

NSW Housing Stock Mapping Phase II

FINAL REPORT

Prepared for NSW Office of Environment and Heritage

March 2018

ABOUT THIS REPORT

Title: NSW Housing Stock Mapping Project: Phase II Final Report

Project Team: SMART Infrastructure Facility (SMART)

Jun Ma, Tim Davies, Pascal Perez

Sustainable Buildings Research Centre (SBRC)

Daniel Daly, Georgios Kokogiannakis, Paul Cooper.

Reference: Ma J, Davies T, Perez P, Daly D, Kokogiannakis G, Cooper P, (2017). *NSW Housing Stock Mapping Project: Phase II Final Report*. University of Wollongong, NSW, Australia.

DISCLAIMER: This report was prepared by the University of Wollongong in good faith exercising all due care and attention, but no representation or warranty, express or implied, is made as to the relevance, accuracy, completeness or fitness for purpose of this document in respect of any particular user's circumstances. Users of this document should satisfy themselves concerning its application to, and where necessary seek expert advice in respect of, their situation. The views expressed within are not necessarily the views of the Office of Environment and Heritage (OEH) and may not represent OEH policy.

© Copyright State of NSW and the Office of Environment and Heritage

CONTACT

SMART Infrastructure Facility
Faculty of Engineering and Information Sciences
University of Wollongong
NSW 2522 Australia
Telephone: +61 (02) 4221 5031
Email: smart-facility@uow.edu.au
Web: smart.uow.edu.au

Executive Summary

The primary focus of Phase II of the NSW Housing Stock Mapping Project (HSMII) was expanding the spatial coverage of the database created in the NSW Housing Stock Mapping Pilot project to cover the whole of NSW. The Housing Stock Mapping Pilot project, which was finalised in March 2016, provided a central database which contained information on the housing stock in the Illawarra region amalgamated from a number disparate sources.

The HSMII project was designed as a collaborative project between the University of Wollongong, the NSW Office of Environment and Heritage (OEH), and the NSW Department of Planning and Environment (DPE), as was successfully undertaken in the pilot phase. As well as expanding the spatial coverage of the database, the current project had several other objectives, namely i) to better understand some of the limitations of the BASIX database, ii) to outline the project team's key lessons learnt from the Pilot and Phase II projects in a discussion paper, and iii) to demonstrate the potential of the centralised database through proof-of-concept visualisations.

The HSMII project has delivered a centralised data repository containing information regarding the NSW housing stock, as well as related contextual information (e.g. demographics, climatic conditions, etc.), sourced from a range of public and closed databases. A needs analysis was undertaken to understand the visualisation requirements for a number of OEH and DPE internal stakeholders. Proof-of-concept visualisations were developed to meet a number of the identified requirements, subject to the current data limitations. Several high priority datasets were highlighted for future acquisition, namely property address level energy consumption data, land value and property sales data, and the existing Land and Housing Corporation property assessment survey database. Accessing these databases would substantially increase the utility of the housing stock database.

A pilot study was undertaken as part of the current project to explore how closely the data contained within the DPE BASIX database matched the data contained with a council approvals dataset for a small sample of approvals in the Wollongong area. Several items of interest were noted, including a significant number of BASIX certificates that did not appear to relate to construction activity (i.e. the proposed development did not proceed), and a number of records where the BASIX certificate would under-represent the energy-efficiency of a constructed dwelling (i.e. the final construction was better than the BASIX requirement).

A discussion paper was also prepared for OEH internal use to summarise some of the key lessons gleaned from the project team involved in this project, and to provide some recommended actions for OEH when planning future data collection exercises. These recommendations are based largely on the major difficulties encountered in the HSM Pilot and Phase II projects.

This Final Report provides a summary of the activities undertaken to meet these objectives of the current project, and directs the reader to the appropriate location for further information. Two stand-alone reports are included as appendices, the BASIX data validation pilot study final report, and the Data Collection & Formatting discussion paper. An electronic version of the Housing Stock Database has been provided to OEH, along with supporting technical documentation. The proof-of-concept visualisations were made available through a UOW hosted web-portal for a period after project completion.

Table of Contents

1. Introduction	7
1.1. Background.....	7
1.2. NSW HSM Phase ii Project Description	7
1.3. Outcomes and Outputs.....	9
2. Deliverable 1 – Expanded HSM database	10
2.1. Data Sourcing	10
2.2. Data Profiling and Processing	11
2.1. Data Modelling and Integration.....	13
3. Deliverable 2 – BASIX comparison pilot study	14
4. Deliverable 3 – Discussion Paper	16
5. Deliverable 4 – Proof of concept visualisations.	18
5.1. Needs Analysis	18
5.2. Visualisation development	21
6. Deliverable 5 – Web embedding	23
7. Recommendations and Opportunities.....	25
7.1. Potential Applications.....	25
7.2. Future and Further Development.....	26
8. References	30

1. Introduction

1.1. BACKGROUND

The development of a five-year long collaboration between the New South Wales Office of Environment and Heritage (OEH) Sustainable Households and Communities Team and the University of Wollongong (UOW) led to the signing of a Memorandum of Understanding and commitment to work together on issues of common interest and related expertise.

The need for the development of a NSW Housing Stock Database (HSD) was identified as a key priority for a wide range of reasons during a visioning workshop held in early 2015. The *NSW Housing Stock Mapping Pilot Project* (HSMPP), finalised in March 2016, delivered a first release of a comprehensive database of fundamental residential housing stock characteristics focused on the Illawarra Region.

The HSMPP involved a substantial data sourcing exercise which included detailed evaluation of the datasets accessed and a gap analysis of the attributes therein. This work was undertaken in close collaboration with staff from OEH and the NSW Department of Planning and Environment (DPE). The disparate datasets were then collated into a central database, and proof-of-concept visualisations were developed to illustrate the analytical potential of the database.

There were a number of opportunities identified and recommended for exploration from the outcomes achieved in the HSMPP. These included: expanding the delivered HSD to cover the broader NSW housing stock; improving understanding of the uncertainty associated with the BASIX databases; and the development of demonstration data visualisation tools to illustrate the benefits of the database for various stakeholders and users.

1.2. NSW HSM PHASE II PROJECT DESCRIPTION

Phase II of the Housing Stock Mapping Project (HSMPII) began in July 2016 with the objective to expand and build upon the outcomes and achievements delivered in the pilot phase. Four specific objectives were defined:

- Expanding the data that currently exists within the HSD delivered in the pilot phase from the Illawarra Region to NSW state-wide.
- Demonstrate the potential uses of the expanded HSD through web-embedded proof-of-concept data visualisations, focusing on two stakeholder group applications.

- Pilot a comparison-planned construction (as per the BASIX data) and final submitted construction details (i.e. from Development Applications of other) from accessible information within a pilot location.
- Providing recommendations and a methodology on how to standardise the collection, formatting and characterisation of fundamental housing stock variables across relevant government and non-government entities data gathering initiatives.

As a part of the Phase II project, the UOW project team joined the International Energy Agency (IEA) Annex 70 (Building Energy Epidemiology: Analysis of Real Building Energy Use at Scale). Annex 70 is an international collaboration of researchers, industry and government from across the globe who are working to develop methods for improving the empirical evidence on energy demand in the building stock. Participation in this Annex can help to ensure ongoing work on Housing Stock data in NSW is undertaken in accordance with leading edge methods for studying and modelling the building stock. Several outputs from the HSM Phase2 project will inform UOW's involvement in Annex 70.

1.3. OUTCOMES AND OUTPUTS

The following specific items were included for delivery in this phase of the project.

D1	Expanded Housing Stock Mapping database, scaling existing datasets from the Illawarra Region to NSW state-wide coverage.	Digital transfer to OEH
D2	Stand-alone pilot study comparing the changes in data from DA Completion Certificates to the corresponding BASIX Submission records.	Stand-alone document; Appendix A
D3	Discussion paper outlining recommendations for standardising the collection and formatting of fundamental attributes for the characterisation of a housing stock.	Stand-alone document; Appendix B
D4	Expansion of the Housing Stock Mapping Dashboard platform including: <ul style="list-style-type: none"> a) Expansion of existing dashboard from Illawarra Region to all NSW (as per D1); b) Development of fit-for-purpose reports and visualisations for 2 stakeholder groups. 	Website
D5	Proof-of-Concept website demonstrating embedded reports and visualisations (D4) for 2 targeted stakeholder groups (as per D4).	Website/Web portal
D6	Technical handover to OEH including: <ul style="list-style-type: none"> a) Updated technical documentation including: Database Connection User Guide; Entity Relationship Diagram; Scripts and Tools User Manual; Source Code User Guide; Data Dictionary; and, b) Handover meeting with nominated OEH staff members (2-4) to walk through database access, structure, source code, scripts and supporting documentation. c) Half day training of nominated OEH staff (2-4) on creation of simple BI reports, using a pre-existing data model and data warehouse. 	Stand-alone document
D7	Mid-point progress presentation	Completed 8/03/2017
D8	Final presentation and demonstration of deliverables (D4 and D5)	12/10/2017
D9	Final report	Current Document

This final report presents a summary of the activities undertaken to prepare these deliverables. In many cases, a stand-alone document has been provided separately for a specific deliverable, as noted above.

2. Deliverable 1 – Expanded HSM database

2.1. DATA SOURCING

In contrast to the pilot phase of the HSM project, the sourcing of new datasets was a less significant activity in Phase II. The majority of the data sourcing activities were focussed on extending the spatial coverage of the existing data in the HSD from the Illawarra region to NSW state-wide. In the pilot phase final report, the UOW project team identified a number of potential useful data sources for inclusion, with OEH/DPE to lead the negotiations for acquiring access to the new datasets. Other than publicly available data, no additional closed datasets were included in this update of the HSD.

Except for those datasets identified in HSMPP, the UOW team identified a number of supplementary datasets to provide relevant contextual information for the housing stock data. These new datasets cover issues such as demographics, housing markets, housing activities, and housing usage. A short list of newly sourced datasets is shown below.

- Australian Statistical Geography Standard (ASGS) 2011 and ASGS 2016 (part) (Australian Bureau of Statistics)
- Quarterly housing rent and sale records (NSW Housing)
- Building activities and approval (Australia Bureau Statistics)
- Population projection (Department of Planning and Environment)
- Regional population growth (ABS)
- Occupation and labour market projection (Department of Employment)
- Housing insulation statistics (ABS)
- Electricity bill benchmarks for residential customers (Australian Energy Regulator)
- Community services record (Centrelink)
- Geocoded National Address File (PSMA)

A full list and summary of newly sourced datasets can be found in the data dictionary of the Housing Stock Database (hsd). A number of datasets from the HSMPP have been updated and altered.

- AURIN property market dataset has been dropped from the HSD as AURIN has removed the dataset from its data repository. The LPI dataset is a potential alternative of the AURIN data.
- The ABS Postal Area (POA) data is imported into the HSD. Due to the agglomeration of Local Government Areas (LGA) in NSW in 2016 and 2017, the boundaries and names of some LGAs have been changed which affects the data aggregation at LGA level.

- BOM weather dataset for temperature and rainfall has been updated to 27/09/2016.
- SA1 and SA2 datasets are aggregated into POA level. Based on the outcomes of the needs analysis, postcode was agreed as the most useful geographic unit for data visualisation. SA1 and SA2 datasets are still kept in the HSD to support reports built in the pilot phase, and for future developments.
- BASIX data has been updated to include fiscal year 2014/15. The updated data is of slightly changed format.

An ongoing effort has also been made to access data from Energy Providers, and the Land and Housing Corporation. Data request letters were developed by UOW, and distributed via OEH to the NSW Department of Family and Community Services, and relevant energy providers (AusGrid and Endeavour Energy). In the timeframe of the project no agreement for data access was reached, however it is recommended that OEH continue to pursue access to these datasets.

Access to property address level (PAL) energy consumption data would substantially improve the utility of the data contained within the current HSM database, by allowing an analysis of the energy consumption implications of different building attributes. Energy consumption data is most valuable when it can be linked to a specific property address, however once linked, the data can be de-identified and aggregated without compromising the usefulness of the information. International experience has suggested that negotiating access to sensitive building information may be a long process, but there are substantial benefits to be realised.

The Land and Housing Corporation manage approximately 113,000 dwellings in NSW, for which regular Property Assessment Surveys are undertaken. Importantly, the LAHC database contains information on the age of dwelling, which could be used to explore links between building age and prevailing construction materials and methods, as well as other building attributes. This information may help to generate building typologies for the broader building stock, and would be invaluable for benchmarking energy ratings.

2.2. DATA PROFILING AND PROCESSING

Data profiling and processing was conducted focusing on POA level (approximate postcode) for the Phase II project. A housing stock mapping visualisation workshop was held in November, 2016. It was agreed that POA level detail is the most useful and preferred level of aggregation for data visualisation.

Data profiling was conducted during and following the data collection. Considering the different temporal coverages of datasets in the Pilot Phase, the team particularly focused on the datasets with similar temporal coverage in Phase II to better match different datasets.

Attributes that related to sustainability features, dwelling structures and energy efficiency were the main focus of data processing. These attributes were extracted, converted, conformed and integrated into the HSD. Due to the diversity of the terms used to refer to these features, the team primarily followed the definition from ABS and/or BASIX in order to align the attributes from different data sources. For example, many datasets defined “dwelling structure” using different categories; the team redefined the dwelling structure following the definitions in ABS’s census (see Table 1).

Table 1: Redefined dwelling structure in HSD.

Unified dwelling type	Data source	Original dwelling type	Original description
Separate House	BASIX-AA	Separate House	Separate House
Semi-detached dwelling	BASIX-AA	Attached House	Attached House
Unit/Flat	BASIX-AA	Unit	Unit
Semi-detached dwelling	BASIX-SD	Attached Dwelling	Attached Dwelling
Unit/Flat	BASIX-SD	Unit	Unit
Separate House	BASIX-SD	Single House	Single House
Separate House	BASIX-MD	Separate houses	[SH] Separate House
Semi-detached dwelling	BASIX-MD	Attached houses	[SD] Semi-detach, townhouse, terrace, villa
Other	BASIX-MD	Other	"
Unit/Flat	BASIX-MD	Unit Building	[HR] high-rise(9+ storeys) flat/unit/apartment*
Unit/Flat	BASIX-MD	Unit Building	[LR] low-rise(1-3 storeys) flat/unit/apartment*
Unit/Flat	BASIX-MD	Unit Building	[MR] mid-rise(4-8 storeys) flat/unit/apartment*
Unit/Flat	BASIX-MD	Unit Building	[MR] mid-rise(4-8 storeys) flat/unit/apartment*
Separate House	NEXIS	Separate house	Separate house
Semi-detached dwelling	NEXIS	Semidetached house	Semidetached house
Unit/Flat	NEXIS	2storey apartment	2storey apartment
Unit/Flat	NEXIS	3storey apartment	3storey apartment
Unit/Flat	NEXIS	4plus storey apartment	4plus storey apartment
Other	ABS	Others	Others
Separate House	ABS	Separated house	Separated house
Semi-detached dwelling	ABS	Semi-detached dwelling	Semi-detached dwelling
Unit/Flat	ABS	Unit/Flat	Unit/Flat
Unit/Flat	HPSP	2 bed unit (60-100 m ²)	2 bed unit (60-100 m ²)
Semi-detached dwelling	HPSP	Small terrace house/3-4 bed unit (100-150 m ²)	Small terrace house/3-4 bed unit (100-150 m ²)
Unit/Flat	HPSP	Large unit (150-200 m ²)	Large unit (150-200 m ²)

* Note: This definition is from the BASIX datasets. Recently, DPE has a new definition for LR (up to 3 storeys), MR (4-5 storeys) and HR (6+ storeys).

Unit/Flat	HPSP	Studio apartment (60 m ²)	Studio apartment (60 m ²)
Separate House	HPSP	Detached 4 bed house (250-400 m ²)	Detached 4 bed house (250-400 m ²)
Separate House	HPSP	Average 2-3 bed house (200-250 m ²)	Average 2-3 bed house (200-250 m ²)
Separate House	HPSP	Large 5+ bed house (>400 m ²)	Large 5+ bed house (>400 m ²)

2.1. DATA MODELLING AND INTEGRATION

Data modelling and integration in Phase II employed the same main structure of that used in the pilot phase, with minor changes to fit the requirements of the stakeholders to focus visualisation on POA level details.

The data in the HSD is organised into seven schemas according to the data sources and usage. The seven schemas are: “abs”, “bom”, “cross_reference”, “public”, “source”, “utility” and “reports”. The structure of those schemas is shown in Figure 1. Among them, the “report” schema is used to store tables for data visualisation. Tables in this schema are updatable. The other six schemas are used to store collected and derived data. In particular the schema “source” is used as a staging and operational data storage(ODS) area. This schema contains the major unique datasets sourced through the OEH and DPE teams, excluding the utility data. The schema “cross_reference” contains tables of data mapping between different datasets. For example, it contains regulated address naming rules. The “utility” schema stores data related to electricity consumption sourced from electricity providers. In future, other utility data such as water and waste processing will be stored in this schema. The “abs” and “bom” schemas store datasets sourced from ABS and BOM respectively. The “public” schema is used to store datasets derived from publically available datasets. These derived datasets are further processed to generate tables for reporting (data visualisation).

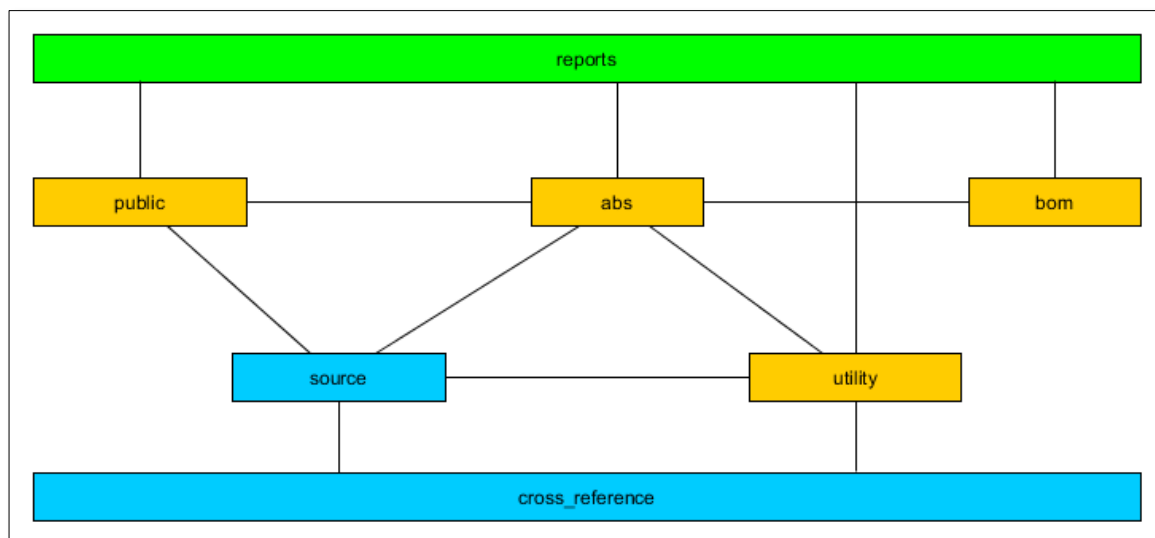


Figure 1: Schema structure in HSD.

3. Deliverable 2 – BASIX comparison pilot study

A significant source of data within the current integrated HSM database is the NSW Department of Planning and Environment BASIX certificate database, containing records of all BASIX certificates generated between 2006 and 2014. BASIX certificates are generated based on owner or builder intentions; i.e. the data is input prior to any work being undertaken, rather than being collected from the building as-built. As more new buildings are constructed under the BASIX regulations, more data will be available from this data source, making this a significant data resource for understanding building energy efficiency in NSW.

Understanding the consistency and reliability of the BASIX data is a crucial step in the ongoing development of the NSW HSM project. This project provided an indication of reliability issues which may exist within the BASIX dataset or elements thereof, identified the best methods to extend the pilot study to a broader sample, and identified possible methods to minimise the impact of any identified issues on the overall reliability of the HSM tool.

Recent studies have highlighted nationwide issues with building permit non-compliance, including with energy-efficiency requirements (Victorian Auditor-General, 2011; Harrington, 2014). A previously completed ‘BASIX compliance audit program’ (Eckstein, 2013) found that compliance between BASIX certificates and plans submitted for Development Applications and Construction Certificates “can generally be described as fair to good”. Eckstein assessed the standard of compliance for 103 applications, across twelve councils in NSW. However, this previous compliance audit compared the BASIX data contained internally within the councils registers (i.e. it compared the BASIX certificate in the council file to the plans in the council file), rather than comparing the information contained within the DPE database with the council documents. Whilst the audit results from Eckstein (2013) can give some confidence that the BASIX data is reflective of the final planned construction, importantly the study did not give an indication of the number of BASIX certificates (present in the DPE database) which are not linked to a DA in a council database (i.e. the planned build did not progress beyond the generation of the BASIX certificate). This pilot study therefore aimed to investigate the consistency of information between the BASIX database provided to UOW by the NSW DPE and the data contained within council planning databases.

The pilot study was limited to the suburb of Port Kembla. A total of 90 BASIX certificates were identified within the NSW Housing Stock Mapping database (being from FY2006/07 to FY2014/15) in Port Kembla. Thirty (30) of these were analysed in detail, with the remaining sixty (60) being scrutinised to determine if a matching record existed in the council register, and if the matching record had been constructed. Several important issues were identified:

1. A relatively high proportion (30-40%) of records in the tested sample of the BASIX database did not have a corresponding record of construction activity in the Wollongong City Council (WCC) planning database
2. In cases where a single address had multiple certificates associated with it, a difficulty was noted in determining whether the multiple certificates referred to dual occupancies or a revision to the original BASIX certificate. In several cases a BASIX single dwelling certificate was found to be a revised certificate (with a certificate number ending in S_02) in the council database, but no corresponding revised certificate was contained within the DPE BASIX database provided to UOW. A search of the BASIX database provided to UOW revealed no revised certificates (number ending in _02) were contained within for single dwellings (In the alterations and additions database, 9.5% of records had a revised certificate number). It should be noted that in a situation where a proponent has generated a revised certificate, both certificates will be valid, and the proponent may decide to proceed with the original design.
3. In many cases, the documentation submitted to council showed a more energy efficient design than that included in the BASIX certificate, meaning that analysis based on the BASIX database may under-estimate the energy-efficiency of constructions (not: there is also uncertainty associated with compliance with BASIX requirements, which was not explored in this study).

The full report (Daly *et al.*, 2017) is included as separate document in this submission.

4. Deliverable 3 – Discussion Paper

Building upon the lessons learned from the HSMPP, and the recently completed Low Income Energy-Efficiency Program (LIEEP), the project team has prepared a brief discussion of some of the key issues related to the collection, formatting and characterisation of fundamental housing stock. The full discussion paper (Perez *et al.*, 2017) is included as a separate document in this submission. The key recommendations from the paper included:

- **Recommendation #1 – Data collection method.** It is recommended that a metadata flag describing the collection method for each attribute be developed and included in future data collection activities, to allow future data users to assess the relative reliability of data collected by different schemes. It would be beneficial to have an agreed ‘data reliability hierarchy’, outlining which method of collection is considered the most reliable for different building and occupancy parameters, and in what circumstances a less reliable measure should be used. This would clearly indicate to future data users the relative uncertainty associated with different data sources.
- **Recommendation #2 – surveyor’s skills.** It is recommended that simple collection protocols be provided, using informative imagery where possible and relevant (see: Insurance Council of Australia Building Resilience Rating Tool), to minimise the reliance on personal judgement.
- **Recommendation #3 – baseline datasets.** It is recommended that the baseline data collected as part of every program be sufficient to place a dwelling into an appropriate typology. Common segmentation criteria for building typologies include dwelling structure, wall construction type, number of storeys, dwelling age, and floor area. Further details can be found in the NSW Housing Typologies Development final report.
- **Recommendation #4 – data attributes and categories.** It is recommended that a list of attribute ‘constants’ be created for key building attributes, and any future programs ensure that the available data points can, as a minimum, aggregate up to one of the constants. This will allow future projects to collect data at the degree of detail appropriate for that project, whilst ensuring that the data collected can be easily integrated with existing data-sets.
- **Recommendation #5 – data storage.** It is recommended to build upon the data storage and database structure developed during the NSW Housing Stock Mapping project. This 7-schema-based database offers a good scalability and allows for the management of Government and non-Government datasets.

- **Recommendation #6 – data manipulation.** It is recommended to build upon the workflow developed for the NSW Housing Stock Mapping project (SMART, 2016). This set up can be deployed on a developer machine, and then migrated to a dedicated virtual machine for the purposes of repeated uploads of data from the same format.
- **Recommendation #7 – data visualisation.** It is recommended to explore latest visualisation techniques like the 2D/3D Envision Scenario Planner supported by the CRC Spatial Information and the Australian Urban Research Infrastructure Network.

5. Deliverable 4 – Proof of concept visualisations.

5.1. NEEDS ANALYSIS

The purpose of this activity was to inform the data collection, reporting and visualisation needs of this second project phase. It was agreed that the visualisation would be developed for the OEH sustainable households team, and the DPE Demography & Economics teams¹. A stakeholders needs analysis workshop was conducted at the OEH offices on the 9th November 2016 with the OEH and DPE team member listed:

- Christopher O’Dell, Team Leader, GDP Demography & Economics, DPE
- Kevin Yee, Technical Contractor Specialist, DPE
- Gino Cavallaro, Manager ePlanning Data & Information Management, DPE
- Rachel Haley, Senior Project Officer, Sustainable Households, ROG Metro, OEH
- Celine Bachelet, Senior Project Officer, Sustainable Households, ROG Metro, OEH
- Erin Harwood, Senior Information Analyst, Data and Evaluation, ROG Metro, OEH
- Drew Duckworth, Data and Reporting Analyst, Data and Evaluation, ROG Metro, OEH
- Meredith Gee, Project Officer, Research and Design, ROG Metro, OEH
- Jakki Trenbath, Senior Project Officer, Research and Design, ROG Metro, OEH

The 2 hour stakeholders workshop explored the information needs pertaining of the OEH and DPE stakeholder groups, particularly to ascertain how the *NSW Housing Stock Mapping Phase 2* project can best deliver housing stock related data and visualisations useful to their work and business objectives. The workshop was organised into 2 key modules covering a summary of the background and objectives of the project to date, as well as an open-ended business discussion brainstorming a wish list of specific outcomes, outputs and indicators relevant to DPE and OEH stakeholders for the Phase 2 project. The workshop structure is shown in Table 1.

¹ For further information regarding the process of this stakeholder selection see ‘Stakeholder visualisation status update’ shared with OEH project team 12/12/2016

Table 1. Needs analysis workshop structure

Time	Module and Format	Objectives of Module
40 mins	Project Summary – Background, Objectives and Progress <i>PowerPoint Presentation</i>	<ul style="list-style-type: none"> • Refresher of background, drivers and objectives of project to stakeholder group.
80 mins	Project Outcomes, Outputs and Indicators <i>White Board Discussion Session</i>	<ul style="list-style-type: none"> • Explore and define specific business outcomes around NSW Housing Stock • Identify key information themes, outputs and indicative measures to form dashboard

The stakeholder group identified a number of key business activities and objectives that would be supported through the delivery of a comprehensive data and information platform of the characteristics of NSW’s housing stock. These items include:

- Allowing **OEH** to target their intervention programs more effectively for different locality, health and socio-economic groups
- Better equipping **planners and policy makers (i.e. DPE)** to understand the adequacy of the existing NSW housing stock, and evaluate whether the stock that is being delivered is in accordance with the needs of the population.
- Allowing **industry** to better understand the market and more effectively target their services and products in order to improve the social, economic and environmental performance of NSW dwellings;
- Allowing **householders** to see the economic and environmental performance of their home against others in the area, and also against regional and state-wide trends in penetration of products / methods / technologies.

The results of the project outcomes, outputs and indicators module of the workshop are displayed in Table 2, outlining the needs of the key stakeholders for a housing stock visualisation tool including purpose of the dashboards, themes of reports, and indicators to display.

Table 2: Stakeholder Visualisation Requirements – Results from Project Outcomes, Outputs and Indicators module of workshop

OUTCOMES (WHAT DO YOU WANT THIS TOOL TO DO FOR YOU?)	OUTPUTS (SEE) (WHAT DO YOU WANT TO SEE VIA REPORTING AND VISUALISATIONS)	INDICATORS (QUANTITATIVE VALUES OR RESULTS INDICATING PROGRESS AGAINST OUTCOMES)
<ul style="list-style-type: none"> Understanding the adequacy of the housing that we have and are delivering and whether they meet the requirements of residents (i.e. demographics, real estate market, job opportunities). 	<ul style="list-style-type: none"> Understanding of liveability of locales – comparison of areas that DPE view as working well vs other areas How to modify other areas to become more liveable; Better understanding of evolution of actual stock Linking housing stock with infrastructure 	<ul style="list-style-type: none"> % apartments/units vs detached vs semi-detached % type of dwelling (# of bedrooms) % secondary dwellings (Granny flats) Size of dwellings (# bedrooms) vs Size of families Demographics (age) vs dwelling size and structure Building footprint vs land area Energy consumption / capita
<ul style="list-style-type: none"> Link to Voluntary Disclosure project (evidence and assumptions) Support marketing of energy efficient buildings for real estate agents. Using the existing data (BASIX, but importantly, pre BASIX) to inform/pre-populate the rating tools developed. 	<ul style="list-style-type: none"> Evaluation of the sustainability of pre-BASIX housing Click on a house and see information of all the sustainability features of that house. Small precinct/street block aggregated information for comparison against single dwellings (e.g. for real estate agents to differentiate and market dwellings based on features). Tracking of Energy Savings Scheme in households 	<ul style="list-style-type: none"> % dwellings with specific “sustainability features” in a given area. (Possibly based on ESS or liveability framework) Energy consumption per house/house type Actual occupation of dwellings.
<ul style="list-style-type: none"> A tool to design targeted interventions for government, linked with demographics, area, health, economics. 	<ul style="list-style-type: none"> Age of house + history of transactions (identify older occupants) Demographic profiles of neighbourhoods Economic, health and social groups challenges. Targeted at disadvantaged social groups, including elderly, poor education, non-English speaking, disability, Centrelink recipients. 	<ul style="list-style-type: none"> Sustainable features (PV, INS, HWS etc.) vs income, location (Regional/metro), climate exposure, other target indicators. Sustainability features (ins) vs health indicators (DOH)

5.2. VISUALISATION DEVELOPMENT

The visualisations for the project have been developed to give some insight into the datasets collected and stored within the project database. The development of the visualisations was based on a number of different criteria. The needs analysis was used as the foundation for the visualisation. From the needs analysis four main themes emerged. These were energy, demographics, housing stock and data sources. We kept these themes in mind while creating the visualisations, however, as the visualisations can only reflect what is in the database many of the visual outputs were also based on opportunistic insights that were discovered on closer inspection of the data. The visualisation themes were:

Major Data sources



The Major Data Sources theme demonstrates the relative coverage of the major data sources provided by the NSW government for the current project. The total number of records contained in each database is listed, and the spatial distribution of records shown at postcode level.

The data contained within the various OEHL database contains many potentially valuable pieces of information. The value of the data has been increased in the current project by linking it at the property address level. The value can be further improved by including an independent variable (e.g. property address level energy consumption or property value) to allow the influence of various rebate schemes to be compared.

Socio-Demographic



The Socio-Demographic theme is largely based largely on publicly available ABS census data, which is only available at aggregate levels. This theme is designed to help to understand how the housing in different areas relates to the demographics of those areas. This can help government ensure that the housing we have, and the housing that is planned, meets the requirements of the residents.

Housing Stock



The Housing Stock theme collates our understanding of the housing stock from different data sources and can help planners understand what type of buildings exist in NSW, what types of buildings are being constructed, and where.

BASIX Energy Score















The BASIX energy theme presents some basic energy consumption information available from the energy providers. Some analysis is provided of the influence of various factors on the BASIX energy score. Although BASIX energy score is related to the reductions of greenhouse gas emissions from a given benchmark, it can be used as a proxy of efficiency outcomes affecting energy consumption.

6. Deliverable 5 – Web embedding

One of the limitations discovered in Phase 1 of the housing stock mapping project was that users were required to log into the business intelligence platform Yellowfin to engage with the visualisations. A licence is required for each user and strictly speaking all users must be named users with no users able to share log in details. For phase 2 of the project UOW has decided to take advantage of the Yellowfin Javascript API.

Using the Javascript API, Yellowfin Reports and Dashboards can be embedded into external webpages, which can be hosted on a separate web server to Yellowfin. The only requirement is that the end user's browser can access the Yellowfin server.

The embedded report contains a Title Bar and the Report itself. The Title Bar contains the name of the report, as well as buttons to control any interactive elements used on the report. The interactive features include:

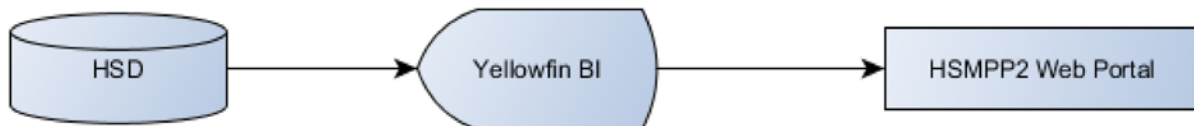
	DESCRIPTION
	Displays basic Report Information including; Name, Description, Last Modified Date, Category, Sub Category, and Current User
	Displays the Javascript used to embed the report
	Allows the user to toggle to Chart view if Table and Chart display is used
	Allows the user to toggle to Table view if Table and Chart display is used
	Opens a panel of User Prompt Filters enabled on the report
	Opens the Export panel, allowing the user to export the report to various formats, including: <ul style="list-style-type: none">  Export to CSV  Export to DOC  Export to PDF  Export to Rich Text Format  Export to Text  Export to XLS

The current limitations of the Javascript API include:

- Only one dashboard is supported per HTML page
- The calendar selector for date filters is not available
- Drill Anywhere is not available
- Saved Filter Sets are not available
- Chart Animation is not available
- The Multi-Chart Canvas is not available
- Drill Through and Drill Down on a Map
-

6.1. WEB PORTAL ARCHITECTURE

The web portal was designed and developed using the WordPress platform on top of a Yellowfin Business Intelligence (Yellowfin BI) and the Housing Stock Database (HSD) developed for the project. The architecture of the web portal is shown below.



The HSD is a PostgreSQL database which is sitting on a PostgreSQL server and contains the sourced and derived data. The Yellowfin BI connects to the HSD and is used to build and host reports and dashboards for visualisation. The HSMPP2 web portal is built in a WordPress platform, which contains embedded reports and dashboards from the Yellowfin BI.

The web portal provides public and subscribed services to different types of users. Users of the HSMPP2 Web Portal can access the reports and dashboards on the Yellowfin BI according to assigned privileges.

7. Recommendations and Opportunities

The HSD in Phase II has extended the geographic coverage from the Illawarra region to the NSW state-wide. It converted and aggregated the available data to postcode level. Throughout this project a number of potential future applications for the NSW housing stock map were identified, as well as several opportunities for future research.

7.1. POTENTIAL APPLICATIONS

In its current format, the centralised database can support a range of applications, for a range of stakeholders. The inclusion of additional data, most importantly property address level utility data can greatly increase the usefulness of the current database. Some potential usages envisaged include:

- **Government** - The data centralised in the database can support and inform effective planning; for example, by illustrating areas with higher, or lower densities of particular attributes. Further, the database can be used to: more effectively target the delivery of new government and market driven activities; provide a benchmark for evaluation; and support OEH's media and communication activities. The inclusion of property address level utility data would vastly increase the usefulness of the data set, for instance by allowing OEH to analyse whether participation in rebate schemes is linked with lower energy consumption, or DPE to assess the factors affecting energy consumption from BASIX-affected dwellings after occupation.
- **Energy Utilities** - The NSW Housing Stock Mapping database can also provide many benefits to energy utility companies. The tool can help identify potential energy saving opportunities, areas of high penetration of distributed generation technologies (e.g. PV panels), and to anticipate demand on the network at a high resolution by understanding the distribution of different appliances. In the future it could be used to track installation of energy storage technologies, and any grid disconnection that may occur. The energy supply companies could also use the database to provide specific, targeted information to customers on their energy performance in comparison with appropriate benchmarks.
- **Householder** - If the database and visualisation were made publicly available, a householder would be able to benchmark the economic and sustainability performance of their home against similar homes in their region. Benchmarked average performance values according to various building features and attributes can also be calculated across the region, to allow homeowners to compare

their property against similar homes. This will give householders an insight into particular features that may improve their home and/or household practices that can reduce their energy consumption.

- **Industry and Businesses** – The database can provide quantitative information on current market conditions and market penetration of a range of technologies, to inform product and service development for relevant businesses and corporations servicing the residential building sector. This information can help provide greater certainty in investment decisions, as well as aid in the development of products and services that most effectively address the shortcomings in the economic, social and environmental performance of NSW dwellings. In the future, linking energy utility data with on-site generation and storage information, and appliance details (e.g. heating and cooling systems) would allow industry to actively identify potential customers for energy efficiency to generation interventions.
- **Researchers** – The HSD can form the basis of further study where academics and analysts can interrogate and analyse different datasets. It will provide a more integrated view of operations and markets, as well as broader access to data enabling encouraging multi-disciplinary research.

7.2. FUTURE AND FURTHER DEVELOPMENT

The following section lists and describes opportunities that can be explored through further development on the outcomes delivered and achieved in the HSMII.

Spatial Expansion and variable spatial scaling

At present the Phase II project has focussed on preparing data for visualisation at the postcode level. To fully realise the utility of such a stock mapping tool it should be expanded to include other available and informative levels for government management and infrastructure planning. These levels may include SA2 and LGA.

Aligning OEH data with forthcoming NSW Digital Market Place requirements

NSW Government has commissioned the development of a centralised data exchange system, called NSW Digital Market Place that will support its ambitious Open Data – Open Government strategy. Digital Market Place will act as a data brokering system whereby all NSW Government departments and agencies will commit relevant data sets through APIs. These datasets will be made accessible through metadata repositories and discoverable through visualisation tools developed by NSW Spatial Services (NSW Live for example). Henceforth, it is essential for NSW-OEH to make sure that data structure and content will meet the requirements imposed by NSW Digital Market Place. Although, the research team has made all

reasonable efforts to develop an efficient data structure and create relevant metadata records, it is worth mentioning that recurring issues encountered with residential addresses across various data sets will need to be addressed as NSW Digital Market Place will use a GNAF based address verification protocol.

Pilot project to align HSM project with other initiatives

There is a strong incentive for NSW-OEH to align its Housing Stock Mapping initiative with equivalent or complementary projects like CSIRO's Energy Use Data Model (EUDM):

<https://www.csiro.au/en/Research/EF/Areas/Electricity-grids-and-systems/Economic-modelling/Energy-Use-Data-Model>) or NSW-DPE's NSW Housing Monitor (<http://www.planning.nsw.gov.au/Research-and-Demography/Research/Housing-Monitor-Reports>)

Defining data standards for future programs

As discussed in the discussion paper (Perez *et al.*, 2017) the definition of standard attribute definitions consistent classification for housing stock data is an important task. This can assist any future OEH programs to collect data using appropriate definitions, and to store the data in an appropriate in an appropriate format to allow it to be integrated easily with other data sources. For instance, the current Home Energy Action appliance and community housing programs have an opportunity to provide additional data, but are likely creating custom data collection methods, which may or may not align with previous program data.

This action has increased urgency as voluntary and mandatory disclosure of home energy rating approaches. The forthcoming Home Energy Rating System presents a unique opportunity to greatly improve the dwelling data that is available to researchers, government and home-owners. Care should be taken to ensure in the final design of the system to ensure the data collection tool align with existing data sets, and where possible collects information to address key data gaps. Importantly for NSW the Home Energy Rating System data should align with definitions used by ABS, BASIX, and other OEH programs, such as the Energy Saving Scheme Home Retrofit Method.

Improving and Enriching the Database

The data contained within the database could be improved by:

- **Sourcing and processing further property data.** It is recommended that OEH focus efforts to access the following datasets, which can help fill key data gaps, in the near term:
 - Energy Utility Data. At present, the usefulness of the HSD is most limited by the lack of property address level utility data. The inclusion of energy utility data would allow analysis of the energy consumption implication of different building attributes, demographics, climatic variables etc.... Energy consumption data is most valuable when it can be linked to a specific property address, however once linked, the data can be de-identified and

aggregated for visualisation without compromising the usefulness of the information.

International experience has suggested that negotiating access to utility data may be a long process, but there is substantial benefits to be realised.

- Land and Housing Corporation Property Assessment Survey data. Initial discussions were held with LAHC regarding this data set. It was not able to be sourced within the timeframes of the current project. However it seems likely that OEHL would be able to gain access in time. The dataset contains relatively detailed property information for approximately 115,000 public housing properties in NSW, updated based on physical assessments by trained surveyors on a three-yearly basis. Importantly, the LAHC dataset includes age of construction, which would be a valuable piece of information to enable inferences to be made on the bulk of the existing building stock for which limited information is available.

A number of additional, useful, datasets for potential future incorporation are listed in Section 3.4.1 of the pilot project final report. However the experience of this project has highlighted the difficulty in preparing a comprehensive list of datasets. It is therefore likely that there are more useful datasets existing, which could be sourced on an ongoing basis.

- **Addressing Data Gaps.** Through the pilot and follow up stages of this project some key gaps in the currently accessible data related to the existing NSW housing stock have been identified. This issue was explored in detail in the pilot project, and a detailed list of data gaps was included therein. As identified above, some of these data gaps can be addressed to a degree through the inclusion of additional data sources (e.g. utility data, building age, and sales data). Large scale data collection exercises could also be designed to fill these gaps, and to provide validation (i.e. BASIX datasets) of the included information.
- **BASIX data.** This project has identified a number of specific actions that could improve the utility of the BASIX data set in the future, including i) collect and integrate revised BASIX certificates of single dwellings to the existing DPE database; ii) explore method for signposting dual occupancy certificates within the DPE database; iii) explore methods for flagging certificates that do not link to a construction certificate record through the planning portal

Development of Sector Specific Tools and Functionalities

As identified in the pilot project, the database is likely to be of use to a range of stakeholders as ascertained by the stakeholder consultation survey process conducted as part of this project phase. A recommendation of this project is to further improve the current tool and develop sector specific tools to fully unlock the benefits of the database for each stakeholder. This may require:

- Adjust the tool to support less technical users, ensuring it is more intuitive e.g. integrated user manuals, data definitions and report descriptions; flexibility to change data visualisation views as preferred by the user flexibility to change visualisation type (pie chart, bar etc.).
- Adjusting the tool to enable further flexibility around data views and creation e.g. ability to change data visualisation between pie/bar/line charts as preferred by the user; ability to create analysis views and visualisations on the fly.
- Developing enhanced and targeted visualisation and user specific tools and/or functionalities. This may include tools for visualisation of market penetration of certain technologies, predictive modelling visualisation tools incorporation synthetic populations to support what-if scenario analysis, and visualisation for the evaluation of specific interventions.
- Incorporating a range of normalisation factors and algorithms to provide benchmarking tools to rigorously compare the impact of initiatives against appropriate control populations.

8. References

- Safadi M, Ma J, Berryman, M, Clancy, D, Perez P. (2016). *NSW Resource Efficiency Baseline Project: Final Report*. University of Wollongong, NSW, Australia.
- Safadi M, Ma J, Daly D, Berryman, M, Clancy, D, Perez P. (2016). *NSW Housing Stock Mapping Pilot Project: Final Report*. University of Wollongong, NSW, Australia.
- Perez P, Daly D, Ma J, Davies T, Kokogiannakis G, Cooper P, (2017). *NSW Housing Stock Mapping Project: Phase II Discussion Paper – Data Collection & Formatting*. University of Wollongong, NSW, Australia.
- Daly D, Kokogiannakis G, Cooper C, Perez P, Ma J, Davies T. (2017). *NSW Housing Stock Mapping Project Phase II: BASIX data validation pilot study*. University of Wollongong, NSW, Australia.