DATA CUSTODIANSHIP GUIDELINES

FOR THE

GOVERNMENT OF BRITISH COLUMBIA

Version 1.0
September 30, 2015
Province of British Columbia

Document History

Version	Description of Change, Review or Approval	Date
0.1	Initial draft to remove extraneous material and create basic structure.	November, 2012
0.11	Updated diagrams and references. Included reference to DataBC Council. Reviewed, edited and updated several subsections. Inserted comments to identify where further work is required.	December, 2012
0.12	Added illustrative example and tidied up some sections.	January 11, 2013
0.13	Edits to reflect review meeting on 15th Jan 2013	January 31, 2013
0.14	Edits to reflect review meeting on 6th Feb 2013	February 14, 2013
0.21	Edits to reflect review meetings in Jan, Feb and March 2013	March, 2013
0.3	Edits to incorporate DataBC Working Group feedback	June 12, 2013
0.4	Updated references; updated license references	June 24, 2013
0.5	Final comments from WG incorporated	June 25, 2013
0.6	Draft document prepared for posting – references to Security Policy added	March 21, 2014
1.0	Final edits – OCIO - clarification on data management roles versus positions.	September 28, 2015

History:

1. The Data Custodianship Guidelines document is modelled with permission after a similar body of work by ANZLIC – the Australian and New Zealand Land Information Council.
2. In 2008, a set of Data Custodianship Guidelines for the Natural Resource Sector was produced under the Natural Resource Sector Information Council and signed off by the Government CIO and Council Members. This document was the foundation for the development of the Government-wide version of the Guidelines.
3. The Guidelines are complimentary to Government Core Policy and Procedures manual - data management policy and augment the Guidelines for Best Practices in Data Management – Roles and Responsibilities, March 2012.

Sponsor

- DataBC Council

DataBC Working Group Custodianship Task Team:

- Glenna Boughton, Lead Data Architect, Corporate Services for the Natural Resource Sector
- Elaine Dawson, Director, DataBC Enterprise Data Services, Technology, Innovation and Citizens' Services
- David Wrate, Director, DataBC Engagement, Technology, Innovation and Citizens' Services
- Charito Elderfield, Manager, Business Intelligence, Education
- Stephen Gidden, Director, Justice Business Intelligence, Justice
- Bonnie Laine Farrell, Director, Client Business Services & Applications, Finance
- Annika Livingston, Senior Spatial DA, DataBC, Technology, Innovation and Citizens' Services
- Kevin Metcalfe, Senior Spatial DA, DataBC, Technology, Innovation and Citizens' Services
- Tim Salkeld, Team Lead, Data Management, Forests, Lands and Natural Resource Operations
- Karen Samuelson, Manager, Mineral Titles, Energy and Mines
- Nainesh Agarwal, Enterprise Architect, Transportation & Infrastructure
- Per Wallenius, Client Business Consultant, Economy Sector, JTST and CSCD
- Tony Baker, Consultant

Document Approvals

Approved by the DataBC Working Group	June 26, 2013 Date
Endorsed by DataBC Council	August 1, 2013 Date
Tabled at the OCIO Architecture Standards and Review Board	December 12, 2013 Date
OCIO Signoff	
	Date
	Date
	Date
	Date

Table of Contents

1 THIS DOCUMENT

1.1 Purpose & Scope

The Data Custodianship function is a vital part of good data governance, and is required by Government policy[1]. These guidelines describe the data custodianship function, the roles and related accountabilities required to manage data, and explain its role in the data lifecycle, relating it to other important data management roles. The purpose of the document is to both inform and guide so as to promote good data custodianship.

While the document introduces several critical roles related to data custodianship, it by no means is advocating that ministries must hire a new set of data management professionals to fulfill the tasks and accountabilities outlined below. Rather, current employees are more than likely already performing these roles and DataBC sees adopting the guidelines as simply formalizing the work that is already being done. In addition, the guidelines may expose gaps in a ministry's data management program but again, the thinking is that these accountabilities will be assigned to those employees already familiar with the set(s) of data in question.

The DataBC Council, as the body responsible for managing data custodianship, has issued these guidelines.

The guidelines apply to custodianship of all sets of data held by the Government of British Columbia. The guidelines conform to applicable government policies laid out in Chapter 12 of the Core Policy and Procedures Manual.

1.2 Audience

The intended audience for the document is:

- Existing and prospective Data Custodians in the BC Government
- Other people who interact with Data Custodians in various capacities – e.g., Data Stewards, Data Standards Managers, Data Resource Managers, Discipline Authorities, and data users
- Those concerned with information management at all levels, including staff from ministries, the Government CIO, and ministry CIOs.

1.3 Owner

These guidelines are published and maintained by DataBC on behalf of the DataBC Council.

[1] Data and corresponding information systems must have an identified Data Custodian. *Core Policy and Procedures Manual – Chapter 12*

2 CONTEXT FOR DATA CUSTODIANSHIP

2.1 The Need for Data Custodianship

A Data Custodian is the person who

- protects and promotes the use of data holdings under his or her care;
- sets policies, and is accountable for defining the appropriate use of the data;
- is provider of the authoritative version of the data;
- is ultimately accountable for issues related to definition, collection, management and authorised use of the data.

The Government of BC, its business partners and the citizens of BC all depend on information. Health care, education, justice, mining, farming, forestry, transport, tourism, planning, all make use of information. But the value of information depends on knowing its provenance – its accuracy, completeness, currency, and how it is managed. Information is a fragile resource. If it's not properly managed it can be reduced to meaningless numbers that lack any value.

Data custodianship plays a vital role in the management equation. The Government manages hundreds of sets of data on behalf of the people of BC. It has a responsibility to preserve the highest value for the information under its care.

It is clearly neither possible nor desirable for a single group to manage all information in Government. However, it is possible and highly desirable for data authorities to follow common practices of good governance and process.

This coordinated data management model depends on the role of *Data Custodians* – senior managers responsible for collecting and maintaining data on behalf of its owners and, where appropriate, making this data readily and publicly available.

2.2 Data Custodianship Objectives

The objectives of data custodianship are to:

- Ensure consistency of data management practices so that goals for integrated data can more readily be achieved.
- Formalize the roles and accountabilities required to manage data holistically
- Maximize the value of investments in data collection and maintenance from a provincial perspective.

- Increase certainty regarding accountabilities for data.
- Minimize data duplication.
- Maximize the business benefit derived from data sharing and widespread use.

2.3 The Benefits of Data Custodianship

Custodianship is at the heart of information management because it establishes accountability for data, and identifies authoritative sources that give users a measure of consistency and certainty. In addition, custodianship is the foundation for:

- providing a trustee and standards bearer for data;
- eliminating unnecessary duplication in the collection and maintenance of data;
- managing data on behalf of the entire enterprise;
- providing a sound data infrastructure;
- assisting the production and management of data products; and
- facilitating the collection of data.

Collectively, Data Custodians manage data as trustees in a partnership with national, provincial, regional and local providers and users to enable the integration of data for the benefit of the entire community.

Custodial activities – including negotiations with other agencies and users and development of data products – must take place for the betterment of the whole community rather than any individual agency. The overriding philosophy in all these activities should be one of a trustee acting in partnership for all participants. Custodianship reinforces the concept of one senior individual being responsible and accountable for the data that others might use. This gives users confidence in the level of integrity, timeliness, precision and completeness of data, and in the quality and soundness of decisions made based on that data.

2.4 Data Custodianship Principles

In pursuing these objectives, the following six principles provide guidance for assessing the appropriateness of possible actions.[2]

2.4.1 Principle 1 – Data Custodian is Corporate Trustee

Data Custodians operate as trustees on behalf of the entire community of data users throughout the Province.

Data Custodians act in the interests of all users of the data for which they are accountable, irrespective of who owns the data. Thus, data custodial activity takes place for the betterment of

[2] The principles and definitions of roles in this document are consistent with Best Practices in Data Management – Roles and Responsibilities, Data Architecture Advisory Committee, Information Architecture & Standards Branch, OCIO.

the whole community, rather than any one agency, including the Data Custodian's own agency. The result is an emphasis upon collaboration, teamwork, cooperation and sharing.

2.4.2 Principle 2 – Data Custodian is Standards Bearer

Data Custodians ensure the development and enforcement of standards for data within their care.

It is important to set standards for the description, format, structure, classification, collection, accuracy, consistency, quality, access and retention of data and associated metadata. Data Custodians must ensure that appropriate standards are set, taking into account the needs of users, and that these standards are then followed.

2.4.3 Principle 3 – Data Custodian is the Authoritative Source for the Province

Data Custodians are the authoritative source for data within their care.

This principle serves to lessen confusion for users and overcomes accuracy and reliability problems that may be encountered when supposedly identical data is held separately by several agencies; where several agencies contribute data to a common database, or where data provided by different agencies is combined.

2.4.4 Principle 4 – Data Custodian is Accountable

Data Custodians are accountable for the data within their care.

Data Custodian accountabilities for the data under their care include:

- determining the information required to meet business goals;
- establishing and maintaining the standards and rules for that data;
- managing the data as a vital resource;
- identifying responsibility for data integrity, security[3], privacy, and quality; and
- ensuring that conflicts in business needs are successfully resolved.

Although a Data Custodian may delegate any of their custodial responsibilities to one or more Data Stewards (see Section 3.2), their *accountability* for the data is retained by the Data Custodian.

Custodianship accountabilities are only extinguished when all retirement obligations of the Data Custodian have been adequately fulfilled (especially concerning records retention, destruction and archive).

2.4.5 Principle 5 – Data Custodian Ensures Availability

Data Custodians will ensure that data are made available to qualified users.

[3] Additional obligations with regards to Information Security are established in the Office of the Chief Information Officer, Information Security Policy. http://www.cio.gov.bc.ca/local/cio/informationsecurity/policy/isp.pdf

To derive the maximum benefit from data through data sharing, Data Custodians will ensure that data within their care are made available and are accessible to all authorized users. Consistent with BC's Open Data policy, this includes making data accessible to the public where possible.

2.4.6 Principle 6 – One and Only One Data Custodian

Each set of data[4] *has a single, designated Data Custodian, without exception.*

Data should always have a designated official Data Custodian. Absence of data custodianship implies absence of accountability, and therefore absence of such things as assurance of quality, integrity, availability and relevance of underlying data.

There can be no more than one Data Custodian for any given data. If data had more than one Data Custodian, it would admit the possibility of more than one source, multiple standards, and unclear accountability for the data.

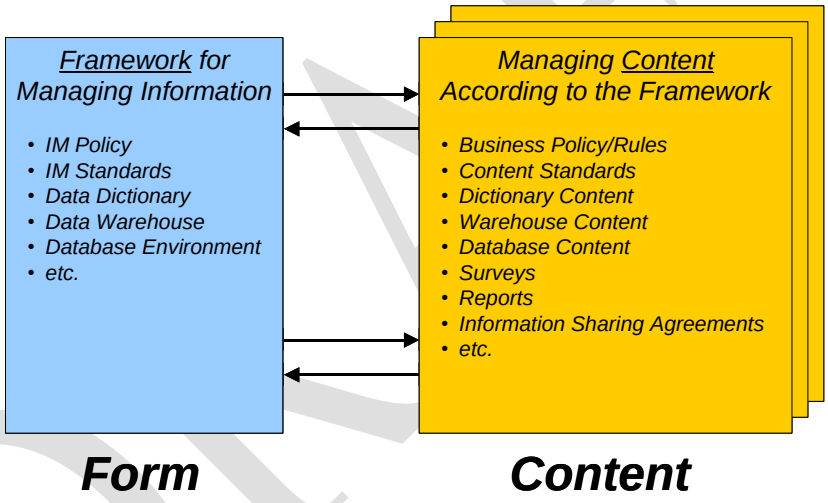
[4] *Set of data* has a specific meaning in this document. It is defined as a discrete corporate information subject that is of a lasting nature, about which relevant data is collected, managed, and used to serve an essential defined business purpose for government. Some may also use the term "corporate data" for similar meaning in the context of a ministry's business (i.e. ministry corporate data) or in the context of government-wide standards (i.e. government corporate data).

3 MANAGEMENT FRAMEWORK

3.1 Accountability Framework

In establishing accountabilities for data, two key accountability principles may be applied:

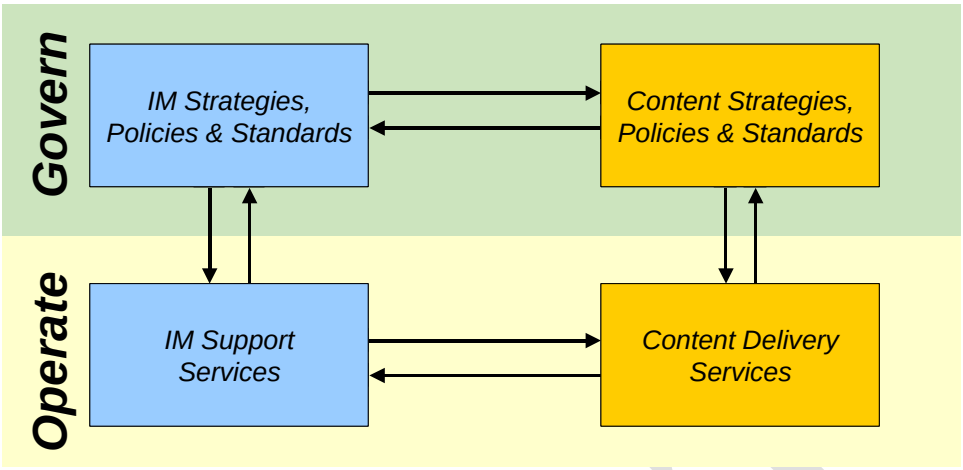
1. ***Separation of Form and Content***:
    a. "Form" refers to the model or framework for managing data. This framework should be established before putting in place the activities for managing a particular set of data. Examples of form-related accountabilities are laws, regulations, IM polices and standards.
    b. "Content" refers to accountabilities that are particular to a set of data, its business environment, its day-to-day management and its use.[5]



2. ***Separation of Governance and Operations***: Within each of the two domains of form and content, there is a further need to distinguish between accountabilities for governance and delivery.[6]

[5] The "water pipeline" analogy. If you were to build and operate a water pipeline, before any water flowed you would first have to make sure you understood and respected applicable laws for the construction and location of the pipeline. You would have to agree standards for pipe diameter, materials used, etc. These are "form" accountabilities. Once the pipeline is build and the water begins to flow, you have to consider water quality, safe flow rates, spill response procedures, and so on. These are "content" accountabilities.

[6] In our analogy, "governance" accountabilities address regulation and standards both of the pipe and the water flowing through it. They are the basis for establishing whether the enterprise is being well run. "Operation" accountabilities are concerns with the day-to-day management and maintenance of the pipe (the form aspect) and the water running through it (the content aspect). Note that even if the water was turned off for a period, the pipe would still need to be maintained.



3.2 Definition of Roles

The next section introduces several critical roles and related accountabilities associated with good data governance. The roles are extensive and the accountabilities comprehensive to ensure that all facets of the data management lifecycle are addressed and there are no breaks in the custodial chain. This information should help ministries to:

- Become informed on the roles and accountabilities required to holistically manage their data
- Identify gaps in their current data management program
- Formally recognize those employees associated with these roles and related accountabilities
- Build out their data management capabilities based on a government endorsed framework

3.2.1 Form-related Roles

Government Chief Information Officer

- Accountable for Government-wide data management frameworks, standards and infrastructures.
- Accountable for ensuring the appropriate standards, structure, content, care, use and disposal of shared or common sets of data.

Ministry Chief Information Officer

- A Ministry CIO (MCIO) may set policies and standards that are appropriate for the Ministry to supplement those established by the GCIO.

Data Administrator (Data Architect)

- Ministry-wide role providing data management leadership, data modeling expertise, and custodianship of the enterprise data models. Provides and promotes the framework for consistency in scope, meaning, and handling of data across the entire organization. Manages the ministry's corporate metadata to support the organization's data related goals and objectives, ensuring timely, accurate, and sharable data across diverse program areas.
- The Data Administrator position is usually at the senior technical specialist level.
- ***Related Roles***: Other roles involved with "form" at the operational level include *Database Administrators* and *Application Architects*.

3.2.2 Standards-related Roles[7]

These roles have both form and content accountabilities.

Data Custodian

- The person accountable for operational policy, definitions, rules, standards, structure, content, use and disposal for data under their responsibility.
- Custodial authority is typically either "statutory authority" as defined in legislation or "administrative authority" as defined in policy (or similar source of mandate).
- Core Data Custodian responsibilities: *Local Leadership*; *Data Standards*; *Consultation*; *Assessing Compliance*; *Business Training and Support*; and *Information and Reporting*.

Data Standards Manager

- Person who develops and sets data standards that reflect the data policies established by the Data Custodian.
- Responsible for the day-to-day management of the data according to the defined data standards and data management plan. Examples of duties include authoring the data management plan, checking for data standards compliance, resolving issues, advising other roles – particularly Data Resource Managers.
- There is one Data Standards Manager for a set of data, and that person is appointed by the Data Custodian.
- ***Variations***: Some ministries may use the term *Data Manager* to refer to this role.

[7] Again, there is no expectation that a ministry must hire new employees to fill these roles. Rather, the guidelines are a framework for ministries to use as a tool for identifying a) employees already performing these roles, or b) identifying gaps in their data management program, and then formalizing the related custodial tasks and accountabilities.

Discipline Authority

- A business expert or specialist who understands the business relevance of the data standards within their scope of work. They must actively use their knowledge in support of the broad scope of business and established data standards.
- Interprets the meaning and appropriate use of detailed data standards to meet organizational needs.
- A primary resource for the Data Standards Manager, though not a direct reporting relationship. May be multiple Discipline Authorities (e.g. one per large geographic organizational area for a particular business mandate). A Discipline Authority for a particular subject area or set of data does not necessarily work in the same ministry as the Data Custodian – if they are business experts, they are a good resource for the Data Custodian to use to provide comments on data standards.

Data Steward

In some situations, Data Custodians may not have the appropriate operational or technical resources to meet their custodial responsibilities. Also, there may be resources that are better able than the Data Custodian to provide certain services (e.g., through aggregation of common needs across multiple Data Custodians).

In these situations, a *Data Steward* can be engaged by the Data Custodian to fulfill specific aspects of custodial responsibilities. Data Stewards provide a set of services on behalf of Data Custodians. These should be formally described in a Data Stewardship Agreement between the Steward and the Custodian. (Occasionally there may be more than one Data Steward if supporting services are provided by more than one organization. See Section 7.)

There may also be situations where legislation mandates a program area to deliver Data Stewardship services for one or more Data Custodians (e.g., the registration requirements of the Integrated Registry). Under these circumstances the engagement becomes obligatory.

Note however, that even though a Data Steward may assume some responsibilities on behalf of the Data Custodian, accountability for the data remains solely with the Data Custodian (in accordance with Principle 4 – Custodian is Accountable).

3.2.3 Content-related Roles[8]

The accountability framework and roles described are focused toward the creation and management of data. These can be extended to *data use*, as follows.

[8] Again, there is no expectation that a ministry must hire new employees to fill these roles. Rather, the guidelines are a framework for ministries to use as a tool for identifying a) employees already performing these roles, or b) identifying gaps in their data management program, and then formalizing the related custodial tasks and accountabilities.

Data Resource Manager

- A designated employee with responsibility for the collection and/or management of corporate data, to the standards set by the Data Custodian. Note that full accountability rests with the most senior person in each office (e.g. a district manager, or branch director) for ensuring corporate data collection and management is conducted to the defined standard.
- Includes responsibility for data collection and update to defined standards as well as the right to give operational input into the business design decision-making process.
- Note that anyone who uses or updates information in any form is accountable to a Data Resource Manager and thereby must ensure that activities are consistent with standards set by the Data Custodian.

Data Usage Contact

- A technical database resource or sophisticated business user who understands the business data and how it has been physically implemented. Manipulates and queries physical database content to support operational information needs.
- May define user views for repeated queries.
- A primary resource for Data Resource Managers and Data Users.
- Representative for a community of data users. (User communities will generally be defined by organizational boundaries or by professional roles.)

Data User

- A consumer of the data. Someone who uses the data to conduct analysis, make decisions or otherwise do their job.

Data Product Provider

- A person responsible for creating and publishing a data product that is derived from one or more sources of primary data. (See Section 4.3.)
- Accountable as a Data User for respecting the standards and policies set out by Data Custodians of the primary data.

3.3 Identification and Registration of Data Custodians

3.3.1 Identification

- As noted above, ministries receive their custodial authority for data from legislation or government policy.
- Core policy requires that Data Custodians be identified for ministry data. The Data Custodian for a given set of data can be identified by applying these criteria:

- Has sole statutory responsibility or other mandated accountability for creation, capture and/or maintenance of the set of data in the Province;
- Has the greatest operational need for the set of data;
- Is in the best position to establish standards and define business needs;
- Heads the office of record for changes to the set of data.

- Data Custodians are usually Branch directors or above.
- Once identified, the Data Custodian is registered and thereby assumes the rights and obligations set forth in this document.

3.3.2 Processes for Managing the Register of Data Custodians

- DataBC will maintain and publish an official register of data and Data Custodians for the BC Government.
- The Ministry should notify DataBC of additions, changes and deletions to the register.

3.4 Data Stewardship Agreements

A Data Stewardship Agreement is intended to formalize stewardship arrangements in which a Custodial agency has authority for all aspects of a given set of data, but desires assistance from a second agency (the Steward) to provide some or all of the associated services. The second agency generally does not have any particular mandate or authority to provide the services, but is willing and able to enter into an arrangement to do so.

Where appropriate, the Agreement also allows for the creation and distribution of interpretive or derivative product sets of data by the Data Steward, based on source data provided by the Data Custodian.

A Data Stewardship Agreement is not a Data Sharing Agreement. The latter is appropriate where a peer-to-peer relationship exists between two or more Custodians who wish to share data for some common purpose (see below).

If required, DataBC can provide a Data Stewardship Agreement template.

3.5 Data Sharing

3.5.1 Why Share Data?

Data Sharing refers to multiple organizations accessing, and possibly contributing to, a common pool of data, independent of applications and technology. The concept of a single source for any particular set of data is implied by this definition. The Data Custodian responsible for the data should consider the implications of data sharing when defining data standards and practices.

These are possible reasons for sharing data:

- The same source of data is used for business decisions. This guarantees consistency and accountability of data.
- Makes data management easier. Only one source needs to be managed – provides efficiencies with respect to security, backup and recovery, data administration, and so on. Thus, data management effort is reduced. (However, propagation of updates of the source data to users has to be considered; otherwise "stale" and inaccurate copies may proliferate.)
- Avoids incidences where multiple derived sets of data exist, each one managed in a different way, with perhaps inconsistent amalgamation with other sets of data, and unclear or incomplete metadata to describe the set of data. These together lead to poor decision-making that is difficult to support.
- Multiple uses of a set of data by many audiences encourages improved quality. The better the data, the more people benefit.

The Data Sharing Considerations

For a set of data that is to be shared, it's important to consider the following questions:

- What standards apply to the use of the data?
- Who is responsible for which aspects of data management, particularly version and currency management?
- What is the relationship between the parties sharing the data?
- Who is the copyright holder?
- Does the data contain intellectual property from a third-party?
- Is the data under an existing licence or agreement? (e.g. BC Open Government Licence; Statistics Canada licence, etc.)
- Is there a cost to obtain the data?

3.5.2 Data Sharing Agreements

Creating an Agreement

The questions outlined above highlight the need to formalize the relationship between the parties sharing data. This can be done by means of a data sharing agreement or license. These set out the terms of use of the data, including any restrictions, the obligations of the data user and provider and outline any financial or data exchange requirements.

An agreement is not necessarily a long or complicated document however it is important to address the rights and obligations and mitigate to the degree possible any risk in sharing the data among the parties involved. An example of a short but comprehensive license is the Open Government License – British Columbia. It addresses all rights, obligations and mitigates risk by identifying specific data which is not licensed.

A data sharing relationship with an industry group or one that that involves several types of data might require a more complex agreement. This may be a paper document that includes sections

for each aspect of the relationship. It allows the data sharing partners to express their requirements and expectations in an organised way. The parties sign the agreement once it's complete, and it becomes the basis for the data sharing relationship. If the relationship changes over time, then the agreement needs to be revisited and updated. Such an agreement should be formally cancelled when the data sharing relationship ends.

Here is an example of a basic data sharing agreement tailored for personal information. Agreements for non-personal data sharing would take a somewhat different form.

Types of Agreements or Licenses

All agreements or licenses are forms of contracts. Examples include:

- Open – free access to data with minimal restrictions.
- Priced – the user must pay some sort of charge for the data – e.g., so many dollars per table, report or mapsheet,
- Data exchange – the user provides some data in exchange for access to source data. Usually, this means the user provides data associated with their use of the source data (e.g., roads or survey data). The user-provided data is used to update the source data, thus improving its accuracy, currency, coverage or general utility. Such data exchange agreements support the concept of the public good of government-held data.

4 MANAGING & IMPLEMENTING THE DATA LIFECYCLE

4.1 The Data Lifecycle

4.1.1 Lifecycle Model

Figure 1 illustrates the data lifecycle. It represents the activities that are undertaken to manage a data collection throughout its lifecycle. In fact, it comprises two cycles:

1. The Governance and Standards Cycle, which represents activities relating to governance of the business relevance and architecture of the data. (Though not shown in the diagram, the Governance and Standards Cycle occurs within the context of the "Form" framework that's defined by broader CIO and data administration policies and standards. See Section 3.)

2. The Content and Usage Cycle represents activities relating to the management of the actual data as it passes from acquisition to consumption by users. The cycle is triggered out of the Governance and Standards Cycle and relates to the "Content" framework described in Section 3.

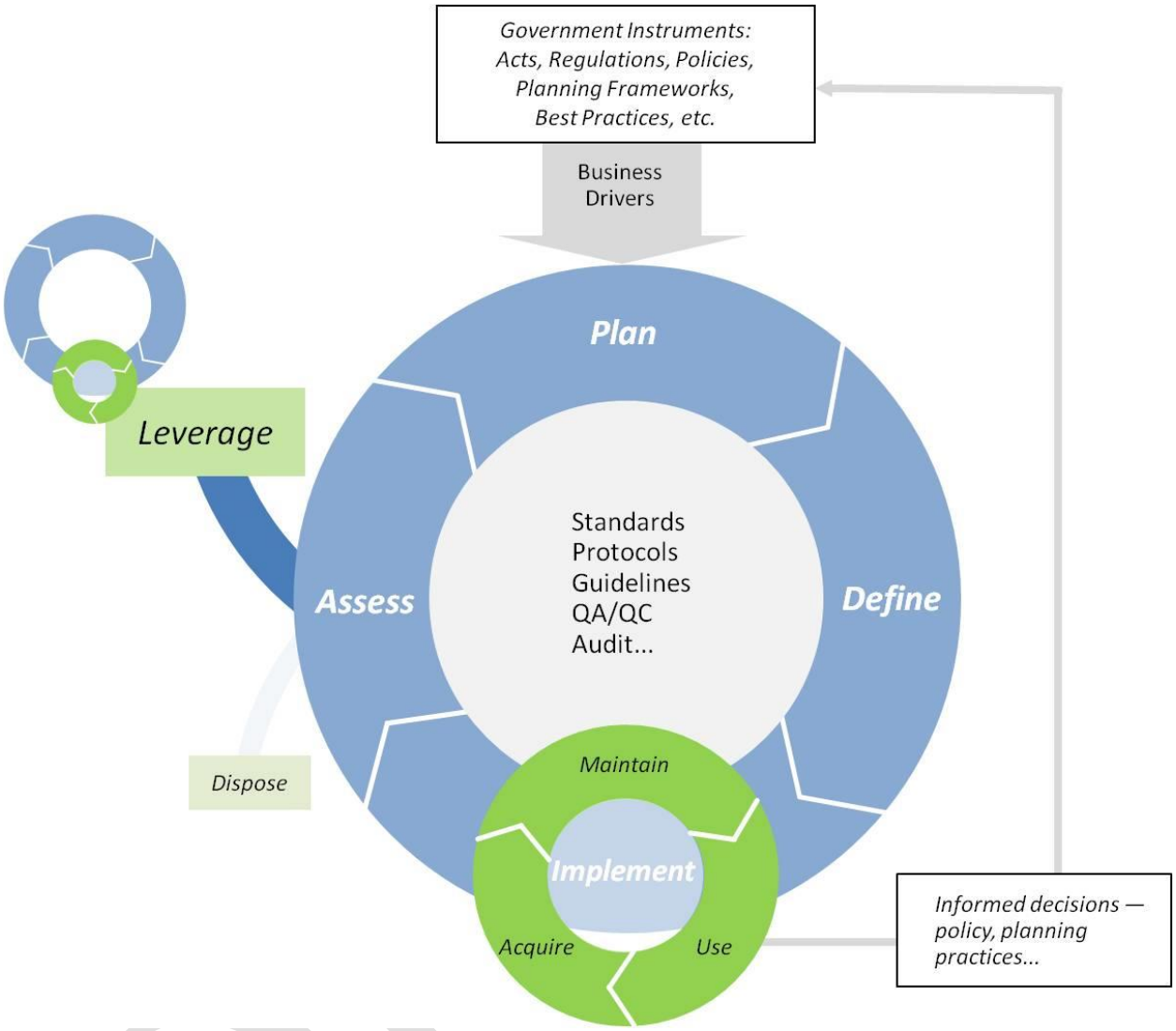


Figure 1. The Data Lifecycle

The activities shown in the diagram are described below.

4.1.2 Governance & Standards Cycle

1. **Plan** how the data will be managed maintained and provided for use. This establishes the framework for the complete Standards Cycle. The plan defines how all other activities will be conducted.
2. **Define** the contents and scope of the data collection, its structure (i.e., conceptual, logical and physical schema), and applicable policies, standards and processes.
3. **Implement** the plan and defined standards in order to create the business and technical environment for the data, which then allows the Content and Usage Cycle to be initiated (see below).
4. **Assess** the value and continued relevance of the data to its intended use. This is the part of the cycle where the custodian reviews how the data collection is being used and whether it is still meeting its business purpose. As a result of the assessment, there may be opportunities to change the scope, structure, acquisition, definition or use of the data. This activity feeds directly back into the planning process, which factors in the improvements in an updated data lifecycle plan.
5. **Leverage** existing uses. This is an open-ended activity whereby the Custodian, users and others seek to increase the value of the data collection, such as by linking it with other data collections, by changing its presentation, by applying post-processing to create data products, or by other means.
6. **Dispose** of the complete set of data finally. If the Data Custodian determines that collection, management and use of a set of data are no longer required, the Data Custodian may initiate retirement of the data by notification to the affected community and DataBC. Disposal must conform to applicable data retention and disposal policies. This causes both the Governance & Standards and Content & Usage cycles to close.

4.1.3 Content & Usage Cycle

A. **Acquire** the data. This is usually a continuous (or at least, regularly repeated) process for collecting new data.

B. **Maintain** (update) the data to keep it current and accurate by applying QA processes, integrating new data, purging old data, and making operational changes to the business and technical environments.

C. **Use** the data. The data are made available to users according to applicable policies and standards established by the Data Custodian.

4.2 Activities of the Governance & Standards Cycle

The table below is a checklist of activities associated with the Governance and Standards part of the data lifecycle. Most of these are related to Form, as described in Section 3. They are the responsibility of the Data Custodian, though in practice will fall to the Data Standards Manager who will generally have more hands-on management of the data. Their engagement is shown in the second column.

The third column describes any engagement by people in roles associated with Content.

Activity	Data Custodian / Data Standards Manager (Form-related)	Involvement by Other Stakeholders (Content-related)
1. Plan how the data will be managed, maintained and provided for use within the context of the strategic direction and business plans. This establishes the framework for the complete Standards Cycle. The plan defines how all other activities will be conducted.		
Define strategic direction	This should describe the strategic purpose of the data and goals associated with making it available to target end-users. It should support enterprise (government-wide) goals. It should be consistent with applicable policies and standards from GCIO and the Ministry CIO.	
Define business plan	This is a more tactical view that derives from the statement of strategic direction. It should describe how the data will be managed over its lifetime. It should be consistent with, and support, the Ministry's business plan.	
Consider all user needs	The custodian should consider the needs of all prospective user groups, but is not necessarily required to meet those needs. For example, the extent and disparate nature of those needs may make it impractical to meet them all for reasons of cost or other resource constraints.	End user representatives should be canvassed and encouraged to express their needs and be involved in the planning process. The Discipline Authority should also be consulted.
Decide whether Open Data or not	Based on the applicable policies and anticipated types of use, the Custodian should determine whether the data meets the requirements for distribution under the Open Government License – British Columbia. Consult DataBC for guidance. See also Open Information and Open Data Policy.	
Establish funding for the program	The custodian must secure adequate funding to execute the business plan. This requires establishing a budget and sourcing the funding internally, or by cost sharing with other parts of Government and/or with external user groups.	If considering charging users for the data (which, by definition, will not be Open Data), more extensive consultation is required with IPP, Treasury Board and others.

Activity	Data Custodian / Data Standards Manager (Form-related)	Involvement by Other Stakeholders (Content-related)
Prepare data management plan	This describes how the data will be managed as a resource and explains the day-to-day activities and roles required to provide the data.	The prospective Data Resource Manager should be consulted, and possibly given the responsibility to create, maintain and update the plan.
Establish Data Stewardship if required	The Data Custodian and Data Standards Manager must decide whether the data will be provisioned through a Data Steward, such as DataBC. If so, the Data Custodian and Data Steward director must sign a formal written Data Stewardship Agreement.	May need to consult Data Users to understand their needs for data discovery and access.
2. Define the contents and scope of the data collection, its structure (i.e., conceptual, logical and physical schema), and applicable policies, standards and processes.		
Define business-level data policies, standards, processes	The Data Standards Manager typically develops these. They must conform to applicable government and ministry standards and policies set by the GCIO and MCIO respectively.	May need to consult with the Data Architect or Data Administrator.
Define data administration policies, standards, processes (including metadata)	The Data Administrator typically develops these.	4.2.1.1.1 May need to consult with DataBC, MCIO or GCIO regarding conformance with applicable metadata and data management standards.
Define data quality standards	4.2.1.1.2 The Data Standards Manager typically develops these. The Custodian should ensure that these have been defined, and they should be accessible to Data Users.	
Establish day-to-day operational procedures for data capture, quality assurance and maintenance	4.2.1.1.3 The Data Standards Manager typically develops these.	Should consult with Data Resource Manager.
Establish metadata management procedures	The Data Standards Manager typically develops these.	Should consult with Data Resource Manager, as well as DataBC if they are to be the Data Steward.
Define data access and sharing policies	The Data Standards Manager typically develops these. The Custodian should ensure that they have been defined, and they should be accessible to Data Users. Note that for Open Data they may be dictated by the Open Information and Open Data Policy.	For non-open data, it may be a good idea to consult with Data Users to understand their expectations for use, sharing and possible creation of derived data products.
Establish data security procedures (including backup, archiving and disaster recovery)	The Data Standards Manager typically develops these. The Custodian should ensure that they have been defined and tested.	May need to consult with WTS or whichever group is responsible for physical storage, security and recovery.

Activity	Data Custodian / Data Standards Manager (Form-related)	Involvement by Other Stakeholders (Content-related)
3. Implement the plan and defined standards in order to create the business and technical environment for the data, which then allows the Content and Usage Cycle to be initiated		
Implement the plan	The Data Custodian and Data Standards Manager should oversee the implementation of the plan. They should periodically review implementation activities to ensure that they comply with the Governance and Standards framework.	The Data Resource Manager should be fully involved.
4. Assess the value and continued relevance of the data to its intended use. This is the part of the cycle where the custodian reviews how the data collection is being used and whether it is still meeting its business purpose. As a result of the assessment, there may be opportunities to change the scope, structure, acquisition, definition or use of the data. This activity feeds directly back into the planning process, which factors in the improvements in an updated data lifecycle plan.		
Assess continued relevance and fitness for purpose of data	The Data Custodian should set regular points for assessing the program. It's easy to forget about this step, which may result in the data's structure becoming out-dated for its intended use or generally of less value.	Engage Data Users and Discipline Authorities to confirm that the data is still needed and that its form is appropriate.
5. Leverage existing uses. This is an open-ended activity whereby the Custodian, users and others seek to increase the value of the data collection, such as by linking it with other data collections, by changing its presentation, by applying post-processing to create data products, or by other means.		
Increase the value of the data by applying secondary integration, processing, analysis and/or by creating derivative products	By assessing the relevance and fitness for purpose, and by consulting Data Users, the Data Custodian may be able to identify additional uses of the data.	Consult Data Users and Discipline Authorities about potential added-value products.
6. Dispose of the data. If the Data Custodian determines that collection, management and use of a set of data are no longer required, the Data Custodian may initiate retirement of the data by notification to the affected community and the DataBC Council. Disposal must conform to applicable data retention and disposal policies. This causes both the Governance & Standards and Content & Usage cycles to close.		
Decommission a set of data when there is no longer a business case for its continued existence	The decision to decommission is that of the Data Custodian (unless the requirement to provide the data is expressed in law or government policy).	Consult Data Users.
Appropriate disposal of data	The Data Custodian must ensure that the data set is properly retired.	The Data Custodian, or designate, must consult the Ministry Records Officer to ensure proper disposal.[9]

[9] Government records must be disposed of in accordance with the *Document Disposal Act*, which stipulates the approvals required before they can be destroyed, transferred to the legal custody of the British Columbia Archives, or alienated from the Crown. The *Document Disposal Act* covers records in all media, including paper, microforms, electronic records, audio-visual records, cartographic records, photographic records and other media formats as

4.3 Activities of the Content & Usage Cycle

The table below describes the activities that comprise the Content and Usage Cycle. Most of these are Content related and the responsibility of the Data Resource Manager, who provides day-to-day management of the data and its use.

A designated Data Steward, under a formal agreement with the Data Custodian, may perform some of the activities described.

The involvement of other stakeholders, including the Data Custodian, is also shown

Activity	Data Custodian / Data Standards Manager (Form-related)	Data Resource Manager / Data Steward (Content-related)
A. Acquire the data. This is usually a continuous (or at least, regularly repeated) process for collecting new data.		
Implement data collection and update processes		The Data Resource Manager, together with the Data Steward if there is one, conducts this activity, reporting back to the Data Custodian
Apply data quality management	The Data Standards Manager should expect quality reports, and may need to assist the Data Resource Manager.	The Data Resource Manager conducts this activity, engaging the Data Standards Manager and/or Data Architect if necessary to resolve quality issues. The Data Steward must apply quality management procedures according to the terms of the Stewardship Agreement.
B. Maintain the data by applying QA processes, integrating new data, purging old data, and making operational changes to the business and technical environments.		
Apply processes for QA and integration of new data, and purging of old data		The Data Resource Manager, and/or Data Steward, manages this activity according to procedures established by the Data Standards Manager.
Make data available on a day-to-day basis according to applicable policies and standards		The Data Resource Manager, and/or Data Steward, manages this activity according to procedures established by the Data Standards Manager.

defined in the *Interpretation Act*. Additional requirements are specified in Core Policy & Procedures Manual. Consult with the Ministry Records Officer for technical guidance on record retention requirements and policy.

Activity	Data Custodian / Data Standards Manager (Form-related)	Data Resource Manager / Data Steward (Content-related)
Conduct periodic audit	The Data Standards Manager receives and reviews audit reports, providing to the Data Custodian if necessary.	This may cover a number of distinct activities, depending on the type of audit (e.g., fiscal, security, risk, data quality). The Data Resource Manager, and/or Data Steward, manages the activities, providing audit reports according to policies established by the Data Standards Manager (which should be consistent with broader government audit requirements).
C. Use the data. The data are made available to users according to applicable policies and standards established by the Data Custodian.		
Provide access to the data according to applicable policies and standards		The Data Resource Manager, and/or Data Steward, manages this activity according to procedures established by the Data Standards Manager.
Provide user support		The Data Resource Manager, and/or Data Steward, is responsible for this activity, though there may be opportunities for promoting self-support through user groups and forums.
(Utilise the BC Geographic Warehouse (BCGW)[10])	The Data Standards Manager must consult with DataBC and ensure appropriate processes, policies and standards are followed.	This is a particular activity that applies only for data sets that are suitable for storage in the Geographic Warehouse and for which, therefore, DataBC is a Data Steward. It will be conducted by DataBC staff working with the Data Resource Manager. If utilizing the Warehouse, applicable standards and procedures to establish data in BCGW must be followed (e.g., metadata definition, publication processes, access control.)

[10] The steps for Data Custodians and Data Standards Managers to establish a data collection in the BCGW are described in a separate document.

5 DATA PRODUCTS

5.1 Definition

Data Products may also be called resultants, derivative products, analytical products, value-added sets of data, integrated sets of data, and more.

A Data Product is created by combining or manipulating one or more sets of data in order to address a particular business need. A data product is a set of data that:

- Interprets, modifies or aggregates other sets of data;
- Incorporates significant parts of the data from which it is derived; and/or
- Is managed as a set of data.

The following are not considered a Data Product:

- Modifications or aggregation of sets of data where the data is not exposed outside of the Data Custodian's business area where the data is used.
- Alternative presentations (e.g., a database view) of a source set of data – the Data Custodian's policies and standards still apply.

5.2 Management Principle

Data Products present a management challenge because they contain elements of source sets of data with Data Custodians defining the associated usage and standards. Data Product standards are determined within the context of the following principle:

> *The policies and standards (e.g., for access, security, availability, integrity, fitness for purpose) for a Data Product must not infringe upon or contravene the most restrictive interpretation of the collective standards and rules of the source sets of data the Data Product is created from.*
>
> *This principle applies unless otherwise specified by written agreement with the Data Custodians for the source sets of data.*

In particular, this principle *must* be respected when creating a Data Product that is a combination of Open Data (i.e., under the Open Government License) and non-Open Data, which will have a more restrictive license.

So suppose, for example, that Data Product X is created from source sets of data Y and Z. The Data Custodian for Y has restricted access to Government staff only, but Z is Open Data freely available to anyone under the Open Government License. Thus, under the principle described above, X may only be accessed by Government staff. (The only caveat to this is if Y and Z are

separable within X – for instance, by layer or geographic extent. If so, then the Z part of X may be made freely available.)

5.3 Role of Data Product Provider

Data Products do not have Data Custodians. Instead, there is a somewhat analogous role called a Data Product Provider who has the following rights and responsibilities only:

- Accountable for the Data Product - may set standards (e.g., terms of use, access rights, quality, pricing), but only insofar that these standards do not contravene those of the source sets of data (consistent with the principle above)
- Must enter into written agreements with Data Custodians of source sets of data regarding standards of the Data Product.

  There are some circumstances in which a data product is created from a number of primary sources, but has a sufficiently distinct business purpose that it needs its own formal management and accountability.

From the perspective of the Data Custodian of a source set of data, the Data Product Provider has the same rights and obligations as any other Data User (unless modified by a written agreement).

6 ILLUSTRATIVE EXAMPLES

6.1 Strategic Land and Resource Planning Data

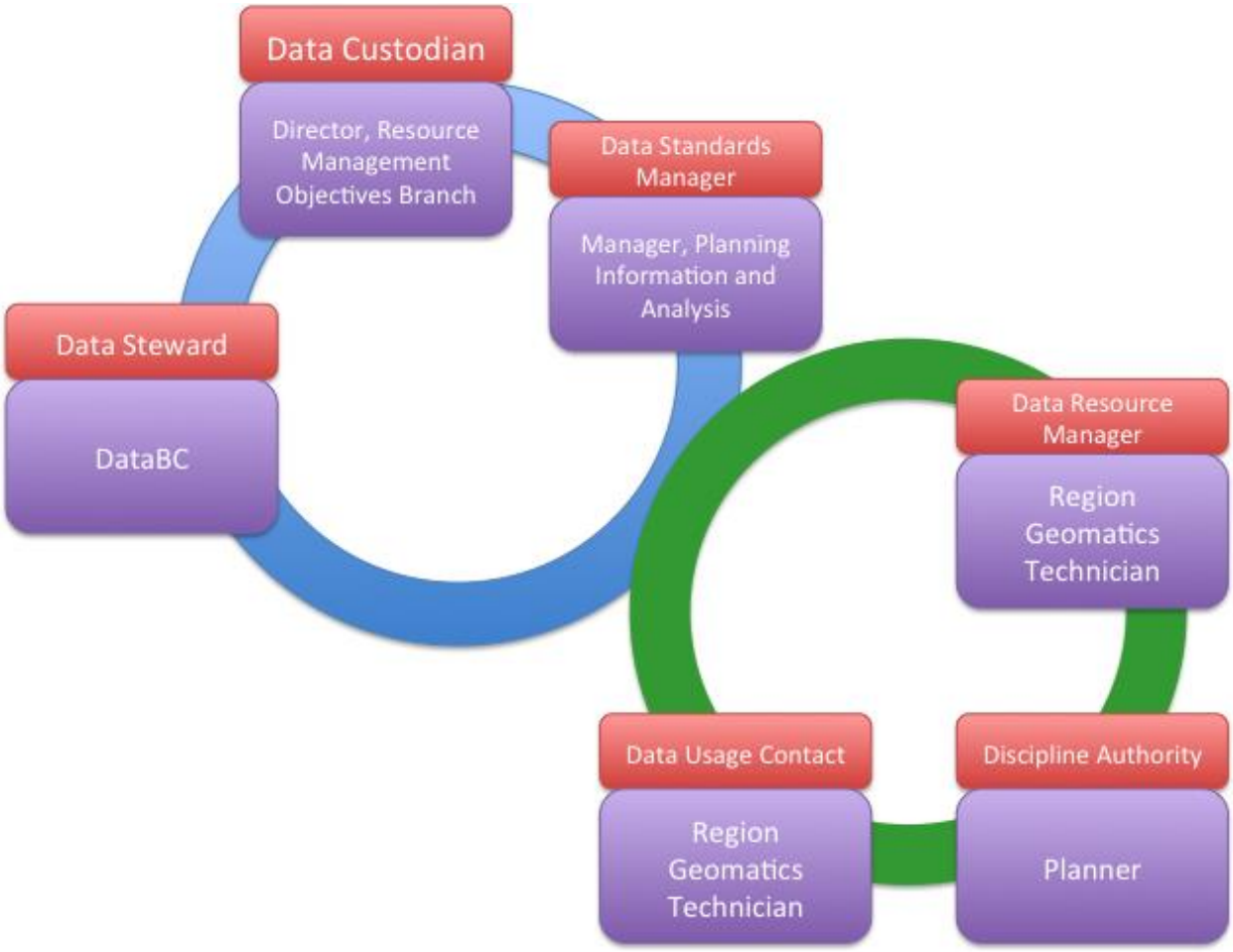


Figure 2. SLRP Data Custodianship Model

Strategic land and resource planning (SLRP) data supports many business functions of the Ministry of Forests, Lands and Natural Resource Operations (FLNRO), as well as other government agencies and external stakeholders.

As the figure above illustrates, the Data Custodian for this set of data is the Director, Resource Management Objectives Branch. He or she is accountable for data and its use. A manager in that branch sets applicable data standards and develops the Data Management Plan. Much of the data is regional in nature, and so other roles are filled by regional staff. The Data Resource Manager is a geomatics technician in the south region. In this and other regions there are one or more Discipline Authorities, who have a deep understanding of the data and its appropriate use, and Data Usage Contacts, who users can turn to for technical expertise in accessing and using the

data. The data are maintained in the Geographic Warehouse, and therefore DataBC acts as a Data Steward on behalf of the Data Custodian.

6.2 Justice BC Court Services Data

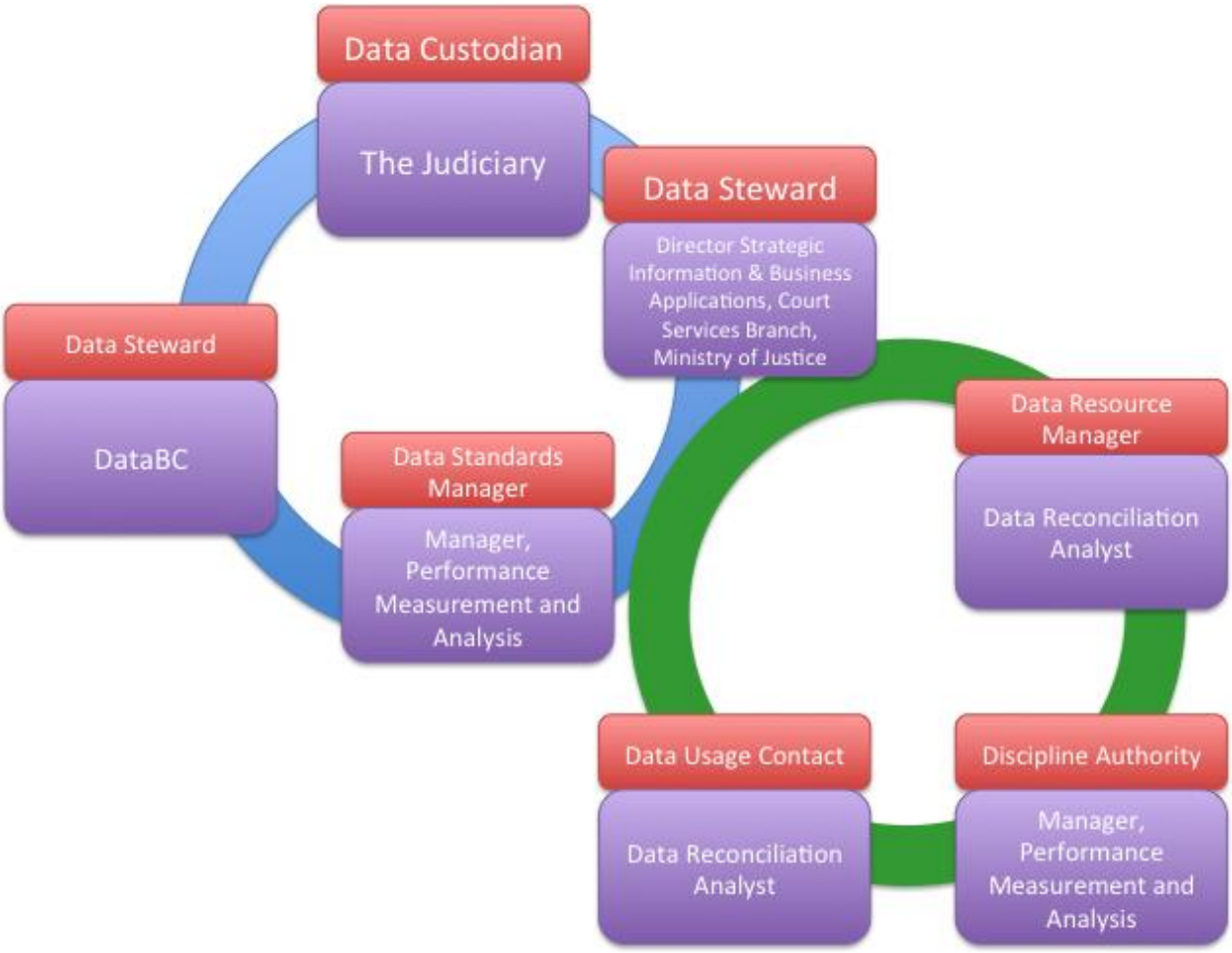


Figure 3. Court Services Data Custodianship Model

The Justice BC Court Services set of data includes number of completed court cases, new cases, sitting hours, appearances, merged cases, documents filed and time to conclusion.

The custodianship model for this set of data is shown in the diagram above. The Data Custodian is the Judiciary. They provide the data to Court Services Branch for management and publication to user groups. Note that Court Services acts as a Data Steward to the Judiciary – they are not the Custodian. DataBC also provides discovery and publication services to Court Services, so DataBC is another Data Steward.

Court Services has designated staff to fill both Form and Content roles as indicated in the diagram.

6.3 Opportunities BC

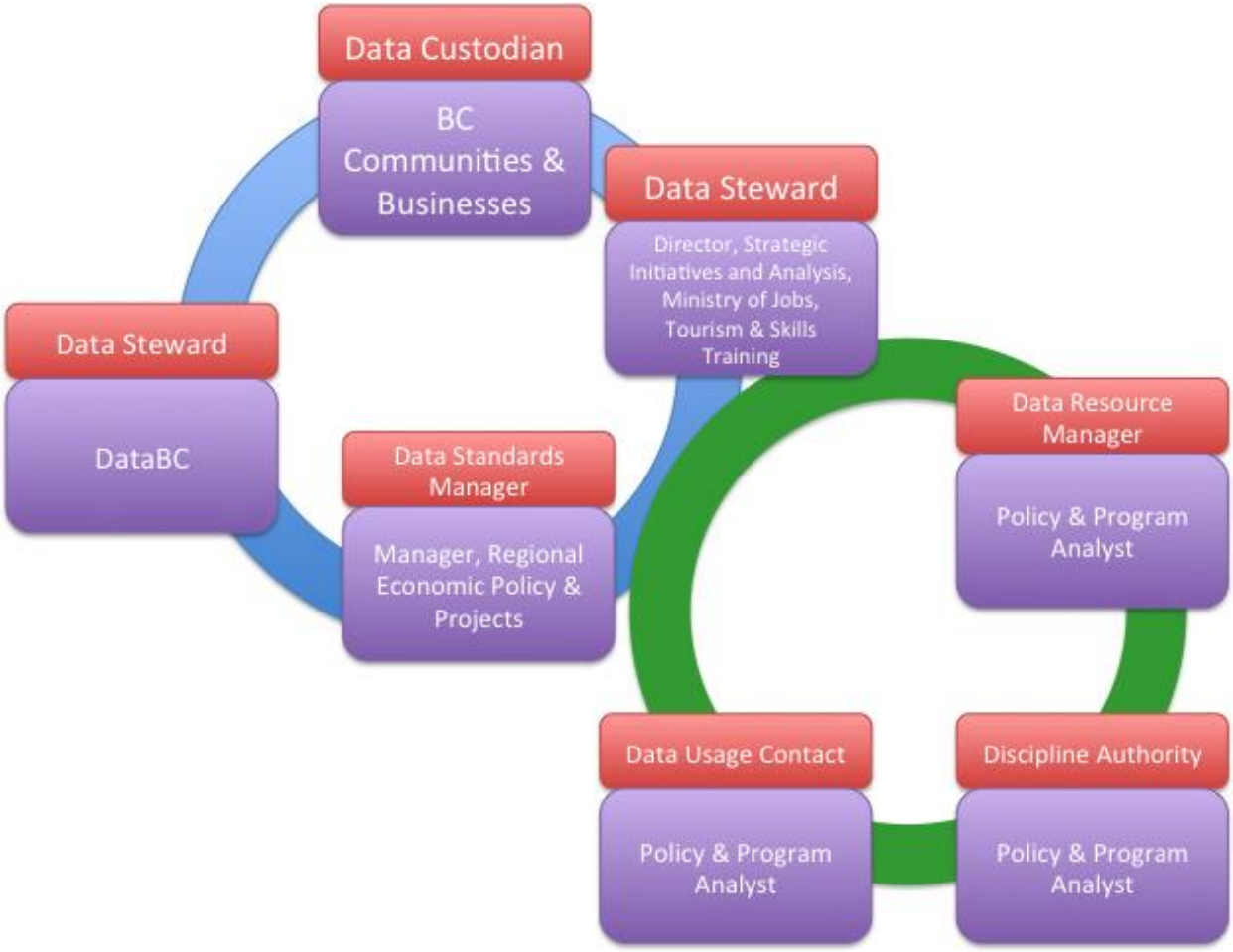


Figure 4. Opportunities BC Data Custodianship Model

Opportunities BC is a spatial database of business and project opportunities within British Columbia suitable for foreign investment. The data set has three functions:

- to provide updated listings of possible of opportunities across the province;
- to serve as a portal for foreign visitors to explore opportunities in B.C. communities; and
- to connect opportunity representatives with potential foreign investors.

Data is provided by BC communities and businesses that have opportunities and wish to attract investment.

Under the Data Custodianship model, the agencies that provide this data are Data Custodians. The Ministry of Jobs, Tourism & Skills Training (JTST) provides a Data Stewardship role by managing the data set on behalf of those providing the data. DataBC is also a Data Steward because it provides discovery and access services for JTST. In addition, JTST assigns staff to fulfil the Data Resource Manager and other Content-related roles.

In some ways, this model is similar to the Court Services structure discussed above. However, it is different in the important respect that the Data Custodian role is *outside of Government and is distributed among many organizations*. These organizations are obviously under no obligation to follow the custodianship practices recommended in these guidelines. Although technically JTST is a Steward, for all practical purposes, they act as the Custodian. They must make management decisions and exercise appropriate care of the data as if they were the Custodians.

6.4 BC Geographic Warehouse

The Geographic Warehouse (BCGW) is managed by DataBC. It is designed and operated for the purposes of hosting and cataloguing sets of data (mostly spatial data) on behalf of Custodial ministries and other government Custodial agencies, and for providing discovery and distribution services to users.

In this capacity DataBC acts as the Data Steward for sets of data stored in the BCGW. The Data Standards Manager (BCGW) is responsible for the development and application of the BCGW metamodel and associated standards. This position currently carries the role of Discipline Authority (BCGW) as well. Note that this role provides expertise related to the data standards applied to the BCGW specifically and spatial data warehousing in general, and is not the Discipline Authority for the individual sets of data under stewardship. Similarly, the Data Usage Contact (BCGW) provides expertise and understanding of the business relevance and appropriate use of Corporate Access Services and not the usage of the sets of data under BCGW stewardship.

Custodial Agencies supply sets of data to the BCGW for stewardship. Typically, the custodial Data Standards Manager also serves as the Data Resource Manager (BCGW). This role is defined as supplying data to the BCGW to the standards set by the BCGW Data Steward.

7 REFERENCES

1. Ministry of Finance, Core Policy and Procedures Manual, Chapter 12, Information Management and Information Technology Management. http://www.fin.gov.bc.ca/ocg/fmb/manuals/CPM/12_Info_Mgmt_and_Info_Tech.htm
2. Guidelines for Best Practices in Data Management – Roles and Responsibilities, Data Architecture Advisory Committee, Information Architecture & Standards Branch, Ministry of Citizens' Services and Open Government; March 2012. http://www.cio.gov.bc.ca/local/cio/standards/documents/standards/data_mgt_roles_responsibilities_guidelines.pdf
3. Guidelines for Custodianship, ANZLIC, The Spatial Information Council; July 1998. http://spatial.gov.au/system/files/public/resources/anzlic/ANZLIC_Custodianship_Guidelines__April_1998.pdf
4. Data Custodianship Guidelines for the Natural Resource Sector, Version 1.1, June 2008. http://www.data.gov.bc.ca/local/dbc/docs/geo/services/standards-procedures/Data_Custodianship_Guidelines_for_the_Natural_Resource_Sector.pdf
5. Management Guide to Custodianship, Guide S35, Ministry of Forests, Lands and Natural Resource Operations, Information Management Group, January 2000. http://www.for.gov.bc.ca/his/datadmin/s35.pdf
6. Ministry of Ministry of Forests, Lands and Natural Resource Operations, Information Management Group Glossary. http://www.for.gov.bc.ca/his/datadmin/
7. BC Geographic Warehouse Guide for Data Custodians & Data Managers. http://www.data.gov.bc.ca/local/dbc/docs/geo/services/standards-procedures/A_Guide_for_Data_Custodians_Data_Managers_Nov_2013.pdf
8. Office of the Chief Information Officer, Information Security Policy. http://www.cio.gov.bc.ca/local/cio/informationsecurity/policy/isp.pdf