2.3 Network utility demand forecasting
- U 2.4 Co-benefits calculator
- U 3.1 Communities, households and businesses will directly or indirectly use outputs from the programs
- U 3.2 Governments at all levels will use the databases, social psychological tools

A LOW CARBON LIVING SPATIAL DATA HUB

Appendix A3 List of Latest CRC Commissioned Projects

Project	Potential Spatial Data Use		
	Building	Precinct	City
RP2006: Action research to examine and demonstrate how to mainstream low-cost and low carbon housing in Western Australia Prof. Peter Newman, Curtin University			
RP2008: Beneficial Reuse of Solids from Wastewater Treatment Operations Project Leader: Professor Richard Stuetz		Possible	Likely
RP2009: Scoping study - Planning a research agenda for low carbon transport. Project Leader: Dr Rocco Zito, University of South Australia			
RP2010: Informing and Trialing the Inclusion of Low Carbon Requirements in State Government Built Environment Sector Tenders Project Leaders: Charlie Hargroves and Prof. Peter Newman, Curtin University Sustainability Polity Institute			
RP3009: High Performance Housing: Monitoring, Evaluating and Communicating the Journey: “Josh’s House” “Lochiel Park” and “CSIRO ZEH” Living Laboratories. Project Leader: Prof. Newman, Curtin University	X	X	
RP3010: Building Low Carbon Communities Project Leader: A/Prof J Merson, UNSW/ Blue Mountains World Heritage Institute (BMWHI)	Possible	X	
RP3011: Community Carbon Reduction and Wellbeing Enhancement Project Project Leader: Prof Yoshi Kashima, Melbourne University		X	
RP3012: Transformation to Low Carbon Living: Social Psychology of Low Carbon Behavioural Practice Project Leader: Prof Yoshi Kashima, Melbourne University		X	
RP3015: Increasing knowledge and motivating collaborative action on Low Carbon Living through team-based and game-based mobile learning Project Leader: Professor Shirley Leitch, Swinburne University of Technology	Likely	X	

Appendix A4 List of Potential Future CRC LCL Projects

Project	Potential Spatial Data Use		
	Building	Precinct	City
Integrated Carbon Metrics – Multiscale Life Cycle Approach to Assessing, Mapping and Tracking Carbon Outcomes for the Built Environment	X	X	
Co-Benefits Calculator – Scoping Project. Assessing Options for Evaluation of the Ancillary Health and Productivity Benefits of Environmental Modifications to Reduce Greenhouse Gas Emissions	X	X	
Facilitating Low Carbon Precincts – Precinct Information Model for Increasingly Complex and Data-Rich Urban Environments	Likely	X	
Brownfields Redevelopment as a Carbon Emissions Reduction Tool	X	X	X
Preparing Australian Schools for a Low Carbon Economy: An Operational Framework for Emissions Abatement in the Education Sector			