CRC for **SPATIAL INFORMATION**

crc•si

A CRCSI Collaboration project in conjunction with

MERCURY PROJECT SOLUTIONS

BUSINESS ASPECT

The Future of Geocoded Addressing in Australia

5 March 2015

Final Version 4.0

Disclaimer

This report has been prepared by Mercury Project Solutions on behalf of the Cooperative Research Centre for Spatial Information (CRCSI) for the benefit of its Government College Partners and ANZLIC. The CRCSI does not accept any responsibility whatsoever for, or in respect of, any use of or reliance upon this report by any third party.

This report contains technical detail and comments from a vast range of stakeholders and jurisdictions which may, or may not, apply to all of those stakeholders or jurisdictions. The report outlines a range of potential opportunities and discussion points for the stakeholders involved in the geocoded address supply chain but in no way dictates a prescribed set of actions. The recommendations provide examples of how the issues raised by stakeholders, may be resolved. The next step in the consultation process will be to assess the recommendations; assign owners and develop sound proposals, where appropriate, and which are aligned them to the strategic objectives of those organisations.

The CRCSI would like to thank the Government College and ANZLIC for the opportunity to deliver this study; its 43PL partners Business Aspect and Mercury Project Solutions and the many stakeholders (including PSMA Australia Limited and Data Policy Branch) who spent a great deal of inkind hours working with the team to deliver this report.

Contents

Executive Summary

Geocoded addressing in Australia has evolved over the last 40 years. In 2004 the Geocoded National Address File (G-NAF) was launched. Since then, G-NAF has taken up a central position in supporting Australia's digital economy, and has become the authoritative benchmark for geocoded address data in Australia.

The landscape continues to change. Continual and significant improvements in information technologies mean ever evolving user expectations regarding quality, accessibility and timeliness. The Open Data movement is unstoppable, and the Federal Government's and several jurisdictions' policy position is that foundation data (including addressing) will become ubiquitous in all sectors of the economy. The private sector is responding, and several alternatives to G-NAF have become available in the market, as provided by for instance Sensis (WhereIs), Google or HERE maps, and left-field initiatives such as Geepers.com.

As our digital economy grows, so does the need to have a trusted, ubiquitous National Geocoded Address dataset that can be used to link an increasing range of complex information to a unique address, location, house, property, premise or residence.

The current state of the national geocoded address supply chain, which underpins G-NAF, will not be able to fully support this vision, due to a number of systemic limitations such as being limited to property-based addresses only.

The CRCSI, in collaboration with ANZLIC, has commissioned this Geocoded Address Optimisation Project (GAOP), a study into the future of Geocoded Addressing in Australia, as one of the key themes addressed by the Foundation Spatial Data Framework (FSDF). GAOP focuses on optimising the Geocoded Address supply chain by base-lining, identifying and optimising the Geocoded Address supply chain.

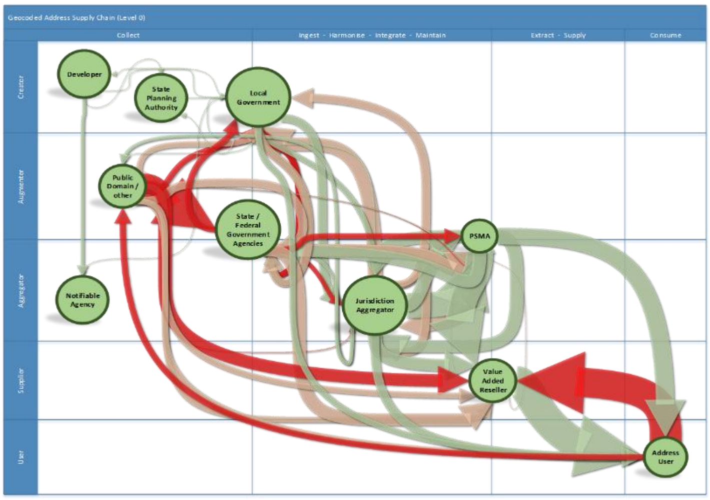
This study focuses on three objectives:

1. Describing the current state of the supply chain;
2. Identifying user needs and short-term improvement opportunities;
3. Developing a long-term future vision for geocoded addressing, and a roadmap towards that vision, identifying proposed changes and additions to the supply-chains, needed to achieve that vision.

Current State

In one sentence, the current state can be summarised as *"many organisations doing similar things with the same data by different address custodians on different timeframes with differing results and, costing a lot."*

The current geocoded address supply chain was not designed, it simply evolved over the past 40 years where local knowledge was often the best resource for identifying location.



Geocoded addressing as currently provided by G-NAF is too limited in scope by constraining itself to property-based street-addresses, is inherently complex, non-linear and in many aspects convoluted. This creates contradictory evidence in applying confidence levels to address verification and geocoding processes, and reduces data currency. Millions of unverified addresses are captured annually and follow various paths through the supply chain, often involving considerable duplication and manual intervention. While most addresses eventually become verified and many ultimately geocoded, a substantial minority will not be verified (how much this last point is an issue, is yet to be determined).

These systemic issues with the current supply chain for geocoded address data in Australia lead to a national dataset that:

- Is cost-inefficient;
- Has significant duplication of effort, inconsistencies, and ambiguities;
- Is based on a supply chain that is structurally unable to provide currency levels that many users expect;
- Doesn't allow users to identify non property-based locations, or locations that don't have a (complete) street address;
- Makes it impossible or cumbersome for citizens to report errors or updates;
- Is not designed to support emerging requirements and use-cases.

User Needs and Short-Term Improvement Opportunities

To get the most benefit out of geocoded addressing, users consistently indicate that they need data that is timely, accurate, robust, comprehensive, accessible, secure, authoritative, shareable and cost-effective.

A gap-analysis between these needs and the current state resulted in a number of recommended improvement opportunities:

1. Undertake a review of national address standards, to deal with the inconsistencies in definitions and terminology being used across the geocoded address supply chain;
2. Provide an unrestricted one-time copy of G-NAF for non-commercial use, to promote and facilitate take-up of the dataset by the development community;
3. Develop a specification for national geocoded address interfaces, to enable exposing APIs for address data verification, update and notification web-services;
4. Initiate a nation-wide survey of geocoded address requirements to a wider group to include all jurisdictions, local councils and user communities; and;
5. Develop a "National Geocoded Address Portal" available to local councils, Notifiable Agencies and federated to the Jurisdictions.

Future Vision: A National Geocoded Location Management Framework

The longer-term future vision (beyond the year 2020) for the use, creation and management of geocoded locations in Australia is presented below:

By the year 2020, geocoded communication of location in Australia will comprise the infrastructure, processes and knowledge that enable accurate, trusted, timely and unambiguous translation of a descriptive tag to a place in time and 3-dimensional space.

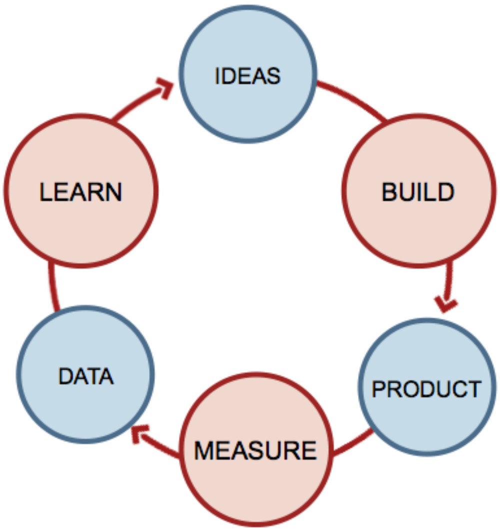
The broad community captures, maintains and shares geocoded location data. Confidence in the data comes from its single source of truth, its transparency and auditable provenance. Government's role is that of a facilitator, ensuring trust, setting relevant standards and (open data) policies, while the market provides value-add, fit for purpose channels.

The opportunity now is to design an efficient and effective **National Geocoded Location Management Framework** with a better understanding of user need, taking advantage of advanced information and communications technologies to provide an enhanced, scalable and sustainable foundation for location based services.

In this framework, users have access to geocoded location data that

- is trusted (which doesn't necessarily mean 'authoritative');
- enables point to point navigation, both indoor and outdoor, in 3D and in time;
- has a greater resolution than is available through 'traditional' notion of address;
- allows greater ease of communication of location (e.g. not just 29 Jardine Street, Kingston ACT, but also Belgian Pub in Kingston, or even 4sq.com/53qr6I); and
- supports fit for purpose geocodes, access points, and tags.

The framework will be: characterised by a disruptive approach, to overcome the systemic issues in the current supply chain; rely heavily on crowd-sourcing for data capture, updates and validation; extend the scope of locations beyond traditional property-based addresses; incorporate contemporary and emerging technology enablers such as the semantic web and linked data; and redefine the role of government from a data provider to a facilitator.



The scope of the future vision goes well beyond the current G-NAF product and its underlying jurisdictional, currently defined as authoritative, supply chain. However, it is expected that G-NAF and the jurisdictional supply chain will continue to form a key component of any future geocoded location management framework.

Given that many of the concepts and technologies are emerging, implementing this future direction will be a largely experimental process towards a solution that will evolve and materialise over time. The roadmap will need to have relatively short development cycles, regular progress reviews and measurable outcomes, incorporating emerging R&D outcomes and international best-practice, and could adopt a 'lean start-up' methodology.

1 Introduction

1.1 Background: Optimisation of the Geocoded Address Supply Chain

ANZLIC—the Spatial Information Council, representing the Australian and New Zealand Governments, and the governments of the States and Territories of Australia, has initiated a project to deliver the Foundation Spatial Data Framework (FSDF). When realised, the use of a common framework, embedded into the everyday business of government and private sector entities, will allow for seamless exchange of spatial information and knowledge across organisational, sector and jurisdictional boundaries.

Geocoded addressing is one of the foundation data themes being addressed by the FSDF. It is a key enabler for many organisations. As our digital economy grows, so does the need to have a trusted National Geocoded Address product that can be used to link an increasing range of complex information to a unique address/location/house/property. Applications for addressing will be more complex in the future (e.g. 3D, increased accuracy, unmanned aerial vehicles, etc.) and the demand for addressing will be greater and vital to the digital economy. It is important that current capabilities are streamlined and structured to be ready for future changes.

A number of projects have been identified to help deliver the CRCSI's Program 3 – Next Generation Spatial Infrastructure's objectives, namely improved search, discovery, access and delivery of spatial knowledge, utilising semantic web technologies (see Spatial Infrastructures Research Strategy http://www.crcsi.com.au/Research/3-Spatial-Infrastructures).

Separate from the Program 3 research projects, this Geocoded Address Optimisation Project (GAOP) is the first of a series of 'short and sharp' high impact projects that are aimed at solving current issues.

GAOP focuses on optimising the Geocoded Address supply chain by base-lining, identifying and optimising the Geocoded Address supply chain. This report presents the results of Stage 1 of GAOP. It examines the current supply chain and scope specific and immediate opportunities for development improvements in a range of areas identified by key stakeholders as well as looking at future needs in the short term, and defining a longer term future vision.

1.2 Project Summary

The recommendations from Stage 1 of this Project– the Review, may be used to implement changes in the Geocoded Address Supply Chain. These changes will be designed to have the greatest potential to improve the efficiency of the production of the Geocoded Address Supply Chain and the user experience. The review will also make recommendations about improving the "future state" (post 2020) of the supply chain. This "future state" will be based on the current FSDF needs analysis that commenced in 2013, CRCSI Supply Chain Research and detailed user analysis obtained from Stage 1 of this project.

This project aims to improve accuracy, timeliness and customer experience for the Geocoded address supply chain. These changes may include technical, business process, utilisation, education, product development, policy or legislative.

1.3 Scope & Objectives

The objectives of this first stage are to:

- Identify, and document the current 'as-is' state of the geocoded addressing supply chain in Australia;
- Outline a defined set of end user needs;
- Define a goal future state of the geocoded physical address supply chain;

- Develop proposals for 'greatest impact' (short term) improvements to the geocoded addressing supply chain;
- Develop a roadmap for moving towards a longer term (post 2020) 'future state'.

Explicitly excluded from the scope of works are:

- The business models of PSMA;
- Improved purchasing processes for G-NAF by the Australian Government; and
- The nature or extent of economic benefit for G-NAF.
- Further, this review will leverage and not duplicate outputs from other projects that have been recently completed or are underway (e.g. NAMF Review, G-NAF Review – KPMG, ICSM Local Government; Jurisdictional Addressing Review and the new ISO addressing Standards).

1.4 Related Documents

This document builds upon the document "Geocoded Address Supply Chain Review, CRCSI – Program 3", (Business Aspect, August 2014), which analyses in detail the current state of the supply chains, identifies user requirements and makes recommendations for short-term improvement opportunities to the existing supply chain.

The outcomes of this review are summarised in section 2 of this report.

1.5 Report Structure

This report is structured as illustrated in Figure 1. It distinguishes between the description of the current state in 2014 (section 2.1), the needs of the current user-base and how the current state can be improved in the short term (section 2.2), a future vision for the year 2020 and beyond in section 3, and a possible roadmap to start working in the future direction in section 4.

Section 5 lists a number of relevant case studies and references that demonstrate specific aspects of the vision.

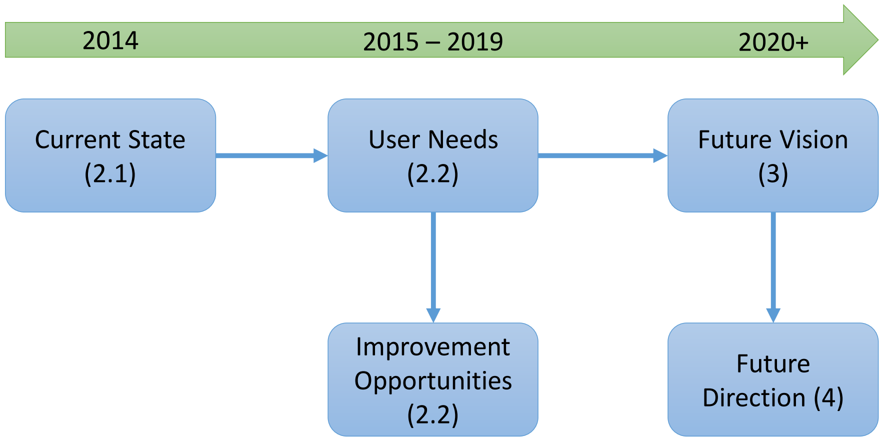


Figure 1 Report Structure

2 Current State and User Needs

As part of this project, an in-depth review was conducted on the current state of the national geocoded address supply chain, to establish a baseline, identify 'immediate' user needs and short-term improvement opportunities.

This chapter presents the highlights from the current state and user needs report[1].

2.1 Key Findings – Current State

In one sentence, the current state can be summarised as **"many organisations doing similar things with the same data by different address custodians on different timeframes with differing results and, costing a lot."**

Essentially, addresses are labels used to identify a location. There can be many labels associated with the same location as well as many locations associated with the same label. The challenge has been to reliably verify the correct address by using as much available information about a location from a variety of sources so that it can be reliably geocoded and used.

The geocoded address data lifecycle represented in this supply chain analysis of the current state emphasises the upstream or supply side activities prior to presentation to PSMA. Downstream activities through resellers and end users represent the demand side characteristics that are captured in the subsequent user needs analysis. In this context, the geocoded addressing supply chain conceptual model, described in the diagram below, comprises five (5) fundamental processes and three (3) value tiers, which is effectively multiple times along the supply chain.

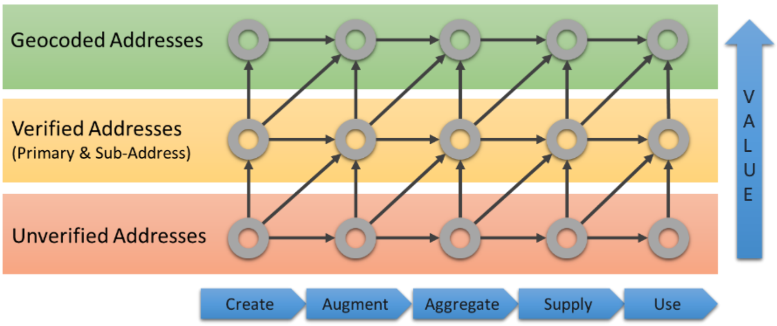


Figure 2 Geocoded Address Supply Chain Conceptual Model

G-NAF, produced and supplied by PSMA, is the subject of this supply chain model and offers high value geocoded addresses for business use. Additional "value" is created in the supply of G-NAF addresses to end users through innovative processing, searching and matching routines provided by Value Added Resellers (VARs).

The current geocoded address supply chain was not designed; it simply evolved over the past 40 years where local knowledge was often the best resource for identifying location. At the same time, continual and significant improvements in information technologies are being applied to accommodate an ever-increasing demand for identifying the correct location of addresses in support of a multitude of government, business and community functions.

Stakeholder interviews, documentation reviews and industry analysis produced the following key findings of the current state (as-is) of the geocoded address supply chain:

---

[1] As described in the report: "Geocoded Address Supply Chain Review, CRCSI – Program 3", Business Aspect, August 2014.

1. There is considerable **variation in the understanding of terms and definitions** of addressing throughout the supply chain. This leads to inconsistencies in how addresses are created, interpreted and managed across the different state jurisdictions as well as variability between local government authorities and their respective state jurisdictions.
2. Local Government and hence the land valuation system is interested in the “property” address yet the surveyor and Land Titling system is interested in the “parcel” for identifying an associated address. This **dichotomy of address references** causes a large majority of the problems in resolving the correct or most likely address in the aggregation processes at state and national levels.
3. The only legislated creators of addresses established through the land development process are Local Government Authorities (LGA). These addresses are either primary addresses (in the case of lot subdivisions) or sub-addresses (in the case where strata titles are created). However, **addresses can also be created by other land planning and administration processes that are not necessarily validated by the LGA process**. These addresses may never appear in G-NAF, or become a duplicate of an address that already exists in G-NAF without anyone even knowing.
4. There is no legislation in place to control address creation through the naming and numbering of streets and properties in private estates and complexes such as retirement villages, universities and hospitals. While these “**private” addresses fall outside the normal council processes** because they don’t affect land ownership, they are never-the-less an important additional address dataset which should be captured and managed as well as possible.
5. PSMA is recognised by all supply chain participants as the de-facto authoritative source of national geocoded addresses, being the G-NAF products. However, some users suggest that **G-NAF needs to be more robust to better meet the needs of key emerging markets**. This implies improving the governance related to data lineage to be recognised as a truly authoritative geocoded national address dataset. Authoritative source also implies that PSMA have more control over the upstream supply chain than is currently the case.
6. The supply chain is complex, non-linear and in many aspects convoluted, which creates **contradictory evidence in applying confidence levels to address verification and geocoding** processes using reference address files compiled from similar sources. Millions of unverified addresses are captured annually and follow various paths through the supply chain, often involving considerable duplication and manual intervention. While most addresses eventually become verified and many ultimately geocoded, a substantial minority will not be verified (how much this last point is an issue, is yet to be determined). Similarly, this leads to duplication of effort by stakeholders re-geocoding and re-validating addresses.
7. **Time delays from address data creation at local government level to then becoming available through G-NAF significantly reduces the usefulness** of G-NAF to many users. This encourages by-pass mechanisms and work-around strategies by VARs and users to acquire addresses, however poorly managed, that meet their business requirements. These delays are usually a by-product of the land titling system where addresses are established when the development application is approved by the planning authority but not incorporated into G-NAF until the new property has been processed in the land titling system, which can be many months later.
8. **Several competing initiatives to create “authoritative” national geocoded address datasets as potential alternatives to G-NAF** are being pursued aggressively by different government agencies and the private sector. Some of these are being justified by the internal business requirements of specific agencies such as: limitations imposed by privacy legislation at the Department of Human Services; operational imperatives for infrastructure connectivity at NBN Co; and potential for revenue generation opportunities at Australia Post.
9. Supply chain dynamics are largely passive and rely on progressive movement of address data along the supply chains from local government level to state and national levels of aggregation before becoming available in G-NAF. This limits the optimisation opportunities since **creation and aggregation processes are invariably linked to the lengthy land administration processes in each jurisdiction**.

10. Sharing of geocoded addresses (sourced from G-NAF) between government agencies is impeded by the current licensing model designed to support the existing funding model and protect the existing PSMA distribution channels. While guaranteeing product sustainability, this licensing regime limits business improvement initiatives for government agencies needing to collaborate in providing more cost effective and efficient community services.
11. **Addresses have no authorised owner**. PSMA develops a comprehensive, national geocoded address dataset but has no control over the addresses themselves which could be changed without notice. The G-NAF address set that it creates could be seen as “optimal“, but is not “authoritative” as there is no authority that is ultimately accountable.
12. The reverse information flows, or **back-channels, from Users to VARs to PSMA to Jurisdictions to Councils are either ad-hoc, informal and undocumented or non-existent**. This weakens the whole geocoded address supply chain, as valuable user feedback which could be used to improve the quality of both the address datasets and the systems that constitute the supply chain is lost.

2.2 User Needs, Recommendations and Improvement Opportunities

User needs are reflected in the findings of the current state analysis and the stakeholder expectations for an improved geocoded national address product or service through supply chain efficiency improvements that will reduce their costs and potentially deliver increased benefits to their business.

The following table summarises the current state issues, user needs and associated short-term improvement recommendations that are based on the gap analysis. These are grouped by issue categories if inconsistencies, supply chain dynamics, technology and governance.

For a detailed overview of the user needs, gap analysis and (short-term) improvement recommendations, please refer to the document “Geocoded Address Supply Chain Review, CRCSI – Program 3”, (Business Aspect, August 2014).

Note that the methodology in this review, by necessity, restricted the user-needs analysis to a limited set of stakeholders (six jurisdictions and two commercial entities). The future recommendations therefore include a much wider user-needs analysis exercise (see section 4.3.1).

Issue Categories	Current State	User Needs	Recommendations	Short-term Improvement Opportunities
Inconsistencies	Variation in the understanding of terms and definitions of addressing throughout the supply chain. Dichotomy of parcel vs. property address references leads to national aggregation problems. Inconsistencies and contradictory evidence in address validation and aggregation. Private addresses not available in G-NAF.	**Accuracy**: Users want the data present in the address dataset to be unambiguously identified and as accurate as possible. **Comprehensive:** users need to have a reliable, comprehensive national addressing system that allows for a wide variety of address types including the relationships between addresses. **Cost-effective**: Better pricing and access conditions for address products and services would lead to increased take-up.	Jurisdictional address verification services must support address information feedback. Councils should have responsibility for managing all addresses. Jurisdictions must implement processes to ensure capture of addresses for all development activity. Make the use of the National Geocoded Addressing service cheaper.	**Leveraging address notifications and in-field validation of addresses.** By using these address to update the jurisdictional geocoded address datasets that flow to PSMA. **Undertake a review of national address standards.** The CRCSI could initiate a review of the ISO standards to deal with the inconsistencies in definitions and terminology being used across and the geocoded address supply chain.
Supply Chain Dynamics	Supply chain is complex and non-linear, leading to duplication of effort, unverified addresses, and time delays. Passive and reactive processes, lacking right incentives, in most nodes along the supply chain. Many non LGA generated addresses, which do not end up in G-NAF (orphaned addresses).	**Timeliness**: updates to be available within days. **Robustness:** users need a robust, stable and auditable address dataset.	Addresses created should immediately be accepted by a jurisdictional address management function. Implement support for address history and traceability.	**Develop a "National Geocoded Address Portal"** Providing the right tools and incentives for Local Councils, State and Territory Jurisdictions, and PSMA to interact with a common geocoded address dataset that can be updated and maintained by authorised address data custodians in respective jurisdictions. The incentives for stakeholders to participate must be clear, and any potential duplication issues must be addressed by technology/business

Issue Categories	Current State	User Needs	Recommendations	Short-term Improvement Opportunities
				processes as resources to participate at local government level are potentially very limited.
Technology	No back channels for user updates available.	**Accessibility:** users need to access foundation G-NAF data, as well as value-added services. **Security:** a more accurate, more accessible, more reliable and comprehensive national address service.	Provide access to G-NAF as a Service; easy to access using standard tools, using standard web-service APIs. Develop a security model to support Councils and jurisdictional aggregators to update and maintain G-NAF.	**Provide an unrestricted one-time copy of G-NAF for non-commercial use** This will stimulate the development community to learn the data structures and start developing apps and web sites with G-NAF capabilities. **Develop a specification for national geocoded address interfaces** For exposing address data verification, update and notification web-services.
Governance	Data lineage not mostly not available. G-NAF is an authoritative dataset, but PSMA as aggregator has no authority over upstream processes. Competing datasets emerging as G-NAF alternatives. G-NAF licence model limits	**Authoritative:** currently no addresses in Australia are truly “authoritative”. **Shareable:** it should be possible for users of the address dataset to communicate address information to each other easily and quickly, without restrictive licensing terms and conditions	Councils should be designated as authoritative source for parcel and property addressing Councils need to be empowered as the authorities for all land parcel or property addresses Develop jurisdictional-level services to hold and maintain authoritative address dataset on behalf of all the councils in a jurisdiction.	**Initiate a nation-wide survey of geocoded address requirements** The CRCSI should extend the stakeholder involvement to a wider group to include all jurisdictions, local councils and user communities

Issue Categories	Current State	User Needs	Recommendations	Short-term Improvement Opportunities
	usage and uptake.		Improve licence agreements to make sharing of G-NAF-based data easier across Government.	

3 Future Vision

Where the previous chapter focussed on the current state of the supply chain, and how it may be improved in the short- to medium term, this section presents a longer-term future vision (beyond the year 2020) for the use, creation and management of geocoded locations in Australia.

In this vision we define a goal 'future state', gathered from information gleaned from various national and international stakeholders, and from existing and emerging initiatives.

The vision focuses on the greatest benefit and utility for the end-user, taking into consideration community, industry and technological trends, in particular semantic-web capabilities.

3.1 Vision Statement

By the year 2020, geocoded communication of location in Australia will comprise the infrastructure, processes and knowledge that enable accurate, trusted, timely and unambiguous translation of a descriptive tag to a place in time and 3-dimensional space.

The broad community captures, maintains and shares geocoded location data. Confidence in the data comes from its single source of truth, its transparency and auditable provenance. Government's role is that of a facilitator, ensuring trust, setting relevant standards and (open data) policies, while the market provides value-add, fit for purpose channels.

3.2 User Stories

The following stories depict some example narratives as an illustration of the how the vision might work in practice from the year 2020 onwards. The user stories should be read as thought provokers, rather than defining the ultimate scope and extent of the future vision.

3.2.1 Users

*It is the year 2020, and in the comfort of her office, **Alex**, data officer at a federal government department, reflects on the ease with which she can use widely available, reliable, current, geocoded location data in her daily work.*

Whether it's for verifying customer details, ensuring correct delivery of critical services, or doing policy research, having easy access to trusted geocoded location datasets that are timely, accurate and easily shareable means that she can focus on solving her core business problems, delivering quality services in a cost-effective way.

Whenever she receives a customer address, it is automatically checked if it exists in the national geocoded location register, and if so, is given the unique ID for that location. If it doesn't exist, she has the option to submit a new alias for the location, or to submit a new location to the register.

How different were things back in 2014: There was an authoritative geocoded dataset available in the form of G-NAF. While a lot of energy and effort was put into creating and maintaining G-NAF, systemic supply-chain issues meant that desired levels of quality, timeliness, and shareability were inherently unattainable. Only when in the following years, the data supply-chains were successfully transformed, was it possible to achieve the vision of a truly fundamental geocoded locations dataset that underpins the digital economy.

*While **Bernadette** is sitting comfortably in her office seeing all the benefits of the modern geocoded location facilities built into her applications, she doesn't see a lot of the work that goes on behind the scenes.*

She didn't know that her organisation subscribed to an address change notification service so that, on a regular basis, the organisation was notified of changed addresses. Typically this happened when houses were demolished or land subdivided. Previously (in 2014), they had no way of knowing that the address was invalid.

As she was sitting there, she hears a fire engine race by. She turns on the radio, and hears that there's a bomb at the Old Town Hall drinking fountain. She doesn't know the address, but types its name into her online map. Up comes the map, showing the fountain, and she realises it doesn't actually have a street address. But there it is, nestled on Main Street between numbers 214 and 216. She knows this because all the land parcels have the street numbers neatly displayed along the correct road frontage. This had been put place several years ago, and everyone found it so much more useful than the old way of having a number vaguely placed in the middle of a property.

*"It is the year 2020 and **Chris** is finishing year 12. He logs onto MyRegistration (or some equally aptly named web-site) where he will create his profile and register for various Government services. During the registration process Chris selects his residential address from the list provided by the web-sites address validation software. Once his profile is created he then selects the Government services he requires – so he registers with the ATO for his tax rolls, DHS for his allowances, Medicare so he can obtain his card, AEC for electorate purposes and so on. Chris is also able to advise "MyRegistration" that he is happy for his address information to be shared with other governments and commercial providers (e.g. Insurance/banks etc.) with whom he currently has an account, or may in the future.*

A year later and Chris decides to move town to start University. He logs onto MyRegistration and, after passing the identity verification process, changes his address. All "down-stream" users of the address data are automatically notified of the change of address. His insurance company is using advanced analytics which detects the change in address and SMS's him with a reminder to check his insurance cover is sufficient for the new residential area he is moving to, as it has a higher incidence of break and enters recorded.

A year later, Chris's residence is unfortunately affected by a bushfire. The emergency service's Rapid Damage Assessment team visits the site the day after the event and notes on their mapping software that the residence has been partially destroyed. This information is automatically sent through MyRegistration to all consumers of the information – so his insurance company is notified of the damage to the residence without Chris having to contact them directly, the Tax Office is advised so that they don't send him reminders that he has a tax debt, Centrelink and EMA identify that Chris is eligible for a disaster payment and so forth.

3.2.2 Contributors

***Fred** is a farmer in the Riverina. He wants to insure his farm equipment against flood. To be able to assess his risk accurately and give him the most competitive price, the insurance company needs to know the exact location of the equipment shed on his property.*

Fred goes online, finds his property record displayed on a recent aerial image, and updates it to include the geocoded location of the shed.

The insurer sees the updated location, determines it is on a higher elevation than the rest of the property and quotes him a discounted premium.

***Gabrielle** lives in an apartment building in St Kilda. She often shops online, but finds that couriers or grocery stores will not deliver because there is a no-stopping zone at her building's street frontage.*

There is parking and a second entry in the back-lane, but this does not have a street number or official address.

Gabrielle goes online to MyRegistration, and updates the 'delivery' access details for her unit. Courier companies who subscribe to the same service, now know where to deliver Gabrielle's new pair of shoes, and she doesn't have to go and pick them up from the post office.

***Holly** and her friends are on a holiday camping trip, when their car breaks down on a road near the surf club on the beach where they went to for a surf. She contacts roadside assistance.*
It's a quiet area, and she doesn't have an address, but knows it's known as 'breaking bad' on the Facebook check-in page of the surf club. Fortunately, the surf club has opted in to publicly share its Facebook place location with the national location register, so the roadside assistance dispatch centre knows where to send the support vehicle.

***Isaac** is a developer who bought a block of land that he wants to subdivide and build a retirement village on. He has the lot surveyed, and the registers the parcel, as well as the sub-addresses with the government's online service, including the geocoded location of the dwellings. This is immediately available as 'provisional' addresses to relevant stakeholders (such as Australia Post), while a formal validation process is ongoing.*

3.3 Systemic Issues

The current state of the national geocoded address supply chain will not be able to fully support the vision and these user and contributor stories. It is too limited in scope by constraining itself to property-based street-addresses, inherently complex, non-linear and in many aspects convoluted. This creates contradictory evidence in applying confidence levels to address verification and geocoding processes, and reduces data currency. Millions of unverified addresses are captured annually and follow various paths through the supply chain, often involving considerable duplication and manual intervention. While most addresses eventually become verified and many ultimately geocoded, a substantial minority will not be verified.

The current state of the geocoded address supply chain was not designed. It simply evolved over the past 40 years. When mapping the underlying data flows and supply chain(s), it looks like this:

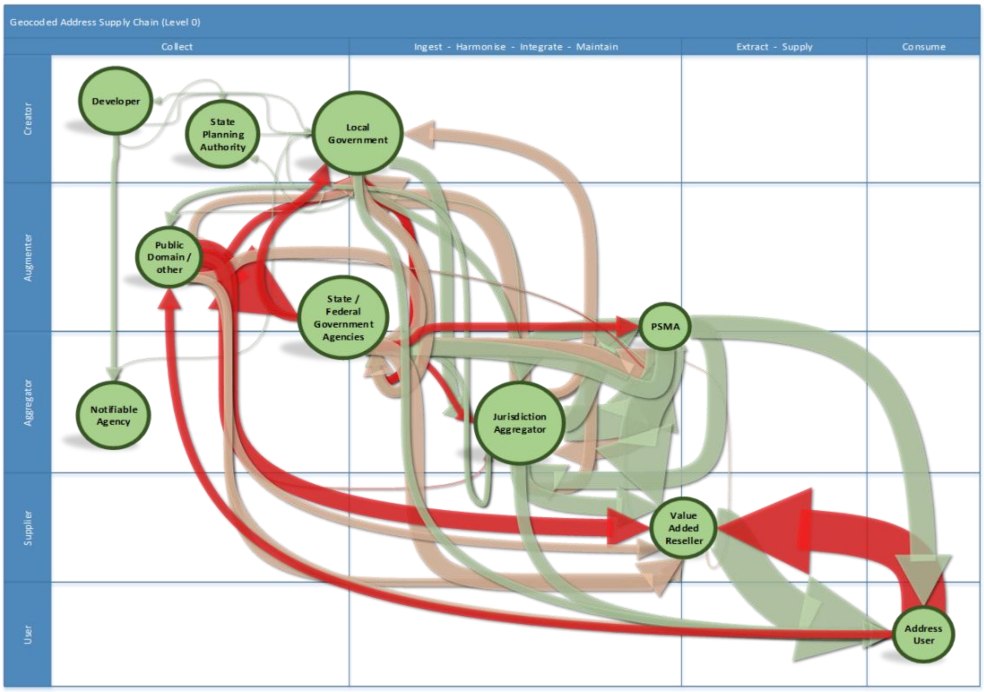


Figure 3 Geocoded Address Supply Chain – Current State Composite. Source: "Geocoded Address Supply Chain Review, CRCSI – Program 3", Business Aspect, August 2014

The current supply chain for geocoded address data in Australia leads to a national dataset that:

- Is cost-inefficient;
- Has significant duplication of effort, inconsistencies, and ambiguities;
- Is based on a supply chain that is structurally unable to provide currency levels that many users expect;
- Doesn't allow to identify non property-based locations, or locations that don't have a (complete) street address.
- Makes it impossible or cumbersome for citizens to report errors or updates.
- Is not designed to support emerging requirements and use-cases.

As an illustration, see an example below for a block of residential and commercial units in Surry Hills, NSW.

Figure 4 One property, many location descriptors

The cadastral unit is known as 99-15 Flinders Street, Strata Plan 48654. It is composed of 24 Strata lots, and 26 'units', as lot 24 has 3 separate tenants or 'shops'. G-NAF lists both the property parcel (99-115 Flinders Street), and 21 units at the same address, all with the same geocode of the parcel centroid.

In reality, to satisfy the myriad of use-cases of residents, tenants and visitors of the building (as illustrated in some of the user stories above), a future system will need to be more fine-grained than parcels or properties. For instance, it will need to identify, label, and geocode the following:

- Postal addresses;
- Visitor addresses, including pedestrian ('frontage') vs. car park access;
- Goods delivery address – which may be different from postal addresses;
- Informal (e.g. Facebook) identifiers;
- Access points that don't have a full address (e.g. Hutchinson Place);
- Multiple descriptions for the same access point (e.g. unit 24/99-115 Flinders St, 1/18 Hutchinson St, or even https://www.facebook.com/suzieqcoffee).

All these are essential for most use-cases, and require labelling and geocoding of locations at a finer resolution than can be derived from property-based addressing.

3.3.1 What do users want?

As has been reported in section 2.2, the initial user requirement analysis established that users want geocoded location data that is:

- Timely (Current);
- Accurate;
- Robust;
- Comprehensive;
- Accessible;
- Secure;

- Extensible;
- Machine readable;
- Shareable;
- Cost-effective.

Furthermore, in our discussions with stakeholders regarding the future vision, it emerged that users also need geocoded location that:

- Is trusted (which doesn't necessarily mean 'authoritative' – see boxed text on page 19);
- Enables point to point navigation, both indoor and outdoor, in 3D and in time (and is 'BIM ready');
- Has a greater resolution than is available through 'traditional' notion of address;
- Allows greater ease of communication of location (e.g. not just 29 Jardine Street, Kingston ACT, but also Belgian Pub in Kingston, or even 4sq.com/53qr6I);
- Supports fitness for purpose geocodes, access points, and tags.

The opportunity now is to design an efficient and effective **National Geocoded Location Management Framework** with a better understanding of user need, taking advantage of advanced information and communications technologies to provide an enhanced, scalable and sustainable foundation for location based services.

What happened to 'Address'?

The reader will notice that the vision for a Geocoded Location Management Framework doesn't mention the word 'address'. In our discussions with the various stakeholders, we found that the meaning of the word can vary greatly between speakers, confusing the discussion. Where ANZLIC defines address quite broadly as *"The descriptive elements describing a fixed location (for example, a plot of land, building, part of a building, way of access or other construction) which is represented by a structured composition of place names and point identifiers"*. Others will interpret it as a street address: the plain English label (albeit hierarchical) describing how to navigate to a (uniquely defined) location by road, thus limiting its scope.

There is a fundamental distinction between an 'address' and a 'location'. Essentially, addresses are labels used to identify a location. There can be many labels associated with the same location as well as many locations associated with the same label. The challenge has been to reliably verify the correct address by using as much available information about a location from a variety of sources so that it can be reliably geocoded.

To avoid confusion, and to ensure the scope of our discussion can accommodate locations that cannot always described by a unique street address, we refer to the term '**Geocoded Location'** for the purposes of the future vision discussion[2].

[2] The scope of the 'location' concept as opposed to address is more powerful, but also potentially attracts more "responsibilities" in terms of deriving a suitable model which is fit for purpose for any 'location'. The conceptual modelling of any location together with all the possible associations between them could turn out to be a substantial undertaking and involve many stakeholders. For the purpose of implementing the vision, the scope could be constrained to types of locations relevant to the use-cases (i.e. for navigating to points of access or service delivery). This would include complex addresses (e.g. retirement homes), landmarks, aliases, and relevant Points of Interest (POIs).

Section 3.4.3 below examines the relation between locations, addresses and geocodes in more depth.

3.4 Key Aspects & Principles

The vision described above can be decomposed into its key aspects relating to the overall approach (or philosophy), people, data, technology and governance.

3.4.1 Disruptive Approach

Systemic issues with the current data supply chain limit the scope of geocoded locations to property-based locations only, and also prevent a number of requirements to be met.

A number of short-term improvements on the current supply chain have been recommended in section 2.2 above. However, these cannot address the inherent systemic issues with the way the supply chain has evolved.

Given that many of the concepts and technologies are emerging, implementing a future, longer-term vision will entail at least in part a disruptive approach that needs to re-assess the entire supply chain, and identify which G-NAF elements can be kept/improved (for property-based addressing), and where additional, non-traditional supply chain elements and approaches are needed as additions.

It will encompass a 'Bazaar instead of the Cathedral'[3] paradigm shift, with a focus on an open, crowd-sourced and distributed model that complements the current jurisdictional supply chain model. The distinction between users and producers of location data will further blur.

Such an approach would also be consistent with the currently emerging and accelerating open data movement.

Hybrid models are already being trialled in many places. For instance in New Zealand, where Land Information New Zealand (LINZ) has made its authoritative road data available to the community for updating in OpenStreetMap (see section 5.1)

With rapidly emerging and evolving user needs and technology advances, it will be critical that any implementation allows for maximum innovation and the adoption of best- and emerging practice, based on open standards for information sharing and linking.

The scope of the future vision goes well beyond the current G-NAF product and its underlying jurisdictional, currently defined as authoritative, supply chain. However, G-NAF and the jurisdictional supply chain will continue to form a key component of any future geocoded location management framework.

3.4.2 People

In the future model, the distinction between users and producers of geocoded location data will become less and less clear. So-called 'prosumers' will use geocoded location data, update data when they find errors, or when data elements change. Prosumers can be individuals, corporate (e.g. developers, insurers), or governments.

[3] Though originally applied to open source software, it equally applies to other open, community based initiatives such as Open Street Map or Wikipedia. See: http://en.wikipedia.org/wiki/The_Cathedral_and_the_Bazaar

Importantly, data contributors will be motivated by market driven incentives: for instance a developer will register a new address and its location so they can more easily or cheaply get utilities connected, a retirement village manager will register access points to its units so emergency services can find them, and a farmer will register the location of their equipment shed to get a discount on their insurance.

To empower the prosumers, using and contributing will be as easy as possible, with user friendly systems and tools being made available at minimal to no cost. Users will also be able to easily share, value add and link the data to other sources. Examples of such (emerging) practice can be found in Victoria (SPEAR – see section 5.2), Western Australia (Local government online requests for road names- see section 5.4), and nationally with PSMA's (non-operational) NAMF national compliance service (section 5.5).

Governments are also known to subsidise targeted community collection initiatives as an efficient and highly cost effective alternative, as for instance demonstrated in Christchurch with the Traveller Information Project (see section 5.8).

People will have confidence in geocoded address data, not because it's 'authoritative' in the traditional sense of the word, but because its provenance is transparent and auditable, and errors are fixed quickly, much like Wikipedia. Likewise, address contributors may be "voted up" if their contributions are regularly validated and recommended, noting that this will require a critical mass of users. Also, it is possible to nominate trusted contributors, whose data contributions can be attributed with a higher confidence value unless proven otherwise. In the case of geocoded locations, postal delivery personnel or meter readers could be examples of such contributors.

What makes location data 'Authoritative'?[4]

People have always relied on 'authoritative' address data for mission critical purposes, most notably emergency management. The underlying assumption being that only authoritative sources can confidently provide data that is current, accurate, and of the best quality possible.

Traditionally, 'authoritative' in this context was (implicitly) defined as collated, validated and distributed by an official entity. And while governments were the primary, if not sole, collectors of location data, the authoritative source became the responsible government agency, and in recent years at a national level, the PSMA.

The Oxford Dictionary however, describes 'authoritative' much more broadly as: "able to be trusted as being accurate or true; reliable", "commanding and self-confident; likely to be respected and obeyed", as well as "proceeding from an official source and requiring compliance or obedience"[5]. For all intents and purposes, what really counts for the users and majority of use-cases is the trust and confidence.

In the 21st century, governments are no longer the monopolist collector or provider of location data, and non-government data sources are often as good, or at least fit-for-purpose, as the 'official' datasets.

Furthermore, the advent of crowdsourcing has given us the concept of 'distributed authority', where the confidence in an (online) information source is derived from the community that maintains it, instead from the reputation or mandate of a central organisation. The success of Wikipedia over its

[4] This section references material originally posted at: http://spatial21.blogspot.com.au/2011/01/psma-sensis-or-openstreetmap-what-makes.html

[5] http://www.oxforddictionaries.com/definition/english/authoritative

traditional counterpart the Encyclopaedia Britannica is maybe the most well-known example. People also happily rely on eBay's review systems to decide if a seller can be trusted, Tripadvisor to choose a hotel, or OpenStreetMap to navigate in an unknown city.

While the so-called 'wisdom-of-the-crowd' approach has its pitfalls, *a priori* there is no reason to assume crowd sourced location data is less trustworthy than 'authoritative' location data in the traditional sense. There are many mechanisms to assert or improve the trustworthiness of crowd sourced data. It is self-correcting in nature, which can be further strengthened through weighting of contributors' reputation, nominating 'trusted' contributors, certification policies, and/or moderation. It is likely that the role of government in enabling future 'authoritative' location data sources will include the facilitation for such mechanisms to operate effectively.

3.4.3 Data

The scope of the data will need to be extended to all locations relevant to the use-cases (i.e. for navigating to points of access or service delivery) that have a descriptive label. That means this can no longer be limited to rateable property units, but will include complex addresses (e.g. retirement homes), landmarks, aliases, and relevant Points of Interest (POIs). This scope cannot be achieved by relying on the jurisdictional, property-based supply chains alone.

The data will be collected as closely to 'source' as possible, leading to much more timely updates and accurate descriptors. Updates are available in a matter of days (if not immediate), rather than months.

The data provenance will be auditable, and simple feed-back mechanisms are in place to capture updates, additions and corrections. Data validation can still take place, either retrospectively, or through the identification of trusted collectors and or trusted reviewers. Some validation processes may be accomplished automatically using rules-based systems. Knowledge bases of how location data works can be refined and extended over time.

Non-validated ('provisional') data is also available and clearly marked as such.

The data model will facilitate separation of feature, location and descriptor objects, include persistent identifiers, and be an 'anchor point' for integration with land use and property information, as well as other reference or associated data. Such data models and the supporting infrastructures are already being trialled and implemented in initiatives such as INSPIRE in Europe, the Open Geospatial Consortium's (OGC) CityGML (see e.g. Gaitanis 2013) and Spatial Identifier Reference Framework (SIRF) in Australia (see section 5.7.1).

The data will be 3D and 4D (time) enabled, and link with emerging BIM initiatives. Linked data techniques would emphasise linking to such externally maintained resources rather than embedding them in the addressing content directly. However, the recognition of the central register of "sharable locations" is the key step. Once this is recognised then two-way linking can begin.

Geocodes may include more than point features (e.g. building footprints). Future applications may well require geocoding object's locations that change over time. Private initiatives such as Geepers (see 5.9) already allow user-defined, moving locations (enabled by their '.Now' mobile app).

A simplified, abstract representation is depicted in Figure 5 below.

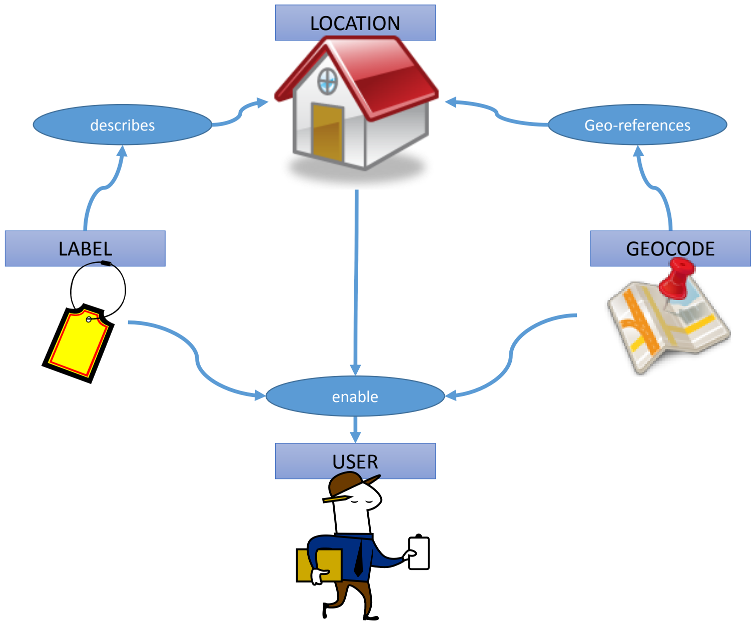


Figure 5 Simplified, conceptual model for geocoded locations

The user requirements and use-cases are enabled by a combination of the attributes of the location feature (is it a postal delivery point, or an insured building?), its descriptive labels (e.g. one or more address strings, a building name), and the geocode(s) that geo-reference the location feature.

3.4.4 Technology

While many claim technology is not the problem, understanding (future) technological capabilities is important, and the future vision will be flexible to adapt to new and emerging technologies if and when needed. It goes without saying that the supply chains will be standards-based and vendor agnostic.

Key technologies that will form part of the future supply chains include Linked Data, Semantic Web technologies and Machine-to-machine automation, enabling machine readable transactions, as well as registration and resolution services for Persistent Identifiers (as e.g. implemented by SIRF, see section 5.7.1)

Several facets of the business process of data collection and the end user interaction model will need to inform the selection of technologies for the project. Of particular importance are:

- federated and distributed nature of the data collection cycle;
- need for open distribution of some or all of the data;
- use of provenance and attribution metadata for the establishment of trust;
- utilisation of “address” data as a source of master data against which records in many external systems may be linked;
- ad hoc query and linkage of data for applications as yet unimagined.

These requirements have been addressed previously in a number of fields and share many common characteristics with the use cases for the Semantic Web (http://www.w3.org/standards/semanticweb/) and Resource Description Framework (RDF) data model (http://www.w3.org/RDF/) proposed by W3C. RDF is a standard model for data interchange on the Web. RDF has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed.

An elegant example of how complex (spatial) relationships between named places can be implemented in a useful and usable manner with the help of semantic web technologies, is provided by the Ordnance Survey's linked data platform (see http://data.ordnancesurvey.co.uk/ and section 5.7).

3.4.5 Governance

The governance aspects will be critical, as will be the role of government in achieving and managing geocoded location data as envisaged.

In debating the role of government in an organic, crowd-sourced, data environment, the analogy of the Internet is useful. The Internet is widely used, trusted and reliable, yet no one organisation or entity 'owns' or controls it, it is more like a living, constantly evolving 'ecosystem'. Yet there are a number of key bits of infrastructure and protocols in place, and organisations who own and manage them. These are critical for making it work, and include elements such as Domain Name Servers (DNS), TCP/IP, HTML and the W3C to name a few[6].

Similarly, the role of government in the future of geocoded location data in such an ecosystem, will be that of a facilitator, setting standards, delivering confidence, providing the minimal set of infrastructure (including e.g. location registration services, identifier management), and stepping in where the community or market fails to deliver. This should include further enabling 'Open Data' policies and licencing, protecting privacy and security and could extend to certification of 'validated' data and/or trusted data collectors that meet minimum confidence standards.

If people are to trust the information (completeness, accuracy and currency) then some minimum data coverage and quality standards will be required and it is likely that in some sectors or geographies or regions, incentives, subsidies or grants will be necessary to achieve good market acceptance. This may include targeted, subsidised data collection initiatives, as for instance demonstrated by the Christchurch Traveller Information Project (section 5.8)

Such a 'tiered' model of data quality will lend itself to a number of commercial models, including 'freemium' models, where the government provides and assures one or more 'minimum' foundation data offerings for free or at marginal cost, and the private sector or even jurisdictions develop value-added offerings, and/or assume aggregation and validation roles. Many aggregation and validation processes involve complex, explicit and implicit rules which can be automated.

[6] See e.g. http://www.internetsociety.org/internet/who-makes-it-work. There are several other useful parallel ideas in the dynamic maintenance of the Internet, such as authoritative name servers for specific IP ranges. It may in some circumstances be possible to structure addresses into namespaces with recognised authorities and to have maintenance processes validated for conformance. This control needs to be balanced against the value/currency of having the open, distributed maintenance provided through crowd-sourcing.

4 Future Direction

4.1 Objectives and Recommended Activities

The ultimate objective is the design and development of a next generation National Geocoded Location Management Framework to compliment G-NAF as part of the FSDF infrastructure. This Framework will comprise the people, data, technology, and governance aspects that will deliver the vision.

Even if there is agreement on the conceptual future vision, the implementation details will need to be developed and are likely to evolve over time. Many elements are innovative, are still being developed or trialled, and therefore there cannot be a single, clearcut roadmap to achieving the vision.

The future direction will support longer term improvements and objectives with regard to People, Data, Technology and Governance. Specifically, we recommend the following outcomes & activities:

- People
    - Identify and test contributor incentives ;
    - Specify tools required for community contributors;
    - Demonstrate feasibility of crowd-sourced geocoded location collection and validation;
- Data
    - Adopt/develop a national location data model that meets future requirements;
    - Demonstrate viability of geocoded location data model, and integration with 3D/4D and BIM;
    - Demonstrate and implement Persistent Identifier services for geocoded location data;
- Technology
    - Design, trial and demonstrate a notional technical architecture for a National Location Management Framework;
    - Research and demonstrate the value of semantic web technologies in such an architecture;
    - Develop and trial a “National Geocoded Location Portal” for registration, updates and validation, available to the community, local councils, notifiable agencies and federated to the Jurisdictions;
    - Develop standards and APIs (e.g. web-services) for real-time machine-to-machine interfacing;
- Governance
    - Design a trust model based on a mix of distributed (crowd-based) authority, validation processes, and trusted contributors;
    - Explore and define the role of government in facilitating the National Location Management Service, and how it fits with the Commonwealth and State and Territory jurisdictions’ Open Government and Open Data policies; and
    - Identify and/or develop appropriate licence models.

4.2 Methodology & Guiding Principles

Given that many of the concepts and technologies are emerging, implementing this future direction will be a largely experimental process towards a solution that will evolve and materialise over time. The roadmap will need to have relatively short development cycles, regular progress reviews and measurable outcomes, incorporating emerging R&D outcomes and international best-practice.

In manufacturing, technology and business, the so-called ‘lean start-up’ approach has been developed as a method to manage such a process. A lean start-up approach will help prevent the investment of time designing features or services that users do not want.

Lean start-up initiatives adopt a 'build, measure, learn' approach that has relatively short Minimum Viable Product (MVP) build cycles with measurable outcomes, learning and adjusting for the next cycle[7].

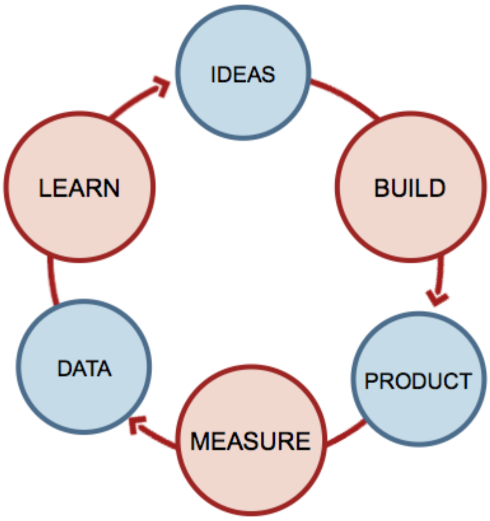


Figure 6 Lean Start-up Cycle (source: Eric Ries, 2011)

Guiding principles for the future direction include:

- Collaborative & inclusive;
- Build succession of MVPs to test, measure and learn from;
- Collect data to measure success;
- Learn from build cycles, but also from similar (inter-) national research and best-practice; and
- Augmenting, not replacing, current products & supply chains.

4.3 Roadmap

4.3.1 Short-term Improvements on Current Supply Chain

The roadmap below includes the short-term improvements presented in section 2.2, and will further include the specific CRC P3 work plan to deliver specific supply-chain improvements. This work plan is expected to be completed in Q4 2014.

[7] http://en.wikipedia.org/wiki/Lean_startup

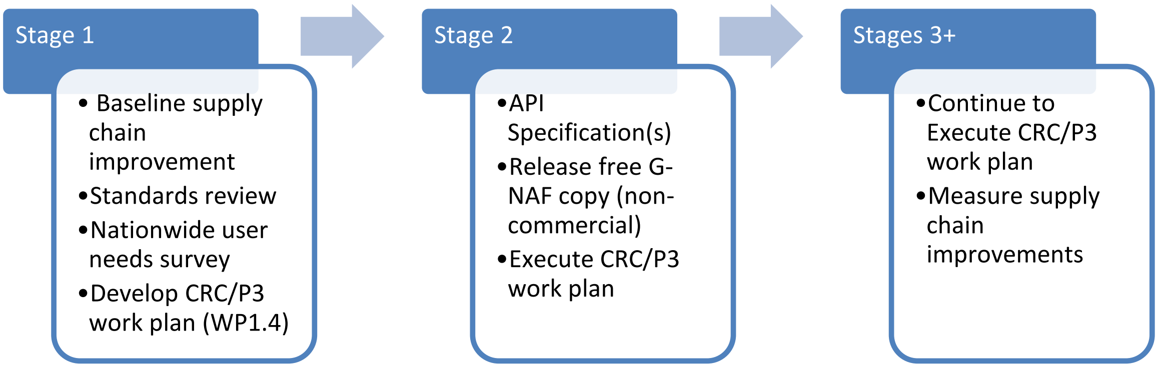


Figure 7 Short Term Improvements roadmap

4.3.2 Implementing the Future Direction

As stated above, the future direction will be a largely experimental process towards a solution that will evolve and materialise over time. The roadmap will need to have relatively short development cycles, regular progress reviews and measurable outcomes, incorporating emerging R&D outcomes and international best-practice, and could adopt a 'lean start-up' methodology.

Key activities will include:

- A series of Collaborative Experiments to trial data, technology and trust models:
    - With a broad range of participants: e.g. FSDF, SIRF, PSMA, Industry participants (through SIBA/43PL), government agencies (e.g. LINZ, Australian Electoral Commission, Australian Bureau of Statistics, Department of Human Services), academia (CRC), selected jurisdictions, user representatives (to ensure user requirements are met).
    - Open to community developers (e.g. Hackathons) at selected project stages.
    - The scope of a first MVP could include a revised national data model, a (jurisdictional) geocoded location registration & validation service with a user feed-back facility, and initial quality assurance mechanisms.
- Develop and evaluate commercial models (and associated cost-benefit analyses) that would underpin future directions. This should also address the question of the underlying incentives of non-commercial contributors and participants.
- Developing and testing options for the policy & governance framework that:
    - Could include the Department of Communications, ANZLIC the Spatial Information Council, SIBA, Intergovernmental Committee on Surveying and Mapping (ICSM), and/or user representatives.
    - Should cross reference other relevant strategies (such as the cadastre 2034 strategy)
- Set-up & identify targeted R&D initiatives in support of the vision (aligning with the CRC P3 research agenda)
- Ongoing assessment of objectives, best-practice and international technology developments, and adjusting the roadmap accordingly.

A possible broad implementation roadmap is present below.

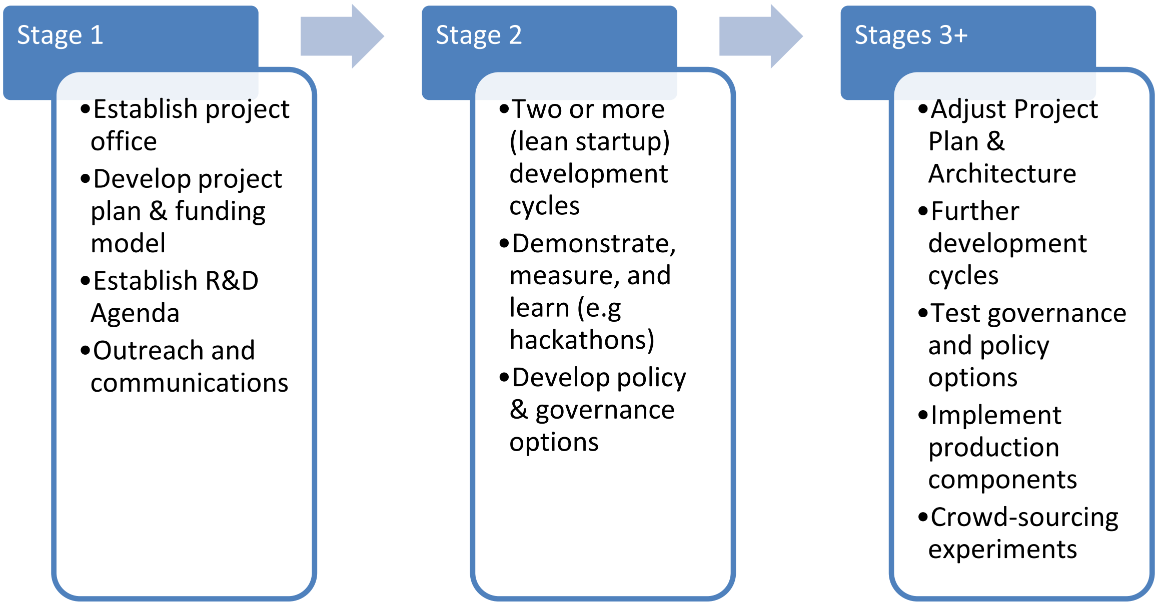


Figure 8 Future Directions roadmap

4.4 R&D opportunities (including Semantic Web Opportunities)

R&D topics and activities to support the future direction include:

- Spatial domain ontology for standardised representation of geocoded locations of various kinds;
- Optimal Representation of different kinds of spatial extent (3D point, 3D volume etc.) for use in representation of geocoded locations;
- Modelling of the relationships and navigation between different geocoded location roles (public entrance, goods entrance, mail delivery point etc.) relating to the same real world entity (e.g. commercial property, recreation area etc.); and
- Suitability of Semantic Web techniques to applications involving geocoded locations from a functional and non-functional perspective.

4.4.1 CRCSI R&D Topics

R&D topics relevant to Spatial Data Supply Chains and Geocoded Addressing, currently in the CRCSI Spatial Infrastructure program, include:

- Machine readable transactions;
- Critical Path Analysis;
- Dynamically controlled supply chains;
- Artificial Intelligence / Semantic Web;
- Automated Data Set Conflation;
- Automated Feature Extraction;
- Automated Update Propagation;
- Standards-based Approaches;
- Linked Data – Semantic Metadata;
- Semantic Queries;
- Crowd-sourcing, Trust models, automated feedback processes; and
- Legal-policy models for easier data sharing.

4.4.2 Other Research Opportunities

Additional R&D opportunities that would underpin the future vision include:

- Options for modelling the "Geocoded Location" (or "Address") concept compared to the introduction of a new 'tag' used for geocoding to higher resolutions than property address; and
- Business models (and incentives) that contribute to altering a collection of independent address custodians into a functioning, integrated supply chain.

5 Case Studies and References

Many of the concepts, technologies and approaches presented in the vision and future direction are already being developed, trialled or implemented. In this chapter, we present references to a number of relevant case-studies and other references.

5.1 LINZ OpenStreetMap Transport Data Import

Since 2011, the New Zealand government has been releasing many (spatial) datasets under open (Creative Commons) licences. Land Information New Zealand (LINZ) has been working with OpenStreetMap to bulk upload the government transport data sets into OpenStreetMap. This provides a baseline 'seed', for future community updates and additions in OpenStreetMap.

It serves as an example of the sort of hybrid model where 'authoritative' government data is augmented with crowd-sourced data, allowing the community to contribute efficiently to completeness, timeliness and quality.

More information: http://wiki.openstreetmap.org/wiki/New_Zealand_data_import_projects

5.2 Canterbury Maps Building Footprints

In another example of augmenting government datasets with community updates, LINZ, Environment Canterbury and the University of Canterbury combined to issue a mapping competition for schools, where students are recruited to digitise building footprints online, and capture basic building attributes.

In effect, this is targeted, subsidised volunteered geographic information collection.

See more: http://canterburymaps.govt.nz/BuildingOurFootprints/

5.3 SPEAR

The Victorian government has initiated the 'SPEAR' (Surveying and Planning through Electronic Applications and Referrals) initiative, as a way to shorten the pathway from address creation to the users. SPEAR is an internet-based system which supports development approval processes in Victoria - it supports all subdivision and planning permit applications. SPEAR provides a common, web-based interface for all parties, which enables a complete electronic process that updates the titles system and the Vicmap digital map.

In the future, SPEAR could enable direct interaction between users and creators of addresses, without the need for jurisdictional involvement.

More information: http://www.spear.land.vic.gov.au/spear/

5.4 WA Road-name Creation

The Western Australian Government has a website for the community to screen road names online, and submit them for 'approval'.

See: http://www.landgate.wa.gov.au/corporate.nsf/web/Geographic+Feature+Names

5.5 ANZLIC National Address Management Framework (NAMF) Compliance Service

The National Address Management Framework (NAMF) was an ANZLIC initiative developed over 2006-2010, defined the need for a set of central, online services to validate, certify and update addresses against a common, authoritative dataset (G-NAF).

In 2009-2010, PSMA, following a request from ANZLIC, developed the NAMF Compliance Service (NCS), incorporating a website, web notification services, database and GUI interface for jurisdictions, iPhone app and underlying business processes to submit, manage and approve notified address updates.

The NCS was decommissioned by ANZLIC pending further investigation and a review. Primarily because the there was no ANZLIC funding to maintain and improve the system, and jurisdictions did not support (and were not incentivised to support) the resource burden associated with processing the notified addresses the service would generate.

More info: http://www.anzlic.gov.au/resources/national_address_management_framework

5.6 NSW Comprehensive Property Addressing System (CPAS)

The overall objectives of the CPAS program are to improve the property addressing system in NSW, to increase the efficiency and effectiveness for the delivery of emergency services, post and utility services, and to support a range of government activities.

The objectives of the program will be achieved by delivering identified improvements to the addressing system made through the capture, storage, maintenance, geocoding, access and distribution of NSW addressing data. These improvements will include:

- strengthened ability to create/allocate, capture at point of creation, store, maintain, geocode and distribute authoritative official addresses;
- the ability to capture, store, maintain and distribute information necessary to locate alias addresses and features of interest information and link these, where appropriate, to authoritative addresses;
- rural addresses that are correctly assigned, captured, stored, maintained, geocoded and distributed;
- complex site and strata site secondary addresses that are correctly assigned, captured, stored, maintained, geocoded and distributed (within five years);
- the creation of metadata that will enable the efficient and accurate discovery, exchange and dissemination of information and increase the value of address information.

More information:

- http://www.gnb.nsw.gov.au/addressing/cpas
- http://www.lpi.nsw.gov.au/about_lpi/comprehensive_property_addressing_system

5.7 Semantic Web & Linked Data

The Resource Description Framework (RDF) extends the linking structure of the Web to use URIs to name the relationship between things as well as the two ends of the link (this is usually referred to as a “triple”). Using this simple model, it allows structured and semi-structured data to be mixed, exposed, and shared across different applications.

This linking structure forms a directed, labelled graph, where the edges represent the named link between two resources, represented by the graph nodes. This *graph view* is the easiest possible mental model for RDF and is often used in easy-to-understand visual explanations.

One of the great strengths of RDF is the flexibility and extensibility of the data model, with the possibility that new data can be added to existing data elements at any time and by any partner. However, it also permits the representation of highly structured and semantically meaningful data should the application require it. These formalised "data contracts" can be expressed using a dedicated schema language such as RDF Schema (RDFS) or Web Ontology Language (OWL).

RDF Schema (http://www.w3.org/TR/rdf-schema/) is a semantic extension of RDF. It provides mechanisms for describing groups of related resources and the relationships between these resources. RDF Schema is itself written in RDF. These resources are used to determine characteristics of other resources, such as the domains and ranges of properties. The RDF Schema class and property system is similar to the type systems of object-oriented programming languages such as Java.

Web Ontology Language (http://www.w3.org/2001/sw/wiki/OWL) is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things. OWL is a computational logic-based language such that knowledge expressed in OWL can be exploited by computer programs, such as to verify the consistency of that knowledge or to make implicit knowledge explicit. OWL documents, known as ontologies, can be published on the internet and may refer to or be referred from other OWL ontologies. OWL is part of the W3C's Semantic Web technology stack, which includes RDF, RDFS, SPARQL, etc.

SPARQL is the query language for RDF. SPARQL can be used to express queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF via middleware. SPARQL contains capabilities for querying required and optional graph patterns along with their conjunctions and disjunctions. The results of SPARQL queries can be results sets or RDF graphs.

The GeoSPARQL standard (http://www.opengeospatial.org/standards/geosparql) supports representing and querying geospatial data on the Semantic Web. GeoSPARQL defines a vocabulary for representing geospatial data in RDF, and it defines an extension to the SPARQL query language for processing geospatial data. In addition, GeoSPARQL is designed to accommodate systems based on qualitative spatial reasoning and systems based on quantitative spatial computations.

The use of the Semantic Web framework to exchange and link data also opens up the possibility of using sophisticated machine reasoning (logic) techniques to derive novel (application specific) data from geocoded locations and the wide variety of data which may be linked to it in future via ontologies dedicated to that application area.

For distributed and dynamic data linking initiatives, perhaps the most promising set of case studies come from the Semantic Web and Linked Data world. Projects such as DBPedia (http://en.wikipedia.org/wiki/DBpedia), the structured data subset of Wikipedia, have shown that it is possible to maintain and distribute a large and continuously evolving dataset to a world-wide audience in a form that enables targeted re-use by applications that were never envisaged by the original compilers of the information. This kind of flexibility against future uses is precisely the characteristic which is desired in the geocoded location datasets of the future. A number of other ongoing projects demonstrate elements of resource identification, linking and query technology that may play a central role in the evolution of a comprehensive geocoded location framework.

5.7.1 Spatial Identifier Reference Framework (SIRF)

The Spatial Identifier Reference Framework (SIRF) initiative is developing an internet-based linked data infrastructure. SIRF provides a means to manage and use spatial identifiers (the names and codes used to reference real world locations) to reliably link and rapidly integrate information (such as socio-economic statistics) about locations, which are stored in multiple distributed systems.

Currently SIRF implements Gazetteer and Administrative Boundaries datasets for Australia, and could be relatively easily expanded to Geocoded Locations more broadly, including street addresses.

More information: http://portal.sirf.net/about-sirf

5.7.2 Ordnance Survey Linked Data

'OS OpenData' is the opening up of Ordnance Survey data as part of the drive to increase innovation and support the "Making Public Data Public" initiative. As part of this initiative, Ordnance Survey has published a number of its products as Linked Data. Linked Data is a growing part of the Web where data is published online and then linked to other published data in much the same way that web pages are interlinked using hypertext.

- http://data.ordnancesurvey.co.uk/
- http://data.ordnancesurvey.co.uk/doc/7000000000001255
- GB Postcode Units (try searching for CB4 0WZ 1Spatial's postcode: http://data.ordnancesurvey.co.uk/datasets/code-point-open
- Placename search (try searching for "Chesterton"): http://data.ordnancesurvey.co.uk/datasets/os-linked-data

5.8 Christchurch Traveller Information Project

Following the 2012 Earthquake, many roads in Christchurch have been changed or rebuilt. Most of these changes were not reflected in trip planning tools.

In August 2013, under the Canterbury SDI Programme's Open Data project, the first pass towards a full routable network for the city was crowd-sourced within 5-days, and created in OpenStreetMap, with the help of 5 volunteers, and minimal logistical expenses.

More information: http://wiki.openstreetmap.org/wiki/Christchurch_Traveller_Information_Project

5.9 Private Sector Addressing Innovations: Geepers

Geepers (http://geepers.com) labels and locates people, places and things in an online register that can be searched in a similar way to a Twitter hashtag. Geepers lets users create a unique digital alpha-numeric user profile or username, and then add multiple locations using latitude and longitude coordinates. These can refer to where they wish to be found, such as home, work, or current on-the-move locations, which is updated via their smartphone or internet-enabled device.

In effect, Geepers is a private registry and resolver of unique, universal location tags, or 'persistent location identifiers', with conceptual similarities to SIRF (see 5.7.1 above).

5.10 Gaitanis, Harry (2013): "Towards a semantically and spatially richer address data model"

Harry Gaitanis (2013) wrote his master's thesis combined at the Technical University of Berlin, and the University of Melbourne.

The key issue he addresses is that: *"Geocoded addresses usually consist of a single point representation with no attached semantics. This hinders the navigation of address users as well as their decisions in the choice of location-based services, since the geocoded point seldom represents a real world feature and is often far from a user's actual destination"*.

In this thesis, he develops a new, semantically and spatially enriched data model, and tests this model to conclude that it will "*significantly reduce the costs of navigation to Points of Interest (POIs) within a property…"*.