NEW ZEALAND GEOSPATIAL OFFICE

Toitu te whenua
Land Information
New Zealand

SPATIAL DATA INFRASTRUCTURE COOKBOOK

V.1.1 – 17 NOVEMBER 2011

www.geospatial.govt.nz | www.linz.govt.nz

New Zealand Government

Chapter 1

DATA STEWARDSHIP AND CUSTODIANSHIP RESPONSIBILITIES
Outlines the key responsibilities of data stewards and custodians for the management of fundamental datasets. These can be applied across a wider geospatial data context.

Chapter 2

INTRODUCTION TO SDI STANDARDS
Contextual review of SDI Standards authorities and a description of 'essential' and 'highly desirable' standards and interoperability specifications for New Zealand's introductory SDI.

Chapter 3

MAKING DATA ACCESSIBLE: CHARACTERISTICS OF A PROVIDER NODE
How organisations can contribute to SDI by providing services of geospatial information.

Chapter 4

MAKING DATA ABLE TO BE FOUND: CHARACTERISTICS OF A CATALOGUE NODE
How organisations can contribute to SDI by providing catalogues of geospatial information.

Chapter 5

USING DATA EFFICIENTLY: HOW CAN MY ORGANISATION USE SDI?
How organisations can benefit from SDI participation, through the utilisation of catalogues and services of geospatial information.

Chapter 6

HOW EXISTING SYSTEMS AND PRODUCTS CAN CONTRIBUTE TO SDI
How supplier geospatial applications use open geospatial standards in support of the New Zealand SDI.

FEEDBACK AND QUERIES

TERMS
Key terms used in the cookbook.

ISBN: 978-0-473-19218-1

SPATIAL DATA INFRASTRUCTURE:

Preface

Preface

This initial version of an SDI cookbook for New Zealand has been developed to provide guidance for the early stages of implementing a national SDI. These chapters are based on what is currently known about the elements typically underpinning SDIs and guided by developments and experiences in jurisdictions internationally.

The approach presented in this cookbook recognises the contemporary geographical information systems in New Zealand organisations and how modern information architectures increasingly support aspects of SDI.

This document has been developed along the lines of other geospatial reference documents like the GSDI cookbook. The chapters can be read individually, or in whichever order suits the reader.

As a minimum, the guidance here provides initial steps and consideratioins so that willing organisations can commence their journey towards active participation in the national SDI.

Invitation to participate and provide feedback

This initial version of the cookbook has been developed with the help of members of the wider geospatial community. As such, it should be considered a work-in-progress. We intend to keep adding and building on this document, with the help of your valuable feedback.

As part of their community engagement activities, New Zealand Geospatial Office staff will present material within this cookbook to a wide range of audiences or as appropriate, arrange training on aspects of the cookbook. A formal consultation process will also be put in place, particularly to address critical elements of the SDI such as the development of community data specifications. The feedback from these engagement opportunities will help shape future versions of this cookbook.

We're really keen to hear your thoughts. You can comment on these pages through the comment function on this website[1], or email nzgo@linz.govt.nz

We're considering what other channels we can establish to capture thoughts and discussion on the content included, so we welcome your suggestions on this also.

1 http://www.geospatial.govt.nz/cooking-up-a-national-sdi-framework

SPATIAL DATA INFRASTRUCTURE:

Introduction

Introduction

Context

Land Information New Zealand (LINZ), through the New Zealand Geospatial Office (NZGO), is committed to working in partnership with central and local government, academia, businesses and the open data community to deliver a formal Spatial Data Infrastructure (SDI) over the next three years.

The development of an SDI also aligns with broader government initiatives such as the open government information and data re-use work programme.[2]

In December 2010 Cabinet agreed[3] that Land Information New Zealand lead the development of a more formalised SDI for New Zealand, with the New Zealand Geospatial Office designated as a consultative resource.

Cabinet also agreed that government agencies be directed to support and be involved in a collaborative way with LINZ and the NZGO as the national SDI framework is developed and realised. This cookbook contains initial guidance to help recognise or implement components that will contribute towards New Zealand's SDI.

There are many benefits of an SDI for New Zealand, including the potential for significant economic gains. Geospatial information is widely used in New Zealand and already contributes over $1.2 billion a year to the economy.[4]

It also forms a key part of New Zealand's knowledge infrastructure and enables innovation and better decision-making. Removing key barriers to connecting this information could add a further $500 million a year in productivity benefits and generate an extra $100 million in government revenue.

In December 2010, LINZ through the New Zealand Geospatial Office released a video which describes SDI in everyday terms and highlights the practical implications of these benefits.

You can watch the **video**[5] on the LINZ website.

A national SDI represents a comprehensive system of inter-related elements involving governance structures, policy, standards, data, hardware, software, and people across all levels of their organisations. This cookbook therefore has been structured with chapters that touch on these various elements to create a context for participation.

Cookbook Objectives

The design and implementation of a national SDI represents an ongoing journey which will move us from an 'introductory' state, with a focus on adoption and increased participation, through to a 'formal' SDI where participation is widespread and the emphasis is on improving the components and benefits of that involvement.

2 www.ict.govt.nz
3 http://www.linz.govt.nz/geospatial-office/about/projects-and-news/spatial-data-infrastructure/
4 http://www.geospatial.govt.nz/productivityreport/
5 http://www.linz.govt.nz/geospatial-office/about/projects-and-news/spatial-data-infrastructure/video/index.aspx

Introduction

This cookbook includes initial technical guidance to help organisations recognise and implement components that can contribute towards New Zealand's SDI. In its current form this document provides context and guidance for the introductory phase of the SDI, describing minimum requirements to help support participation at this early stage.

Intended audience

The primary audience for this initial version of the cookbook comprises technical management leads in government agencies, local councils and other organisations seeking to participate in the national SDI and who are currently at the early stages of that journey.

This includes a range of disciplines; GIS professionals, data analysts, business analysts, enterprise architects, information and solution architects, and programme and procurement managers.

While the material within this cookbook will be particularly relevant for those responsible for the implementation of geospatial systems within their organisation, others will also find it valuable. Because an SDI can generate positive changes to organisational structure, influence business plans, and enable significant workflow improvements, managers are also likely to benefit from this guidance. The material describing data stewardship and custodianship responsibilities, and how organisations can generally gain from SDI participation, will for instance resonate with those concerned with business outcomes. Components of this cookbook are also intended to facilitate procurement by providing information that can be readily included in tender documents.

Once realised, a formal national SDI will reflect participation and input from the entire New Zealand geospatial industry. So too should the guidance in support of that SDI represent a collaborative effort. In that spirit, the material included in this release of the cookbook has been developed by NZGO staff together with input from other central government agencies, local authorities, academia, members of the open data community and private sector organisations. For that valuable assistance we are most grateful.

As the New Zealand SDI evolves to a more formal state, this cookbook will also change to reflect guidance befitting a more mature implementation.

Along with the rest of the team at the New Zealand Geospatial Office, I look forward to continuing to work together with you and your organisations to establish and grow an effective infrastructure for our country's geospatial information.

Kevin Sweeney, GISP
Geospatial Custodian
ksweeney@linz.govt.nz

New Zealand Geospatial Office
Land Information New Zealand

SPATIAL DATA INFRASTRUCTURE:

1 Data stewardship and custodianship

1 — Data stewardship and custodianship

1.1 Introduction

Stewardship ensures that there is an organisation with formally appointed accountability for the management and maintenance of a geospatial dataset so that it meets the needs of most users. Stewardship includes all aspects of managing spatial information including data architecture, quality, maintenance, metadata, pricing, licensing, access and release.

The **steward** is accountable for maintaining the quality, integrity, availability and security of the data.

The **custodian** is responsible for the continued physical existence, maintenance, availability and dissemination of the data.

1.2 Data stewardship and custodianship

The stewardship and custodianship model is applicable to the management of all datasets whether they are fundamental or non-fundamental.

Stewardship is the act of ensuring that appropriate data management policies and standards are developed and maintained on behalf of the Crown. Stewardship covers all aspects of managing spatial information. This includes data architecture, quality, metadata, pricing, licensing, access and dissemination, bearing in mind any statutory responsibilities and government data management policies.

Custodianship of spatial information is the act of ensuring appropriate care in the collection, storage, maintenance and supply of the information.

The **Chief Geospatial Steward**[6] ensures that appropriate data management policies and standards are developed and maintained on behalf of the Crown, including the management of the custodial environment. The New Zealand Geospatial Office (NZGO) is best placed to fill the Chief Geospatial Steward role for geospatial data in New Zealand, as delegated from the Chief Executive of LINZ.

The over-arching principle is that stewards are formally appointed and responsible for engaging with the user community to ensure that the dataset is created and maintained to an agreed quality, using an agreed data specification and that the data is easy to find and use.

6 The Chief Geospatial Steward is a role proposed at the time of writing this cookbook. The role is included in a discussion paper submitted to the Geospatial Strategy governance groups (Geospatial Steering Committee and Geospatial Executive Group). The points from this discussion paper are reproduced here and are still subject to final approval and endorsement by the geospatial governance bodies and will be published as part of a formal consultation process.

1 — Data stewardship and custodianship

1.3 Criteria for stewardship selection

Formal stewardship is assigned in one of two ways:

1. The organisation has a statutory responsibility for the collection of specific data to meet certain obligations for the Crown. This is the primary criteria.

2. The organisation is invited to be a steward by the Chief Executive of LINZ under the auspices of the Chief Geospatial Steward.

If statutory responsibility is shared or unclear, or if there is no legislative authority, then the selection of the most appropriate organisation to undertake the stewardship responsibilities for any particular dataset will need to include the consideration of criteria (in no priority order) such as:

- does the organisation have the greatest business or operational need for the data?
- does the organisation have or is it best placed to assume a national perspective or sense of the broadest good for the data, across all potential stakeholders?
- does the organisation already maintain the data?
- is the organisation the best able to capture and/or maintain the data?
- is the organisation well placed to obtain the resources needed for the collection and maintenance of the data?
- is the organisation in the best economic position to justify the collection of the data?
- does the organisation have the highest quality standards for the data and is the organisation committed to continued improvement of the quality of the data?
- is the organisation willing to collaborate with others on the overall user requirements for the data, subject to the necessary funding for extra requirements?
- is the organisation willing and able to accept stewardship responsibilities?

1 — Data stewardship and custodianship

1.4 Responsibilities for stewards

The following are considered to be the essential responsibilities for stewards (whether an organisation has statutory responsibility for a dataset, or they have been assigned the stewardship role by the Chief Geospatial Steward).

1.4.1 Data collection, maintenance and revision:

- consult with the Chief Geospatial Steward and the user community[7] to determine data needs and priorities of spatial information in their custody.

1.4.2 Standards development:

- determine appropriate custodian/s and develop formal contracts/SLAs as necessary
- develop or revise existing data collection and maintenance requirements/processes (with all custodians of the data as necessary) to meet the community's agreed data needs and priorities
- coordinate with custodians for the data to ensure a deliverable that best meets the needs of the user community
- avoid duplication of capture, by ensuring, in conjunction with the Chief Geospatial Steward, that data to be captured is not already held elsewhere
- develop with the Chief Geospatial Steward and user community appropriate access and interoperability standards for the management and use of the datasets in their care
- develop a community schema for the definition of the data
- ensure that the datasets under their stewardship conform to appropriate national, international or agreed standards.

1.4.3 Quality:

- define the quality requirements of the dataset in collaboration with the user community.

1.4.4 Access:

- ensure the spatial information under their custodianship is both accessible and readily available
- ensure appropriate storage, maintenance, security and archival procedures for their spatial information
- safeguard the government's interest in the use of its information through licensing arrangements or letters of understanding to protect privacy and confidentiality
- act as the authoritative source for the information in their care

7 User ccommunity may include central/local government, CRIs, SOEs, business, academia and representatives from the open data community, as appropriate for that dataset.

1 — Data stewardship and custodianship

1.4 Responsibilities for stewards *continued*

- encourage the proper use of spatial information to discourage duplication through ignorance
- nominate a point of contact for enquiries about the datasets under their care.

1.4.5 Metadata:

- develop and maintain metadata for their data that complies with the ANZLIC[8] Metadata Profile and Guidelines
- provide metadata statements describing the data in their care to the SDI catalogue
- provide full and frank quality statements within the metadata regarding source, reliability, accuracy, completeness and currency.

1.4.6 Privacy:

- provide a level of appropriate security to protect the privacy of any personal data.

1.4.7 Negotiations:

- Stewards shall not negotiate unilaterally with any party either on an exclusive basis, or for the exclusive use of datasets under their care. Agreements should wherever possible be negotiated to benefit the user community as a whole, not any individual steward.

Once an organisation has agreed to accept the responsibility for stewardship of particular data, the responsibility and accountability for the stewardship will rest with the head of that organisation. Within the organisation, the various stewardship roles and responsibilities will need to be explicitly defined and documented, and then be formally delegated to the most appropriate group or individual. The key 'players' will need to be adequately briefed and resourced to perform the roles allocated to them. The responsibilities associated with each of these roles will need to be defined and agreed.

When implementing their responsibilities for a particular dataset, the steward will need to establish and maintain an appropriate mechanism to determine the needs of the data users within both their organisation and the wider community.

Appropriate funding mechanisms will need to be developed to encompass the needs of the wider community, particularly if the dataset meets the fundamental dataset specifications. It is not expected that the steward meet all of the funding needs for user requirements that are outside their normal business requirements, however they may consider this is in the best interest of New Zealand. Other agencies that need additional requirements should contribute to the funding, or agencies should be prepared to fund their own individual, specific needs which are over and above the base requirements of the user community.

8 ANZLIC – the Spatial Information Council http://www.anzlic.org.au/

1 — Data stewardship and custodianship

1.5 Responsibilities for custodians

The custodian will be responsible for the continued physical existence, maintenance, availability and dissemination of the data for as long as is required by the steward. The role of the custodian would include the following specific responsibilities:

- collect and maintain the datasets under their custodianship to appropriate national, international or agreed standards
- ensure appropriate storage, maintenance, security and archival procedures for their spatial information
- maintain appropriate and agreed metadata
- maintain appropriate geoprocessing specifications required to conform with those agreed for the SDI
- maintain a mechanism for data discrepancies to be registered and their resolution tracked
- ensure accessibility and availability of the data by agreement with the steward, to appropriate national, international or agreed standards
- review the data quality and provide quality statements to the steward on at least an annual basis
- nominate a point of contact to coordinate with the steward and represent their organisational perspective in relation to the data.

A steward may fulfil the role of custodian for the same dataset or they may choose to outsource this function. If functions are outsourced, the accountability remains with the steward for maintaining the overall quality, integrity, security, and release of the data.

SPATIAL DATA INFRASTRUCTURE:

2 SDI standards

2 — SDI standards

2.1 Introduction

Standards are essential for underpinning consistent connectivity and user experience within an infrastructure framework.

Internationally, the notion of a spatial data infrastructure has evolved in jurisdictions with ambitions to improve data quality and access, and achieve interoperability. The geospatial *infrastructure* concept is comparable to a road infrastructure (consistent road design, driving rules, etc) or telecommunications network (transmission standards, switching, multiple device types, etc).

Standards are necessary for facilitating robust, open transfer of packages of geographic information between platforms, anywhere, anytime. This is especially true in a varied network of computers that are managing a diverse range of geospatial data stores and data types.

The vision for New Zealand's Spatial Data Infrastructure (SDI), as outlined in the New Zealand Geospatial Strategy[9], encompasses a series of goals to improve data quality, access, and interoperability. This will provide a consistent basis for discovery, evaluation, and use of geographic information for everyone; governments, industry, non-profit organisations, academia, and citizens in general.

New Zealand's SDI will evolve over time. Initially, the focus is on providing relatively simple functionality and services, but it must also support the growth of more advanced capabilities in the future.

Open web map services (OGC WMS) will facilitate basic view and query access in the SDI. Some stakeholders, at least initially, will also want to interact with the geographic data made available through the SDI. The best way to achieve this is through web feature services (WFS).

Beyond these basic considerations, an SDI can enable a considerable number of additional open services or software to support applications of geographic information. A fully interoperable SDI will enable advanced applications to discover resources through a catalogue, combine diverse services and geographic data, and automate processes, all in a platform-independent way.

This chapter focuses on entry level standards that are considered essential, but also encourages adoption of other 'highly desirable' interoperability specifications for New Zealand's initial steps toward SDI.

Some additional components are introduced in this chapter, however advanced components that are found in more highly developed SDI implementations are not covered in this version of the Cookbook.

The New Zealand Geospatial Office (NZGO) can provide supplementary advice on standards, and is very interested to work alongside communities who wish to advance any characteristics of New Zealand's SDI. The office maintains linkages with geospatial standards bodies, including ISO, OGC, ANZLIC, Standards NZ, etc.

9 http://www.geospatial.govt.nz/geospatial-strategy/

2 — SDI standards

2.2 SDI standards bodies

Geographic information standards that will be of the most value to the New Zealand SDI are those based upon standards developed and maintained for the wider information technology industry by the International Organisation for Standardisation (ISO). It is important that the SDI uses these more widely applicable standards since this approach facilitates interoperability with other IT systems and reduces effort for our communities.

Geographic information standards are primarily developed and maintained by the ISO Technical Committee 211 (ISO/TC211)[10] and the Open Geospatial Consortium (OGC)[11]. The development and implementation of standards for geospatial information started in 1994 when the ISO/TC211 and the OGC were formed. From the outset the geographic information standards were conceived as dependent on other ICT industry standards.

ISO/TC211 and OGC collaborate closely, and their objectives are defined in a cooperative agreement. Essentially they define and maintain two types of standards in the field of geographic information management:

- **service invocation standards** – these define the *interfaces* that allow different systems to work together, or the expected behaviour of geoprocessing software systems
- **information transactional standards** – these are used to define the content of geospatial information or its encoding for interchange between different processing systems.

Through its membership of ISO, Standards New Zealand[12] facilitates New Zealand's participation in the international standards-setting and maintenance programme of ISO/TC211. Standards NZ and Standards Australia jointly monitor standards that may be relevant for the Australasia region. Many ISO geographic information standards have been adopted without modification as Standards Australia/Standards New Zealand (AS/NZS) standards.

A number of the foundation geographic information standards are recommended within the New Zealand e-Government Interoperability Framework (e-GIF)[12]. In 2010 the ANZLIC Metadata Profile[14] was added to the set of recommended geospatial standards.

10 http://www.isotc211.org/
11 http://www.opengeospatial.org/
12 http://www.standards.co.nz/default.htm
13 http://www.e.govt.nz/standards
14 http://www.e.govt.nz/news/anzlic-metadata-profile-replaces-nzgms

2 — SDI standards

2.3 Core SDI standards

Including the various standards that sit within information and communications technology, nearly 100 standards can be identified that can be considered as part of the architecture and deployment of interoperable geospatial solutions.

Moving too swiftly into the selection of an appropriate technical architecture can be daunting, and yet if independent selections of standards were to occur this could lead to incompatibilities within a national SDI.

The GSDI Cookbook (a resource collectively developed by the Global Spatial Data Infrastructure community)[15] recommends that the *"… definition of a relatively small suite of standards allows a shorthand reference for nominal capabilities in an SDI environment, with provision for identifying optional supplemental standards."*

The geospatial standards referenced here represent an initial list developed as part of work to establish a small set of standards suitable for New Zealand's first steps in the journey toward formal national SDI. The recommended first step is OGC WMS as well as maintaining adequate metadata (at least conforming to the few core elements of the ANZLIC Metadata Profile).

Essential	**Abbreviation**
OGC Web Map Service v1.3	WMS
ANZLIC Metadata Profile v1.1	ANZLIC Profile
Highly Desirable	
OGC Web Feature Service v1.1	WFS
OGC Filter Encoding v2.0	Filter
OGC Geography Markup Language v3.2.1	GML
OGC KML v2.2	KML
OGC Catalogue Service 2.0.2	CSW
Additionally Useful[16]	
OGC Web Coverage Service 1.1.2	WCS
OGC Styled Layer Descriptor Profile v1.0	SLD
OGC WFS (Transactional) v1.1	WFS(T)
OGC Web Processing Service	WPS
OGC Web Map Tile Service (includes a REST option)[17]	WMTS
Other Standard Interfaces	
OGC Web Map Context v1.1	WMC
OpenGIS Sensor Model Language, Version 1.0.0	SensorML
OpenGIS Sensor Observation Service, Version 1.0.0	SOS

15 http://www.gsdidocs.org/GSDIWiki/index.php/Chapter_10
16 These OGC specifications are accessible from http://www.opengeospatial.org/standards
17 Implementation standard to serve digital maps using predefined image tiles. NB includes an approach to the resource oriented architecture style: http://www.opengeospatial.org/standards/wmts

2 — SDI standards

2.3 Core SDI standards *continued*

XML and GML are referenced heavily in the following section so a brief explanation of them is included here:

XML – *Extensible Markup Language* – this is the universal format or 'language' for expressing information in a way that can be understood and processed by computers and web services.

GML – *Geography Markup Language* – this is an extension of XML that specifically relates to describing geographic information, for example shapes, points and coordinates.

2.3.1 Essential

2.3.1.1 WMS (Web Map Service)[18]

OGC WMS enables open sharing of information in the form of online map images. It is the simplest form of open geospatial interoperability. Internationally it is by far the most popular and widely implemented of the open geospatial standards. The most recent version of the OGC Web Map Service is WMS 1.3; AS/NZS ISO 19128:2006.

OGC WMS is considered the easiest way to begin to participate in New Zealand's SDI. The service generates graphical images and most client applications can read these.[19]

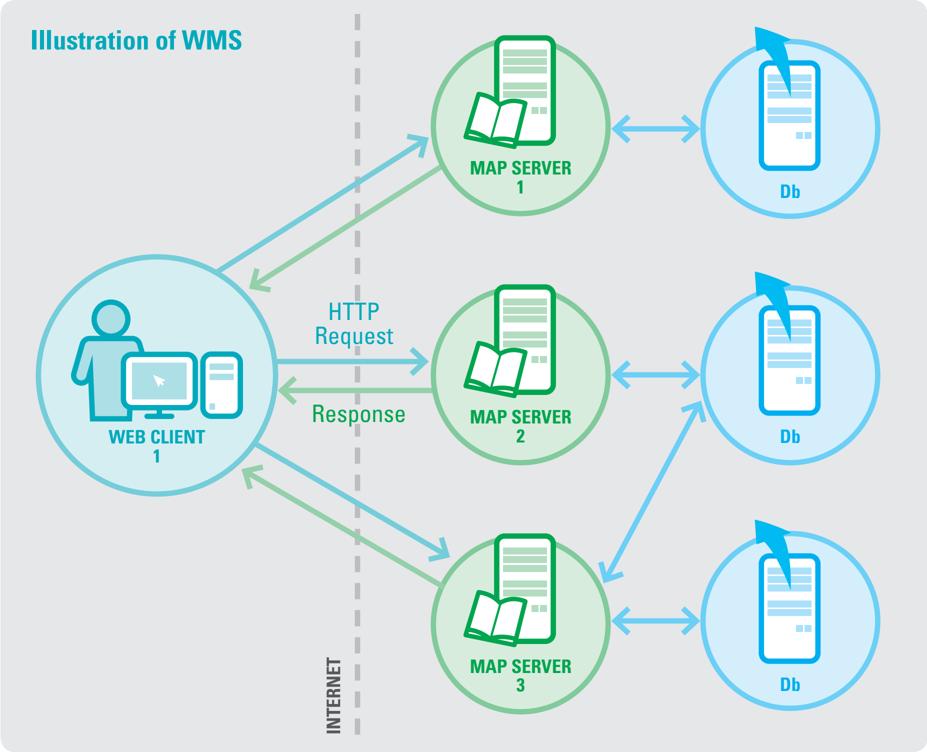
Wide use of this specification can 'kick-start' the basic component of New Zealand's SDI. For many scenarios it is likely to be completely adequate for helping make geospatial information widely accessible and useful.

The Web Map Service standard mandates two operations (GetCapabilities and GetMap). The GetCapabilities operation determines what's available. OGC WMS makes a request to the database of the agency serving the data and delivers to the user a graphical image (GIF, JPEG, TIFF, etc.). Requests may be sent to one or more WMS and the returned images are overlaid in browsers or client applications. This enables the creation and display of map-like views of information that arrive simultaneously from multiple sources.

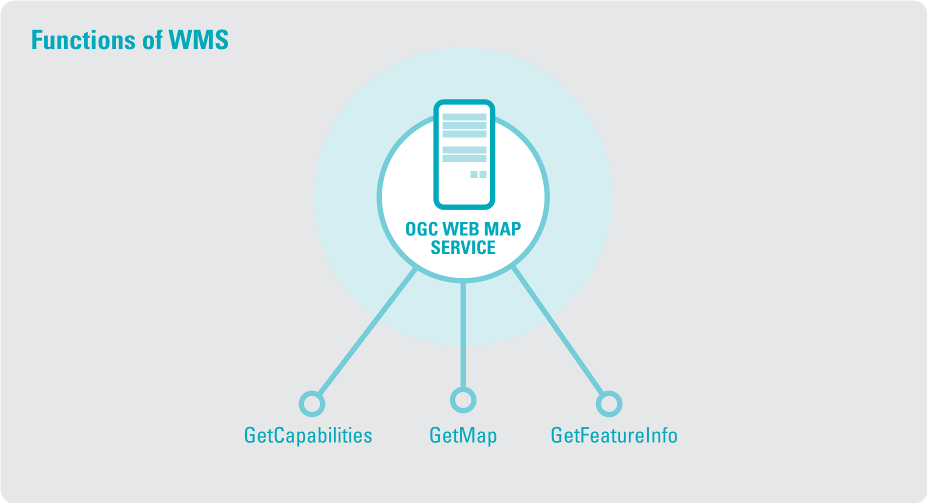
18 http://www.opengeospatial.org/standards/wms
19 Whereas when implementing every other geospatial web service (e.g.WFS) these are more likely to require a substantial increase in considerations for server performance and for clients

2 — SDI standards

2.3
Core SDI standards
continued



An optional third operation (GetFeatureInfo) provides access to information about displayed map features. With this option the features 'behind' the map can be queried, and their properties can be returned to a requesting client i.e. the user selects a point, line or polygon feature on the map to display further information about that feature.



2 — SDI standards

2.3
Core SDI standards
continued

An additional Styled Layer Descriptor Profile[20] is available. This enables users to control the visual portrayal (symbology) of layers and provision for legends.

The service provider determines what information to include, exclude or aggregate in each WMS layer served. Each layer should be documented in accordance with the ANZLIC Metadata Profile (ISO 19115) and provide a link to that metadata. The WMS standard provides guidance, including the relationship between WMS Layer Properties and the relevant metadata elements.

2.3.1.2 ANZLIC Metadata Profile[21]

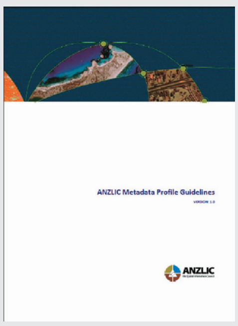
Metadata, simply defined, provides information about a dataset or the data within it. The ANZLIC Metadata Profile is applicable to the cataloguing of geographic datasets, the full description of feature characteristics, and services. It is a recommended standard in the New Zealand e-Government Interoperability Framework (e-GIF)[22].

An organisation's information and services are valuable assets, and substantial amounts of time, money and effort may be invested in them. Metadata contributes a significant role towards managing and putting these assets to good use, re-use or even re-purposing. Metadata benefits include:

- enabling effective cataloguing of what exists
- facilitating confident discovery
- enabling others (unfamiliar with the resource) to assess fitness for their purpose
- supporting confident information sharing and effective collaboration.

The standard for this is an Australian/New Zealand Profile of AS/NZS ISO 19115:2005, Geographic information – Metadata (implemented using ISO/TS 19139:2007, Geographic information – Metadata – XML schema implementation). The Profile defines the schema required for describing geographic information and services. It provides for comprehensive, structured information about the identification, extent, quality, spatial and temporal schema, spatial reference, and distribution of digital geographic data.

ANZLIC Metadata Profile Guidelines



20 http://www.opengeospatial.org/standards/sld
21 http://www.linz.govt.nz/geospatial-office/about/projects-and-news/anzlic-metadata-profile/index.aspx
22 http://www.e.govt.nz/news/anzlic-metadata-profile-replaces-nzgms

2 — SDI standards

2.3 Core SDI standards *continued*

The Profile embraces all of the AS/NZS ISO 19115:2005 elements. However, only a small set of core metadata elements must be completed in order to conform to the standard.[23] The vast majority of elements remain optional. When adding metadata, authors are encouraged to consider how they may be advantaged (and can empower others) by adding extra content that is readily accessible to help users confidently interpret and use the resource.

Metadata can be entered and encoded using a metadata entry tool. Extensive online resources are available about the ANZLIC Profile, including access to a free metadata entry tool.[24]

Metadata that conforms to the standard enables information resources to be consistently found, shared and re-used using the network of web-based catalogues. An example of a national catalogue that operates this way is geodata.govt.nz; New Zealand's catalogue of publicly-funded geospatial data. The catalogue harvests metadata compliant with the ANZLIC Metadata Profile.[25]

2.3.2 Highly Desirable

2.3.2.1 WFS (Web Feature Service)[26]

For some users, access to graphical images (maps) of spatial datasets alone (via WMS) may not be adequate.

Advancing to the OGC WFS 1.1.0 service exposes the actual data e.g. the spatial feature geometries and related properties that it is made up of. This data is expressed in XML schema form. Users typically interact with WFS through browser-based or desktop geospatial clients, which allow them to access *vector data*[27].

WFS allow clients to retrieve geospatial data from multiple WFS instances encoded in Geography Markup Language (GML). GML is the expected packaging (XML dialect) for features requested from a WFS. The request is generated on the client and is posted to a web feature server using HTTP. The web feature server then reads the request and returns a set of results in GML, as defined through GML application schemas referenced by the service. These can be GML schema defined by an information community (see section on Community Data Specifications).

WFS enables users to access just the pieces of information that they want, and combine them with additional specific pieces of data sourced from other WFS. The WFS makes a request to the database of the agency serving the data and delivers to the user's client the preferred data from the database which fits within the bounds of their map on the screen.

23 Minimum elements required to create an ANZLIC-compliant metadata record http://www.osdm.gov.au/Metadata/ANZLIC+metadata+resources/ANZMet+Toolkit+(final+draft+-+07.2009)/09_ANZLIC_MetaProfile_Short[table].pdf/?id=1003

24 http://www.linz.govt.nz/geospatial-office/about/projects-and-news/anzlic-metadata-profile/index.aspx

25 http://www.geodata.govt.nz/

26 http://www.opengeospatial.org/standards/wfs

27 Geographic features represented with vectors use a geometry type of point, line or polygon. NB: refer to WCS for interoperability with coverage features e.g. a set of features, or features whose geometries are of type set of cells or set of pixels (surfaces), such as imagery collections derived from satellites, multi-beam, LiDAR, and numerous other sensors.

2 — SDI standards

2.3 Core SDI standards *continued*

Initially, and as a component of introductory NZ SDI, WFS will use a basic form of XML encoding i.e. transact data 'as is' from existing data stores. Over time, community data specifications will enable a commonly agreed structure of the data to be shared and well documented content (conforming to international standards) will enable wider re-use, re-purposing and semantic interoperability. These domain specific data specifications will be expressed in GML. Such specifications typically require extra effort to agree (initially), so these will appear in a more mature SDI (see GML).

A WFS works in a transactional way – for example, users can describe the specific data they need and the WFS uses this description to query the database holding that data. If the information meets the agreed standards, it's much easier for the WFS to return what the user wants providing a much more efficient service.

The WFS standard defines operations that enable clients to:

- Discover which feature collections are available (GetCapabilities)
- Describe the attribute fields available for the features (DescribeFeatureType)
- Query a collection for a subset of features based on a provided filter (GetFeature)

An extension to WFS can also be configured to enable clients to update features (WFS-Transactions) i.e. add, edit or delete geographic features.

2.3.2.2 Filter Encoding[28]

Filter encoding is applicable to the development of systems that **use** the interfaces specified by WFS (or a number of other OGC interfaces).

This is a neutral way, based on XML encoding, to manage transactions (searches, retrieval, operations on data, etc). XML can be validated and transformed into whatever target language is necessary to retrieve or modify features.

2.3.2.3 GML (Geography Markup Language)[29]

GML is an implementation of XML, the standard language for encoding information in a machine-readable form. GML enables the specific characteristics of geospatial data, such as geometries and other specialised characteristics, to be represented.

The ability to integrate all forms of geographic information is key to the utility of GML. GML serves as a modelling language for geographic information as well as an open transactional interchange format. Data encoded in GML can be validated using XML schemas, published with the standard and maintained in a schema repository by the OGC.

The OGC Geography Markup Language (GML) v3.2.1 is also known as AS/NZS ISO 19136:2008.

28 http://www.opengeospatial.org/standards/filter
29 http://www.opengeospatial.org/standards/gml

2 — SDI standards

2.3 Core SDI standards *continued*

This standard defines the XML schema syntax, mechanisms, and conventions that:

1. provide an open, vendor-neutral framework for the description of geospatial application schemas for the transport and storage of geographic information in XML
2. allow profiles that support proper subsets of framework descriptive capabilities
3. support the description of geospatial application schemas for specialized domains and information communities
4. enable the creation and maintenance of linked geographic application schemas and datasets
5. support the storage and transport of application schemas and data sets
6. increase the ability of organisations to share geographic application schemas and the information they describe.

Regarding point (3) above – advanced interoperability can be achieved when domain specific semantics are agreed e.g. vocabularies, code lists and ontologies. GML enables collaborative information communities to achieve open data content specifications that are sophisticated, model-driven, platform-independent, machine-readable and verifiable (refer to the section on Community Standards).

2.3.2.4 KML[30]

OGC KML is used predominantly for visualising geospatial data in Google mapping applications. It is an OGC standard.

KML is an XML language focused on geographic visualisation and used to encode and transport representations of geographic data for display in a web browser, including annotation of maps and images. Geographic visualisation includes not only the presentation of graphical geographic data, but also the control of the user's navigation.

KML, like GML, uses a tag-based structure with nested elements and attributes. KML documents and their related images (if any) may be compressed using the ZIP format into KMZ archives. KML documents and KMZ archives may be shared by e-mail, hosted locally for sharing within a private intranet, or hosted on a web server.

2.3.2.5 Catalogue Services[31]

Catalogues are an essential component of SDI. Catalogue nodes rely on open standards to connect users, providers and other catalogue nodes. The OGC has a specification for catalogue services. It defines common interfaces to discover, browse, and query metadata about data and services.

For more information on this please refer to the chapter entitled *'Characteristics of a Catalogue Node'*, within this Cookbook.

30 http://www.opengeospatial.org/standards/kml
31 Unified Modeling Language is used to create standardised geographic information and service models; AS/NZS ISO 19103:2006 Geographic information – Conceptual schema language

2 — SDI standards

2.4 Community standards

2.4.1 Community Data Specifications (data content)

Geographic information transactional standards, such as WFS (see above), can use agreed community content standards for the open exchange of data between different processing systems i.e. machine-to-machine interoperability of geospatial data across a network of diverse systems.

Such content standards are structured according to the conceptual modelling framework used in the ISO/TC211 standards suite. This enables an information community to conform to a well-known structure and to derive benefits inherent with international consensus. It also facilitates comprehensive documentation and agreement on all the requirements for all the spatial and non-spatial characteristics of the geographic features that are of interest to that information community.

Development of these content standards is usually driven by an information/knowledge community (sometimes called a '*domain*'). The community defines the information outcomes it desires and orchestrates a project to design an information model to amply express all the characteristics of all the geospatial data needed to be openly shared, re-used, or re-purposed. The community engineers its information requirements in a standards conforming data specification (an application schema in graphical form, using UML notation). This is converted into GML to achieve a platform neutral machine-readable application schema.

This level of standardisation hinges on a culture of collaboration and sharing. The community needs a desire to conform to open standards and resolve semantics, and the ability to reach consensus on data models that satisfy the collective need. Generating a community model (data specification) has a few prerequisites. It can be relatively straightforward to achieve, providing that the community:

- is active and engaged
- has a well-defined and accepted conceptual model to start from
- knows how to use the software tools for developing such models
- is prepared to engage with the other 'standards' communities, and
- accepts that it is not about being right or wrong, but about agreement.

2 — SDI standards

2.4 Community standards *continued*

Some significant examples of community standards for transacting geographic information include:

- GeoSciML is an application schema that can be used to transfer information about geology of the type that was conventionally portrayed on geologic maps.
- INSPIRE Data Specifications provide for geospatial interoperability of the significant administrative and environmental datasets that need to be interoperated among the 27 nations of the European Union.
- IHO S-100 was developed by the International Hydrographic Organisation (IHO). It enables development of advanced global navigation products and services, and provides the platform for using and integrating official hydrographic data with other geospatial data.

2.4.2 Standards for Standards (informing the definition of data content)

A suite of *standards for standards* supports the development of community data specifications. The standards provide a framework of patterns for information architects to inherit in their models. This achieves well documented, openly structured data to help achieve the highest level of across-the-board re-use and interoperability. The suite of standards is especially relevant for fundamental data or other nationally significant datasets intended to be widely re-used or re-purposed.

ISO Technical Committee 211 leads development of the family of international standards that have been designed for the purpose of openly transacting geographic information content. These are commonly referred to as the ISO/TC211 suite of standards.

AS/NZS ISO 19101:2003 Geographic information – Reference model.
This provides a family of abstract conceptual schemas for describing the fundamental components of geographic features as elements of geographic information.

AS/NZS ISO 19109:2006 Geographic information – Rules for application schema
This specifies a general feature model for integrating these components into features and provides rules for doing so in an application schema i.e. to comprehensively record an open data specification.

AS/NZS ISO 19107:2005 Geographic information – Spatial schema
This specifies UML (Unified Modelling Language) classes for representing the spatial characteristics of features as composites of geometric and/or topological primitives.

AS/NZS ISO 19108:2003 Geographic information – Temporal schema
This specifies the UML classes for representing the temporal characteristics of features and also specifies classes for describing relevant temporal reference systems.

AS/NZS ISO 19123:2006 Geographic information – Schema for coverage geometry and functions
This provides a schema for an alternative representation of spatial information as a coverage, in which non-spatial attributes are assigned directly to geometric objects rather than to features composed of such objects.

2 — SDI standards

2.4 Community standards *continued*

AS/NZS ISO 19135:2006 Geographic information – Procedures for item registration
This is a standard for governance of reusable items, models, etc; a procedure for establishing, maintaining and publishing registers of reusable artefacts (such as aspects of data models).

AS/NZS ISO 19137:2008 Geographic information – Core profile of the spatial schema
This provides a profile of ISO 19107 that is limited to describing features as simple geometric primitives of 0, 1, or 2 dimensions.

ISO 19141:2008 extends ISO 19107 to support the description of moving geometric objects.

ISO 19146:2010 Geographic information – Cross-domain vocabularies
This defines a methodology for cross-mapping technical vocabularies that have been adopted by industry-specific geospatial communities. It also specifies an implementation of ISO 19135 for the registration of geographic information concepts for the purpose of integrating multiple domain-based vocabularies.

2.4.3 Coordinate Reference System

What differentiates geospatial information from other information domains, and helps derive and visualise additional insight from existing data, is the ability to combine different datasets using the common denominator of location. A common spatial framework is needed.

A series of spatial reference frames are applicable to New Zealand, offshore islands, and dependencies for the geodetic datums, vertical datums, and projections.

- New Zealand Geodetic Datum 2000 (NZGD2000) is the official geodetic datum for New Zealand and its offshore islands
- New Zealand Vertical Datum 2009 (NZVD2009) is the official vertical datum for New Zealand and offshore islands
- New Zealand Transverse Mercator 2000 (NZTM2000) is the projection used for New Zealand's national topographical mapping (scale 1:50,000) and other small scale mapping.

Refer to the **Land Information New Zealand** website[32] for details.

International codes for coordinate reference systems are maintained by the Geomatics Committee of the International Association of Oil and Gas Producers. Its Geodesy Sub-committee maintains and publishes a dataset of parameters for coordinate reference system and coordinate transformation description.[33]

32 http://www.linz.govt.nz
33 http://www.epsg.org/

2 — SDI standards

2.5 Standards compliant applications

The OGC provides a testing programme with the objective of increasing systems interoperability while reducing technology risks. Solutions can become 'Certified OGC Compliant'. This is a mechanism by which users and buyers of software can increase certainty that the software follows the mandatory rules of implementation as specified in the standards.

The aim is that buyers gain confidence that a compliant product[34] will work with another compliant product based on the same OGC standard, regardless of which company developed the product.

In Chapter 6 there is a list of applications supported in New Zealand which the submitters claim have implement specified standards to contribute to SDI. Where additional assurance is desired, concerning compliance with the standards, the OGC Compliance Products list may assist.

2.6 Resources

Open Source Geospatial Foundation (OSGeo)
http://live.osgeo.org/en/standards/standards.html

USA Federal Geographic Data Committee (FGDC)
http://www.fgdc.gov/standards/fgdc-endorsed-external-standards/index_html

The Open Geospatial Consortium (OGC)
http://www.opengeospatial.org/resource/cookbooks

Scotland's SDI guidance for Web Map Service
http://www.scotland.gov.uk/Publications/2010/05/06161701/0

ISO/TC 211 Geographic information/Geomatics
http://www.isotc211.org/

Standards New Zealand
http://www.standards.co.nz

34 http://www.opengeospatial.org/resource/products/compliant

SPATIAL DATA INFRASTRUCTURE:

3

Making data able to be accessed: characteristics of a provider node

3 — Making data able to be accessed: characteristics of a provider node

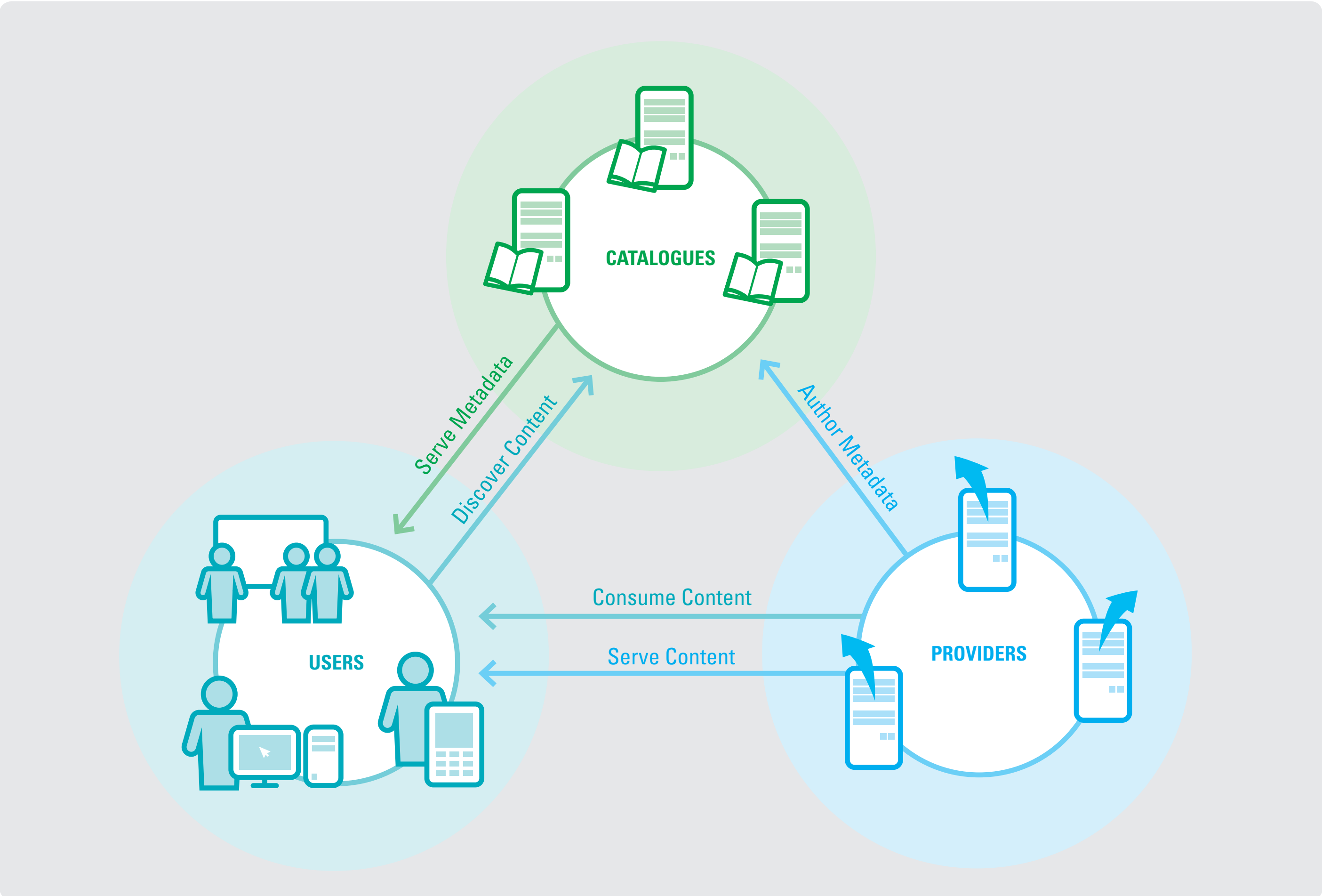
3.1 Participating in Spatial Data Infrastructure

This chapter is one of three (Chapters 3-5) that explain how organisations and their supporting systems can participate in Spatial Data Infrastructure (SDI). The focus is on systems because all other aspects of SDI participation require a deep understanding of the business focus of the organisation; which will be different in every case.

Systems participate in one of three ways.

1. They can contribute to SDI by providing services of geospatial information. These systems are then acting as 'Provider Nodes'.
2. They can contribute to SDI by providing catalogues of geospatial information. These systems are then acting as 'Catalogue Nodes'.
3. They can benefit from SDI by utilising the catalogues and services of geospatial information. These systems are 'User Nodes'.

The relationship between different nodes is shown in the diagram below:



3 — Making data able to be accessed: characteristics of a provider node

3.1 Participating in Spatial Data Infrastructure *continued*

Systems participate in SDI by having certain characteristics that create interoperability with other nodes. That interoperability is enabled through the use of well-defined and open international standards.

The guidance provided as part of this cookbook is intended to define SDI participation characteristics in vendor-neutral terms. As such, this and the subsequent two chapters could be appended to requests for proposal (RFPs) for any geospatial system which needs to participate in SDI.

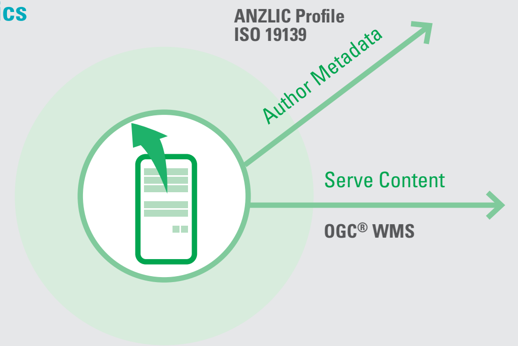
This chapter explains how an organisation can contribute to the New Zealand SDI by establishing a provider node. Provider nodes are an essential foundation for SDI as they provide greatly improved accessibility to geospatial information. Most organisations are already accessing fundamental data so this document explains why improved accessibility is so important.

It is also important to understand the context of provider node use. SDI is not intended to be used directly by a member of the public since the layperson's needs to access geospatial information is typically satisfied using third-party applications and mapping tools such as Google Maps and Bing Maps. SDI provides a foundation for public-facing government data and subsequent applications leveraging that data, but the foundation itself is entirely transparent to the public.

3.2 A provider node

In order to contribute to SDI, an organisation with geospatial information will need to create a provider node to serve that information. This chapter describes the characteristics of a 'well behaved' provider node and provides a guide for establishing a provider node.

Minimum characteristics of a provider node



- Create and Maintain Content
- Create and Maintain Metadata
- Maintain Web Services

3 — Making data able to be accessed: characteristics of a provider node

3.2 A provider node *continued*

A key purpose of SDI is to make spatial information easily discoverable, easily accessible and easily usable. It is important to define 'easily' in terms of the user's expectations rather than the provider's constraints. Ease of use is becoming paramount and ideas such as 'one click to content' are becoming common means of expressing growing expectations amongst the user community.

Using today's technology it costs little extra to provide a rich range of services. Doing so satisfies a wide range of user needs (and expectations) and avoids the risk of parallel provider nodes being established to serve specific user communities. It is very important to appreciate the business focus that lies at the heart of SDI; the aim is to make business uses of spatial information more efficient and effective.

Described here are the essential characteristics of a provider node that an organisation could implement to participate in SDI immediately. Characteristics that would require more substantial time, risk and resource to deliver are shown as 'highly desirable' or 'additionally useful'. This document does not consider emerging technology trends such as cloud-based computing which may offer different approaches in the near future; suffice to say that any system that is participating in an infrastructure needs to be as capable of evolution as is possible.

3.3 Characteristics of a provider node

The following sections describe the characteristics that define a provider node that will fully contribute to the New Zealand SDI by providing a range of services to achieve maximum business benefit for users. In each case the characteristic is described in terms of a priority of requirements:

Essential – the absolute minimum requirement in order for the system to participate in SDI. An implementation that does not satisfy these requirements will not contribute usefully to SDI.

Highly Desirable – requirements important to include in the implementation of this feature but not so important as to be considered absolute requirements.

Additionally Useful – requirements which may provide additional levels of benefit but are not critical to the implementation of this feature. They may be included in the implementation of this feature if time permits, but priority must be given to the 'essential' and 'highly desirable' requirements. In some cases, these additionally useful services might include proprietary services which should be provided where they are achieving business outcomes for users that cannot immediately be satisfied by the use of open standards.

3.3.1 Content

In considering content that is to be provided through the node, it is important to appreciate different use scenarios. Some users simply need a map background against which to overlay their work. In this case, a service of scanned maps provides an authoritative depiction of the background map. Since it is only contextual, there is no need to query this service.

3 — Making data able to be accessed: characteristics of a provider node

3.3 Characteristics of a provider node *continued*

Other users may need to query data. These users require access to features through a web mapping service. In this case, the utility of the service is greatly enhanced by separating groups of layers into individual services.

Above all, SDI will see unanticipated uses of data and it is this unforeseen use that offers the greatest reward for New Zealand. Some organisations might see risk in the release of data because of concerns such as liability or reputation. Other jurisdictions' experiences with SDI have shown that reward greatly outstrips risk, and that the potential risks are often greatly over-exaggerated during pre-SDI discussions. If risk is over-emphasised and mitigation involves restricting access to spatial information, SDI is more likely to fail to provide a broadly useful infrastructure.

SDI can only be based on the spatial information that organisations already have. To stall SDI until all data are inventoried, documented and made perfect (or even 'good' in some cases) risks delaying the implementation of SDI. It is important to appreciate that SDI is a powerful tool for improving spatial data – increased visibility will result in more feedback to correct errors.

An organisation will need to perform a data audit in order to determine what geospatial data they have and what layers of information they can provide to the national SDI. New Zealand Open Government[35] initiatives suggest that all government data should be made accessible, except those containing information about identifiable individuals. In addition, government agencies with geospatial information are now required to release that information for broader use, as specified in the Cabinet paper '*Capturing the economic benefits of better connecting our location-based information*' agreed by Cabinet in December 2010.

In the context of applicable government mandates, organisations participating in SDI will need to prioritise geospatial data to be made available through the SDI provider node. While decisions regarding that prioritisation will need to factor in costs, available resources and organisational business needs, they will also need to incorporate a consultation component to capture the broader perspective supporting needs of the wider user community. Guidance in regards to these roles and responsibilities for data providers is covered in the chapter Data Stewardship and Custodianship within this cookbook.

Essential

- Those geospatial data sets that the organisation is already mandated to make accessible and that are part of the organisation's core business.

Highly Desirable

- All other geospatial datasets that the organisation can make available.

35 http://www.ict.govt.nz/programme/opening-government-data-and-information/declaration-open-and-transparent-government

3 — Making data able to be accessed: characteristics of a provider node

3.3 Characteristics of a provider node *continued*

3.3.2 Service Types

A key outcome of SDI is that geospatial data become more easily accessed. It is important to understand the traditional workflow involved in using geospatial data in order to appreciate why this outcome offers considerable efficiencies.

A GIS user or business system traditionally receives geospatial data as a file on a DVD or other media. In order to use any of that geospatial data, the whole file has to be loaded onto the user's system. One of the characteristics of spatial information is that it tends to be large so this could involve the transfer of several Gigabytes of data. Moreover, the data may have to be translated from the supplier's format into a format understood by the user's system. This can be a time-consuming, processor-hungry activity. All of this activity is a precursor to the actual use of this data in the system.

From this, it can be seen that a file download, using a mechanism such as file transfer protocol (FTP), offers little advantage over DVD delivery. Indeed, from an ICT perspective, the surge in bandwidth required to download whole files of geospatial information can be seen as a significant disadvantage. It should be noted however that alternative methods of user-definable file download in the form of clip-zip-ship services may offer significant user benefits.

Using a modern web services approach, the GIS user or business system can directly connect to the service and thus directly consume the content into their system. This offers a number of advantages. Perhaps most important is that there is no precursor to using the data. This offers the opportunity for the GIS or business system to be used more directly as a tool, and as such the effectiveness is improved. In systems terms, the user is able to access just the extent of data required to conduct their task. This reduces processing and bandwidth terms, improving efficiency.

The Open Geospatial Consortium (OGC)[36] has worked for many years to establish a range of open standards for the web service delivery of spatial content. A wide range of rich interface specifications are available for use, and are explained in the previous chapter.

There are numerous types of web services, some open and some proprietary, that offer differing degrees of richness to users. It is tempting to regard services based on open standards as being the only mechanism for delivering content to users. However, proprietary standards may offer significant advantages to specific user communities by taking earlier advantage of rapidly evolving underlying web standards. The recommendation then is to prioritise those services that will deliver the greatest value to the user community within the context of the provider organisation and its capabilities.

As providers of the country's reference or fundamental data, which is used by the greatest number of geospatial consumers, it is particularly important that government agencies make that data available in a way that supports user needs to the greatest degree possible. If for example a government agency fails to provide a full range of service types, there is a significant risk that a third party will fill the gap by providing authoritative content using these other service types. That creates unnecessary confusion in the marketplace.

36 www.opengeospatial.org

3 — Making data able to be accessed: characteristics of a provider node

3.3 Characteristics of a provider node *continued*

Therefore it is highly recommended that data providers add as many service types as possible in addition to the essential open standards that must provide the baseline service.

The open standards that define a provider node are often misunderstood. There are three key standards: Web Mapping Service (OGC WMS)[37] ; Web Feature Service (WFS)[38]; and Web Coverage Service (WCS)[39]. These are described in the preceding chapter but the high level characteristics are:

- WMS returns a map image (a picture of the map) based on a client request. As such, WMS is sometimes misinterpreted as a 'dumb' standard as 'it just returns an image'. That is incorrect since WMS supports optional request types that can query underlying features. WMS is the most commonly implemented OGC standard.

- WFS returns a stream of features into the client. This allows a map to be rendered in the client software and will also support client-side query and analysis of the features. The transactional form of WFS will also allow the editing of features. WFS demands more sophisticated client-side capabilities and it is a far less common implementation.

- WCS provides access to geospatial 'coverages' which typically refer to content such as imagery and elevation data. Although the outcome of a WCS query might result in a map, it is important to understand that WCS is intended to support richer spatial analysis.

Essential

The minimum standard required for a provider node is:

- OGC WMS – Web Mapping Service

Highly Desirable

- WFS – important for analytic use
- KML[40] – important for Google users

Additionally Useful

- WCS – important for analytic use, particularly for raster datasets
- REST API[41] – important for interoperability with other types of enterprise information systems
- Clip-zip-ship service – to allow download of any layers in many formats, many projections. This is significant to permit any user of any proprietary system to receive data in the optimum format for their use.

37 Web Mapping Service – http://www.opengeospatial.org/standards/wms
38 Web Feature Service – http://www.opengeospatial.org/standards/wfs
39 Web Coverage Service – http://www.opengeospatial.org/standards/wc
40 Formerly Keyhole Markup Language – http://www.opengeospatial.org/standards/kml
41 http://en.wikipedia.org/wiki/Representational_State_Transfer – potentially useful architecture for interoperability with some enterprise systems

3 — Making data able to be accessed: characteristics of a provider node

3.3 Characteristics of a provider node *continued*

3.3.3 Metadata

A key requirement of SDI is that geospatial data is discoverable. Metadata provides a description of data in a structured way that allows for useful discovery. Although metadata records are most useful as part of a catalogue node, it is essential that metadata accompanies each service as well. So metadata form an essential element of a provider node.[42]

Essential

- ANZLIC Metadata Profile ISO 19139[43] record associated with each service (aggregated layers)
- Mandatory fields populated

Highly Desirable

- ANZLIC Metadata Profile ISO 19139 record associated with each layer that is exposed as a service
- More than mandatory fields populated

Additionally Useful

- ANZLIC Metadata Profile ISO 19139 record associated with each layer that is exposed as a service
- All relevant fields populated

3.3.4 Licensing

Geospatial information is expensive to produce and inevitably contains valuable intellectual property (IP). It is important to appropriately protect this IP. Government agencies currently use a diverse range of licensing arrangements to protect their IP. As part of broader open government initiatives, the New Zealand Government Open Access Licensing framework (NZGOAL)[44] provides Creative Commons licences to protect the release of non-personal government data[45]. It recommends the use of the most open licence, the Creative Commons Attribution (BY) licence, unless a restriction applies.

All Creative Commons New Zealand licences contain attribution requirements, although these can be waived by the licensor (ie the licensing agency).

Content provided through a web service also needs to be protected. It is essential to note that there is no application service associated with an SDI provider node. In instances when data from multiple datasets may feed into an end application, or be combined, NZGOAL recommends that agencies keep attribution requirements (if any) to a minimum – see paragraphs 37-44 in the NZGOAL framework document referenced above. The licence conditions should be added into the licensing field of the metadata.

42 For more info see http://www.geospatial.govt.nz/linz-coordinates-geospatial-metadata-workshops-and-provides-ongoing-support-for-users/
43 Australia New Zealand Land Information Council profile of the International Standards Organisation's 19139 standard: http://www.anzlic.org.au/metadata/
44 http://www.nzgoal.info/
45 An example of a New Zealand website using the NZGOAL framework is the Ocean Survey 20/20 Portal for the Bay of Islands Coastal Survey Project – http://www.os2020.org.nz/. Creative Commons licensing is also used for Wellington City Council GIS data at http://www.wellington.govt.nz/maps/gis/gis-terms.html

3 — Making data able to be accessed: characteristics of a provider node

3.3 Characteristics of a provider node *continued*

It should be noted that additional security provisions (such as the need for a government log in) or e-commerce mechanisms should be implemented at the IT level, not the SDI level. Some programmes of work participating in SDI may impose these limitations, not as part of an SDI implementation, rather as part of the IT infrastructure.

Essential

The minimum expectation for the licence agreement is:

- Existing licence agreement in licensing field of service Metadata (minimal attribution requirements based on NZGOAL)

Highly Desirable

- NZGOAL-recommended Creative Commons licence in place

3.3.5 Resilience

Since SDI enables more effective and efficient use of geospatial information, it is important to appreciate that users will come to rely on services of geospatial information. This creates an expectation of a service level for each provider node.

As a minimum it should be expected that a service is available 24x7. This would be the normal service level for any internet-based web capability. As such, users should expect unforeseen interruptions to that service and it is a user responsibility to make provision for continuity of business if the service is not available. This provision might include maintaining a cache of geospatial data in their systems as a fail-over mechanism.

Where users require a higher service level to sustain business- or safety-critical functions, it is the user responsibility to negotiate that service level as a bilateral arrangement between user and provider. This is part of the broader resilience measures that must be part of any business- or safety-critical system. This is not a part of the core definition of SDI.

Therefore the minimum expectation for a provider node is:

Essential

- Usual 24x7 service but with no statement of service level

Highly Desirable

- Usual 24x7 service with statement of service level

Additionally Useful

- Highly resilient node with automatic fail-over with statement of service level.

3 — Making data able to be accessed: characteristics of a provider node

3.3 Characteristics of a provider node *continued*

3.3.6 Updates

As the world changes, geospatial information has to be constantly updated to reflect those changes. This constant maintenance is a challenging aspect of geospatial information.

Within the traditional model of file delivery of geospatial information, the onus for using the most current data fell to the user. With web service delivery, the user has an expectation that the service is being maintained for currency by the provider.

The ideal situation is that there is a direct connection between the provider's production system and the web service. So as an edit is made to a geospatial feature in the production system, that edit is immediately available in the web service. This ensures that business decisions employing the web services are made on the very latest view of the world.

There are very significant risks if a provider node is disconnected from the data maintenance environment. Changes being made in the maintenance environment will not immediately be available in the web service and yet the user of the service will have no way of understanding that the web service has become 'stale'.
As a consequence, there is a real risk of decisions being based on incorrect data.

This direct connection involves establishing a replication mechanism between the production system and the web service. This cannot be made a prerequisite for participation in SDI and so a routine bulk update mechanism with an appropriate update cycle should be regarded as the minimum essential requirement.

Essential

- A routine bulk update mechanism with an appropriate update cycle to connect the relevant organisational production system with geospatial web services.

Highly Desirable

- A replication mechanism that provides a direct, real-time connection between the relevant organisational production system and geospatial web services.

3.3.7 Physical Environment

SDI must be based on a physical ICT infrastructure. Fortunately, representing a geospatial extension to the Internet, the essential physical characteristics of any SDI node are similar to any other web server in terms of hardware, network, bandwidth, firewalls, etc.

However, an important characteristic of geospatial information and technology is that it can stress the underlying ICT infrastructure.

- Geospatial information is typically large and the data structures are complex. This imposes potentially significant burdens on hardware and bandwidth. Moreover, small nuances of tuning can create orders of magnitude changes in loading.

3 — Making data able to be accessed: characteristics of a provider node

3.3 Characteristics of a provider node *continued*

- Geospatial technology has evolved quickly and the pace of change is increasing. This makes the need to evolve inevitable.

- The experience of other jurisdictions shows that once spatial information becomes easy to discover and consume, its use will grow dramatically. This means that the computing infrastructure has to be able to scale linearly with the growth in demand.

This creates a level of risk in programme terms since the metrics are difficult to estimate ahead of implementation. If poorly performing or limited services are made available, the demand will be limited. If high performance, rich services are made available, demand could potentially be significant.

The implementation of outward-facing web services poses significant implementation challenges. It is difficult to anticipate demand during the planning stages. There are many variables that need to be considered during implementation planning which may lead to wide estimates of server scaling.

Some web service types scale in unpredictable ways: this is particularly true of Web Feature Services (WFS). Great care should be taken to set up the service in a way that controls how much data a user can access in a single transaction.

Care also needs to be taken to tune every aspect of a web service eg database, scale dependency, image transfer type, image transfer extent. Very small differences in tuning will have a significant impact on processor loading and bandwidth consumption. This tuning needs to be undertaken by a skilled GIS Server professional.

Given all these factors, it is strongly recommended that any web server implementation be based on a scalable ICT infrastructure that can quickly respond to changing demands. Consideration should be given to establishing the initial iteration of web services from a hosted environment where 'real world' scaling lessons can be learned.

Essential

- At the very least, a provider node needs to be based on a scalable physical ICT infrastructure that is based on realistic assumptions of usage.

Highly Desirable

- Initial iteration of a web service is provided from a hosted environment so that scaling can be understood prior to hardware and bandwidth acquisition.

3 — Making data able to be accessed: characteristics of a provider node

3.3 Characteristics of a provider node *continued*

3.3.8 Ongoing Management

Implementing the technology involved in a provider node is comparatively easy – the standards are already implemented in many proprietary and open source products. Much more challenging are the leadership and management requirements. Decision makers need to understand the connection between current and emerging technical capability and business impact.

When implementing web services, an organisation is making a leap from information as a product to information as a service. This represents a surprisingly profound change that can have a significant impact on organisational strategy. If web service implementation is treated as simply a technology issue, there is a risk of disconnect between organisational leadership and operational reality.

Modern web services (often labelled Web 2.0) offer much more than a new mechanism for delivering information. The ability for users to post information back to the provider offers new ways for stakeholders to interact with providers. If the potential for that interaction is ignored and poor data services go uncorrected, it is possible that the host organisation's credibility will be diminished within the user community.

SDI concepts are not new but implementations in technology platforms are constantly changing. It is important that organisations constantly invest in their information architects and geospatial professionals so that implementations are based on up-to-date technology realities.

SPATIAL DATA INFRASTRUCTURE:

4

Making data able to be found: characteristics of a catalogue node

4 — Making data able to be found: characteristics of a catalogue node

4.1 Participating in Spatial Data Infrastructure

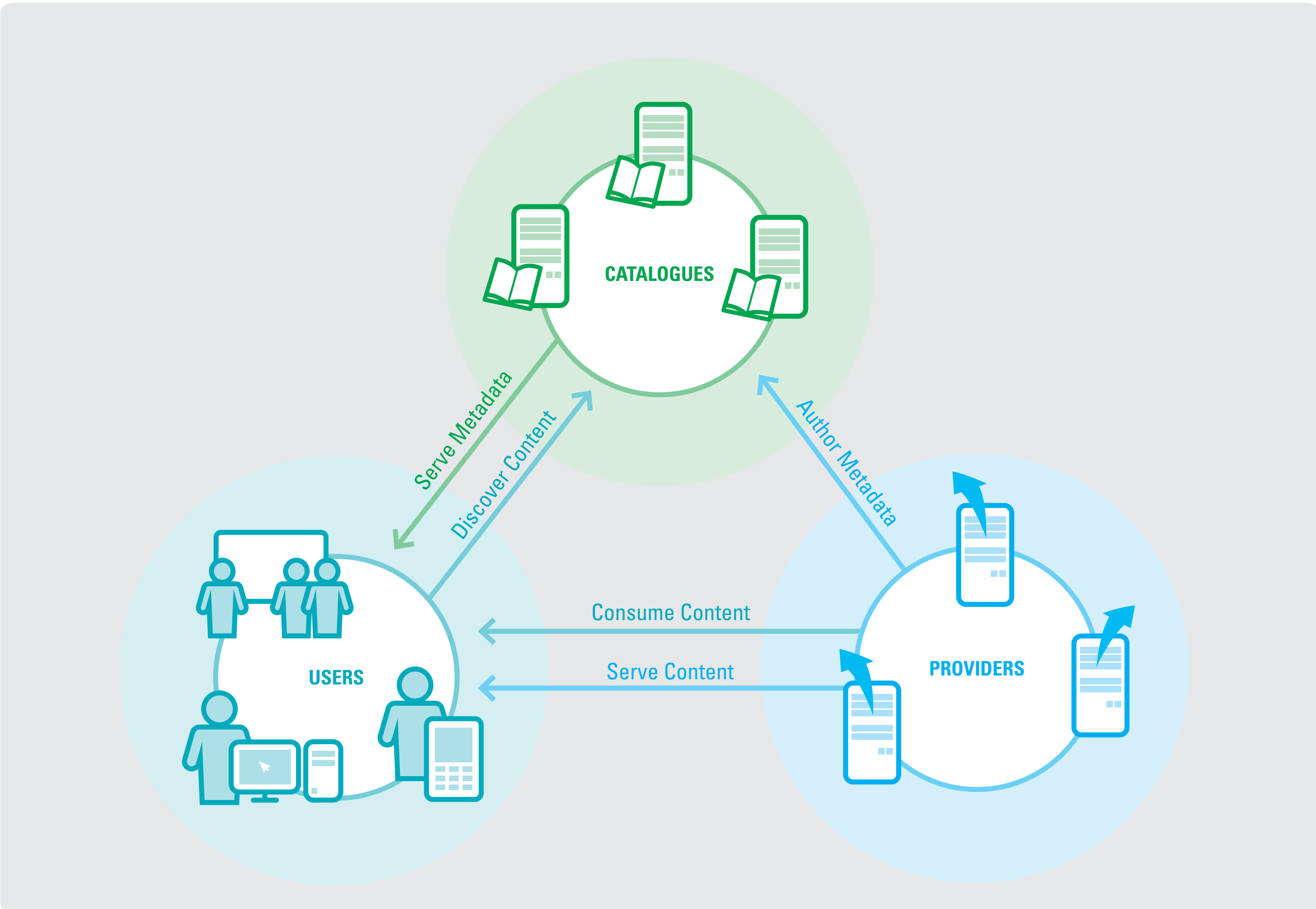
This chapter is one of three (Chapters 3-5) that explain how organisations and their supporting systems can participate in Spatial Data Infrastructure (SDI). Systems participate in one of three ways.

1. They can contribute to SDI by providing services of geospatial information. These systems are then acting as 'Provider Nodes'.

2. They can contribute to SDI by providing catalogues of geospatial information. These systems are then acting as 'Catalogue Nodes'.

3. They can benefit from SDI by utilising the catalogues and services of geospatial information. These systems are 'User Nodes'.

The relationship between different nodes is show in the diagram below:



4 — Making data able to be found: characteristics of a catalogue node

4.1 Participating in Spatial Data Infrastructure *continued*

Systems participate in SDI by having certain characteristics that create interoperability with other nodes. That interoperability is enabled through the use of well-defined and open international standards.

The guidance provided as part of this cookbook is intended to define SDI participation characteristics in vendor-neutral terms. As such, this chapter, the previous chapter and the subsequent chapter could be appended to requests for proposal (RFPs) for any geospatial system which needs to participate in SDI.

This chapter defines the characteristics of an SDI catalogue node. Of the three participation roles, the catalogue node can be the most challenging. SDI provider nodes or user nodes are typically implemented as by-products of systems that serve other business functions (such as providing public-facing web applications or consuming other forms of web services).

Although catalogue nodes are a critical part of an SDI, there is no other primary business rationale for standing up a catalogue node. So a catalogue node is a more specialist undertaking than either a provider or user node.

It is also important to understand the context of catalogue node use. SDI is not intended to be used directly by the consumer – a consumer's need to access geospatial information is typically satisfied using third-party applications and consumer mapping tools such as Google Maps and Bing Maps. SDI can provide a foundation for public-facing government data and subsequent applications leveraging that data, but the foundation itself is entirely invisible to the public.

The public should never have to use an SDI catalogue node – these catalogue nodes are for the use of geospatial professionals to discover and then consume geospatial information in professional applications. Those professional applications might serve an individual's need to understand property information for example, but they are built upon the foundation of SDI and thus isolate the individual from direct interaction with SDI.

A catalogue node is increasingly seen as an integral part of a provider node – in that combined provider/catalogue context it is sometimes termed a 'geoportal'.

4.1.1 Catalogue Nodes

Catalogues are an essential part of SDI. Catalogues allow users to easily discover the geospatial data that they need using metadata (compact data that describes datasets). This document focuses on a geospatial data catalogue as distinct from catalogues that might describe other types of data.

This document further describes the characteristics of a 'well behaved' catalogue node. It is intended to guide any agency in establishing a catalogue node.

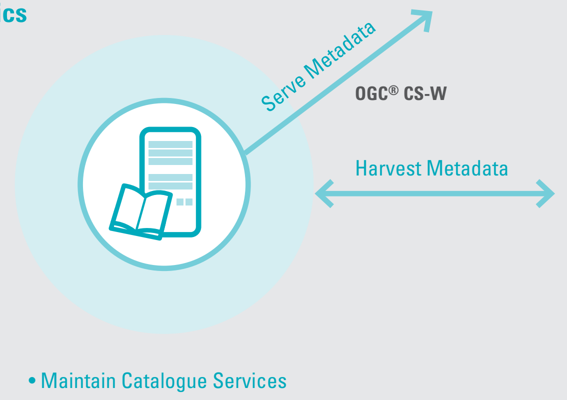
A key purpose of SDI is to make spatial information easily discoverable, accessible and usable. It is important to define 'easily' in terms of the user's expectations rather than the provider's constraints. Those expectations are increasingly based on a user's experience with consumer mapping applications such as Google Maps. That poses some challenges for catalogue nodes since professional/business users need to be able to discover a much richer resource of spatial information than the consumer user.

4 — Making data able to be found: characteristics of a catalogue node

4.1 Participating in Spatial Data Infrastructure *continued*

Catalogue nodes rely absolutely on open standards to connect users, providers and other catalogue nodes. Each catalogue node can be built from either open source or proprietary technologies based on the specific needs of the user organisation. It is important to appreciate this separation between open standards used to connect nodes and underpin interoperability within an SDI, and the choice of technology to implement the nodes themselves.

Minimum characteristics of a catalogue node



4.1.2. Many Catalogue Nodes

It may be tempting to envision only one all-of-government geospatial catalogue node – a single place where all geospatial data can be discovered and thus reject the need for any other catalogue nodes. There are a number of problems with the notion of one catalogue which contains everything.

Some provider nodes have literally millions of datasets. The NASA Landsat archive is an example; Digital Globe's archive is another. Having all metadata from all nodes placed into one master node overwhelms the central node.

Metadata is not a static resource. It has to be constantly updated as data is updated. Some provider nodes are updated daily (and automatically) necessitating a daily (and automatic) updating of metadata. This is difficult to achieve with one central catalogue node.

There are additional reasons why most organisations might choose to implement their own catalogues:

- discovery of data can be focused on that which is of relevance to the organisation's business needs
- the node can be used to control access to datasets: the agency node might be the only mechanism to connect agency systems to external services of data
- a direct connection to production systems can be established, facilitating the posting of updated metadata

4 — Making data able to be found: characteristics of a catalogue node

4.1 Participating in Spatial Data Infrastructure *continued*

- the node can be built according to internal architectural and platform needs
- an agency catalogue node can conform to internal user interface guidelines and agency branding.

Moreover, the relationship between geospatial catalogues and other sector catalogues needs to be considered. Geospatial catalogues are based on a metadata standard (ISO 19139) that has 42 fields of information. The Darwin Core metadata standard that is used by biologists, in contrast, has 172 fields of information.[46] Many other domain metadata standards exist. It would be unreasonable to expect a geospatial catalogue client to be able to query these other types of metadata records.

Increasingly, there is interest in creating generic data catalogues. The data.gov site in the US is one example. In New Zealand, the Department of Internal Affairs has established an all of government site, http://www.data.govt.nz. These generic catalogues need to interoperate with other types of catalogue. It is unrealistic to expect that every catalogue will understand the specific catalogue flavour of every other catalogue. A reasonable approach is to use domain catalogues as gateways to interoperate with generic data catalogues.

If an agency is setting out to construct a catalogue node, there must be a clear understanding of how this node will relate to other nodes.

4.2 Characteristics of a catalogue node

The following sections describe the characteristics that define a catalogue node that will operate well in the New Zealand SDI. In each case the characteristic is described in terms of a priority of requirements:

Essential – the absolute minimum requirement in order for the system to participate in SDI. An implementation that does not satisfy these requirements is not acceptable.

Highly Desirable – these requirements are those which are very important to include in the implementation of this feature, but are not so important as to be considered absolute requirements.

Additionally Useful - these requirements are those which may provide some level of benefit but are not critical to the implementation of this feature. They may be included in the implementation of this feature if time permits, but priority must be given to the 'essential' and 'highly desirable' requirements.

4.2.1 Metadata

Metadata are structured facts that describe information, or information services. Metadata facilitates many things beyond enabling searching and cataloguing; it also informs appropriate use of products and services. It is very important to understand this broader use of metadata since it avoids the misconception that an organisation doesn't need metadata if it doesn't have a catalogue.

46 http://www.tdwg.org/activities/darwincore/

4 — Making data able to be found: characteristics of a catalogue node

4.2 Characteristics of a catalogue node *continued*

Every organisation requires up to date metadata to provide a full understanding of the complicated and nuanced nature of geospatial data. Metadata for example describes the age of data and its accuracy – without the understanding communicated by metadata, decisions might be founded on old, inaccurate data.

Few organisations in New Zealand currently maintain geospatial metadata. It is seen as an expensive and unnecessary activity, especially since many people only see metadata as valuable in the context of catalogues, and few organisations have catalogues.

There is another growing excuse for not creating and maintaining metadata: 'Google search doesn't use metadata, it uses web crawlers'. That is a correct statement but it once again associates the value of metadata only in terms of search. It is indeed possible to find geospatial data using Google search – but without metadata it is extremely challenging to make safe use of that data.

4.2.2 Standards

The International Organisation for Standardisation (ISO) includes ISO/TC211[47], which is an international, technical committee for geographic information. TC211 has created a strong, globally implemented set of standards for geospatial metadata: the baseline ISO 19115; and ISO 19139 for implementation. These open standards define the structure and content of metadata records and are essential for any catalogue implementation.

ANZLIC – the Spatial Information Council connects Australian Federal and State government and the New Zealand government. A key goal of ANZLIC is to develop consistent policies and guidelines (using technical working groups and committees) to minimise barriers to spatial data and services wherever possible.

ANZLIC has developed a specific profile of ISO 19139[48] that is now a New Zealand government e-GIF recommended standard for geospatial information. This specific profile only differs very slightly from the core ISO 19139 standard; primarily by making one element mandatory for Australasia, rather than optional.

The ANZLIC profile's minimum elements for consideration total 46 fields[49]. Of these, 12 are regarded as mandatory. For implementation of a Catalogue node, it is essential to have the 12 mandatory fields populated. It is highly desirable to have as many relevant fields as possible populated and additionally useful to have all 46 fields populated. This is very soft guidance: experience clearly demonstrates that something is better than nothing and striving for completeness often results in nothing.

Essential

- ANZLIC Metadata Profile – 12 mandatory fields populated

47 http://www.isotc211.org/
48 http://www.linz.govt.nz/geospatial-office/about/projects-and-news/anzlic-metadata-profile/index.aspx
49 Custodians can choose from other optional metadata elements described in AS/NZS ISO 19115:2005 Geographic information – Metadata (these are all encompassed by the ANZLIC profile)

4 — Making data able to be found: characteristics of a catalogue node

4.2 Characteristics of a catalogue node *continued*

Highly Desirable

- ANZLIC Metadata Profile – 12 mandatory, and as many additional fields as possible populated

Additionally Useful

- ANZLIC Metadata Profile – all appropriate metadata fields populated

4.2.3 Metadata Hierarchy

Metadata can be applied to hierarchies of information service and source data. It is important to create metadata records at the appropriate level to facilitate meaningful discovery.

This can be illustrated using a topographic map service as an example. The topographic map comprises many layers of information (such as contours, roads, lakes and trig points). In web service terms, this can be provided as many different types of composite and separated data service. Each case needs a different metadata descriptor.

Service type	Metadata	Example abstract
Scanned topographic map contained as a single web map service (WMS). All layers are displayed together and can't be separately controlled.	Describes the complete topographic map.	This service is derived from the national topographic map series at 1:50k and contains all map face content. It is intended to support contextual backdrop use only and cannot be queried.
Topographic map served as a single rendered web map service. Individual layers can be switched on and off.	Describes the complete topographic map.	This service is derived from the national topographic map series at 1:50k and contains all map face content. The service can be queried to obtain feature attributes according to the data dictionary at http//:URL
Groups of like layers are served out as aggregated services (for example hydrology, containing rivers, lakes, springs etc). Individual layers can be switched on and off.	Describes just the hydrology content of the topographic map.	The hydrologic features in this service were captured to support the national topographic map series at 1:50k. As such, some cartographic generalisation can be expected. The service can be queried.
Individual layers are served out as individual services (for example, power lines) in a web map service. Individual layers can be switched on and off but cannot be re-symbolised.	Describes just the content of the layer of the topographic map.	The power line features in this service were captured to support the national topographic map series at 1:50k. Some cartographic generalisation will have occurred that might eliminate some power lines and move others. The authoritative source of power line features is the relevant utility company.

4 — Making data able to be found: characteristics of a catalogue node

4.2 Characteristics of a catalogue node *continued*

Each layer of data in the provider's organisation needs to have metadata to facilitate internal discovery. Where these layers of data are aggregated into a web map service, a new metadata record must be created that describes this aggregated service. Since the layers are not individually accessible to an external service user, their metadata records are not directly relevant in discovering the service.

Essential

- Metadata associated with each service available for external consumption.

Highly Desirable

- Metadata associated with each layer of data that contributes to the service, permitting internal discovery.

Additionally Useful

- Separate services for each layer of information, each with its own metadata record. This can create some challenges in terms of scaling and performance of the server.

4.2.4 Catalogue Services

A catalogue node needs to be able to provide metadata records that a user can query using open standards. The Open Geospatial Consortium[50] (OGC) has created the Catalogue Service for Web (CSW) standard to enable discovery from a catalogue node.[51]

Catalogue services support the ability to publish and search metadata for data, services, and related information. Metadata in catalogues can be queried and presented for evaluation and further processing by both humans and software. Catalogue services are required to support the discovery and binding to published web map services.

The CSW standard is extremely rich. In addition to supporting a query from a user, it can support distributed queries (one query that searches many catalogues) and the harvesting of metadata from node to node.

This richness creates flexibility in implementation that is both a strength and a weakness of the standard. A software client that is implemented to rigidly handle a subset of CSW capabilities might fail when connecting to a catalogue server that has implemented a richer set of capabilities, and vice-versa.

For that reason, in the context of SDI, it is recommended that in the first instance catalogue services conform to the core profile (OGCCORE). Where specific organisational needs are identified that go beyond the core, those should be implemented carefully and tested to ensure that all connected clients can cope with the additional richness.

The outcome of this is that catalogue clients should be implemented to be as flexible as possible in terms of the profiles of CSW that can be consumed.

An example of a significant CSW implementation is the Geospatial One Stop site, based in the United States.[52]

50 http://www.opengeospatial.org
51 http://www.opengeospatial.org/standards/cat
52 http://www.geodata.gov

4 — Making data able to be found: characteristics of a catalogue node

4.2 Characteristics of a catalogue node *continued*

Essential

- OGC's CSW standard as the basis for catalogue services. At the least, a catalogue server needs to implement the OGCCORE profile in order to guarantee a minimum service level for client queries.

Highly Desirable

- Extend the use of the CSW standard to support metadata harvesting.

4.2.5 Harvesting

A mature or formal national SDI might involve hundreds of geospatial catalogue nodes, including:

- national, regional and local government catalogues
- individual agency catalogues
- sector-wide catalogues
- university-based catalogues
- private-sector catalogues.

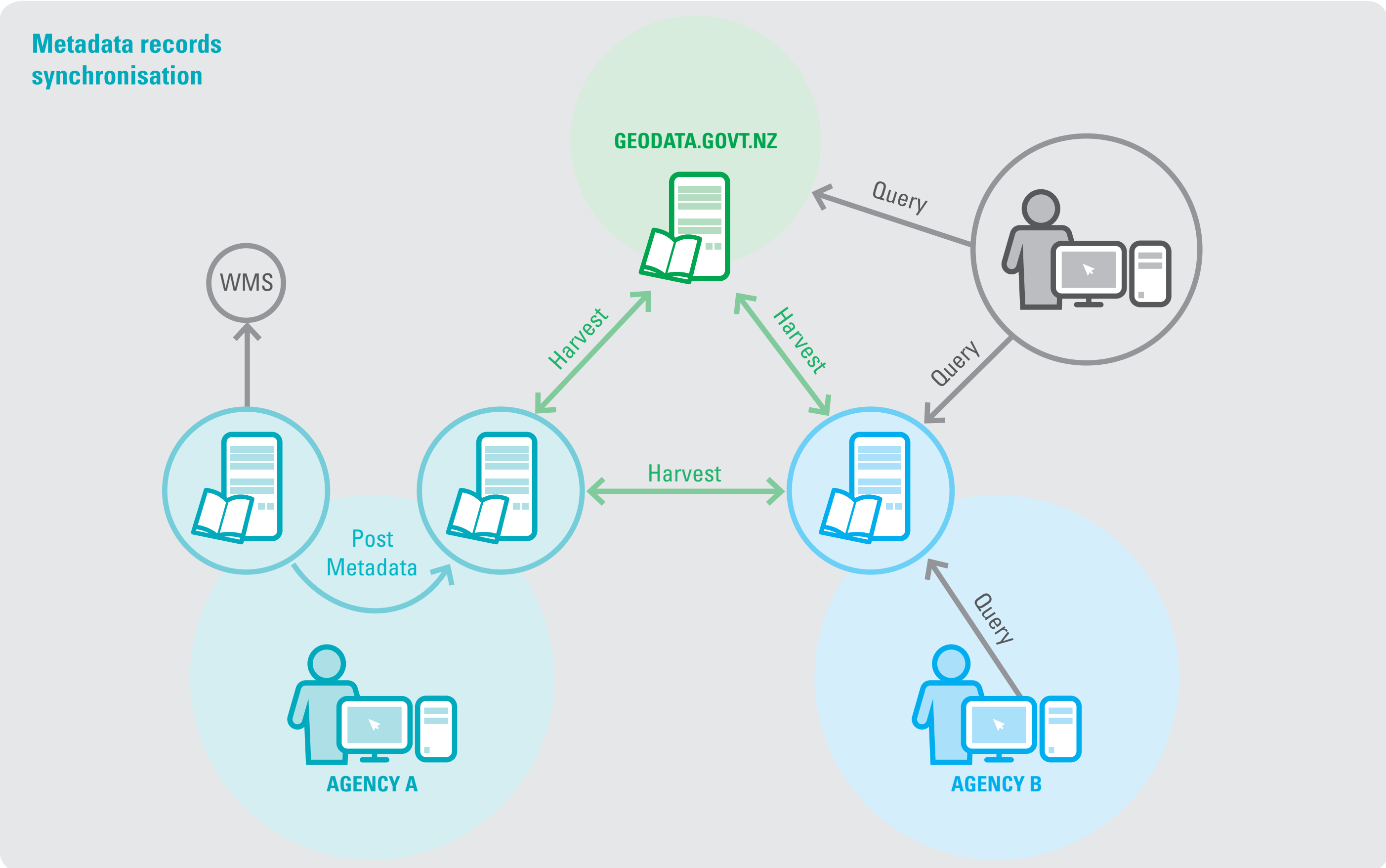
It would be extremely inefficient if metadata had to be manually posted onto each node separately. Harvesting synchronises metadata records across many nodes and allows metadata to be created once, posted once and then reused many times.

Harvesting is enabled by the OGC CS-W standard outlined above. Most catalogue node solutions will include harvesting as a latent capability.

Implementing a harvesting solution remains challenging. It involves complex system implementation issues including scalability and performance. Ongoing operations and maintenance also involve complex systems administration.

Harvesting is the mechanism for connecting various catalogue nodes to ensure that metadata records can be synchronised. This is illustrated in the diagram overleaf.

4 — Making data able to be found: characteristics of a catalogue node



During the introductory stage in the development of the New Zealand SDI, it is recommended that harvesting is regarded as useful but not essential. There is a risk that by insisting on harvesting, this might act as a deterrent to standing up catalogue nodes.

Since it is anticipated that harvesting would become a requirement within the typical five year life span of a system, it is recommended that catalogue nodes have a latent capability to implement harvesting.

Highly Desirable

- Catalogue nodes are capable of implementing a harvesting mechanism.

4.2.6 Feature Level Metadata

Some mission-critical environments have a requirement to record metadata at the individual feature level: every feature has metadata describing the feature. This might be used to record when a feature was edited and by whom.

Feature level metadata is challenging to implement and imposes very real limitations in terms of database size, system scalability and display and query performance. There are very few implementations currently in New Zealand and so feature level metadata is currently outside the scope of the New Zealand SDI.

SPATIAL DATA INFRASTRUCTURE:

5

Efficiently using data: how can my organisation use SDI?

5 — Efficiently using data: how can my organisation use SDI?

5.1 Participating in Spatial Data Infrastructure

This chapter is one of three (Chapters 3-5) that explain how organisations and their supporting systems can participate in Spatial Data Infrastructure (SDI).

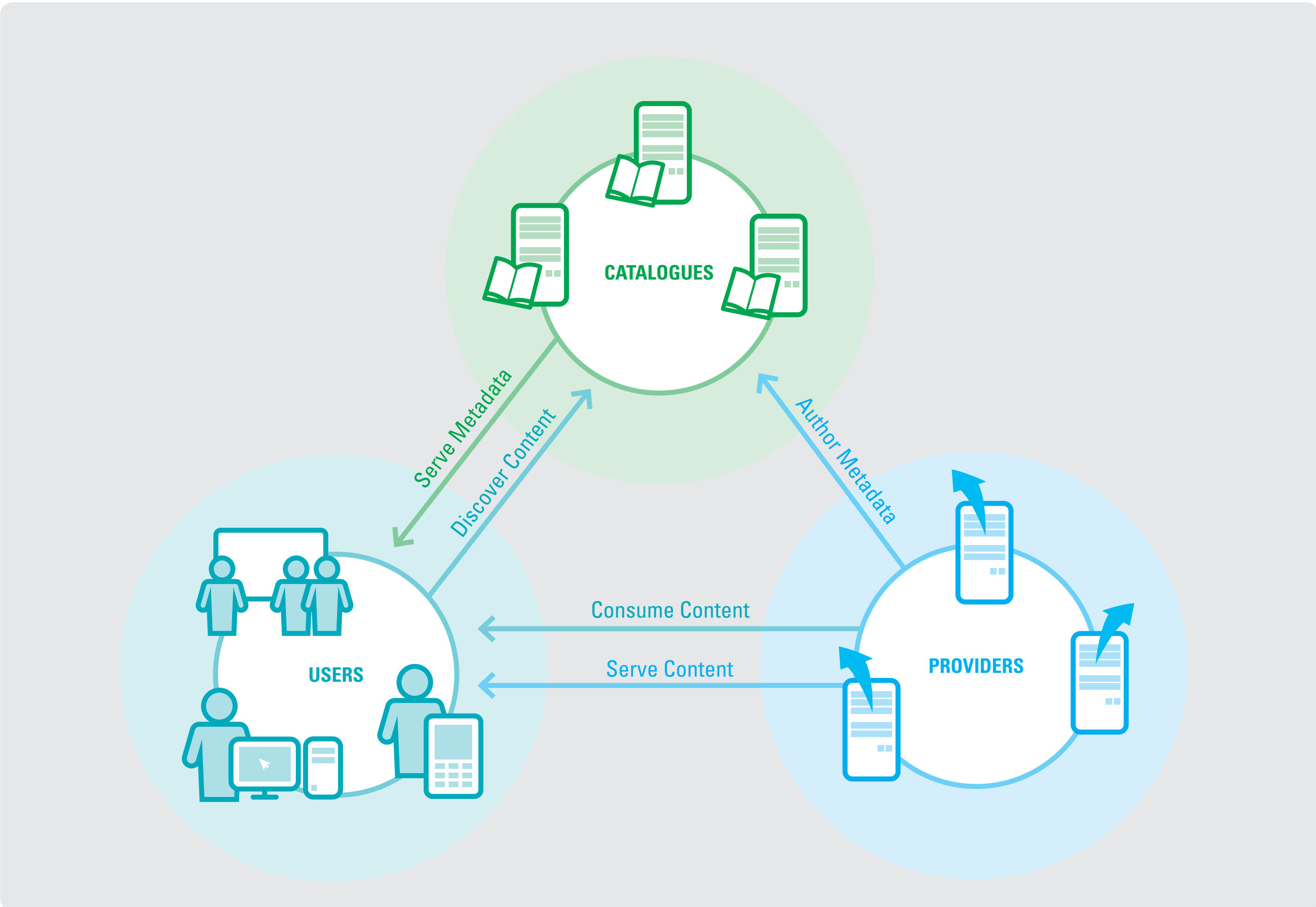
Systems participate in one of three ways:

1. They can contribute to SDI by providing services of geospatial information. These systems are then acting as 'Provider Nodes'.

2. They can contribute to SDI by providing catalogues of geospatial information. These systems are then acting as 'Catalogue Nodes'.

3. They can benefit from SDI by utilising the catalogues and services of geospatial information. These systems are 'User Nodes'.

The relationship between different nodes is show in the diagram below:



5 — Efficiently using data: how can my organisation use SDI?

5.1 Participating in Spatial Data Infrastructure *continued*

This chapter outlines how an organisation can benefit from participating in the New Zealand SDI. It is not intended for those who are already experts in the use of geospatial data, although they may find certain parts useful in the preparation of RFPs. It is intended more for those who understand the concepts and might benefit from the first stage of participation – a raised level of awareness of the benefits and first level practical advice for developing user systems that can consume the evolving infrastructure being provided.

The success of a spatial data infrastructure flows from it being widely used. If it is not used, it is literally useless and there is a risk that provider and catalogue nodes become irrelevant. This chapter therefore outlines the benefits that flow from an organisation's use of SDI in business terms. It then provides a structure for the types of applications and users of the SDI, and then finally outlines the functionality required and sources of further advice.

The guidance provided as part of this cookbook is intended to define SDI participation characteristics in technology-agnostic terms.

5.1.1 The Benefits of participation in SDI

In generic terms the benefits of an SDI can be identified as:

- awareness: enabling users to discover spatial data that they were not previously aware existed. Indeed, it will not only help them to discover data but also assess whether it is fit for purpose (thus enabling decisions to be based on better understanding)
- data sharing: lessen the amount of work behind the scenes to make that geospatial information available (thus saving resource for the organisation)
- currency: provide the user with more up-to-date data (thus improving the decision support the system provides)
- consistency: ensure that everyone is using the same geospatial information (thus avoiding confusion)
- integration: often it is difficult to combine datasets together because there are no common attributes. As most events occur in a given location or area, geospatial information has a unique ability to act as a 'foreign key' to link enterprise systems such as asset management systems, enterprise resource planning systems and financial systems.

In today's economic climate, few organisations will be prepared to significantly increase capital spending for the 'greater good' of SDI, especially if business outcomes are compromised. For this reason, it is advocated that organisations adopt an evolutionary approach to participation.

5 — Efficiently using data: how can my organisation use SDI?

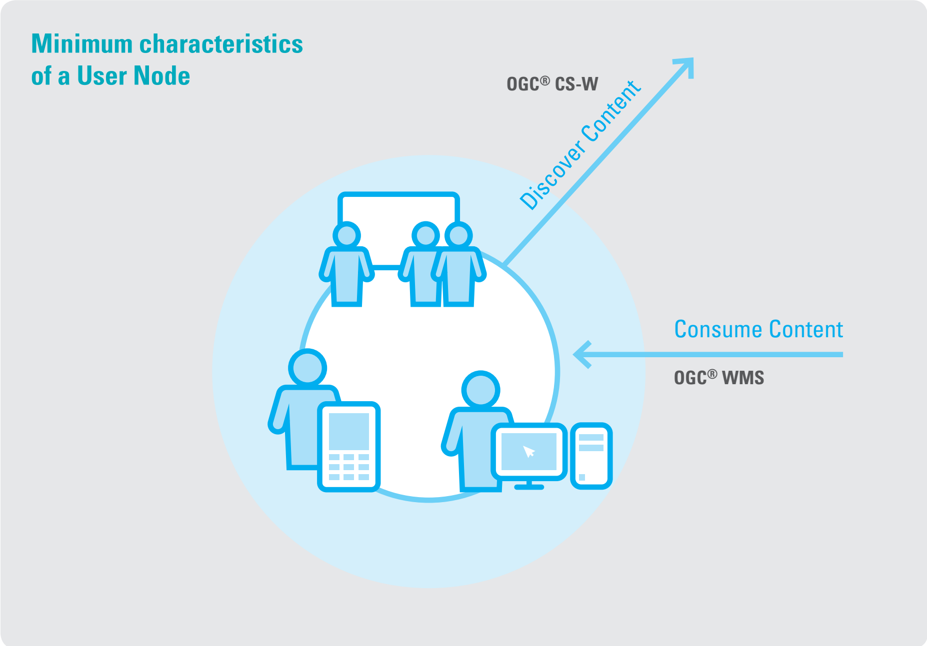
5.1 Participating in Spatial Data Infrastructure *continued*

SDI is not intended to be directly used by the consumer – the consumer's needs to access geospatial information are typically satisfied using third-party applications and consumer mapping tools such as Google Maps and Bing Maps. SDI can provide a foundation for public-facing government data and subsequent applications leveraging that data, but the foundation itself is entirely transparent to the public. The public will typically not query catalogue nodes or directly connect to provider nodes – they are not barred from doing so but their needs are not a focus for establishing an SDI.

It is important however to appreciate that Business to Consumer (B2C) providers will benefit from SDI in being able to innovate on top of SDI services. The consumer indirectly benefits through the increased availability of map-based consumer applications. So B2C systems do need to be considered in the use cases for an SDI user node.

5.1.2 User Nodes

User nodes need to provide for a very wide range of organisational, business, individual and technology needs. The standards that underpin SDI allow for flexible implementations to support this wide range of use cases.



5 — Efficiently using data: how can my organisation use SDI?

5.1 Participating in Spatial Data Infrastructure *continued*

A 'user node' wrongly implies a dedicated SDI system. In reality, SDI usage will occur in existing business systems throughout the organisation. In a successful SDI, most systems that currently display or use geospatial information will transition from existing external sources of geospatial information to consume the new geospatial web services from SDI provider nodes. Typically, the SDI geospatial web services will be mashed up with local information to meet internal user needs.

This cookbook is about the **implementation** of SDI, and this chapter is about the **implementation** of SDI user nodes. Although the user of SDI may require no knowledge of geospatial information, the designer and implementer of the user's system may need a deep understanding of spatial data. This is to ensure that the system is scalable and performs well but most importantly that the user obtains the right outcome from the use of their system.

5.2 Using SDI

5.2.1 Geospatial Information

Many organisations make use of geospatial information to support a wide range of business outcomes and almost all organisations make some use of geospatial information. A national SDI can offer real efficiencies for organisations' use of that geospatial information.

In order to understand the benefits of SDI, it is first necessary to understand the nature of geospatial information. Geospatial information has a number of characteristics which make it a unique and challenging type of information.

First and most important, geospatial information is an abstraction of our world. Our world is constantly changing and so geospatial information must be constantly maintained to reflect that change. Since the business outcomes being supported often include policy making and operational decision support, it is very important that those decisions are based on up-to-date information.

Geospatial information is also expensive to produce. That's particularly the case for the geometry that might represent the shape of a road or building. This has to be laboriously produced by digitising in a specialist workstation. To maximise value and the return on what can be a significant investment therefore, it is very important that as far as possible, information is created once and used many times.

Geospatial information is often large in occupying significant storage space. This makes it challenging to transfer as a file since that file is often too large to transfer across typical internet connections. It is also important to avoid duplicating data holdings within an organisation, unless required to preserve business continuity in the event of disastrous loss of internet connectivity.

5 — Efficiently using data: how can my organisation use SDI?

5.2 Using SDI *continued*

5.2.2 Applications

An SDI is merely a foundation. The data infrastructure must be augmented by a wider range of tools and applications that are needed to deliver business capabilities. These applications range from simple business applications through to sophisticated enterprise Geographic Information Systems (GIS).

Many organisations will implement a wide array of business applications and sophisticated GIS applications. Other organisations just need business applications and have no requirement for GIS.

This has a relevance to the use of SDI since those organisations already making extensive use of enterprise GIS will typically find it much easier to participate in SDI: they already have the expertise in-house and may already have a spatial web service infrastructure that only needs minor adjustment to participate in SDI.

The maturity level of GIS within an organisation is often a reflection of how important geography is as a part of policy making, decision support and operations. Much of local government's work is based on geography (i.e. they are 'geo-centric') and so many local government organisations exhibit a high GIS maturity level. However, other agencies are very vertically oriented, have little interest in 'place' and so might be 'geo-enabled' but will not be dominated by spatial systems. In these cases the SDI will need to fit into existing IT infrastructures rather than defining it.

In New Zealand, most regional and local government organisations are very mature users of geospatial information. They have specialist teams, a number of existing generic applications that consume geospatial web services and will find the transition to participation in SDI fairly straightforward.

The situation in central government is a lot more patchy. Many central government organisations have no geospatial capabilities in house and will need considerable assistance to participate in SDI. Other organisations, particularly in the natural resources sector have a higher level of development.

5.2.2.1 Generic Applications

Geospatial data contributes to a wide range of applications and it is important to understand this range, as it will define how best to harness the power of SDI.
At the top level of any organisation it can be applied to:

- Policy and strategy formulation
- Customer engagement
- Planning and design
- Operations
- Support
- Information management.

5 — Efficiently using data: how can my organisation use SDI?

5.2 Using SDI *continued*

The table below provides a greater level of detail of each of these primary activities and shows one example of each. In each case, the application and the example are not geospatial applications – they are business applications that benefit from the use of geospatial data.

Primary activity	Generic application type	Examples
Policy and Strategy	Business intelligence	Directors dashboard
	Decision support	Common operational picture
	Performance management	Measuring crime prevention outcomes
	Scenario modelling	Presenting options for a new road
	Social engineering	Analysis supporting social inclusion
	Pattern analysis	Understanding disease spread
	Risk reduction	Insurance premium assessment
Customer Engagement	Marketing	Planning where to market a product
	Sales	Analysing an optimum store location
	Advertising	Determining billboard placement
	Consumer interaction	Getting feedback on District Plans
	Local democracy	Revision of electoral boundaries
	Fault reporting	Electrical outage reporting and analysis
	Public record access	Land registration, local land charges
	Customer profiling	Identifying target markets
	Public consultations	Planning enquiries
	Customer relationship management	Call Centre customer mapping
Planning and Design	Viewshed analysis	Siting mobile phone masts
	Site assessment and selection	Planning a new school location
	Urban and rural design	3D aerial imagery
	Environmental impact assessment	Evaluate a Resource Consent
	Contingency planning	Emergency planning exercises
	Enterprise Resource Planning (ERP)	Planning utility asset maintenance

5 — Efficiently using data: how can my organisation use SDI?

5.2 Using SDI *continued*

Primary activity	Generic application type	Examples
Operations	Asset management	Utilities pipe network locations
	Facilities management	Airport information systems
	Works order management	Optimising work allocation
	Construction	Road opening notices
	Monitoring	Structural movements
	Emergency response	Ambulance despatch
	Command and control	Maintaining situational awareness
	Asset tracking	Electricity network monitoring
	Field Force management	Plant pest reporting from the field
	Route optimisation	Field worker despatch
	Navigation	Marine vessels
	Regulatory compliance	Electricity outage statistics
	Supply-chain management	Delivery tracking of parcels
Support	Map production	Producing a District Plan map
	Revenue collection	Identifying potential tax fraud
	Financial assessment	Asset valuations/insurance
	Procurement	Identify location of potential suppliers
Information management	Data acquisition and maintenance	Collecting road surface information
	Data modelling	Defining a data model for spatial relationships
	Database building	From geospatial concepts
	Quality assessment	Based on location intelligence
	Integration	Unions between disparate data

To determine where SDI can add the most value it is recommended you map these generic applications against your organisation's core objectives. This provides a framework for conducting a cost benefit analysis.

5 — Efficiently using data: how can my organisation use SDI?

5.2 Using SDI *continued*

5.2.3 User Tasks

It is useful to think about different kinds of SDI use – reflecting the different things that users want to do. Each of these user tasks plays a role in the generic applications identified above. There are four broad categories of use that are amplified below:

- View
- Analyse
- Maintain
- Manage/Administer

Some users will obviously fit into more than one of these types over time or even simultaneously. However, they are useful to define the functions necessary to participate in the SDI.

In the amplification below, each of these tasks are examined in the context of use. More importantly in light of the focus of this cookbook, the implementation inferences are discussed.

5.2.3.1 View

This is the simplest category of use involving a user 'mashing up' (or layering) different sources of spatial information into a map. It is often useful to think about the resulting map being built from two broad categories of information:

- The background or base mapping that provides a contextual backdrop. This will typically come from an external source and may be a good candidate for replacement by SDI web services.
- Operational or overlay data that displays information needed to support policy making, support operational decisions or gain an understanding of a situation. This may involve a mix of external data and internal data. Not all of this overlay data will be spatial in nature – much of it might be tabular information from other enterprise systems.

The user conducting a view task will not require any geospatial expertise provided their application successfully shields them from the complexities of spatial information. They should not be able to access inappropriate spatial data: data that is in the wrong projection system, of an unsuitable scale, or that is out of date.

In order to shield the user from this complexity, the designer of their viewer does need to understand the characteristics of geospatial information. As the viewer is implemented in the organisation, a geospatial specialist will need to ensure that the non-spatial information can be presented correctly into the viewer; is the spatial reference using the correct map projection for example? A failure to address these issues may result in a poorly performing, inaccurate view experience.

5 — Efficiently using data: how can my organisation use SDI?

5.2 Using SDI *continued*

5.2.3.2 Analyse

Analysis involves a user obtaining additional information from the spatial information. The task might be a simple query (clicking on the map to obtain information about an asset for example) or a richer geoprocessing activity (select all schools within 200m of an urban motorway for example).

Some analysis tasks will be undertaken as part of business applications and require no knowledge of geospatial information. Most people are now aware of route-finding capabilities through their experience of Google Maps and in-vehicle navigation systems. These are good illustrations of how rich geospatial functionality can be delivered as easy-to-use capabilities to generalist users.

Other analysis tasks can only be done by geospatial specialists. These are the 'ad hoc' analysis tasks that might be required to support a specific decision support requirement. In 'geocentric' organisations which encounter these requirements frequently, these might be addressed by a dedicated geospatial team. In organisations where these tasks occur less often, these specialist tasks might be conducted by external consultants.

Analysis can be performed in two places in an SDI: in a client-side application or on the server. Each has different implications in terms of the standards that need to be considered and the flexibility of the resulting analysis.

Client-side Analysis

Client-side analysis requires a sophisticated client application that can support the required geoprocessing. In most cases this will involve a desktop geospatial application in a 'thick' client as opposed to a browser. Browser technologies are becoming more sophisticated and can increasingly support very large plug-ins to enable client-side analysis, but the resulting applications remain a very niche capability.

The implementer of a client-side analysis application needs to have a full understanding of the participating data and the geoprocessing task. This will typically require an implementer well-grounded in geospatial principles and applications. The implementer also needs to take considerable care to ensure that the analysis task is implemented to suit the capacity of the client-side computer; some geoprocessing tasks require access to large volumes of data and are very compute-intensive.

Client-side analysis requires access to the geospatial information in a form that enables geoprocessing; typically access to the feature geometries. WMS is not sufficient for this purpose; typically this requires a feature service. WFS is an example of a suitable open standards feature service.

One of the characteristics of client-side processing is the flexibility that is afforded. This may be seen as an advantage when a system is being implemented by someone who has a good understanding of geospatial information and geoprocessing. This might be a disadvantage if a more generalist implementer just wants access to a well-defined function such as a routing algorithm or a geocoding operator.

5 — Efficiently using data: how can my organisation use SDI?

5.2 Using SDI *continued*

Server-side Analysis

Server-side analysis is a more common implementation because of the many advantages involved. Server-side analysis allows for much simpler clients (such as a straightforward browser) to be used and makes the implementation of the client much easier.

The implementer of the client application only needs to know about the inputs required to start the analysis (for example a click of the mouse on the map to indicate where the analysis is to be conducted and a numeric input to specify a radius of search) and the characteristics of the returned information (the locations of all required features within the specified radius).

The client application will typically not be very compute-intensive. The server is storing the potentially large data sets and is doing the 'heavy-lifting' processing.

If the processing is performed on the server, the client does not need sophisticated access to the raw data. In many cases, a simple WMS service can provide the contextual map information and the geoprocessing request can be made using either an Application Programming Interface (API) call or a Web Processing Service (WPS) request.

The implementer of the server-side functionality will need a deep understanding of the nature of the geospatial data inputs and the geoprocessing tasks needed to perform the analysis. They will need to appreciate the scaling issues involved in allowing large numbers of users to access the analysis functionality.

The disadvantage of server-side analysis is the potential loss of flexibility. The analysis is created to satisfy the needs of multiple users and only analysis parameters can be changed using the client application.

5.2.3.3 Maintain

Data needs to be maintained in order to reflect continuous change in the real world. Data maintenance can be thought of in several categories:

- simple location notification – eg a citizen reporting graffiti at a particular location
- simple change in a location – eg a water meter being moved from one side of a building to another
- change in an attribute – eg a transformer on a utility pole has been replaced, necessitating a change in the equipment ID number
- complex spatial editing – eg a council adding information about a new park and all its facilities.

5 — Efficiently using data: how can my organisation use SDI?

5.2 Using SDI *continued*

Many of the simple maintenance tasks will be conducted within business applications and require no knowledge of geospatial information on the part of the user. The system designer and implementer will need to have some understanding of how the spatial reference is stored to ensure that the system performs and scales well and correctly displays the location of the information.

Attribute changes can be done without needing any map context at all – and in a mobile context, a field worker might be able to update an attribute using a simple text message. Attribute changes need no spatial understanding on the part of either the system designer or the implementer although care is needed when attributes might be performing a subtle role in a behind-the-scenes geospatial function.

More complex spatial editing such as digitising from imagery or integrating survey measurements should be approached with more care. Although these tasks can be performed by users with limited understanding of geospatial information, they should be guided by processes and procedures that ensure adequate quality of collection.

The designers and implementers of the richer systems needed to capture and edit complicated spatial information must either be geospatial professionals or be guided by external consultants that understand the nuances of geospatial information.

5.2.3.4 Manage (Administer)

The business applications need to be managed. This introduces the final user task – that of administering the business applications.

This might involve connecting a user's application to a new source of information. An example of this might be adding a service of the newly agreed District Plan map into a customer service application. The customer service representative can't be expected to know that a new District Plan has been agreed and won't know how to add that new information to their application.

The administrative function will typically be performed by a system administrator within the organisation. It will either involve the use of a dedicated administration application or administrative access to the business application. Either way, that administrator must understand the business context of the change and know that the source of information is appropriate to that business function.

In some cases, the administrative function may require knowledge of geospatial information. That's particularly the case when connecting to new sources of geospatial information; such as updating a service of cadastral or utility information. In this case, a full evaluation of the quality, reliability and resilience of the new service may be required to be performed by a geospatial professional.

Just knowing that geospatial information comes from an 'authoritative source' is not sufficient – it is important to understand the original purpose of that authoritative information and confirm that it matches the purpose that it is planned to be applied to.

5 — Efficiently using data: how can my organisation use SDI?

5.2 Using SDI *continued*

The administrative function might make extensive use of catalogue services to see whether better services of geospatial information might have become available. They will be able to query metadata to understand whether the new service of information might satisfy business requirements.

5.2.4 Enterprise GIS

Increasingly, GIS is deployed as an enterprise asset in organisations that make extensive use of geospatial information. This means that the geospatial information is stored in enterprise DBMS, served throughout the organisation as geospatial web services and integrated with other enterprise systems such as asset management, customer relationship management and financial systems.

Many organisations are also providing external-facing GIS applications to engage with the public and customers. There are many examples in government where sophisticated web GIS applications can be used for querying property information such as rateable values and rubbish collection days. Examples of these external-facing GIS applications include:

- Auckland Council's viewer
  http://maps.auckland.govt.nz/AucklandRegionViewer/
- Carterton District Council's viewer
  http://xplorer.xgl.co.nz/CartDC_xplview/Default.aspx/
- Kapiti Coast District Council's viewer
  http://www.mapimage.net/KCDC/
- Ministry of Fisheries viewer
  http://www.nabis.govt.nz/Pages/default.aspx

In a mature enterprise GIS implementation, geospatial capabilities permeate the organisation. Almost all business systems will be powered by geospatial information – in simple terms, there will be a map in almost every application.

These extensive geospatial capabilities will be enabled by a team of geospatial professionals who will centrally administer the enterprise GIS system, ensure the provision of externally sourced geospatial information, maintain the organisation's operational geospatial information and design the integration with the business systems.

A mature enterprise GIS will be built as a Service Oriented Architecture. It will have servers that will provide geospatial information, it will have some form of geospatial catalogue and it will have widespread users of geospatial information. Indeed, a mature enterprise GIS can be thought of as an organisation's SDI. Therefore connecting to a national SDI will likely be technically straightforward.

5 — Efficiently using data: how can my organisation use SDI?

5.3 User system characteristics

From the preceding discussions, it can be seen that user systems can be characterised in four ways:

- View Systems
- Analysis Systems
- Maintenance Systems
- Administration Systems

In each case it will be recognised that all systems consuming SDI services need to satisfy the business needs of the organisation. Their acquisition will be justified based on the internal business case of the organisation. It will not be likely that the greater good consideration of SDI will be a valid part of the business case. For this reason and reflecting the early state of SDI in New Zealand, the standards needed to deliver the characteristics are kept deliberately minimalist.

5.3.1 View Systems

Essential

The following capabilities should be regarded as essential:

- the ability to view Web Mapping Service (OGC WMS)

Highly Desirable

It is highly desirable that systems also have the following capabilities:

- the ability to query Web Mapping Service (OGC WMS)
- the ability to consume Web Feature Service (WFS).

Additionally Useful

It may be additionally useful to consume and interact with a wide range of other open standards such as:

- Web Coverage Service (WCS)
- KML
- REST-based services.

5.3.2 Analysis Systems

The capabilities below only refer to the system's ability to participate in SDI. The first priority for the system must be to satisfy the business requirements of the organisation. Many analysis functions may require access to enterprise systems other than GIS and these may require the use of proprietary standards in order to access these proprietary enterprise systems.

5 — Efficiently using data: how can my organisation use SDI?

5.3 User system characteristics *continued*

At the early stages of the development of SDI in New Zealand, it may be premature to implement analysis systems as a part of SDI. The open standards needed to implement analysis systems are challenging to employ successfully and are not widely available in geospatial products. For those reasons, no essential characteristics are listed.

Highly Desirable
The following capability should be regarded as highly desirable if it contributes to the business needs of the organisation:

- For a client-side analysis system, the ability to consume Web Feature Services (WFS)

OR

- For a server-side analysis system, the ability to view Web Mapping Service (WMS) to provide a contextual backdrop for analysis and the ability to connect to a Web Processing Service (WPS).

Additionally Useful
It may be additionally useful to perform the following:

- The ability to connect to REST-based analysis functionality that might include connections to non-geospatial systems.

5.3.3 Maintenance Systems

As discussed above, maintenance systems potentially encompass a broad range of geospatial capability. Simple edit activities such as adding point information or editing attribute information may be more easily performed using IT standards such as writing SQL commands to update an underlying DBMS. Geospatial standards might only be needed for more complex editing tasks.

Whilst SDI in New Zealand is at an early stage of development, putting any maintenance activities down as 'essential' is premature. The geospatial standards required to support data maintenance are not widely implemented and are very dependent on the specific implementation in order to achieve interoperability.

Highly Desirable
For simple data maintenance tasks:

- Use appropriate IT standards such as SQL, or REST

For more complex editing tasks:

- Transactional Web Feature Service (WFS(T))

5 — Efficiently using data: how can my organisation use SDI?

5.3 User system characteristics *continued*

5.3.4 Administrative Systems

As discussed above, administrative systems potentially encompass a broad range of geospatial capability.

Whilst SDI in New Zealand is at an early stage of development, putting any administrative activities down as 'essential' is premature. Implementing this class of system using open standards can be challenging.

Highly Desirable

It is highly desirable to implement the following standards:

- the ability for an administrative user to be able to view and query OGC WMS
- the ability for an administrative user to be able to query catalogues using CSW

Additionally Useful

It may be additionally useful for an administrative user to be able to view and query WFS.

5.4 Where to obtain specialist assistance during implementation

Participating in SDI as a user should be straightforward– the need for specialist advice should be very limited and provided from within an organisation's own support resources.

Implementing a user system should be approached with more caution and it may be appropriate to seek advice from a geospatial specialist for some aspects of implementation, particularly where knowledge of matters such as geospatial data quality, projections and datums may be required.

Where can that advice be found?

- In some cases, the user organisation might have a geospatial or GIS team. That should always be a first port of call since it is important to ensure that internal geospatial capabilities are well aligned to external SDI capabilities.
- There are many geospatial consultants in New Zealand. The industry body representing the geospatial industry is the Spatial Industry Business Association (SIBA) NZ. SIBA can be contacted by e-mail at secretary@siba.org.nz and will be able to provide a list of geospatial consultants.
- The New Zealand Geospatial Office will be able to address any questions.

SPATIAL DATA INFRASTRUCTURE:

6 How existing systems and products can contribute to SDI

6 — How existing systems and products can contribute to SDI

This chapter is a collaboration with the spatial industry. The contents of this chapter have been volunteered, using a template developed to help achieve a consistent and broad overview of available SDI conforming applications. The NZGO and LINZ appreciate the support of the spatial industries and its communities in participating in this way.

Chapter 6 is accessible online:

http://www.geospatial.govt.nz/sdi-cookbook-home

Information for users:

Content for this chapter is provided by vendors or open source communities and the currency of the information is reliant on NZGO receiving updates from there vendors or communities.

Readers are advised to use due diligence. Evaluate accuracy, completeness and relevance of the information for their own purposes, and to confirm details with original sources as the data surrounding products can change rapidly. New product releases, support for new standards and other factors can mean that the information contained in this document can become out of date.

It's important to note that use of the material in this chapter is subject to different creative commons licenses. These are notes on the submitter contact details page.

Due to the volume and nature of information contained in this chapter, it is only available online and not as a PDF.

Information for vendors:

We are interested in learning about any new products that contribute to participation in SDI. Please contact nzgo@linz.govt.nz if you would like to submit an application.

For suppliers of SDI conforming applications already contained in this chapter, please contact nzgo@linz.govt.nz to advise of any changes or updates to your material.

Feedback and queries

We’re really keen to hear your thoughts. You can comment on these pages through the comment function on this website, or email nzgo@linz.govt.nz

We’re considering what other channels we can establish to capture thoughts and discussion on the content included, so we welcome your suggestions on this also.

The New Zealand Geospatial Office will also be engaging with the New Zealand geospatial community to develop a ‘user forum’ type environment for ongoing discussion and revision of Cookbook content to ensure its relevance and currency.

Queries about the Cookbook can also be sent to nzgo@linz.govt.nz

Terms

ANZLIC – the Spatial Information Council http://www.anzlic.org.au/

Catalogue Service for Web (CSW) – http://www.opengeospatial.org/standards/cat

Custodian is responsible for the continued physical existence, update and availability of a geospatial dataset data. In some cases the steward and the custodian will be the same organisation.

Geography Markup Language (GML) – http://www.opengeospatial.org/standards/gml

ISO Technical Committee 211 (ISO/TC211) – a technical committee of the International Organisation for Standardisation that deals with geographic information standards http://www.isotc211.org/

KML – an XML language used predominantly for Google-based applications http://www.opengeospatial.org/standards/kml

NZGOAL – New Zealand Government Open Access Licensing framework http://www.nzgoal.info/

Spatial Data Infrastructure (SDI) can be broadly defined as a network of components that allows people to find, share and use spatial data. Key SDI components have been identified as: policy, access networks, standards, data and people.[53]

Steward is accountable for maintaining the overall quality, integrity, security and dissemination of a geospatial data. Responsibility and accountability for the stewardship rests with the head of the organisation appointed for a particular dataset.

OGC Web Coverage Service (WCS) – http://www.opengeospatial.org/standards/wcs/

OGC Web Feature Service (WFS) – http://www.opengeospatial.org/standards/wfs/

OGC Web Mapping Service (WMS) – http://www.opengeospatial.org/standards/wms/

53 Developing Spatial Data Infrastructures: From concept to reality, Taylor & Francis Group