Application of Controlled Vocabularies to Foundational Spatial Information

Michael ELLIOTT-SMITH and Ashlee POEPPMANN, Australia

Key words: Cadastre, Capacity Building, Digital Cadastre, e-Governance, Geoinformation/GI, Land management, Legislation, Positioning, Real estate development, Reference systems, Spatial planning, Standards

SUMMARY

This paper explores the modernisation of foundational spatial data, with a specific lens on location addressing and the use of controlled vocabularies to ensure data consistency and accuracy. The paper emphasises the need for a structured semantic approach to data management, advocating for the transition to the Queensland Addressing and Location Information (QALI) system, a sophisticated address system that leverages the ISO 19160.1. A key focus has been the integration of machine-readable vocabularies, a process involving the development of authoritative, dynamic vocabulary lists maintained within a public repository. This paper underscores the ongoing need for enhanced data management practices, reminding us that to have effective and reliable data analysis and data systems, foundational spatial data needs to be consistent, accurate and reliable.

Paper Application of Controlled Vocabularies to Foundational Spatial Information

Michael ELLIOTT-SMITH and Ashlee POEPPMANN, Australia

BACKGROUND

What is an Address?

An address serves as a compound label for a place, acting as a critical marker within our spatial surroundings. It enables the identification and location of a site, whether that be a building or a specific section within it. Essentially, an address is a string of unique information structured in a way that can be easily remembered and conveyed, allowing a person or a service to pinpoint a location with precision. In the context of Queensland, Australia, while the creation and adherence to address standards are not mandatory, practices have been informed by guidelines such as the AS/NZS 4819:2011 and the earlier AS/NZS 4819:2003. These standards outline essential components for an address, which include the address number, the road name, locality name, and the name of the state or territory. They also include optional details that can enhance specificity, such as unit type and number, address site name, postcode, and country, thereby fulfilling an essential need for a wide range of service delivery organisations.

Defining an Address Spatially

Addressing is a formalised description that contextualises a location within physical space. The term 'address site' or 'addressable object' refers to the tangible spatial area that an address defines. A 'geocode' is the numerical expression of an address, translating its location into precise, mappable coordinates that allow for pinpointing specific parts of the address site.

When defining an address spatially, cadastral land parcels are often used due to their convenience in defining the extents of many sites. However, the correspondence between land parcels and addressable objects isn't always direct or 1:1. There are instances where an address might exist independently from a land relationship, such as an addressable object within a road corridor (i.e. no real property description for the location).

Implementing the standard requires cooperation between all parties, including addressing authorities and addressing custodians. Our joint efforts ensure that addresses are not only accurate but also relevant to other spatial systems that utilise them.

Who is Responsible for Addressing?

In Queensland, Local Governments are the recognised authority for assigning addresses, each operating within their respective Local Government Area (LGA). Under the provisions of the *Local Government Act 2009*, these local authorities hold the power to sanction road names and allocate road numbers, an authority that extends to both public infrastructure and private developments. The Department of Natural Resources and Mines, Manufacturing, and Regional and Rural Development (NRMMRRD), functioning as the state's custodian of foundational spatial data, plays a pivotal role by collating address information from the LGAs to manage and maintain the state's comprehensive address database. This database comprises over 2.7 million

addresses, and since 2015 has had approximately 86,000 updates annually, reflecting a dynamic and continuously evolving dataset. At collation, data from 78 Local Governments is standardised and merged, subsequently feeding into a national addressing product, and ensuring that address information is coherent, accessible, and up to date nationally.

The Legacy System

Addressing data for Queensland is currently managed by the department's addressing team in a relational database known as the Lands Address Location Framework (LALF) using bespoke with Using desktop applications. LALF, along with derived products and services made openly available, is collectively known as the Queensland Address Management Framework (QAMF). A key piece of this framework is adherence to the Australia/New Zealand Standard for Rural and Urban Addressing, last updated in 2011. While this standard was modelled on property-centric street addresses, it did not adequately prescribe methods for addressing non-property-based locations, nor methods for addressing locations that don't have some mandatory address components such as landmarks with no street number or locations in external territories without named localities.

QAMF was designed primarily to cater for the storage and maintenance of standards compliant addresses. The rigid nature of relational data models therefore introduces challenges to representing real-world address scenarios which are either non-standard or are edge cases which had not been considered during development of the QAMF.

The situation is further complicated by the addressing standard's non-mandatory nature and the lack of retroactive application, which means addresses that pre-date or were created without the standard's considerations may not fit within the existing data framework. Moreover, challenges arise from several other factors, such as:

- Data supply chain: Few address creators generate proposed addressing due to the absence of this 'future' lifecycle stage information from governing legislation and standards. Pre-emptive address assignment would improve the timely supply of accurate address data.
- Accuracy: With advancements in consumer technology, there is a growing expectation for highly accurate and more comprehensive address data, such as specific entry points for service delivery.
- Consistency: The necessity for uniform addresses across government and commercial entities mean that discrepancies can hinder service provision and affect public trust.
- Emerging use cases: Legacy systems and data models struggle to integrate with new applications such as 3D property modelling, Building Information Modelling (BIM), spatial digital twins, and multi-floor buildings.
- Complex addressing scenarios: Legacy systems often lack the flexibility to detail the precise location data for complex sites like retirement villages, shopping centres, and hospitals.

However, significant strides are being made towards modernisation. The development of new strategies and models, driven by the Intergovernmental Committee on Surveying and Mapping (ICSM) Addressing 2035 strategy and its accompanying technical addendum, describes a vision

that encapsulates a dynamic, integrated government address ecosystem. This system aims to support service delivery within a 3D and 4D digital economy context.

The Australian Address Model (AAM), accessible through https://linked.data.gov.au/def/addr, represents a semantic web data model tailored to the specific requirements of Australian address representation. Semantic data models differ from traditional models, as the nature of relationships between semantically modelled objects are expressly defined. This aligns with the ICSM's Addressing 2035 strategy, which identifies key 'pain points' in current addressing systems and frameworks and anticipates future modelling requirements. In Mid-2024, the AAM received ICSM endorsement as the national data model for addressing.

These and other ongoing modernisation efforts aim to eradicate the fragmentation of address data siloed by local government authorities, tackle the complexities of intricate sites, minimise reliance on ETL processes, and reduce duplication of human effort, and data storage, management, and retrieval. Doing so enables the creation of a single authoritative source of open addressing data, which is free to use and allows third parties to reuse or directly link to foundational data.

THE LINKED DATA STRATEGY FOR QUEENSLAND

Address Modernisation

The catalyst for Queensland's progression towards address modernisation is the transition to a semantic web-based approach. The new Queensland Addressing and Location Information (QALI) system is a sophisticated address model that leverages the AAM. QALI is designed to be highly reconfigurable and reusable, ensuring it can accommodate the dynamic and varied needs of address data management. This forward-thinking design will position Queensland to more easily respond to new addressing and location information service demands and opportunities in a digitally driven world.

Controlled Vocabularies

Controlled vocabularies are curated lists of terms used for organising information. Vocabularies ensure that data across various domains is described with uniformity, while retaining the ability to accommodate specific requirements. Controlled vocabularies enforce clarity and consistency, in that terms are defined, and use of the defined terms is unambiguous.

These vocabularies are hierarchical and provide both broader and narrower terms to enrich the semantic relationships within data. QUALI relies heavily on machine-readable controlled vocabularies, which serve as pre-defined lists of addressing concepts. These lists include road types (e.g., road, drive, street, lane) and address geocode types (e.g., parcel centroid, building centroid, access point), among others.

Every term within a controlled vocabulary is assigned a persistent unique identifier. For example, road types may be denoted with a URL such as https://linked.data.gov.au/def/road-types, ensuring it is persistently findable and identifiable. These identifiers are registered by authorised entities, in this case the Australian Government Linked Data Working Group. Addresses then reference these unique identifiers for various components like road or water feature names, localities, property names, and building names, which are ultimately integral to the composition of an accurate address.

Certain address elements may need to be subclassified or associated with specific terms drawn from controlled vocabularies. The intentional reuse of terms from these structured lists promotes consistency and precision within the data ecosystem. These vocabularies are maintained within an authoritatively managed code repository, openly accessible and maintained under the guidance of ICSM.

Distinct from static code lists typically found within standards documents, these vocabulary repositories are dynamic. They are open to public suggestions for updates, which undergo a review process and are administered by authorised users. This dynamic nature allows an auditable trail of changes, encourages community collaboration, and provides both machine- and human-readable access to the vocabularies, usually via by APIs.

For example, abbreviations for 'Avenue', such as 'Av' and 'Ave', are recognised by humans as synonymous, but a machine-readable system also needs to interpret them as equivalent. References to these road types within the data would link to a specific controlled vocabulary identifier, such as https://vocabs.gsq.digital/object?uri=https://linked.data.gov.au/def/road-types/avenue. This interlinking negates the need for static code lists and ensures any updates to the controlled vocabulary will automatically reflect across the linked data.

QALI is not confined to an isolated vocabulary list; it has the capability to connect to other web-based vocabularies. This enables the reuse of other relevant vocabularies, eliminating the need to recreate or duplicate terms. Datasets gain efficiency through this interconnectedness, opening pathways for both humans and machines to discover additional information through these semantic links.

This disruptive approach signifies that the migration of foundational datasets to semantic data models reveals an overarching need for broader modernisation. As foundational datasets adopt this structure, the benefits of semantic model adoption for related datasets become too great to ignore. By doing so, a system is created whereby dependencies are directly referenced through persistent links, rather than by unnecessary duplication of information.

CONCLUSION

The integration of the new addressing model with controlled vocabularies ensures the creation of consistent, accurate, and reliable data, which are fundamental to making effective and efficient decisions. By streamlining the flow of information, this consistency will significantly reduce delays among government departments, enabling AI applications to operate with heightened efficiency. The enhanced collaboration leads to reductions in data duplication, conserves human effort, and utilises fewer resources, all the while improving reliability. These improvements will enhance public trust and strengthen information security, as protecting one central dataset is more secure than safeguarding multiple datasets with varying levels of security. In the broader picture, this equates to significant cost savings and paves the way for a healthy digital economy, where concepts like digital twins can thrive, driven by a foundation of reliable and secure data.

In conclusion, the modernisation of Queensland's addressing system through the QALI system, and its integration with controlled vocabularies, lays the groundwork for more connected and intelligent data. With an emphasis on data quality and structure, this progression in data management represents a significant leap towards a consistent, accurate, and collaborative future for Queensland's foundation spatial data.

REFERENCES

Intergovernmental Committee on Surveying and Mapping (ICSM), 2021, Addressing 2035: Addressing Reform and Innovation for Australia and New Zealand. https://www.icsm.gov.au/sites/default/files/ICSM%20Addressing%202035%20Strategy%20%28Final%29.pdf

International Organization for Standardization, 2015, ISO 19160.1: Addressing - Part 1: Conceptual model. ISO. https://www.iso.org/standard/61710.html

Queensland Government, 2009. Local Government Act 2009 (No. 17). https://www.legislation.qld.gov.au/view/html/inforce/current/act-2009-017

Standards Australia / Standards New Zealand, 2011, AS/NZS 4819:2011 - Geographic information - Rural and urban addressing. https://www.standards.org.au/standards-catalogue/sa-snz/publicsafety/it-004/as-slash-nzs--4819-2011

Standards Australia/Standards New Zealand, 2003, Geographic information – Rural and urban addressing (AS/NZS 4819:2003). https://www.standards.org.au/standards-catalogue

BIOGRAPHICAL NOTES

Ashlee Poeppmann is a Team Leader in the Spatial Information branch of the Queensland Department of Natural Resources and Mines, Manufacturing, and Regional and Rural Development. Ashlee leads the team who maintain state-wide addressing and roads foundational datasets. Ashlee has a keen interest in leveraging new technologies to enhance the accuracy and accessibility of foundational spatial datasets for improved public service delivery and decision-making Ashlee has been a member of the Geospatial Council of Australia since 2024.

Michael Elliott-Smith is a Manager in the Spatial Information branch of the Queensland Department of Natural Resources and Mines, Manufacturing, and Regional and Rural Development. Michael oversees the teams maintaining state-wide datasets of Addresses, Roads and Tracks, Place Names and Administrative Boundaries. Michael has a keen interest in process modernisation with a vision for unlocking data’s potential for the benefit of all Queenslanders.

CONTACTS

Miss Ashlee Poeppmann
The Department of Natural Resources and Mines, Manufacturing, and Regional and Rural Development (NRMMRRD)
Level 6, 1 William Street
Brisbane City QLD, 4000
AUSTRALIA
Email: ashlee.poeppmann@resources.qld.gov.au
Web site: https://www.nrmmrrd.qld.gov.au/

Mr Michael Elliott-Smith
The Department of Natural Resources and Mines, Manufacturing, and Regional and Rural Development (NRMMRRD)
Level 6, 1 William Street
Brisbane City QLD, 4000
AUSTRALIA
Email: michael.elliott-smith@resources.qld.gov.au
Web site: https://www.nrmmrrd.qld.gov.au/