AUTHORITATIVE ADDRESS DATA SET

Endorsed by the Ministerial Online and Communications Council
on 12 December 2008

ANZLIC
the Spatial Information Council

Document Control

Author: John Gottschalk

CTG Consulting

Document No: 01   **Date of Issue:**   12 December 2008

Version: 1.0

Document Release Information and Distribution

This is a controlled document. The Quality Document Register is maintained within CTG to manage the distribution of controlled documents and the issue of any revisions. Printed versions are uncontrolled.

Date	Version	Released To
8th August 2007	0.4	Bruce Thompson, Karina Donaire, Liz Marchant
20th August 2007	0.7	NAMF Project Team
4th Sept 2007	0.8	NAMF Project Team
11th Sept 2007	1.1	NAMF Project Team
August 2008	1.1	Approved by the NAMF Program Control Board
November 2008	1.1	Endorsed by the Cross Jurisdictional CIO Committee
12 December 2008	1.0	Official Version – endorsed by the Ministerial Online and Communications Council
May 2009	1.0	For publication on ANZLIC web site

Table of Contents

1 Glossary of Terms

Address	An address is a textual representation of a physical or virtual point for the purposes of either identifying a physical parcel of land on the land surface of the earth or as a communication channel to a recipient. An address may refer to items other than a physical point such as an email address.
	A single point may have many textual addresses associated with it. These can be different ways of representing the official address or, in the case of Private Estates, a way of addressing the internal structure of the Estate.
	Many attributes may also be associated with an address. For example, Australian Bureau of Statistics (ABS) Mesh Block/ Collector District, Business or Private use, Electoral district(s), Latitude and Longitude of are also attributes to an address.
Address Consumer	Any organisation or individual requiring accurate addressing data as part of their normal operation.
AMAS	The Address Matching Approval System (AMAS®) is a certification program that has been developed by Australia Post to improve the accuracy of postal addresses. The software prepares addresses for barcode creation, ensures quality addressing, and enables mailers to qualify for postal discounts for Pre-Sort letter lodgements. More information can be found on the Australia Post website (www.auspost.com.au).
AGIMO	The Australian Government Information Management Office. See www.finance.gov.au for more information.
ANZLIC	ANZLIC — the Spatial Information Council — is the peak inter-governmental council for the coordination of spatial information policy and strategic issues in Australia and New Zealand. See www.anzlic.org.au for more information.
API	Applications Programming Interface. An interface to a software application for access by other computer systems.
DPID	Delivery Point Identifier (DPID) is an eight-digit number which uniquely identifies a postal delivery point in Australia.
Location Address	A physical address which would typically include an address label of a real area, which includes street number, road name, locality and state, address site name, centroid, datum point (for rural), link to a real property description, alias address and geo-code that spatially defines the address point. A location address point should be defined in three dimensions, where appropriate.
	Location addresses are used by many people and applications to provide clear, unambiguous information to common users in order to assist them to identify a point on the earth's surface where a good or service is usually required to be delivered.
Postal Address	A physical or virtual address used by a recipient for the purposes of receiving correspondence.
	Postal Address is a specific form of location address developed for

	postal delivery. Australia Post has progressively refined and developed postal address to meet its operational requirements, including the creation of a unique identifier (Delivery Point Identifier, or DPID). Any government agency or business wanting to manage postal operations efficiently and effectively will be concerned with postal address.
PSMA Australia	PSMA Australia Limited, formerly known as Public Sector Mapping Agencies, is an unlisted public company wholly owned by the State, Territory and Australian Governments. The core business of PSMA Australia is facilitating the creation of and access to seamless national spatial data sets for government, industry and community use. See www.psma.com.au for more information.
XML	eXtensible Markup Language. XML is a standards controlled framework for the interchange of data. Its primary purpose is to facilitate the sharing of data across different information systems, particularly via the Internet.
XML Schema	A way to define the structure, content and, to some extent, the semantics of XML documents.

2 Introduction

2.1 *Purpose of the document*

This document provides an overview of current address management issues and a recommendation for the development of a single authoritative Address Data Set for Australia. Both postal and location address are required to support the business operations of government, the private sector and the community. The recommended approach means both postal and location address capability will be available in a simple, efficient arrangement, to support the full range of required business activities.

2.2 *Current Situation*

Address data, in all its forms, is vital for the efficient operation of many public and private organisations whilst being a particularly difficult form of data to validate, store and use for purposes such as data matching. These difficulties derive from the unstructured and frequently ad-hoc approach to address storage and dissemination.

Many organisations find that access to high quality address data for address validation, parsing and address capture is difficult and costly. No single source of address truth exists and no formal processes exist to enable a feedback of addressing issues in order to increase address accuracy.

2.3 *Scope*

Figure 1 shows the supply chain for address data. The creation of addresses and address quality improvement initiatives, that is, the supply side shown in the upper part of Figure 1, is *not in scope* of the NAMF.

The processes to make address data available to users such as Government and emergency services organisations, that is, the demand side shown in the lower half of Figure 1, is *in scope* of the NAMF.

NAMF is a demand, or user, side strategy. That is, NAMF will provide a framework to make address management and address operations more efficient and effective, principally through the provision of:

- A single authoritative address data set;

- A national standard for address data storage and interchange; and

- National standards for web services that support common address-related functions (such as address parsing and validation).

These three aspects of NAMF are shown in Figure 1. *The subject of this report is the Address Data Set.*

NAMF is not a supply side strategy. That is, NAMF will have no direct impact on existing address creation and maintenance processes[1]. There is a range of ongoing

---

[1] While NAMF will not directly alter existing maintenance processes, its implementation will provide the means to significantly improve address quality. The single most important factor will be the automatic, real-time notification (from NAMF address user to address providers) of address parsing, validation and geo-code failures,

address quality (currency, timeliness, completeness, accuracy) improvement initiatives underway in States and Territories, in Australia Post and in many other organisations.

Figure 1: Address Data Supply Chain

Lack of early advice of address errors or omissions is one of the fundamental deficiencies of existing address creation and maintenance processes.

3   Addressing Requirements

This section lists the requirements for addressing that were considered for the NAMF.

3.1 Accuracy

In addition to the usual client mailing and account information storage requirements, address consumers increasingly use addressing data for sophisticated data matching, business intelligence and to support identity management. This is placing greater demands on the accuracy requirements of address data.

3.2 Timeliness

Address data should be made generally available as close to the time of its creation as possible. The delay between address creation and its visibility should be measured in hours or days rather than the current months or, in some cases, years.

3.3 Completeness

All addresses likely to be used should be included.

Within the overall address domain there is a range of broadly compatible subsets, delineated by specific user needs and circumstances. For example, the relationship between location addresses and postal addresses is not simple — not all location addresses (such as parks or other public places) have mail delivered to them, and not all postal addresses (such as Post Office boxes and rural mail services correspond directly to a location address. Incomplete administrative and institutional arrangements mean that not all addresses are readily available. For example, addresses in some aboriginal communities, and in some private complexes (nursing homes, retirement villages, private communities) are not readily available.

3.4 Availability

Address data should be available in a manner that enables most organisations easy and speedy access. The data should be available in a format that is standards compliant and flexible enough to meet the changing needs of users.

3.5 Access

Address users require a single framework that supplies a comprehensive set of services designed to meet their address management needs and reduce the complexity of maintaining their address data.

Conceptually this could be viewed as a single source of address truth accessed via a convenient interface while physically the data may be stored in a distributed manner, and be sourced from one or more providers depending on specific requirements.

4   Australian Address Data Sets

This section provides information on Address Data Sets considered as candidates for the NAMF.

4.1   *Australia Post address resources, including the Postal Address File (PAF)*

In terms of this document the PAF is used generically to mean Australia Post maintained address data, although the actual address data resource provided by Australia Post may be a specific resource appropriately specified to meet NAMF requirements, rather than the existing PAF product.

Australia Post maintains and distributes the PAF through a set of third party providers. Access to this data is for the purposes of validating a customer addresses and ultimately supplying a Delivery Point Identifier (DPID). The ability to print a DPID on mail gives the customer access to postage discounts. The use of the PAF, or other Australia Post designated resource, is essential for government agencies wanting to provide efficient postal operations.

The PAF is an authoritative source of postal delivery addresses for Australia. Only AMAS certified matching software is allowed to allocate DPIDs. Australia Post controls the allocation and use of DPIDs. The Australia Post data set:

- Is the only source of DPIDs.

- Is concerned with postal delivery address only.

- Is updated quarterly.

- Has well developed, understood and documented postal delivery address validation processes.

4.2   *Australian Electoral Commission (AEC)*

The Australian Electoral Commission (AEC) maintains an address database for its own operational purposes. This data is largely contained in the GNAF and PAF thus will not be considered further.

4.3   *Geocoded National Address File (GNAF)*

The GNAF is sourced from Australia's State and Territory government mapping agencies and land registries, from Australia Post, and from the Australian Electoral Commission. PSMA Australia Limited has the responsibility for building G-NAF. PSMA Australia Limited, formerly known as Public Sector Mapping Agencies, is an unlisted public company wholly owned by the State, Territory and Australian Governments.

The GNAF data, in addition to the normal Suburb, Street and Street Number data contains a specific and unique real world location — the geocode. All addresses from each of the three major sources (government land registries, Australia Post and AEC) are cross-referenced, enabling a confidence level to be assigned for each address.

The GNAF, to a large extent, can be considered a super set of the PAF and AEC data. However, due to their differing purposes there are addresses in both the PAF and AEC data that do not appear in the GNAF data, and addresses in G-NAF that may not appear in the PAF or AEC data.

4.4 Other Data Sets

Whilst many other public and private organisations, for example Sensis, Australian Taxation Office and Centrelink hold large amounts of address data they cannot provide address data useful for address validation or lookup purposes and are limited in scope to addresses of interest to their business. They do not possess a comprehensive collection of all Australian addresses that meets the requirements of the NAMF. For example the Australian Bureau of Statistics holds a significant amount of address data that is fit for their purpose but could not be considered for inclusion in a NAMF authoritative Address Data Set.

5 Findings and Conclusions

5.1 Data Sources

Many of the deficiencies with the current addressing processes are caused not by the absence of address data but rather by its dispersal across many data sources. NAMF will require a single Authoritative Address Data Set based on the data contained in the G-NAF and PAF.

Using GNAF and PAF as the source for the Authoritative Address Data Set provides both the most complete set of address data and through access to the data created by numerous address experts within these organisations and their contributing organisations leverages a considerable pool of addressing expertise and proven capability. This data set will provide the accuracy, consistency and ease of access fundamental to broad acceptance and use of the services.

While the postal and location address components of the single authoritative Address Data Set will remain fully under the control of Australia Post and PSMA Australia respectively, it will be important for the data set to be a single functional tool for the purposes of users. A single, simple, effective tool providing both postal and locational capability will minimise costs and maximise benefits for users. It should also expand the range of uses address data is put to, as well as expanding the market, providing benefits to both Australia Post and PSMA Australia.

5.2 Address Data Feedback

Although modifications to the processes used for generation of Address Data Sets and the quality of the generated data are not in scope of the NAMF, the NAMF can assist with the improvement of address data quality by enabling address users to provide feedback on failures detected during address management operations (e.g. address parsing, validation and geo-coding).

5.3 Commercial and Legal Considerations

5.3.1 Legal

No legal impediments have been identified to the implementation of NAMF.

5.3.2 Privacy

NAMF will deal only with address data and will not include personal information. Privacy Commissioners at State/Territory and Australian Government levels are part of the stakeholder and consultation group for NAMF.

5.3.3 Intellectual Property (IP) and Licensing

The Intellectual Property and related rights of both Australia Post and PSMA Australia will be protected in developing NAMF. Licensing arrangements for the IP of both parties will need to be developed and agreed.

5.3.4 Commercial

A simpler, more effective and more capable address management framework should mean that postal and location address are used more often and more effectively by existing users, and that use will be expanded to a broader range of users. In this context both Australia Post and PSMA Australia stand to benefit through increased use of their data, and that their combined offering will serve to drive the market to a greater extent than separate, specific purpose addressing offerings.

There is also a range of complementary address data products and services that, while not directly within scope for NAMF in terms of address parsing and validation, will assist government and business in meeting their address management and business objectives. One example is Australia Post's Address Reference File (ARF) which provides, among other information, an attribute identifying whether an address is considered to be business or residential in nature. It will be important for the NAMF model to be developed to allow the simple integration of these additional information resources in business processes and operations.

6 Recommendations

The candidate Address Data Sets have been considered and their suitability determined with respect to the NAMF requirements. This has lead to the following recommendations:

1   The requirement for completeness leads to the recommendation that the NAMF Authoritative Address Data Set be based on G-NAF from PSMA Australia, for location address operations and functions and, via accredited AMAS providers, PAF from Australia Post for postal delivery operations and functions. GNAF and PAF are the only complete authoritative sources of addressing data for Australia and together fully represent both postal and location addresses.

2   The NAMF Authoritative Address Data Set should support a unique identifier. This will assist with the implementation of an address data feedback service in the NAMF, discussed in Section 5.2 of this report, as will as supporting the extensibility necessary to effectively manage complementary address data products and services not part of NAMF but required for government and business operations.

The requirements of accuracy, timeliness, availability and access have less influence the choice of address data set, but are further considered during the examination of the Address Data Interchange Format and the Web Services interface specification for the NAMF.