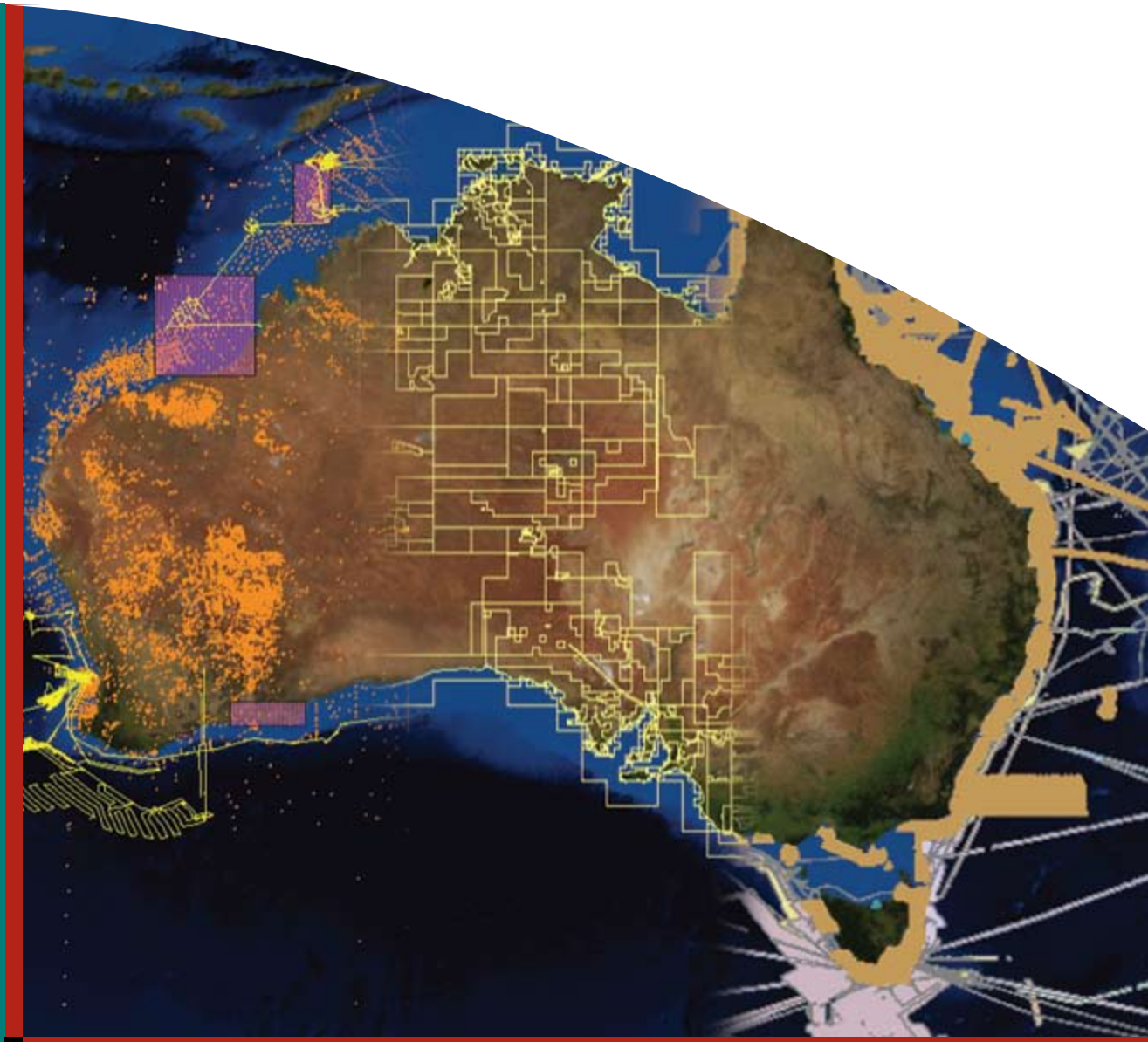
Australian Government
Geoscience Australia

Single Point of Truth (SPOT) Methodology

S. Ross, N. Evans, H. Anderson, M. Jenkins, G. Lawford, D. Simon, F. Brassil, S. Lang and L. Wyborn

Record
2007/17



Single Point of Truth (SPOT) Methodology

GEOSCIENCE AUSTRALIA
RECORD 2007/17

by

S. Ross, N. Evans, H. Anderson, M. Jenkins, G. Lawford, D. Simon, F. Brassil,

S. Lang and L. Wyborn

Australian Government
Geoscience Australia

1. All authors were staff members of Geoscience Australia at time of document creation

Department of Industry, Tourism & Resources

Minister for Industry, Tourism & Resources: The Hon. Ian Macfarlane, MP

Parliamentary Secretary: The Hon. Bob Baldwin, MP

Secretary: Mark Paterson

Geoscience Australia

Chief Executive Officer: Dr Neil Williams

ISSN 1448-2177

ISBN Web 978-1-921236-59-4

ISBN HardCopy 978-1-921236-56-3

GeoCat # 65443

Bibliographic reference: Ross, S.I., Evans, N.D., Anderson, H.A., Jenkins, M., Lawford, G., Simon, D., Brassil, F., Lang, S. and Wyborn, L., 2007. Single Point of Truth (SPOT) Methodology. Geoscience Australia, Record 2007/17. 40pp.

Table of contents

Introduction

What is a SPOT?

Geoscience Australia (GA) produces geoscientific and geospatial data for the benefit of the Australian government and community, to inform public policy, to promote development of Australia's economy, to assist environmental management and to help manage and mitigate natural hazards. Users of GA's data want to know that data are produced to known standards using open and accountable processes and come from a unique and reliable source.

Single Point of Truth (SPOT) is Geoscience Australia's standard for processes that produce data. Data produced to the SPOT standard is "Authentic, Authoritative and Accurate":

- Authentic
    - Geoscience Australia's best available data
    - Sustainable systems support and protect the data and its metadata
- Authoritative
    - GA has a legitimate role to provide and authorise the distribution of the data.
    - Intellectual Property is certain and documented
- Accurate
    - The accuracy, precision and quality, lineage and fitness for purpose of the data are documented, accessible and traceable.

A SPOT Data set:

- Is unique within Geoscience Australia (GA)
- Is managed in a corporate system
- Is maintained and up to date
- Has full and correct metadata published in a corporate metadata system
- Has a documented governance process that includes
    - A fully described and modelled data system
    - Defined roles for data management and qualified staff in those roles
- Meets defined quality standards
- Is discoverable, accessible and interoperable and
- Complies with relevant international, national and agency standards

Most data sets produced by GA arise from some sequence of processes in which original measurements or observations (raw data) are transformed to make usable information. This transformation can be as simple as editing and reformatting or it may require complex algorithms and scientific interpretation. This sequence of processes forms a "*Data System*", in which each new form of data depends on the quality of the input data and the competence of the "value-adding" processes.

A Data System can be called a "*Data Theme*". The SPOT concept requires that the entire data system or data theme is managed to SPOT principles.

Sometimes people describe a data system as transforming data into information and knowledge. The SPOT principles apply to all stages and for simplicity in this document the word 'data' refers any point in this continuum.

The business case for SPOTs : Leadership, efficiency and risk management

GA is a national leader in providing a very diverse range of geoscience and geospatial data to an equally diverse range of users within the Australian community and internationally. GA's data and information form its key business asset and must be managed strategically.

Without a rigorous process for data management and delivery, data and information holdings can become compartmentalised, fragmented, and hidden. This leads to multiple copies of data, uncertainty about provenance and authenticity, higher costs of running projects and operations and reduced confidence of staff and customers in the quality of data. Without easily discoverable authentic data sources, large amounts of time are used inefficiently attempting to discover and access data.

GA continually acquires new geoscientific and geospatial data. The rapidly increasing quantities of data compound data management problems. This increases the need to have well described and accountable processes for data management and delivery.

With GA's increasing profile, clients expect excellence in its outputs and that GA can substantiate this with transparent documented processes. The SPOT methodology will apply a consistent discipline to GA's data and preserve and enhance GA's reputation as a source of high quality data and as a reference for good data and information management practice.

The principles of risk management apply to data and information management and hence to SPOTs. The resources applied to managing GA's duty of care to its clients must be balanced and proportionate and account for the likelihood and consequences of the associated risks. Ensuring that data and information are authoritatively based is therefore an essential step in risk management for Geoscience Australia.

SPOTS and Geoscience Australia's Projects

Most of GA's business program uses a project management method. Most projects use some existing data and acquire new data to which they apply advanced science and information technologies to produce new information and knowledge.

Projects will use SPOT data sets as their source of existing information. New information which a project acquires or produces which can be added to a SPOT data set will be added to that SPOT. This will be reflected in the Project's Information Management Plan which is required by GA's Chief Executive Instruction on Information Management.

All data and value added outputs and products from projects are preserved in corporate systems with appropriate metadata for delivery to clients.

Projects have space for storing data which is being worked on by the project. Data within project storage is not subject to the SPOT methodology, but outputs and products must conform to the requirements of the relevant SPOT data system.

Access to SPOTS

SPOT data is discoverable though a single page on GA's Intranet which has links to each SPOT data theme. Each theme has Intranet pages which give users all the information needed to access and use a SPOT data set.

How the SPOT Methodology works

The SPOT methodology describes a consistent approach to transforming an existing data theme into a SPOT. The same methodology can be used for developing a SPOT for a new data theme.

Establishing a SPOT should be managed as a project and follow GA's Project Management Methodology, which is described on GA's Intranet.

The SPOT methodology has four stages:

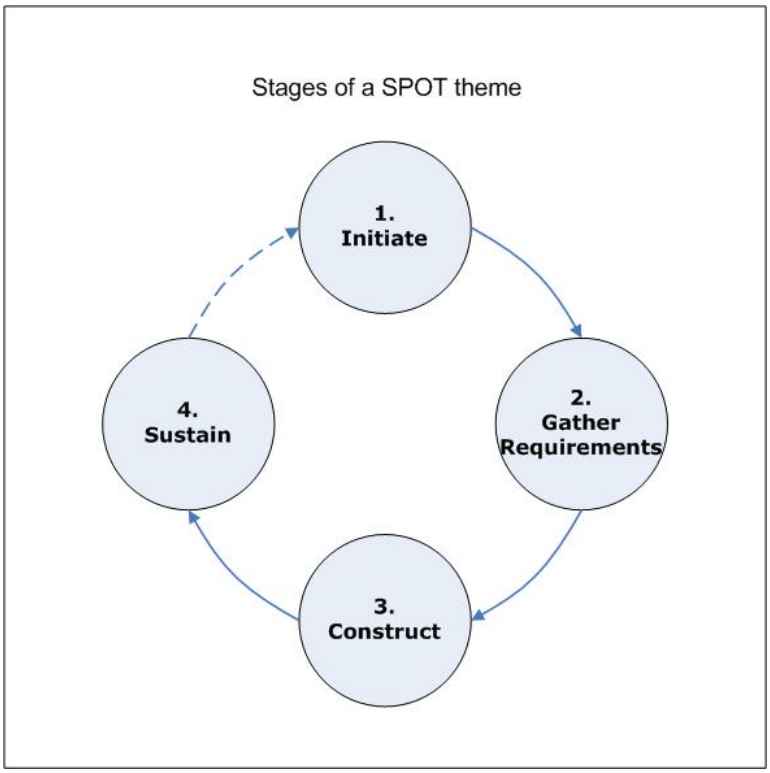
1. *Initiate*
   Scoping the SPOT project. This relatively short step establishes the approval for commencing the SPOT.

2. *Gather Requirements*
   Design and plan the process of establishing a SPOT data system

3. *Construct*
   Building the SPOT according to the plan

4. *Sustain*
   Ongoing maintenance, support and review of the SPOT data system.

The figure shows the sequence of events, and implies that after some time it may be necessary to reassess the SPOT to ensure it is continuing to meet GA's needs.



Maintenance and Delivery of SPOT Data

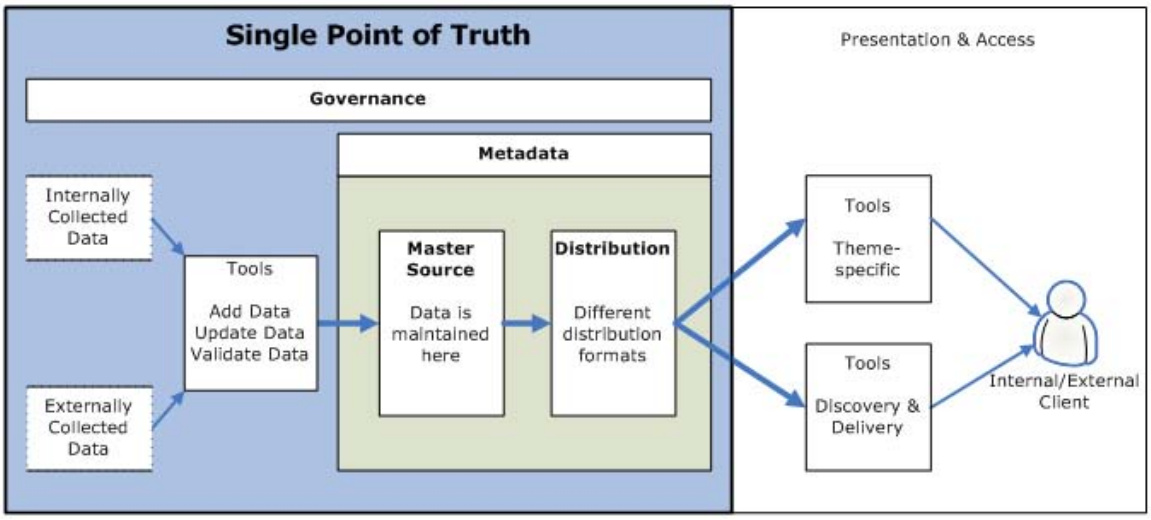
GA data are maintained in and delivered from one of three environments:

- A Relational Database Management System (RDBMS) or an application framework built on a RDBMS. GA's RDBMS is Oracle™.
- A Hierarchical File System (HFS). GA's corporate on-line HFS is called the Corporate Data Store (CDS).
- Off-line on physical media such as tapes or optical disks. These must be physically handled and duplicated when being delivered to users. Over time all but the most voluminous data sets will be moved to on-line access.

The SPOT concept requires that data are maintained in a unique "*Master Source*" environment which is usually separate from the environment from which data are supplied to users. These are called the *Maintenance Environment* and the *Distribution Environment*. This separation is necessary because the business processes which keep data up to date are different from the business processes by which data are made available to clients.

SPOT data can be discovered and delivered in many forms through many channels. The SPOT Methodology requires that Distribution Environments are updated from Maintenance Environments in a well defined way. More than one Distribution Environment may use a Maintenance Environment. The SPOT remains valid provided all data are traceable to a unique Master Source.

The diagram shows the structure of information flows in a SPOT data set and the relationship between a SPOT dataset and its delivery.



Data in a SPOT can come from a single or many sources that may be internal or external to GA.

Tools are the resources for entering, editing, updating and validating data.

The Master Source is the unique location where the data will be maintained. It may be in a RDBMS or the CDS.

A Distribution Environment will contain and deliver data in formats that meet the business needs of users. The Distribution Environment may be implemented in a RDBMS or the CDS.

Metadata for the SPOT will be stored within the Corporate Metadata Repository.

Governance arrangements including custodianship and defined processes for updating are needed to sustain the SPOT.

Distribution Tools can be many and varied, depending on the needs of users and the kinds of data.

Using standards based systems for metadata and interoperability, GA has an agency-wide toolset to enable discovery and delivery of information and knowledge for our external and internal clients. Standards based systems minimise development and maintenance overheads and permit users to combine data from many sources interoperably.

Theme specific tools may be developed for specialised requirements for certain classes of data or users.

Distribution tools and channels may change over time as client, business and technology needs and capabilities change. The SPOT methodology separates the SPOT Master Source data from the environment from which data are distributed. This “loosely coupled” structure helps ensure the currency and consistency of data over time and the agility of the agency to respond to changing business needs.

Specifications for building tools for Discovery and Delivery of a dataset are outside the scope of the SPOT methodology but when establishing a new SPOT it may be necessary to develop new distribution tools to provide data to clients.

Stage 1: Initiate

Objective

The *Initiate* Stage determines the scope and feasibility of establishing the SPOT data theme and documents the basis for approval to establish the SPOT.

A SPOT project is a major investment and requires an *Executive Sponsor* who in GA should be a Group Leader or Chief of Division with organisational responsibility for the data theme. The Executive Sponsor will assign staff to do the analysis necessary to scope the SPOT project.

On completion of this stage, there will be sufficient information to develop a Project Plan to establish the SPOT.

Description

The following activities occur in this stage:

- *Coordination with Division Information Managers*
- *Coordination with Information Services Branch*
- *Stakeholder Analysis*
- *Identify the Data and Data Systems*
- *Write the Project Scope*

- **Coordination with Division Information Managers**
  GA's division structure devolves responsibility for geoscience and geospatial discipline subject matter. Each division has a Division Information Manager who coordinates Information Management capabilities and responsibilities within the Division. The Division Information Manager is a key contact for planning the establishment of a SPOT. Frequently it will be necessary to coordinate with Division Information Managers in all divisions to ensure all relevant stakeholders in a SPOT are engaged.
- **Coordination with Information Services Branch**
  Information Services Branch manages GA's corporate Information Management systems, standards, processes and technology. Coordination with the corporate Information Services Branch will ensure that
    - the SPOT will be developed in a corporately consistent manner,
    - necessary corporate capabilities can be established and
    - the SPOT plan aligns with corporate information management policies and practices.

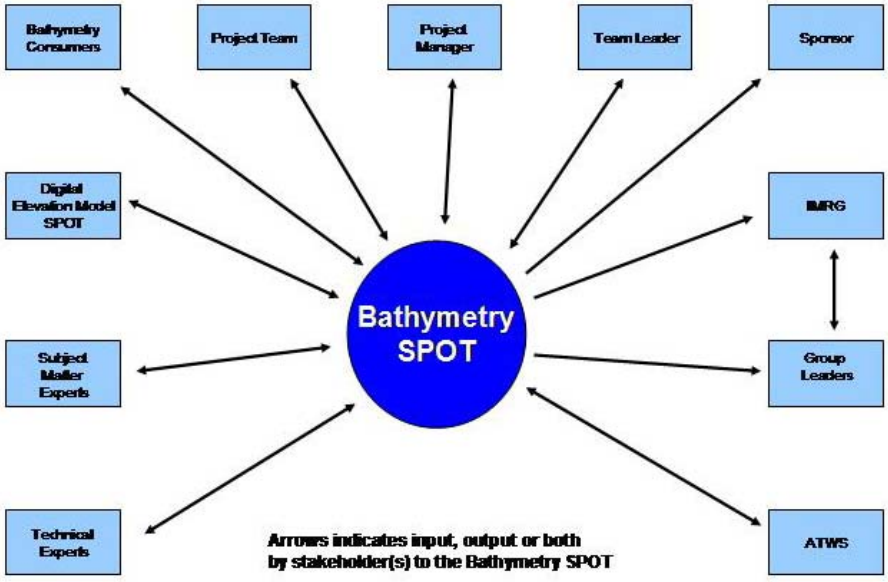
- **Stakeholder Analysis**
  A "stakeholder analysis" seeks input from all significant internal and external users of the data. Each stakeholder will have a particular interest which determines the significance of the stakeholding. Therefore it is necessary to gain an understanding of the range of interests and their relationships early in planning the SPOT.

  An important output from this activity is the documentation of the "critical success factors" for each stakeholder from the SPOT. This information is used to assess the priority of establishing the SPOT and for developing a reference against which the establishment of the SPOT can be tested. These critical success factors will reflect the business critical elements of the SPOT data system.

  Records of client feedback about the data should be reviewed to inform the Stakeholder analysis.

  The Stakeholder Questionnaire template can assist gathering this information. An example is included in Appendix 3 and this may be used to develop a Questionnaire for each SPOT project.

  A Context Diagram is a useful tool to identify stakeholders and document their relationships. This example from the Bathymetry SPOT illustrates the concept.



Example Context Diagram

- **Identify the Data and Data Systems**
  Identify the range of data types and themes relevant to the proposed SPOT. This is a high level review to estimate the range, volume and complexity of the data associated with the SPOT sufficient for scoping the project. A detailed analysis or audit will be done in the next stage. It is necessary to identify the *Custodians* for each data type and theme. Any data that lacks or has weak custodianship is a significant increase in the risk associated with establishing the SPOT.

- **Write the Project Scope**
  A project scope document for the SPOT should be written using GA's Project Scope Template.

Estimates of all significant resource requirements including subject matter experts, development skills and technology resources should be in the Scope. While these may be approximate, they are necessary for realistic costing of the implementation and for prioritising the SPOT.

When developing the Project Scope make allowance for risk, particularly if development of tools is required or the understanding of the data to be incorporated is poor. High risk situations can more than double the time expected to implement the SPOT.

A major element of the Project Scope is identifying the business benefits of establishing the data theme as a SPOT. Inevitably there will be competition for resources and only the themes with the greatest business value will be supported.

Record dependencies between this project and any other projects, whether they are SPOT projects or projects in GA's normal business program. Where appropriate allow for risk associated with dependencies.

Consider carefully the assumptions and constraints associated with the project. Assumptions are things that must be in place for the project to commence. Review carefully what is needed to ensure that the project can commence without being hampered from the beginning.

Constraints are limitations within which the project must work and that can be altered only with difficulty. These may be resource constraints such as availability of essential skills, time constraints relating to key delivery commitments, or business constraints, relating to specific requirements of key stakeholders.

Identify in the project scope the key milestones by which progress through the project can be measured.

Outputs

The following documents will be produced from this stage:

- Stakeholder Analysis
- Preliminary assessment of the range and volumes of data
- Project Scope document

Review and Approval of This Stage

Division Information Managers and custodians should be asked to review the Scope. The Scope should be endorsed by the CIO prior to seeking approval for the SPOT project from the Executive Sponsor. This will ensure that there are the correct linkages to corporate information systems, and Information Management policies, plans and guidelines

The project scope will be approved by the Executive Sponsor, who will nominate the Project Leader for establishing the SPOT.

Resources and Skills

The following resources are required for this stage:

- Task Leader for this stage
- Business Analyst
- Access to Key Stakeholders including Project Sponsor and Custodians

Initiate Stage Checklist

- Project Sponsor
    - Have you an Executive Sponsor who agrees to initiate the SPOT project and who will champion the SPOT at senior Division and agency levels?
- Stakeholder Analysis:
    - Have you validated that all stakeholders have been considered?
    - Have you documented the interest of each stakeholder?
    - Is external stakeholder input appropriate?
      How will you identify and communicate with them?
    - Have you examined customer feedback?
    - Have you categorised stakeholders as key or non-key?
    - Do you have agreement on critical success factors?
- Resource Requirements:
    - Have you identified staff with specific skills and other resources that are needed to develop the SPOT?
    - Have you estimated the quantity of resources to develop the SPOT?
- Related Projects:
    - Have you noted any dependencies between this and other projects?
    - Have you assessed the impact of these dependencies on the success of the SPOT?
- Project Scope:
    - Are there key dates that must be met or worked around?
    - Have you consulted your Division IM Manager about the significance of this SPOT data theme and the intended direction?
    - Have you identified strategic priorities and significant business outcomes?
    - Have the stakeholders had an opportunity to review the scope prior to sign-off?
    - Have you considered the key fields in the Project Scope document, particularly the risk, assumptions, constraints, related projects, and the milestones by which project achievements can be measured?

Stage 2: Gather Requirements

Objective

Once the SPOT Project Scope is approved the *Gather Requirements* stage will:

- Gather User Requirements for the SPOT
- Analyse and document in detail existing data types, locations, work processes and standards
- Model the data system conceptually and develop a logical model and
- Plan the construction of the SPOT data system inline with SPOT principles

At the end of this stage you will know what is required to convert the data theme to a SPOT and have developed the plans to construct it.

For a significant SPOT or when there is a wide range of stakeholders or significant risk a Technical Reference Group (TRG) should be established. TRG members should be subject matter experts or represent key users and should include a Division Information Officer. The TRG can provide independent advice about SPOT planning and construction and can support the SPOT project by assisting to validate and refine data models and proposed standards and outputs.

Description

The following activities occur in this stage:

- Document current systems, data flows and processes.
- Document user requirements
- Determine metadata requirements
- Determine relevant standards
- Produce a logical data model
- Develop acceptance criteria
- Determine distribution formats
- Develop metrics
- Develop the governance model
- Develop the operating model
- Plan the documentation to sustain the SPOT
- Develop a promotion strategy
- Produce a project plan for the SPOT Construction

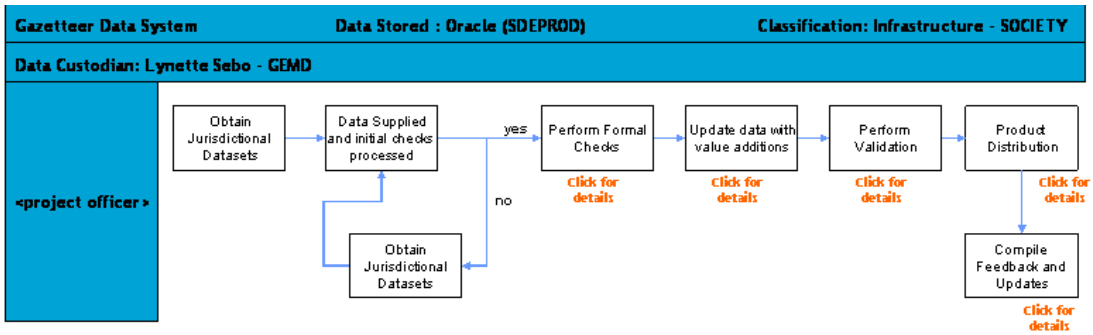
Document Current Systems, Data flows and Processes

The project needs to discover:

- The business drivers for the existing data, processes, outputs and products
- Where the data comes from, how it gets into GA and what processes exist to manage its quality and acceptance.
- How much data exists, where it is stored, what formats and standards are used, and its rate of growth
- The processes GA performs to transform and add value to the data
- How users access the data
- The gaps or deficiencies in the existing processes
- Relevant International, national, industry or GA data standards.

The Stakeholder Analysis from the *Initiate* phase is the starting point for this task. Custodians of existing systems are a principal resource, as are internal and external users of the data. Data Audit and User Questionnaire templates as shown in Appendix 4 and 5 can assist the project team to gather this information.

The data flows, processes and roles should be described in a conceptual model. The conceptual model helps to ensure a common agreement about the high level data flows and transformations the SPOT will encompass. UML is useful for this. An example from Gazetteer data SPOT is shown.



The Conceptual Model may initially reflect the existing arrangements but it will provide a tool for developing a high level concept of the data flows and relationships in the new SPOT. It is vital that some “lateral thinking” be introduced to encourage innovation and creativity in planning the SPOT.

Document User Requirements

User Requirements record the SPOT data system must deliver for it to be an effective and efficient business output.

Tools to assist gathering user requirements are:

- A *User Requirements Questionnaire* to provide a consistent set of information.
- A *workshop* for gathering user requirements; this is an opportunity for many stakeholders to communicate their interest. This helps to refine user expectations by ensuring they appreciate the complexity and range of interests to establish and deliver the SPOT.
- Consultation using face to face meetings with key stakeholders.

The User Requirements describe what users of the data need, expect and prefer from the data system. There is likely to be a range of users of the data and their various expectations should be compiled and prioritised. The TRG and Custodians should validate the requirements and priorities.

Geoscience Australia as owner of the data system has a requirement to maintain the data as a strategic business asset and this must be reflected in the documentation of requirements.

Determine Metadata Requirements

Review existing metadata practices. Determine the required metadata and the appropriate level such as attribute or data set level and any attributes required in addition to GA's standard metadata profile, including IP information.

In determining the metadata requirements, consider the metadata capture process, whether it meets GA's metadata profile and what processes need to be implemented to ensure valid metadata is captured.

To the greatest extent practicable, metadata capture should be automated.

Determining relevant standards

Standards may apply at the international, national, industry or GA level.

A relevant standard is one that:

- Is relevant to the business sector
- Assists in the delivery of business outcomes and
- Is actively developed and supported

There may be competing relevant standards but not all standards may be appropriate. The relevant standards should be determined in consultation with subject matter experts and with Information Services Branch.

As a public body, GA has a responsibility to promote open standards. If an open interoperable standard has been developed, data should be published in that standard. *De facto,* proprietary or application specific standards may be produced or expected by clients and should be generated from the Distribution Environment.

Producing a Logical Data Model

The *Logical Data Model* is the abstract description of the SPOT. It is based on the user requirements and information about existing data systems and processes. It represents all sources of data, all data transformations and all output generation including metadata. It should commence from the Conceptual Model but there should be no presumption that existing processes will be maintained.

Consider the *Information Life Cycle* of the data. Some data has a very short period of interest and can then be put away or deleted. Many types of geoscience and geospatial data have very long life cycles and may have value for decades. Because data storage and maintenance is a long term and significant cost, it is necessary to evaluate carefully the long term strategies for storage and delivery of data and to reflect the ILC in the Logical Data Model.

The logical model is a critical document for the SPOT and will form part of the publicly available information. The most useful way to present it is as a UML diagram. It will be presented to the TRG for consultation and validation, and all interested stakeholders for information.

Agreement on the Logical Data Model is a key milestone for the SPOT project.

Develop Acceptance Criteria

Based on information provided by stakeholders, develop criteria to determine if the SPOT will meet the agreed user requirements. These Acceptance Criteria should be reviewed by the TRG.

Determine Distribution Formats

The project will decide which data formats will be provided from the Distribution environment.

Distribution formats will be determined by:

- Client needs
- Accepted industry practice
- Geoscience Australia and Commonwealth government policy
- Relevant international, national, industry and GA standards

Distribution formats may be proprietary or open. Whenever a proprietary format is used, an open format should also be provided.

Develop Metrics

Metrics enable Custodians and Data Managers to know how much and by whom the SPOT is being used, the rate of growth of the data and whether it is meeting the business needs of stakeholders.

Metrics include:

- Operational
  Rate of growth, frequency of change and update.
- Financial
  The costs of sustaining the data: the cost of governance roles, cost of storing the data or the cost of updating the environments.
- Customer
  Number of users, frequency of use, quantities accessed, time taken to service users, feedback from users.
- Business Process
  Performance of key business processes such as the time spent fixing problems in the QA process.

Determine the Governance Model

Note any existing governance arrangements and develop a governance model for the SPOT. The governance model must include:

- Business Function and objectives of the SPOT
- Policy objectives and related policies for management of the SPOT
- Roles & Responsibilities including Executive Sponsor, Custodian and Data Managers
- Update policies for master source and distribution
- Quality Assurance objectives, standards and procedures
- Intellectual Property considerations, including relevant government policies, external interests in the data and protecting the Commonwealth's title to its information assets.

The Executive Sponsor ensures that the key roles are filled.

Develop the Operating Model

Based on the user requirements, the Logical Model and the resource requirements, develop an operating model for the SPOT. This describes the business processes which will sustain the SPOT once established. It should describe the process through which all sources of data will be captured, all outputs and products to be produced, the data transformations and policies for the generation of outputs. The Operating Model may be expressed as a "Concept of Operations" which provides overall guidance for the ongoing custodians and manager but leaves operational details to the managers. Questions to be considered include:

- How will new data be captured? Where will it be preserved? What policies will apply to access to the data?
- What range of outputs and products will be produced? This includes all value added products and any access to raw or intermediate processing products.
- What policies will apply to making new data available for each output? For example will new data be immediately web accessible, or will compilation products such as maps, images or gridded data sets be subject to regular or occasional update?

The Operating Model should enable a knowledgeable person to take over responsibility for the SPOT data system and be quickly able to manage the operation effectively.

The Operating Model requires sufficient detail to ensure that all user requirements are properly accounted for in ongoing operations. Note: Over-specifying this model could lead to inflexibility as business needs change.

A significant element is the implications for work practices. Are new positions required, or is there a need to reclassify or restructure existing positions? What skills are required, and how should those skills be acquired and maintained?

In developing the operating model it is essential that a business efficiency approach is taken and all options including contracting, outsourcing and internal service provision should be considered.

Plan the Documentation to sustain the SPOT

There are three purposes for documentation:

1. To inform data custodians and managers of the requirements for sustaining the SPOT
2. To provide a backup set of information so that a knowledgeable person could assume custodianship of the SPOT and be able to take operational responsibility in a reasonable time; and
3. To provide a reference against which ongoing operations of the SPOT can be audited.

The documentation plan should consider the best medium for the documentation including where the documentation is maintained and how it is accessed. Intranet pages are useful and accessible. The preferred approach is to save all documents in GA's EDMS and to have links to those documents from the Intranet pages which inform internal users about the SPOT.

Suggested documents needed to sustain the SPOT are:

- High level SPOT Description including data models and standards implementation reference documents.
- System Architecture for technical support and maintenance staff.
- Intranet pages to inform users about the SPOT. These should follow a consistent framework and be linked to a master page which enables discovery of all SPOT data from a singe entry point.
- Help Guides for Users
- Data Administrators Guide
- Operational Procedures reflecting the Operating Model

Significant judgement is required to balance the cost of producing and maintaining documentation against its usefulness for ongoing operations.

Develop the Promotion Strategy for the SPOT

Development of a SPOT is a considerable investment and should represent a significant improvement in data access and capability for clients. It is important that clients, especially external and potential new clients are informed that the SPOT is available. Promotion of SPOT data systems advances reputation of the agency as a source of high quality data and data management practice.

Promotion opportunities encompass the domain discipline of the SPOT data system and those professional groups interested in data and information management.

A promotion strategy for the SPOT will be implemented once the SPOT is established, but there may be reason to do some of the promotion work earlier.

Promotion opportunities include:

- Papers in relevant journals. This may include Peer reviewed journals and the lighter journals produced by professional societies that are often more widely read.
- Presentation of papers and posters at relevant conferences.
- Advertising.
- Web notices
- Promotional visits to key clients.

The Promotion Plan must have the agreement of the executive sponsor.

Producing a project plan for the Construction stage

A detailed project plan for constructing the SPOT should use the GA's Project Plan Template.

The Project Plan will include all the elements that have been identified in this stage. It needs to include:

- Designing the physical model.
- Building the physical structure
- Writing or modifying documentation
- Testing

- Developing the decommissioning strategy
- Handover of the SPOT

Outputs

The following will be produced from this stage:

- Documentation of current systems
- User Requirements Metadata, output formats and standards
- Logical Model of the SPOT Data System.
- Acceptance Criteria for SPOT implementation
- Governance Document
- Documentation of the Operating Model of the SPOT and metrics
- Documentation Plan for the SPOT.
- Promotion Plan
- Project Plan for the *Construct* stage.

To Complete This Stage

All stakeholders including the TRG and the CIO should have an opportunity to provide input into and review the requirements document, the Conceptual and Logical Models and the project plan. If there is significant disagreement about the content of these documents then the matter may be referred to the CIO. The CIO may consult relevant executive staff to determine the resolution.

This stage is complete when the Executive Sponsor approves the project plan for construction of the SPOT.

Resource / Skills

The following resources or skill-sets are required for this stage:

- Project Sponsor
- Project Leader
- Business Analyst
- Subject matter Experts including Custodians and Data Managers
- Stakeholders, including knowledge of external stakeholder requirements
- Data Modeller

Gather Requirements Stage Checklist

- User Requirements:
    - Have you received and evaluated requirements from all stakeholders?
    - Are there any corporate IM or IT considerations?
    - Have you considered external stakeholder requirements?
    - Have you identified a corrective action for each gap or deficiency, or documented reasons for not responding?
    - Has the Technical Reference Group reviewed the requirements?
    - What standards do the stakeholders require?
    - Have you validated the data model and metadata with each stakeholder group and the TRG?
    - Have you identified the consequences for existing job / role responsibilities? Does the project need to produce a Job Impact Statement?
- Existing data and processes:
    - Have you identified, located and quantified all data maintained by data managers and other individuals for the entire data system?
    - Have you verified the Data Audit results?
    - Have you identified existing or potential additional data sources?
    - Have you located and assessed existing process documentation?
    - Does the data conform to the standards that will apply to the SPOT?
    - Does each data source have an associated work flow?
    - Do you know the age of the data and its estimated lifecycle?
    - Have you identified any performance requirements, for example real time data capture or delivery
    - Have you collected any capacity estimates or usage data?
    - If GA is not the custodian, have you documented the entry points into GA, IP and usage?
- Governance:
    - Are the custodian and data manager(s) identified? What training or support do they need?
    - Have you assessed the job implications for custodial and data management duties?
    - Does the custodian have links to the governance bodies for the relevant external standards?
    - Are Intellectual Property matters fully addressed?

- Metrics:
    - Have you developed metrics to sustain the SPOT and included implementing them in the relevant documents?
    - Are these metrics easy to implement and interpret for the Custodian and Data Managers?
- Operating Model:
    - Have you developed a fully described operating model that will support custodians and managers of the SPOT in ongoing operations?
    - Have you identified the range of outputs and products to be produced?
    - Have you defined the policies that govern maintenance and update of the "master source" and distribution environments?
    - Have you evaluated all business models including in-house and outsourced service models?
- Logical Model:
    - Have you validated the model with the key stakeholders?
    - Have you used UML?
    - Does the model represent an abstract rather than a technology specific description of the data system?
    - Does this data model integrate with other models
- Documentation Plan:
    - Do you have a supporting structure for the documentation?
    - Have you analysed how and where the documentation will be accessed?
    - Have you appropriately balanced the usefulness and maintainability of the documentation?
- Promotion Plan
    - Is there a promotion plan that will ensure that GA gains the maximum benefit in increased use of its data and information assets?
- Project Plan:
    - Are the Divisional Information Managers aware of the consequences for their Division?

Stage 3: Construct

Objective

The *Construct* stage constructs the SPOT using the plans developed in the previous stage.

Changes to data structure, format, data store or metadata are implemented. Documentation needed to sustain the SPOT is developed. The Operating Model is tested and the Governance Model is established.

At the end of this stage, the data conforms to the SPOT requirements. Custody of the SPOT is handed to the staff responsible for sustaining it.

Description

The following activities occur in this stage:

- Design the physical model.
- Build the systems from the physical model
- Write supporting documentation
- Test the SPOT system
- Develop the decommissioning strategy
- Handover of the SPOT
- Promoting the SPOT
- Post Implementation Review

Design the Physical Model

The Physical Model describes the specific implementation of the SPOT in the agency's technology framework. This physical model will be implemented when the SPOT is constructed. It must accurately implement the Logical Model and be feasible and effective in the current and anticipated technology architecture.

All requirements for resources, development and technology must be fully described in the physical model.

A physical model includes the structures for the master source and the distribution environments. Changes may be required to:

- The data capture systems
- The metadata capture systems
- The data storage systems
- The data transformation and value adding processes
- The data discovery systems

- The data delivery systems.

The new design may require changed work processes to align with the new physical model.

New or modified presentation or user interface tools may be needed.

The physical model must be reviewed by the TRG, by the custodians of all affected data systems and by the custodians of the corporate information infrastructure where the physical model will be implemented.

At this point all policies and practices for updating master source and distribution environments must be resolved.

Build the Physical Structure

This will comprise:

- building the master source & distribution data structures
- Migrating data to the new structures, including making any special tools that may be needed
- populating the metadata in GA's Metadata Repository
- Establishing new systems to capture new data in the new structures
- Building tools to enable users to discover and access the data.

Write Supporting Documentation

Documentation for the SPOT has been identified in the planning stage. Developing documentation is slow and expensive and must be commenced at an early stage. It is vital to have an independent editor for the documentation.

All documentation must be captured in the agency's EDMS and be linked to the Intranet pages that describe the documentation.

Documentation must be reviewed by the custodians of the SPOT.

Test the SPOT System

The nature of the requirements will dictate the type of testing of the SPOT. This may be unit testing, system testing, user acceptance testing, performance testing and stress testing.

Subject matter experts and data managers should perform the testing. This may involve:

- using existing applications, or
- simulating the use of the data theme or
- using the corporate standard Discovery & Delivery Tools.

The testing will require a written test plan which is validated against the agreed system requirements. Where possible, tests should be automated so they can be re-used during the Sustain phase.

Develop the Decommissioning Strategy

The SPOT planning and construction may identify systems and applications that will become obsolete. Leaving obsolete systems may inhibit acceptance and use of the SPOT and may cause the wrong data to be used or supplied to clients.

Decommissioning obsolete systems ensure that users are accessing the right data sources and will have confidence in the data they are using.

This task is both essential and to be done with extreme care. The testing regime developed in the previous step should be re-run after the decommissioning to validate that no links to previous data systems remain.

Handover the SPOT

The handover signals the completion of the Construct stage, and the release of the SPOT into Stage 4: Sustain.

The Custodian must agree to the handover, validate governance arrangements and have appropriate supporting documentation and overviews provided. If there is a change of roles for people involved in Governance, then training or support must be considered.

Promote the SPOT

Considerable resources will have been invested in establishing the data theme as a SPOT so that internal users and external clients have access to data they can be confident is authentic, authoritative and accurate. To maximise the value to the agency and the community of the investment in the SPOT, the promotion strategy should seek to ensure that the widest range of potential clients is informed about the SPOT and that any opportunities to promote the achievement of its implementation in professional and technical forums are taken.

Post Implementation Review

A Post Implementation Review is an opportunity to glean information from the project which will be useful for future projects. There are many ways the review can be done, both formal and informal. The simplest approach is a workshop for couple of hours where all who took part in the project can discuss openly and freely what was learnt. A summary of the discussions should be included in the project file, with the emphasis on what can be done to improve SPOT projects in future. Some questions to consider are:

- Did the SPOT methodology work? How can it be improved?
- Did stakeholder input provide sufficient information to develop the “Gather Requirements” stage plan?
- Were the conceptual and logical models valid and effective?
- Were the resource estimates accurate? If not what lessons have been learnt?
- What compromises were made in developing and implementing the physical model that might diminish the future effectiveness of the SPOT?
- Was the Testing effective and useful?
- Is the data being delivered to clients really meeting the expectations of the SPOT concept?
- Did the SPOT project make a real difference and did it represent a “value for money” investment of the agency’s resources?

These are indicative and the organiser of the Post Implementation Review should consider additional questions relevant to the SPOT. The review should be a “warts and all” discussion focussing on how the agency can improve its processes in future. While focussing on fault finding and blaming individuals is

never useful, open discussion of problems and limitations are an essential characteristic of high performance projects.

Having done the review, the organiser is responsible to ensure that the learning from the review is communicated to the Organisation. Therefore the summary of the outcomes should be sent to the Executive Sponsor and the Chief Information Officer.

Outputs

The following will be produced from this stage:

- Released SPOT
- Supporting Documentation
- Transfer Scripts
- Initial Collection of metrics.
- Decommissioning report
- Post Implementation Review evaluating the outcome against this methodology.

To Complete This Stage

The Custodian will ensure that what has been delivered meets agreed user requirements. The Custodian will provide the Project Manager with a written acceptance of the SPOT implementation project.

Resource / Skills

The following resources or skill-sets are required for this stage:

- Project Manager
- Developer/s
- Systems Architect
- Stakeholders or subject matter experts, for testing
- Project Sponsor
- Custodian
- Data Managers

Construct Stage Checklist

- Does the SPOT meet the agreed user requirements?
- Is the SPOT consistent with the identified international standards?
- If a database, is the distribution structure adhering to defined Guidelines?
- Is the SPOT data discoverable using a data discovery tool?
- Have you decommissioned any redundant systems?
- Have you identified if there are support projects or tasks needed to complete this stage?
- Testing:
    - Are testing simulations needed?
    - Does the SPOT require volume or stress testing?
    - Have written test plans been produced?
    - Does the test plan reflect user requirements or use cases?
    - Will the critical path be tested?
    - Is testing of external interfaces required?
    - Has the user acceptance testing included representatives from all GA and have you considered using external stakeholders?
    - Have the external stakeholders verified the modified data?
- Handover:
    - Has the SPOT met the acceptance criteria?
    - Has the Custodian accepted the SPOT?
    - Have the Data Managers accepted the SPOT?
    - Has the job impact been assessed?
    - Have external stakeholders been notified?
    - Is further communication needed?
    - Has a process been established to monitor changes to any relevant standard?
- Documentation:
    - Have the user guides been updated or written?
    - Has the system administrator's guide been updated or written?
    - Have you documented the QA/QC testing procedures
    - Have you generated a SPOT lessons learnt document or are updates required to the methodology?

Objective

After the SPOT has been constructed and handed over, this stage sustains the SPOT data theme as a dynamic element of the Information Infrastructure. The Custodian should review the SPOT regularly, at least annually, to ensure that it continues to meet business needs and that the governance model remains appropriate.

If a review assesses that the data theme no longer meets the SPOT standard, then corrective action is required, and the SPOT cycle may be re-initiated.

Description

The following occur in this stage:

- Operating the SPOT
- Assessing the SPOT
- Verifying the governance
- Initiating corrective action.

Operating the SPOT

Operating the SPOT sustains the business processes to continuously deliver to our clients, giving them access to valuable business assets. This requires active management and custodianship to ensure that the SPOT data remains current, meets defined quality standards and adapts over time as technology, users requirements and expectations evolve.

Assessing the SPOT

The Custodian is responsible for the governance and assesses if the SPOT data theme continues to meet the SPOT standard using the metrics and other resources. Included in this assessment are whether the SPOT continues to meet stakeholder requirements and calculating the cost of maintaining the SPOT.

Verifying the governance

Under Geoscience Australia's Chief Executive Instruction for Information Management each Group Leader is responsible for an identified set of data themes. The Group Leader must ensure that each data theme has an identified Custodian.

If there is no Custodian or Data Managers for the data theme, then the responsible Group Leader needs to take action.

Governance policies include the update frequency of the distribution, service level agreements, support mechanisms and notification of where the relevant standards have been modified.

Initiating Corrective Action

If the review finds that the data doesn't conform to the SPOT standard then corrective action must be taken. This action may require re-generating the

SPOT data theme by following the stages in the SPOT methodology or by decommissioning the data.

Outputs

The following items will be produced from this stage:

- SPOT Data System continuously delivered to clients
- Reporting of ongoing performance and cost
- SPOT Review completed

To Complete This Stage

Regular assessments for the SPOT data theme will be approved by:

- Group Leader.

Resource / Skills

The following resources or skill-sets are required for this stage:

- Custodian
- Subject Matter Experts
- SPOT Validation person

Sustain Stage Checklist

- Is the custodian role filled and active?
- Are development resources available if required?
- Is the data still considered the “dataset of choice”?
- Are the available distribution formats still suitable?
- Is the metadata accessible and current in a corporate system?
- Is the governance process still current?
- Is the system accessible through the corporate discovery tools?
- Are the metrics still available?
- Do the metrics help in the assessment of the SPOT?
- Does the SPOT still comply with the relevant standards?
- Does the SPOT still meet user requirements and expectations? Is it feasible to modify or regenerate the SPOT?
- Is the underlying technology capable of continuing to support the SPOT?

Appendix 1 - Roles and responsibilities

Role	Responsibility
Executive Sponsor	A senior executive, in GA typically a Group Leader or Chief of Division, whose area of responsibility includes the data associated with the SPOT and who will support establishing and sustaining the SPOT at senior executive level.
Business Analyst	A person who follows a structured methodology to understand business needs for a data system, identifies how best to meet those needs and designs the set of processes to meet those needs.
Custodian	A Data Custodian has the accountable responsibility for managing and maintaining the data theme. All SPOT datasets must have a Custodian. The Custodian is responsible for ensuring the data adheres to the requirements of the SPOT standard for that data.
Data Manager	A person with specific responsibility for managing and updating the data theme and the specific systems that support its discovery and delivery.
Technical Reference Group (TRG)	The Technical Reference Group is a group of key stakeholders who have a strong interest in the implementation of a SPOT data theme. The Technical Reference Group provides technical or scientific advice to the project team.
Subject Matter Expert (SME)	A person with data theme expertise.
Software Developer	A person who does tasks in the software development lifecycle including design, coding, testing and maintenance.
Information / Data Modeller	A person who can model and analyse data to represent it as an abstract set of relationships and flows.

Appendix 2 - Glossary

.

Term	Description
ANZLIC	http://www.anzlic.org.au
Balanced Metrics	Performance measures that look at a system from many perspectives rather than one. Perspectives include financial, customer, business, and learning and growth.
CDCS	Corporate Data Classification Scheme. A taxonomy of geoscience and geospatial data used in Geoscience Australia to help organise its data and information.
CDM	Corporate Data Model – an Oracle Schema used in Geoscience Australia to build databases with a consistent and interoperable structure.
CDS	Corporate Data Store – A hierarchical file structure in Geoscience Australia to store large volume data. The CDS structure follows the CDCS.
CIO	Chief Information Officer
Conceptual Data Model	A high level data model which can be represented by words and pictures to describe the major elements of a data them or data system.
Data System	A set of data types linked through a value chain to collect, transform, manage and deliver data.
Fitness for purpose	The fitness of the data for a particular purpose will depend on that purpose, and responsibility for that assessment will lie with the user. A basic responsibility on the part of the user will be to consult with the provider on fitness for purpose.
Geomet	A repository within GA that manages metadata to the ANZLIC V2 standard
Logical Data Model	A representation of information relationships. It represents all elements of a data system and shows flows, transformations and dependencies, but avoids technology specific aspects.
Physical Data Model	Representation of a Logical Data Model in a specific set of technologies.
Stakeholder	Persons who have an interest in a data system as sources, users, value adders or who gain business benefit in some significant way from the data system. Stakeholders are: • Key: Those whose interests must be met. • Non-key: Those whose interests should be considered but may be met in a limited way.
UML	Unified Modelling Language. A widely used non-proprietary modelling language.

Output	Anything produced by a project which will be preserved after the project is completed.
Product	Any output or combination of outputs which is transferred to clients outside the agency. This includes Packaged Products and User Selectable Products.
Packaged Product	A product that is presented to a client as a discrete entity, such as a map, a DVD or a fixed file download.
User Selectable Product	A product, generally available as a web system, which allows a user to select a portion of data from one or more data sets.
Information Life Cycle	The changing value of information over time. Value is usually an estimated future value.
User Requirements	User Requirements record what must be delivered by the SPOT data system for it to be an effective and efficient business output.
Discovery and Delivery toolset	A series of applications and tools available to GA staff that facilitates the discovery and delivery of GA's information and data.

Appendix 3 – SPOT Stakeholder Questionnaire

SPOT Stakeholder Questionnaire

Current and Future Interest

This template questionnaire is used during the *Initiate* stage of the SPOT methodology to record the interest of stakeholders in the proposed SPOT and to gain an initial view of the nature and extent of the interest. This information will be used to develop the Project Scope and to develop acceptance criteria for the proposed SPOT. More detailed information may be sought in a later phase if it is required.

The questionnaire should be tailored for the specific needs of the data system.

SPOT Project	
Date	
Stakeholder	
Interest	*Summarise the interest of this stakeholder in the data system for which a SPOT is proposed*
Contact Details	

1. Business Requirements

Requirement	Brief Details
Business Driver	*What is your connection with this data?* *Supplier, processor, translator, user, value adder etc*
Performance and Access Requirements	*What availability expectations or requirements do you have?* *What currency expectations or requirements do you have?* *What distribution expectations or requirements do you have?*
Governance	*What custodianship responsibilities to you have?*

2. Functional Requirements – Current

Requirement	Brief Details
Relevance	How significant is this data to your business? Critical / helpful / limited / background
Standards	What standards and formats do you need the data to meet?
Access	How frequently do you access the data? Is this constant of variable?
Volumes	What quantities of data to you access? What channel do you use to access the data?
Applications	What software applications or tools do you use to access this data?
Satisfaction	Overall are you satisfied with this data at present? Very / moderately / neutral / dissatisfied / very dissatisfied
Strengths	What are the satisfactory aspects of the existing data and associated systems?
Weaknesses	What about the existing data and systems needs improving?

3. Future Requirements

Consider answers to these questions for time frames of 1, 3 and 5 years

Requirement	Brief Details
Relevance	Do you expect the importance of the data to change? *Much less / less / same / more / much more*
Standards	What standards and formats do you expect to need the data to meet?
Access	How frequently do you expect access the data? Is this constant of variable?
Volumes	What quantities of data to you expect access?

	What channel do you use to expect to use to access the data?
Applications	What software applications or tools do you expect to use to access this data?

4. Success Criteria

What criteria you would use to be fully satisfied with this data as a service**?**

Feature	**Attribute Preference**
Quality	
Timeliness	
Accessibility	
Standards	
Other	

Appendix 4:SPOT User Requirements Questionnaire

User Requirements Questionnaire

Current and Future Requirements

This template questionnaire is used to collect User Requirements for a proposed SPOT data set in the *Gather Requirements* stage of the SPOT methodology. This information will be used to model the SPOT data system and to develop an implementation plan for the proposed SPOT. This information should be sufficiently detailed to enable planning of the maintenance and distribution environments for the SPOT. Information from the previous stage should be brought into this questionnaire if it is available.

The questionnaire should be tailored for the specific needs of the data system.

SPOT Project	
Date	
Name	
Contact Information	

1. Present Business Requirements

Requirement	Details
Business Driver	What do you use this data for? *Background, Informing policy, Value added products* How important is this data to you? *Useful background, supporting data, critical* What alternative sources do you have for this kind of data?
Client Requirements	How often and how rapidly to you need to access the data? *Real time; daily; weekly; monthly; annually* *How up to date do you need the data to be?* *Real time; daily; weekly; monthly; annually, best available* *What data formats do you want to receive?* Standards, application formats, flat files etc.

	How do you want to select the data? *All data; database query; web mapping tool* How do you want to receive the data? *Machine to machine, Web page, download, email, CD/DVD, tape* What applications and tools to you use for this type of data? *Specific applications should be recorded* Are your needs being met adequately? *Satisfied with existing arrangements or in what way not?*
Governance	*What do you want to know about how the data are managed?* Access to documentation, Access to metadata, Access to custodian

2. Future Requirements

Requirement	Brief Details
Mode of Access	How do you expect to access these data in the future? *Machine to machine, web page, download, email, CD/DVD, tape*
Growth	Do you expect to increase the quantities of this data you use or the frequency with which you access it? *How much and when*
Formats	Do you expect to change the format in which you receive data? Do you expect the existing formats to continue to be available?
Quality	Do you expect your requirements for data quality to change? *precision, accuracy, resolution, metadata*

Appendix 5: SPOT Data Holdings Audit

Data Holdings Audit

Current/Existing Status

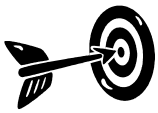
This audit is sent out to all data managers and custodians to identify datasets and their condition. The template may be modified to gather information relevant to a specific data system.

SPOT Project	
Date of Audit	
Item	**Brief Details**
Dataset Name	*Fill in this form for every dataset repository for which you can provide information.*
Brief Description	
Geographic Extent	*Eg. Australia-wide*
Custodian	*May be more than one custodian (externally sourced data)*
Data Managers and administrators	
Who uses this dataset? (Clients/Users)	*What do they use it for?* *How do they use it?* *Do current procedures meet their requirements/expectations?*
How do users access this data?	
How do you access this data set?	*Eg. URL, database name, form, report*
When was the data set last updated?	*Continually, weekly etc*

How is the dataset updated?	
Are there documents for management and maintenance of the data set?	
Where is the data set stored?	*Eg. Oracle/CARIS/Binary/Access db/SDE/ etc*
Dataset Size & Rate of Growth	*Eg. 500Gb of text files, Estimated growth 80 Gb per year*
Is there a written Data Model?	*If there are Entity-Relationship or UML models, please attach or provide links. How current is it?*
What standards apply to the data?	*International/National/Agency, ISO etc* *Please describe standards for structure (tables) or content (data/lookups)*
Workload of data set	*How much effort per year by how many people is required to maintain and deliver the data set in its present condition?*
Availability or access requirements	*Eg. Need 99% availability, data can only be viewed in a secure room*
IP information	Unknown or undocumented (Red) ☐ Not owned by GA but known and documented (Amber) ☐ Details: _____ Owned by GA with confidentiality restrictions (Amber) ☐ Details: _____ Owned by GA and publicly accessible Open (Green) ☐
What metadata is stored and where?	*Eg. URL, Is it in Geomet?* *If no metadata is stored, what metadata attributes are required/desired?* *Consider dataset and feature level metadata.*

What metadata standards apply?	*Eg. ANZLIC, ASDD, Geomet* *What Geomet standards apply?*
Discoverability	*What mechanisms /registries are used to enable discoverability?*
Other Items of Interest **(eg. Scale, Area of Coverage, Other useful information etc)**	
Who to ask for more information	
Name	
Position	
Division	
Contact Number	
Alternative Contact	

Appendix 6: "Is it a SPOT" Checklist

Is it a SPOT? 

Instructions:

- This checklist is designed to assist the SPOT Project Leader to assess if a dataset is a SPOT.
- All requirements must be checked unless stated otherwise.

SPOT Name: _____

Date: _____

Person: _____

Division/Group: _____

Reviewed by: _____

Review Date: _____

Is the dataset unique within GA?

☐ Dataset is unique with GA

☐ Business Case documented; TRIM reference: _____

☐ Governance documented; TRIM reference: _____

- Accountable Group Leader identified & recorded in metadata and Individual Work Plans (IWP)
- Accountable custodian identified & recorded in metadata and IWP
- Accountable data manager/s identified & recorded in metadata and IWP

Master source

☐ Master source is maintained and up to date

Metadata

☐ In corporate metadata system

☐ Metadata is maintained and up-to-date

Is it discoverable internally & externally via the corporate D&D application?

☐ In text based corporate catalogue tool

☐ In map based corporate catalogue tool

☐ Distribution formats and method of discovery documented; TRIM reference: _____

Distribution data

☐ Distribution process is documented and data is up-to-date; TRIM reference: _____

Where is the distribution data? (check all that apply)

☐ Oracle Distribution environment (DISTPROD)

☐ Corporate Data Store (CDS)

Is the distribution data available in common formats? (check all that apply)

- ☐ Ascii File (eg. csv, txt, xyz)
- ☐ Binary File, Please specify _____
- ☐ Oracle Spatial
- ☐ ESRI ArcSDE
- ☐ Other, Please specify _____

Does the distribution data comply with

- **OGC and interoperability standards?**

Mandatory

☐ WMS (Web Map Server)

Optional

☐ If vector data, available as WFS (Web Feature Service)

☐ If raster data, available as WCS (Web Coverage Service)

- **Content Standards?**

☐ Common Standards (eg. GML, Observations & Measurements)

☐ Other Content Standards (eg. GeoSciML, ASEG, POSC)

Where GA is originator of the data or adds value	**Data is externally sourced and maintained (GA has the right to hold/use)**
Does the master source reside in a corporate system? (check all that apply)	
☐ Oracle Production environment (ORAPROD) ☐ ARC/SDE Enterprise Geodatabase ☐ If file based in a corporately supported file system, provide location: _____	☐ Dataset's sourcing process is documented and discoverable? TRIM Reference: _____
Is the data secure?	
☐ Security in place in Master Source (maintenance) ☐ Security in place in Distribution ☐ Dataset's Intellectual Property/Licence requirements documented; TRIM reference: _____	☐ Appropriate security in place in Distribution ☐ Dataset's Intellectual Property/Licence requirements documented; TRIM reference: _____
Is it documented?	
☐ Project Scope; TRIM reference: _____ ☐ Requirements (Business rules & functionality); TRIM reference: _____ • Up-to-date Stakeholder requirements • Logical Data models • Workflow processes (business process mapping) • QA/QC procedures & tests ☐ System architecture (business/technology/ application); TRIM reference: _____ ☐ SPOT Intranet page	☐ Requirements (Business rules & functionality); TRIM reference: _____ • Workflow processes (business process mapping)