Business Plan to Improve Geospatial Data Value



Submitted May 30, 2024



Applied Geographics (AppGeo), a division of The Sanborn Map Company, Inc.
33 Broad Street, Floor 4 | Boston, MA 02109
T. 617-447-2400 | information@sanborn.com
The Sanborn Map Company, Inc. | www.sanborn.com



Contents

1 Executive Summary	1
2 Program Goal	2
2.1 Objectives	3
2.2 Achieving Objectives	3
Objective 1. Define datasets and data themes that are within the scope of this business plan	3
Objective 2. Define measures of geospatial data value for datasets and data themes	7
Objective 3. Assess datasets and data themes using the data quality measures	9
Objective 4. Determine actions needed to improve the value of the selected datasets and data themes	10
Objective 5. Generalize the findings in Objective 4 to a program of data quality improvement	12
3. Requirements and Costs	13
3.1 Assumptions	13
3.2 Resource Requirements	14
3.3 Funding Requirements	15
4. Implementation Plan	17
4.1 Managing Implementation	20
4.1.1 Project Management	20
4.1.2 Risk Management	20
5 Measuring Success	23
5.1 Monitoring Progress	23
5.2 Refinement	24
Appendix A. Data Themes Quality Assessment Workbook	26
Appendix B. Data Value Improvement - Elevation	26
Appendix C. Data Value Improvement - Structures and Addresses	26
Appendix D. Template Dataset Assessment Workbook	26
Appendix E. Template Data Value Improvement Plan Matrix	26
Appendix F. Data Value Program Implementation Plan	26



1 Executive Summary

The Montana State Library (MSL) completed a *GIS Coordination Strategic Plan* in 2022. The strategic plan included a goal for MSL to "Continue to improve the collection, maintenance, and dissemination of authoritative land information". This *Business Plan to Improve Geospatial Data Value* describes goals and actions to reach the goal described in the strategic plan. Geospatial data value, in the context of a statewide geospatial data and service agency like the MSL, largely consists of maintaining and improving data quality. The plan has five objectives, summarized here:

- Determine data themes within the purview of improving geospatial data value and select example data themes to use in the plan
- Define criteria for data value improvement
- Assess the example data themes using the criteria
- Create data improvement plan for the example data themes
- Propose an implementation approach for assessment, plan creation, and acting on the plans for the example data themes and for all other data themes

The resources and funds needed for the business plan are identified after the objectives are discussed. An overall implementation plan and schedule, including acting on the two plans already created through this work, is presented. Data value improvements for one or more data themes can be taken up at any time though, so the schedule shown in the implementation plan is based on an arbitrary start date. Appendices contain templates that can be used in stepping through the actions needed for planning data quality improvements.

Creating a data quality improvement plan for a given data theme is estimated to take less than 3 personweeks of effort in total. If three people are involved in creating a plan, it would take less than 1 personweek of each person's effort. Part of the plan creation is estimating the resources needed to perform the actions called for in it. The level of effort and resources needed will, of course, vary and are not included in this business plan since they are specific to a given data theme. Moreover, MSL will need to manage its data value improvement efforts as MSL projects, just like any other part of its work.

Data value and data quality improvement is always a work in progress because fresh data are often collected and need to be integrated into existing datasets. This business plan formalizes how MSL can create a long-term program of data quality improvement which increases the value of the data it provides. Since MSL is the steward of only some statewide Montana Spatial Data Infrastructure data themes, the program is necessarily a collaboration with other agencies and partners.



2 Program Goal

In 2022, the Montana State Library ("MSL" or "Library") created the *GIS Coordination Strategic Plan.* Goal 4 of the strategic plan is titled "Continue to improve the collection, maintenance, and dissemination of authoritative land information". The goal is described as:

The Library does excellent work collecting and maintaining data. The Library also has an excellent suite of datasets, web services, web maps, online applications, and direct data sharing with partners to disseminate data. MSL should continue improving its products and related coordination efforts as the use of geospatial data continues to expand, expectations about GIS data and services grow, and information system technologies change over time.

The strategic plan describes this goal as ensuring that the Library continues to meet, or even anticipate, the needs of all data consumers. Thus, the goal is to make sure geospatial information has the greatest possible *value* to the people of Montana, both directly (e.g., to geospatial professionals working directly with data) and indirectly (e.g., by providing policymakers with spatial information leading to wise decisions). A more succinct goal statement is:

Improve the collection, maintenance, and dissemination of authoritative geospatial information, aiding the creation of better policies, more informed decisions, and providing value to Montana.

Considering this revised goal statement, the recommendations about improving data collection, maintenance, and dissemination all revolve around data value and data quality. Data value is determined by the information needed by users and whether that information meets those needs. Data value must start from data quality suitable for users' needs. Quality is a necessary foundation but insufficient for overall data value. In addition to quality, or suitability for use, other important elements of data value are discoverability, availability, and curation.

Data must be discoverable for it to be used. Making data discoverable may involve publicizing its existence to stakeholders and ensuring it can be found through contemporary search techniques. Data must be available in usable formats. Depending on user needs, appropriate availability may be retrievable files, on-line services, application programming interfaces (APIs), internet-based applications like mapping web sites, or combinations of these such as web sites on which one selects an area and then downloads data for it as files. Data must be curated, allowing retrieval of historical data values. Curated data values support retrospective studies, e.g., analyzing change over time.

Data suitability, discoverability, availability, and curation are already core activities of the Library. For example, business plans for communications and data governance – addressing discoverability, availability of data – were outcomes of the 2022 strategic planning process. This business plan addresses



the data value elements of data quality or suitability for use and curation. Since all elements of data value are related, it also touches upon availability and discoverability.

2.1 Objectives

Five objectives are steppingstones to achieving the program goal (Table 1). The aim of the plan as a whole is to create a model process MSL can use for any geospatial information dataset or theme. The business plan achieves this by first choosing a sample of data themes (each consisting of one or more datasets) that span the range of data types and themes (Objective 1). This sample, along with a literature review of geospatial data value measures, is used to build a general evaluation framework for data value (Objective 2). In Objective 3, the selected datasets are evaluated with this framework. Actions needed to improve the value of the selected datasets are determined (Objective 4). Objective 5 generalizes from assessments of the sample themes and datasets and provides actions for a data value approach that can be used on any dataset or data theme. Each objective is discussed in more detail, including how the objective will be achieved, in the second part of this section.

Program Goal Improve Geospatial Data Value						
Objective 1	Define datasets and data themes that are within the scope of this business plan					
Objective 2	Define measures of geospatial data value for datasets and data themes					
Objective 3	Assess datasets and data themes using the data quality measures					
Objective 4	Determine actions needed to improve the value of the selected datasets and data themes					
Objective 5	Generalize the findings in Objective 4 to a program of data quality improvement					

Table 1. Business plan objectives

2.2 Achieving Objectives

Objective 1. Define datasets and data themes that are within the scope of this business plan.

This objective involves choosing datasets and data themes that should be examined for value-related data quality improvements. The concept of geospatial data value combines its value for some purpose and the quality of the data within the dataset needed for that purpose. Ideally, every dataset would be perfectly suited for any use and have perfect quality. This, of course, assumes that one can completely

MSL Business Plan to Improve Geospatial Data Value



define the uses of a dataset. The requirements for data quality would follow from this definition. Maximizing data value then involves improving data quality to meet those requirements.

The 15 data themes in the Montana Spatial Data Infrastructure (MSDI) data themes are obvious candidates for consideration because they are considered essential statewide geospatial data (Table 2). Since this business plan involves defining a method for improving data value through data quality, two representative datasets were chosen to work on within the scope of this business plan: Elevation and Structures & Addresses. The Library is the steward organization for both themes, so the plan development involved the data stewards directly. For data themes with other steward organizations, data value and quality improvement will involve using the means available to them. For example, the Soils theme uses the well-established data collection, quality control, publication, dissemination, and archiving standards established by the Natural Resources Conservation Service. Nonetheless, the general ideas presented in this plan apply to any data theme.

Data Theme	Steward Organization					
Administrative Boundaries	Montana State Library					
Cadastral	Montana State Library					
Climate	University of Montana					
Elevation	Montana State Library					
Geographic Names	Montana State Library					
Geology	Bureau of Mines & Geology					
Hydrography	Montana State Library					
Hydrologic Units	Natural Resources Conservation Service					
Land Cover	Montana State Library (Montana Natural Heritage Program (NHP))					
Mapping Control	Montana State Library					
Imagery	Montana State Library					
Soils	Natural Resources Conservation Service					
Structures & Addresses	Montana State Library					
Transportation	Montana State Library					
Wetlands & Riparian	Montana State Library (NHP) Montana Department of Environmental Quality					

 Table 2. Montana Spatial Data Infrastructure (MSDI) data themes and steward organizations
 (https://msl.mt.gov/geoinfo/data/msdi).

The project team selected two themes to study in this business plan, though each MSDI theme is worthy of evaluation and should be evaluated. Overall factors affecting the selection included being able to control performing improvements (i.e., acting on the findings of the business plan) and including a wide range of data sources, types, and characteristics. Elevation and Structures & Addresses were chosen as model data themes for this plan.

Elevation data theme is composed of multiple datasets, many of which are derived from source data



acquired across several projects and timeframes. The elevation data is used in many applications across Montana such as visualizing topography through "shaded relief," modeling water flow, line of sight analysis, and terrain analysis. The following are additional reasons why the Elevation data theme is a good study topic.

1. The Elevation data theme consists of multiple datasets. most of which are derived from "raw" elevation observations made by Lidar collection projects across the state. MSL aggregates the different projects and source datasets to create as seamless a statewide coverage as possible. Lidar projects have been created by many different projects over time, using several collection and data quality standards. Heterogeneous data sources often make for data quality challenges.



Figure 1. A Lidar Use Example from MSL's Montana Lidar Inventory

- 2. The processes used to bring disparate datasets into a statewide whole are complicated and involve considerable technical expertise both in building an elevation dataset and performing quality control on it.
- 3. Elevation data themselves are then used to create many derived, related, products in the data theme: hydrography, contours, bare earth digital elevation models (DEM/DTM), digital surface models (DSM), hillshade map layers, and tree canopy coverage. DEMs, DTMs, and tree canopy coverage are dependent on lidar observations and other datasets are not. The datasets within the theme are dependent on different kinds of source data and thus potentially different requirements for the quality of source elevation data themselves. For instance, the different products have different sensitivity to change over time. Bare-earth DEM and contours are generally more stable over time compared to digital surface models that show vegetation and buildings.
- 4. The datasets in the Elevation data theme, or directly related to it, are in several geospatial formats: vectors (e.g., contours, stream channels), points (processed lidar point cloud observations, e.g., spot elevations), and rasters (also called grids, e.g., the statewide 10m DEM).



- 5. Lidar data are being improved every year. What was formerly the best lidar data available is displaced by even better data with each subsequent project. Constant improvements in data accuracy mean that the definition of "data quality" must incorporate the certainty that data will improve over time.
- 6. Finally, improving lidar coverage of the state is a long-term goal for the MSDI, first formalized in 2019 and actively worked upon since.

The Structures & Addresses data theme is composed of a single dataset of point data. Like the Elevation data theme, the Structures & Addresses theme is used for many important purposes throughout Montana. These include Next Generation 9-1-1 (NG9-1-1), a critical safety improvement initiative with

well-defined data requirements. Other standards that involve this data theme include standardized addressing for other uses (e.g., elections). Structures & Addresses data come from jurisdictions throughout Montana, aggregated and standardized by MSL. The data theme was selected because:

 There are multiple providers with differing technical abilities to meet national and state data standards, so data arrive with various possible problems,



Figure 2. Montana Structures and Addresses Web Map

making it a good candidate for studying how one improves a statewide product built upon varied sources.

- 2. The existence of a well-defined statewide NG9-1-1 standard provides a target for data quality improvements.
- 3. The sources of structure and address data are continually improving over time.

As a program of data value improvement progresses, themes and datasets included in such a program will change. There can be many other reasons for selecting data themes for improvement such as changing demands or priorities for specific themes and datasets or new technologies that improve collection, cleaning, and publication of data. Some value considerations in selecting datasets and data themes for data quality improvement analysis are:



Significance of the data theme and datasets

- Are the data theme and datasets critical to public welfare?
- Are the data themes and datasets required by law or policy?

Demand for the data theme and its datasets

- Current usage
- Likely future usage
- Dependencies of other datasets or themes on the datasets in a theme

Feasibility of improving the datasets within a theme

- Nature of the data (sources)
- Level of effort to achieve meaningful improvement
- Benefit versus cost of improving data value

Objective 1 is completed within this business plan, because two data themes have been selected for study within the scope of this business plan. The criteria used to select these themes can be used to build a long-term program in which improvements to data themes are assigned priorities.

Objective 2. Define measures of geospatial data value for datasets and data themes

As discussed above, data *value* is supported by the quality of data needed for a given purpose. One measure of data value is simply how often a data theme or dataset is used. However, the use of a data theme or dataset may not be something that can be directly influenced in a program of data value improvement. The use of a dataset, if it is available in an appropriate form, is determined by its consumers, not by its publisher.

Data quality, on the other hand, is a component of the value of data. Unlike the use of a dataset, data quality is more amenable to control by data creators and maintainers. Because data quality directly affects usefulness, improving quality improves data value. Data quality can be examined in several dimensions (Table 3).



Table 3. Data Quality Assessment Dimensions

Data Quality Dimension	Description	Example
Precision & Accuracy	Accuracy is whether data values correctly represent the phenomenon observed (subject to usage needs and measurement technologies), precision is how the value is stored and expressed	Each address represented by a point is within 2 meters of the structure it represents (accuracy) and coordinates are stored and expressed as integer meter x-y pairs
Consistency	Similarities and differences between data stored in multiple datasets or databases	A statewide address database having the same set of addresses as each county's address databases is consistent
Completeness	A measure of whether the data in datasets are comprehensive with all phenomena represented and all required observations populated	A 10m DEM that covers the entirety of a state is complete; an address database that does not contain all addresses is incomplete; a DEM dataset that covers an entire state but has some gaps or lacks values in some cells is incomplete
Integrity	An assessment of whether the data's structure, schema, and maintenance workflows maintain consistency, completeness, accuracy, precision, to meet user needs and interoperability	Address data are verified and processed to a common standard so that all addresses have equal validity; raw lidar data is transformed using standardized techniques with consistent documentation of issues for users to consider
Timeliness	The update cycles of a dataset or set of datasets in a theme, supporting internal and external production processes, reporting and usage needs. This dimension includes the ability to use historical versions of a dataset (archiving) and the frequency with which updates are performed and older data are archived	Counties provide new and edited address datasets on a monthly schedule and statewide data are updated twice per year; published elevation data are updated annually
Relevance	Whether data are available, accessible, and having the completeness, consistency, precision/accuracy, integrity, and timeliness for a given business use	Address data support NG9-1-1 systems; statewide DEM data can be used for modeling flood risk; superseded DEM datasets are available to support study of floodplain changes over time



Each dimension speaks to some aspect of data quality that can be tested in some fashion. Some kinds of data are more amenable to assessment in particular dimensions, and this is a consideration in deciding on the weight one might choose to give to these measures. For example, one can easily assess the completeness of something like a DEM: examine the area covered and check whether all cells have a non-null value. Continuing the example, completeness is more difficult to assess for addresses because new addresses are created frequently so there is not a perfectly "known" population to use in checking completeness of a dataset. Nonetheless, the dimensions of data quality are useful measures even if, in some cases, they must be interpreted cautiously.

Objective 2 is completed within this business plan, because data quality measures have been defined for the two example data themes selected for study. The data quality dimensions are sufficiently general that they can be used for almost any data theme or dataset.

Objective 3. Assess datasets and data themes using the data quality measures

This objective uses the data quality dimensions to assess the two sample data themes selected in Objective 1. A simple worksheet was created (Figure 3) for each data theme. Datasets within each theme are listed in rows, and data quality dimensions are listed in columns. MSL staff responsible for each data theme then filled in this assessment, focusing on three aspects of each dataset: status, near-term quality improvement, and long-term quality improvement.

Data Theme (italic) and								
dataset	Data Value and Quality Goals							
	Update	Archive						
	Frequency	Frequency	Spatial Accuracy	Consistency	Completeness	Precision	Integrity	Timeliness
Structures/Addresses (datase	et is same)							
			1				Current major	
							uses: NG911,	Variation in
			1		1Spatial Validation		elections,	update frequency
		Yearly snapshot,	1		Portal used by		broadband.	from sources
		monthly local	1		about 50% of		Integrity is	affects use of
	Monthly	backup copy as	Located on or	County updates	counties; other		adequate but in	dataset for some
Structures and Addresses -	Publication of	data are	close to structure	arrive when they	records are spot-		process of	purposes (e.g.
Current Practice	Data (as-is)	published	being addressed	arrive	checked		improvement.	NG911)
			1					
				MSL data and		Structures and		
			1	county data	All data go through	addresses are		
				contain same	an automated	mapped with best		
				records: data	checking process	precision (and		
				meets standards in	that verifies	accuracy)		
			1	both county and	completeness (and	technologies		
		Ability to get to		MSI datasets:	correctness) Data	(examples: Lidar-		
		dataset as it	1	geography matches	provider potified	derived or high-	Data monte	
Structures and Addresses		ovisted on any	1	in both data	of records that fail	procision GPS	poods of all	
Ideal	Near Deal Time	given day	On the structure	in both data	with markuns	precision GF5	maior usos	Timohu
Structures and Addresses - Ideal	Near Real-Time	dataset as it existed on any given day	On the structure	geography matches in both data sources.	provider notified of records that fail with markups.	derived or high- precision GPS location)	Data meets needs of all major uses	Timely

Figure 3. Snippet from Structures & Addresses dataset assessment worksheet

The two data themes selected as examples for this business plan were evaluated using this matrix (Appendix A). For each theme, MSL's MSDI Data Theme Leads filled in the matrix. For these two themes, the MSDI Data Theme Leads are experts in the current datasets and workflows. They also had already



considered future improvements and needs. This was efficient for this study, because for each data theme a single person was both a subject matter expert and a data steward. For other data themes, MSL might need more than one individual because there could be several subject matter and business usage experts, as well as a data steward. As a general best practice, defining data quality measures for assessment is best done with a group of subject matter experts because suitability for user needs is the determinant of data quality requirements.

Objective 3 is completed within this business plan, because the two sample data themes have been assessed using the defined data quality measures. This assessment provides valuable input to the data quality improvement plan discussed in Objective 4.

Objective 4. Determine actions needed to improve the value of the selected datasets and data themes

After the value and data quality assessment step is completed for a data theme and its datasets, the next step is to determine the actions that must be taken to address deficiencies and make improvements. Typical actions associated with data improvement for the Elevation data theme example in this study are shown in Figure 4. Additionally, Appendix B and Appendix C contain the full data quality matrices for the Elevation and Structures and Addresses data themes respectively.



Increasing Data Value through Data Quality Matrix

THEME: Elevation (Lidar, Contour, Digital Elevation Model, Digital Terrain Model, Digital Terrain Model, related derived products)

DATA VALUE and QUALITY OVERALL GOALS: Five year goals: Process (and plan) in place to generate a complete statewide coverage every 5 to 8 years, with more frequent updates (see Timeliness) in areas of dynamic change (floodplains, landslides, large burn areas, development areas). Have clarity in communications with partners about delivery dates. Any future lidar data is collected at QL1 (8 points per sq. m). Ideal goals: Entire state is covered by QL1 data, replacing areas of current QL2 data. Turnaround time is less than one year from acquisition to publication. Complete statewide point cloud classification, including buildings and vegetation. A pool of funding is available for Lidar updates, shortening the elapsed time from identifying a need for data, funding data acquisition, and actually acquiring the data. Core products – DSM, bare earth DEM, hydro-flattening, hydro-enforcement, certified contours, fully classified point cloud, canopy height, hillshade -- are available statewide, are created with consistent statewide methods, and are updated regularly as new data are acquired.

Archiving and Retrieval Capabilities: One goal is to have 10-20% of statewide coverage archived and replaced by new data each year, achieving a <u>5-8 year</u> statewide update cycle. A second goal is to have more sophisticated retrieval methods and provide storage and self-retrieval of datasets from projects and statewide data.

Outcome	Action	Progress Assessment
Current: User can retrieve version of any product dataset as it existed for any given month and year (may just be annual)	Continue current approach to archiving (annual). Original collections are available as projects.	In place
Near-term: User can retrieve authoritative data for any given month		

Figure 4. Example data value and quality improvement plan

The means by which data quality actions will be accomplished will differ by dataset and data theme. A data value and quality improvement plan may go into the technical means and tools that will be used or such specifics might be decided as part of the work. Technical means vary from those that are very close to the data in a dataset, such as looking for illegal values, to programmatic technical means like having a data management plan for a dataset. Some **common technical means** include:



Performing consistent (i.e., repeatable) data profiling analyses to determine missing values, outliers, inconsistencies, and illegal values in constrained or domain-based value sets.

Implementing proactive data validation in entry and editing workflows and software. Using a data management plan for each dataset that specifies who, when, and how datasets are updated and incorporated into a data theme in easily understood formats like workflow charts and checklists.

Establishing and enforcing standards for a dataset, including constraints, domains, naming conventions, and data formats.

Promulgating standards and educating data creators and maintainers to build a culture of data quality awareness.

The list of technical means is not exhaustive; certainly, each dataset requires some thought as to how best to implement an action.

Objective 5. Generalize the findings in Objective 4 to a program of data quality improvement

This objective's goal is to provide a general program of data quality improvement that can be applied to additional data themes going forward, using the results of the process defined in Objectives 1 through 4.

MSL must first assign priorities to the data themes for which data quality improvement plans should be developed. For each data them, the following steps need to be taken:

- 1. Convene a data theme working group
 - a. Include an MSDI Data Theme Lead/State GIS Coordinator/PM, SMEs, Stakeholders, and other roles as needed
 - b. Set a recurring schedule of working meetings
 - c. Set a recurring schedule of data quality check-ins to begin once a data quality improvement plan has been developed
 - d. Define the datasets included in the data theme that will participate in the assessment and improvement plan
 - e. Identify and document data owners/ maintainers/ consumers for selected dataset/ data theme.
- Confirm the generic quality dimensions identified in this plan are sufficient; if dataset-specific dimensions are needed, develop the list of dimensions to be assessed and, if possible, the means by which they will be evaluated
- 3. Assess the data theme's datasets using the defined list of quality dimensions.



- a. Identify a facilitator/advisor to help the group assess the datasets. This may be MSL staff that have previously gone through this process, or an outside consultant.
- b. Assign parts of the assessment to the working group participants or work together as a group to complete the assessment.
- c. Use the Template Dataset Assessment Workbook (Appendix D).
- d. For each dataset, consider and document the datasets using the conditions discussed in Objective 3: current status, near-term quality improvement, and long-term quality improvement.
- 4. Develop a data value and quality improvement plan for the data theme.
 - a. Use the Template Data Value Improvement Plan Matrix (Appendix E).
 - b. Define overall goals for selected dataset/ theme based on the assessment.
 - c. Based on the assessment and deficiencies identified, define the outcomes and actions needed to achieve the goal.
 - d. Highlight repeatable improvement actions (e.g., data validation, cleansing, quality monitoring, quality checks, etc.).
- 5. Identify resource needs (staff time, licensing, consultant support, funding) for each plan outcome/action sequence using the template.
- 6. Incorporate improvement actions into appropriate staff responsibilities/work plans, MSL biennial budget plan, etc. Adjust staff responsibilities to incorporate recurring improvement tasks.
 - a. Track the progress of each outcome/action sequence in the Data Value Improvement Plan matrix or wherever MSL deems appropriate.

While developing this objective, steps 5 and 6 were added to the process of data quality improvement, as they are necessary for any general program of data quality improvement. Thus, these steps were not done for the 2 example themes, Structures/Addresses and Elevation, during the development of Objective 4. These steps are included as tasks to be completed for Objective 4 in the implementation plan.

3. Requirements and Costs

Business plans almost always require some resources to implement, such as funding, labor, and time. Details on how funding, labor, and time combine to put the plan in place are part of its implementation (Section 4).

3.1 Assumptions

Business plans generally rest upon some assumptions. Recognizing these assumptions early helps avoid problems later in the plan's implementation. The need for this business plan became clear during MSL's 2022 Strategic Plan formulation (see the discussion of program goals in the introduction). The main



assumption concerning the business plan's importance is that the situation described in the strategic plan, where this plan was called for as a recommended action, is essentially unchanged. In addition, it is assumed that required resources are available to meet the timeline and that outside events, such as legislative changes, do not impact the planned order of data set implementation. Furthermore, it is assumed the MSL will consider the work defined in this plan as a high priority. This assumption is necessary to construct a project timeline. Situational realities may give other activities higher priority thereby extending the timeline for this work.

As discussed in the introduction to this plan, all aspects of data value have always been a high priority within the Library. The resources needed to further improve data quality might be hard to find in teams that are already stretched to maintain status quo. Furthermore, increasing data value will lead to increased use which in turn may lead to increasing demands on staff for still further improvements. This can lead to a vicious cycle, as staff get overstretched and data quality (i.e., value) declines. However, the role of an improvement plan is to allow one to manage the resources needed to maintain and improve data value overall, rather than performing singular quality improvements in fits and starts.

3.2 Resource Requirements

Implementing this business plan requires technical and subject matter expertise, as well as coordination or Project Management. For the implementation of this plan, it is assumed that the State GIS Coordinator will provide the project management. Table 4 shows the human and non-human resource requirements for implementing this business plan.



		ach objective in the business pie	
	MSL Roles	Other Agency Roles	Non-human Resources
Objective 1. Define datasets and data themes that are within the scope of this business plan.	 State GIS Coordinator MSDI Data Theme Leads* Consultant 	NA	NA
Objective 2. Define measures of geospatial data value for datasets and data themes	 State GIS Coordinator MSDI Data Theme Leads Consultant 	NA	NA
Objective 3. Assess datasets and data themes using the value measures	 MSDI Data Theme Leads Facilitator/Advisor 	NA	NA
Objective 4. Determine actions needed to improve the value of the selected datasets and data themes	 State GIS Coordinator MSDI Data Theme Leads Facilitator/Advisor 	NA	NA
Objective 5. Generalize the findings in Objective 4 to a program of data quality improvement	 State GIS Coordinator MSDI Data Theme Leads Data SMEs Data Stakeholders Facilitator(s)/ Advisor(s) 	 MSDI Data Theme Leads Data SMEs Data Stakeholders 	 Microsoft Office or equivalent productivity tools Project management and tracking software tools, if appropriate Travel for non- virtual meetings (vehicle, per diam)

Table 4. Resources, both human and non-human, needed for each objective in the business plan.

*MSDI Data Theme Lead role is assumed to be a subject matter expert (SME), but other SMEs may need to be involved.

3.3 Funding Requirements

Considering Objectives 1 through 3 were completed through the development of this business plan. For the two data themes selected as examples Objectives 4 and 5 remain to be implemented. Objective 4 is



mostly complete for the Structures and Addresses and Elevation data themes; however, a few tasks remain to be completed as defined in Objective 5. These include identifying resource needs, incorporating the improvement actions into work plans, and tracking progress on data quality improvements. For these two themes, the level of effort estimated for MSL staff to complete this work is 48 hours per theme, or 96 hours total.

Table 5 shows the estimated level of effort and funding required to implement this business plan per additional data theme. Most of the outlined tasks can be completed by MSL with support from a working group, comprised of stakeholders and subject matter experts, that is convened for a particular data theme at the beginning of the process. The implementation plan, shown in Table 6 and included in spreadsheet form as Appendix F, offers more detail about which tasks are intended to be supported by data theme working groups. Estimated hours in Table 5 and in the implementation plan represent MSL staff time required to complete tasks. These estimates do not reflect stakeholder time required to participate in working group sessions.

Level of effort estimates for tasks involving the working group include time for group working sessions as well as independent work time for MSL staff. The total estimated MSL staff time required to complete all the outlined tasks for a data theme is 220 hours. The LOE in Tables 5 and 6 reflects the total number of staff hours for a particular task. While some tasks may be completed by a single staff member, others may require multiple staff roles to contribute hours toward the task total. Outside assistance with facilitation and advising may be beneficial to enable MSL staff to fully participate in those activities. The cost estimate for an outside facilitator assumes that they will spend roughly the same amount of time as the MSL staff on the applicable tasks. This may be a conservative estimate as MSL staff may spend independent time working on a data theme. The estimated total cost of outside services for a data theme to go through this process is \$17,900.



Table 5. Level of Effort and Funding Requirements per Data Theme

Task	Level of Effort- MSL (Hours)	Outside Services Cost
Identify resource needs for each plan outcome and action (Addresses/Structures and Elevation Themes)	32	
Incorporate improvement actions into appropriate staff responsibilities and work plans, MSL budget (Addresses/Structures and Elevation Themes)	48	
Track progress and achievements (Addresses/Structures and Elevation Themes)	16	
Convene a data theme working group	24	
Confirm generic data quality dimensions as used in this plan, add or refine as appropriate	4	
Assess the data theme's datasets using the defined list of quality dimensions	24	\$6,700*
Develop a data value improvement plan for the data theme based on the assessment	24	\$6,700*
Identify resource needs for each plan outcome and action	16	\$4,500*
Incorporate improvement actions into appropriate staff responsibilities and work plans, MSL budget, etc.	24	
Track progress and achievements	8	
TOTALS	220	\$17,900

*Estimate assumes a consultant facilitator/ advisor supports task at same level of effort, using April 2024 costs.

4. Implementation Plan

The implementation plan consists of the five program goal objectives with tasks identified as necessary. Each task has a timeline. Some of the steps needed to implement this plan were completed by MSL and Sanborn during the development of this business plan. Therefore, only the implementation steps remaining to complete are presented with a timeline duration. The implementation plan assumes that Objectives 1 through 4 are substantially completed. That leaves Objectives 4 and 5 with tasks still to be implemented as shown in Table 6 (also included in Appendix F).



The timeline is based on allocating the resources recommended in this business plan. If fewer resources are allocated or the start of a task is deferred, the timeline necessarily stretches out. Conversely, if more resources are brought to bear, this will usually speed up the timeline. Tasks may be pursued concurrently for different data themes as resource availability allows, however, spreading resources across too many data themes at one time will limit progress.

Table 6 is a timeline (in months) and is a baseline for this business plan. Data value improvements for one or more data themes can be taken up at any time though, so the schedule shown in the implementation plan is based on an arbitrary start date. As shown in the table, the development of a single data theme quality improvement plan could take as long as 12 months. Some of the tasks may overlap, allowing the project team to shorten their duration for a given data theme. For example, while the data theme working group is assessing the quality of the selected datasets, the group may also be able to identify improvement actions that should be taken to improve data value. The final task of tracking progress and achievements, which is an ongoing activity, can also be performed while initiating the improvement planning cycle for the next data theme.



Table 6. Implementation Plan Timeline.

							Cost Estimate									
Objectives and Tasks	Timing Notes	8	Timeline in Months			Hours	(Outside Svcs)	Budget Notes								
		1	2	3	4	5	6	7	8	9	10	11	12			
Objective 1: Define datasets and data themes that are within the scope of this business																
Discuss and decide upon two Data Themes to use as examples for the development of this																
plan	Completed															
Identify datasets to be included in the assessment of each Data Theme	Completed															
Identify data subject matter experts for selected dataset/ data theme	Completed															
Objective 2. Define measures of geospatial data value for datasets and data themes.																
Research geospatial data value and quality dimensions	Completed															
Develop draft list of dimensions	Completed															
Review/revise list of dimensions with project team	Completed															
Define the dimensions with project team	Completed															
Objective 3: Assess datasets and data themes using the identified value measures.																
Assess Addresses/Structures data for current status, near-term quality improvement, and																
long-term quality improvement, for each data quality dimension	Completed															
Iterate on worksheet to develop a template	Completed			-								+				
Assess Elevation data for current status, near-term quality improvement, and long-term												+				
quality improvement, for each data quality dimension	Completed															
Objective 4: Determine actions needed to improve the value of the selected datasets and														96		Objective 4 Total (0.046 FTE)
Develop a draft data quality improvement plan matrix template	Completed															
Review/revise matrix template with project team	Completed															
Evaluate assessment of Addresses/Structures theme and populate data quality																
improvement plan matrix	Completed															
Evaluate assessment of Elevation theme and populate data quality improvement plan																
matrix	Completed															
Refine matrix through using the template for Addresses/Structures and Elevation Theme																
plans	Completed															
Identify resource needs for each plan outcome and action (Addresses/Structures and																
Elevation Themes)														32		
Incorporate improvement actions into appropriate staff responsibilities and work plans,																
MSL budget (Addresses/Structures and Elevation Themes)													_	48	-	
Track progress and achievements (Addresses/Structures and Elevation Themes)	Ongoing												_	16		
Objective 5: Generalize the findings in Objective 4 to a program of data quality														124	\$17,900	Objective 5 Total (0.06 FTE)
Convene a data theme working group				_	_	_	_					\rightarrow	_	24		
Confirm generic data quality dimensions as used in this plan, add or refine as appropriate				_	_	_	\rightarrow					\rightarrow	\rightarrow	4	Ac 700	
Assess the data theme's datasets using the defined list of quality dimensions		-		_					-			-	-	24	\$6,700	Cost estimate assumes a consultant facilitator/advisor
Develop a data value improvement plan for the data theme based on the assessment														24	\$6,700	Cost estimate assumes a consultant facilitator/advisor
Identify resource needs for each plan outcome and action														16	\$4,500	Cost estimate assumes a consultant facilitator/advisor
Incorporate improvement actions into appropriate staff responsibilities and work plans,																
MSL budget														24		
Track progress and achievements	Ongoing													8		
														220	\$17,900	Goal Grand Total (0.106 FTE)

MSL Business Plan to Improve Geospatial Data Value

Page 19



4.1 Managing Implementation

4.1.1 Project Management

A key resource for the successful implementation of any data value improvement activities is a Project Manager. The Project Manager should treat this business plan implementation like a project. Solid project management will result in the best outcomes and experience for MSL and all other stakeholders involved.

It is recommended to use business systems to support Project Management best practices. At the outset of the project, MSL should tailor the optimal combination of these systems for providing project transparency and streamlining communications:

- Task management
- Video conferencing
- Web-based data and document libraries
- Time management and reporting systems
- Gantt charting and project planning tools

The resources involved will vary depending on which dataset or theme is being worked on. An agreed upon plan for communication with the implementation team should be established. Establish a cadence of regular check-ins on the status of the project, addressing the frequency, scope and content of routine communications and task level reporting. For example, a monthly status meeting for the internal team, if applicable, and a quarterly check-in with the working group/SMEs.

Use the implementation plan as a guide. Measuring progress along this schedule should be a core objective of the regular project management meetings.

Project success is closely tied to staying on schedule and preventing or mitigating unforeseen problems. MSL's project manager should proactively manage schedule and risk in the following ways:

- Potential project risks and causes for delay are identified at the outset of the project.
- If unforeseen roadblocks arise, they are immediately reported to the project manager for swift discussion and action.
- The project schedule is kept up to date throughout the project, to keep everyone on the team aware of the current schedule.

4.1.2 Risk Management

The elements of a business plan are subject to risk. The more that these risks can be contemplated and understood in advance, the better the chances are to effectively mitigate each risk. Table 7 below outlines several risks, evaluates the potential level of impact from those risks, and proposes measures to mitigate each risk.



Table7. Assessing the business plan risks, impacts, and mitigation.

	Severity of		
	Impact (Low,		
Risk	Med, High)	Potential Impact	Mitigation
Risks other than the ones listed here may be identified at any point in the timeline.	Severity will range from low to high depending on the specific risk and the ability to mitigate the risk.	The nature of the impact will depend on the specific risk.	Proactive risk management is required as a project management responsibility with support from the project team.
Current events may change data set priority	Low	The planned order of implementation of data sets may change due to outside events, such as legislative changes. This may result in modest inefficiencies from interrupted workflow.	The mitigation actions will depend on the specific situation causing the disruption but could include any or all of the following: 1) stop work on current data set and change resources to complete workflow of new dataset 2) add resources as needed to work on current and new datasets at the same time, or 3) adjust timeline as necessary to accommodate plan changes.
Work cannot be accomplished in the time allocated	Medium	Work that depends on the work that is delayed is affected. More cost and effort may be used than originally planned. Required resources may no longer be available when needed. Credibility of the work itself and reputation may be damaged.	Project management will help to minimize negative impacts by identifying project delays early in the process while mitigation measures can be implemented to meet the timeline. Additional resources, if available, can be added. Minimize project outcomes so that some non- critical outcomes are dropped. Realign project expectations based on the new estimate of completion schedule.



	Severity of		
	Impact (Low,	.	
Risk	Med, High)	Potential Impact	Mitigation
The necessary	Medium	This is a risk to the	The mitigation is to adjust the
resources are not		completion timeline and	timeline to allow for the
available in the		may also adversely impact	availability of resources later.
required timeframe		dependent en timelu	Ideally, another data set could
		completion	be substituted if resources
		completion.	were available for it. Affected
			stakeholders should be
			advised of the change in plan
Darthors data	Madium	Partners stowards and	This is a coordination issue
stowards and	IVIEUIUIII	stakeholders nushback	Meet with parties who are
stakeholders disagree		against the proposed	nushing back to understand
before or during the		actions to increase the	their concerns. Find common
implementation		value of the data set.	ground, if possible, by
process about the		thereby decreasing their	addressing those concerns,
actions proposed to		support of the planned	while also explaining the
increase data value		actions. Resolution could	benefits of the actions as
		impact timeline.	planned. Adjust timeline, if
			necessary, as outlined above.
Partners, data	High	Potential impacts include:	The optimal mitigation is
stewards, and		1) Losing the trust of some	prevention. Make sure to
stakeholders disagree		partners and/or	include partners, data
with plan after		stakeholders; 2) Starting	stewards, and stakeholders
implementation		over, completely or in part,	sooner rather than later in the
actions are		with the implementation	implementation process.
implemented		process for the data set in	Otherwise, revise resource
		question to address	needs estimate and timeline
		concerns; 3) Data may be	to address plan changes.
		perceived as less valuable	to be transparent about the
			situation and to minimize loss
			of trust.



	Severity of Impact (Low.		
Risk	Med, High)	Potential Impact	Mitigation
Technical barriers prevent data quality improvements from being made	Medium	Data may not be able to improve at all or data may only be partially improved	Identify technical barriers in planning how data improvement will be performed, including understanding the technical ways in which partners, stakeholders, and other necessary parties can collaborate. Identify changes that would be needed to make data quality improvements feasible.

5 Measuring Success

As with all planned activities, it is important to measure success and adjust the plan and its implementation as necessary. Four levels of success should be monitored and measured, and two types of refinement should be considered.

5.1 Monitoring Progress

The first level of success is deciding to take actions that improve data value and quality. MSL has already made progress toward this because two data themes have improvement plans created as part of this business plan. These two plans now need to be put into effect. Overall though, the first level of success will be achieved when MSL examines the entire data theme portfolio, deciding on the themes that will be high priority for improvement (see Objective 5). In essence, this step is moving data value and quality improvement from an implicit to an explicit activity that is tracked in some way.

MSL geospatial managers are in the most appropriate role to track overall commitment to data quality improvement. For example, one could create a chart of the 15 MSDI Data Themes showing whether they are a priority for improvement, if an assessment has been performed, a plan written, and then progress is measured against the actions called for in the plan.

The second level of success is tracking progress for each data theme selected for improvement. The data quality improvement plan for each data theme template contains a status column for each recommended action. One could take all the near-term actions and start by building a timeline for them. Mid-term and long-term actions would be later in the timeline. Then the MSDI Data Theme Leads or the State GIS Coordinator would track progress on this achievement chart.



Data quality improvement is a long-term effort, so future improvements (e.g., the availability of QL1 lidar statewide for the Elevation data theme) may remain incomplete for a long time. The key though is not losing track of the outcome or actions, even if they cannot be started yet.

The third level of success is confirming that data value has improved for a data theme. Ultimately, data quality improvements that increase data value are best measured by the satisfaction of data users. One could devise some form of report card for data consumers. Consumers could be asked to fill these out periodically or at appropriate points in the improvement process. An online survey or follow-up email generated when someone accesses a data theme could give the user an evaluation form to fill out.

The fourth level of success is ensuring that data quality does not deteriorate over time. Data quality improvement should be an ongoing action. Having gone through the effort of improving data quality, it is also important to maintain the standards achieved. As discussed in Objectives 3 and 4 of this plan, and shown in the example plans (Appendix B, Appendix C), data quality improvements should not be singular efforts. The MSDI Data Theme Leads can gather technical and subject matter experts for periodic informal reviews, in which team members are asked whether they think data quality is improving, not changing, or deteriorating. Depending on the responses, the MSDI Data Theme Leads may decide to add actions to the improvement plan and then act on them.

5.2 Refinement

These monitoring points, whether proactive or reactive, represent opportunities to modify the data improvement strategy to produce a greater level of success across the state and can take the form of:

- Working with partners to clarify authority, responsibility, and expectations regarding data improvement
- Identifying successful strategies from other states as well as other levels of government
- Creating new policies, standards, and best practices related to data improvement
- Modify existing policies, standards, and best practices
- Ensure that data improvement activities align with the strategic plan and other business plans

While some refinement can occur through discussion and agreement, other refinement will require following a process such as the one provided in Objective 2, above.

Business plan monitoring is a combination of reviews at regular intervals, such as annually, and reviews because of situational changes. Regular reviews are proactive while reviews caused by situational changes are reactionary but normal given the many moving parts of data governance, an abundance of stakeholders and partners, an ever-changing political environment at all levels of government, funding changes, and shifting statewide priorities.



Hence, refinement can occur in two ways, necessitated by their cause and their timing. *Ad hoc* refinement is caused by an unforeseen event or set of events that require rapid intervention. A situation is presented which requires adjustment to Objectives, Tasks, Resources, or Timeline.

The other type of refinement is routine and planned. MSL should review its objectives, tasks, and timelines for refinement on a regular, recurring basis such as annually. This review should include the addition of new Objectives and Tasks to replace the Objectives and Tasks defined in this document as they are accomplished or completed. It also includes the self-assessment of mission success described above as an annual activity. Ideally, this refinement opportunity would follow MSL's annual review of the GIS Coordination Strategic Plan so that it could reflect adjustments to that document. In any event, refinement usually includes changes to one or more of the following areas:

- **Strategies.** Has the big picture changed? How do the changes affect planned courses of action, including activities from strategic and other business plans?
- **Priorities.** Perhaps events require that objectives or tasks be realigned in time, or that more (or fewer) resources are required due to complexity or a new understanding of criticality.
- **Resource levels.** Resource levels often include human resources, but financial and technical resources may also need to be refined.
- **Objectives.** Are the planned objectives still the right ones to pursue? Should an objective be added or removed, or simply realigned? Should changes be made to the strategic plan or other business plans to better accommodate the objectives of this plan?
- **Tasks.** Tasks are associated with objectives and may require adjustment if an objective is changed.
- **Schedules.** Is the length of time that has been planned to implement an action or accomplish an objective still appropriate given the current environment?



Appendix A. Data Themes Quality Assessment Workbook

Appendix B. Data Value Improvement -Elevation

Appendix C. Data Value Improvement -

Structures and Addresses

Appendix D. Template Dataset Assessment Workbook

Appendix E. Template Data Value Improvement Plan Matrix

Appendix F. Data Value Program

Implementation Plan