



MSDI Best Practices Proposal **DRAFT**

Introduction:

This document is to inform members of the MAGIP Technical Committee and MAGIP Board of Directors of the developing data distribution best practices. Previous years' MAGIP Technical Committees and Technical Committee Chairpersons created best practices focused on Metadata, Montana State Plane Coordinate System HARN Spatial Reference, and the use of Persistent Identifiers. Best practices *“represent efficient and effective ways to accomplish many common tasks that have been tested and proven by a community”*. Best practices may lead to adoption of standards through official channels with authority for enforcement; however they are valuable in and of themselves when accepted by practicing GIS professionals.

While ultimately intended as data distribution best practices for all GIS data, initial implementation will be tested by focusing on the Montana Spatial Data Infrastructure (MSDI) framework GIS layers. MSDI is comprised of fourteen framework themes ranging from Administrative Boundaries to Wetlands (described in detail [here](#)). Some themes, such as Boundaries, are comprised of six individual layers, while others, including Elevation and Orthoimagery, are inherently large, multifaceted raster datasets, both of which result in a large number of hefty, complex datasets, thereby presenting unique storage and distribution scenarios for theme stewards and theme leads. GIS layers housed within MSDI themes are constantly being updated, edited, and changed in some form, all at different rates. To compound the rapidity of MSDI data changes, GIS technology as a whole is continually transforming and evolving. Cumulatively, this situation presents challenging circumstances in serving these datasets in a consistent, efficient, and effective manner. The Montana State Library (MSL), in partnership with the MAGIP Technical Committee, has agreed to test the implementation of these best practices on all fourteen MSDI themes. Implementation scoping will begin upon release of this document, with actual theme by theme implementation as part of the FY13 MSDI work plan. MSL will document the level of effort and costs associated with implementation and report back to the MAGIP Technical Committee to either confirm that the best practices are valid as written or possibly benefit from revision.

The MAGIP Technical Committee conducted a very thorough and critical analysis of the MSDI GIS layers to accurately ascertain the current state and organization of the entire MSDI framework in relation to the proposed best practices. The following best practices, while focused specifically at MSDI GIS layers for the purposes of this document, are designed to streamline and make consistent GIS data obtainment, data use, and enhance communication between data users and data stewards. Said another way, the proposed best practices are intended to make GIS data more available, easier to find, easier to use, and make GIS data more transparent, all of which result in more effective, accurate, and consistently utilized MSDI and other GIS data by the end user. These key concepts, generated from a data user's perspective, should increase the exposure and ultimately the overall use of MSDI GIS layers.

Proposed MSDI Best Practices:

Below is the current list of proposed MSDI Best Practices. Most of the proposed best practices are focused on data distribution and data format. For each proposed best practice, a few generalized bullets document any noteworthy details or outliers corresponding to each MSDI theme, along with the purpose of each proposed best practice.

1. All MSDI layers available in most current ArcGIS geodatabase (GDB) version, most current ArcGIS GDB version minus one, and an open source format.

- For vector data: ArcGISv10 GDB, ArcGISv9.3 GDB, and shapefile, including accompanying tables.
- For raster data: ArcGISv10 GDB, ArcGISv9.3 GDB, and BIL, ASCII, GeoTIFF, or other Geospatial Data Abstraction Library ([GDAL](#)) -supported format
- **Purpose:** Two different ArcGIS GDB versions ensures most if not all functionality stored in the GDB format, such as Geometric Networks, Topology Rules, Relationship classes, etc, is available to the majority of ArcGIS users. For individuals and organizations not using ArcGIS products, the shapefile and BIL or ASCII raster formats can be consumed by nearly any other GIS platform, proprietary or open source. Although the open source formats may lack some functionality, all related data is available and an open source user can still rebuild a similar structure in their respective GIS platform. This will allow MSDI theme stewards and leads to create their desired data output while still providing reasonable and consistent output to all users of the data, while still making all MSDI data available to all GIS users.

2. All MSDI layers available with complete and embedded metadata along with a standalone metadata file in XML format.

- Embedded metadata allows ESRI and most other GIS software packages to read at least one format of the data in question.
- A standalone XML file meets requirements for registration with the MT NRIS GIS portal.
- **Purpose:** “Metadata describes data using terminology that defines potentially disparate data and facilitates consistent collection, indexing, querying, and publishing. Metadata documents content, quality, source organizations, data format and organization, collection schedule, uses, data currency, spatial references, and distribution mechanisms for the data. Keeping spatial metadata records is important. From a data management perspective, metadata is important for maintaining an organization's investment in spatial data. Data users need metadata to locate appropriate datasets. Metadata provides information about the data available within an organization or from catalog services, clearinghouses, or other external sources. Metadata not only helps find data, but once data has been found, it also tells how to interpret and use data. Publishing metadata facilitates data sharing. Sharing data between organizations stimulates cooperation and a coordinated, integrated approach to spatially related policy issues” ([ESRI ArcUser News](#)).

3. All MSDI layers available for download at a statewide extent:

- For all vector GIS data, this is possible.
- For most raster datasets, this is not possible due to size.
 - Although MSDI raster datasets cannot be made available in one statewide extent download, individually parsed extents can be downloaded and mosaiced together.
 - Both MSDI raster and vector datasets should not be explicitly clipped to the Montana state border.

- Natural phenomena such as Hydrography, Hydrologic Units, and Elevation, and others, are unaware of natural borders. Therefore, buffered, or ‘contributing’ areas around Montana should be included in MSDI extents if not already.
- **Purpose:** Providing data at a statewide extent facilitates easier data distribution and lessens confusion on available datasets per framework layer. This also increases the usability of the data since users do not have to download multiple datasets and perform geoprocessing operations against the data. Therefore, if a dataset undergoes changes, only one input must be updated and distributed on the user’s end. Services can also be distributed via a .lyr (layer) file.

4. *Create Web Mapping Services (WMS) or a Geoprocessing service for each MSDI layer.*

- This is possible for all MSDI layers.
- Some MSDI layers, such as elevation, are available as geoprocessing services which enable small geoprocessing jobs to be run via the internet.
- Web mapping services should adhere to OpenGIS Web Map Service (WMS ([OGC](#))) specifications, with complete metadata.
 - ‘Where appropriate’ infers that specific parameter is dependent on type of service. Some service properties will return ‘NULL’ dependant on type.

Purpose: Web mapping services are very easy to share and always (or should) represent the most current version/state of a QA/QC’d dataset. Of course, most web mapping services are designed for cartography or for consumption in another web mapping application. Web mapping services also have the potential to promote the overall, consistent, and correct use of MSDI layers.

5. *Create defined symbology*

- A graphic or document visually showing the preferred symbology, such as Styled Layer Descriptors (SLD) ([ArcGIS Resource Center – ArcGIS Server](#)).
- An ESRI layer file (.lyr) available in the MSDI download package.
 - This layer file should be a grouped layer with at least one sublayer associated with the downloaded data and one sublayer sourcing the MSDI web mapping service.
- Symbology for WMS’s, layer fields, and SLD’s should also include defined scales.

Purpose: Defined layer symbology provides a starting off place for cartographic purposes and encourages consistent use of the data amongst users. This also provides an opportunity for the theme stewards and leads to promote and distribute their data for how they think it should be used. Users can still symbolize and customize the data and its representation for their specific need, though this gives every user a place to start and adds consistency to distribution and overall use.

6. *Have a defined update schedule for each MSDI layer.*

- Defined update schedules are critical for users of the data since many state, federal, and private entities utilizing MSDI data to perform analyses, create policy, and drive decisions.
- Incorporate the model of versioned releases.
 - Include in the distribution file a version number to identify the release.
 - Include a brief log file that generally identifies the difference from the current to previous releases, and or include this log file in the metadata.

Purpose: A defined update schedule per MSDI theme enables more efficient management and distribution, and promotes more efficient, effective, and consistent use of MSDI data. Basically, users know when they need to obtain new data downloads if necessary and check for data changes.

7. *Create ONE download location for all MSDI layers and treat that location as the definitive download source.*

- One download source for all MSDI themes is critical to support and promote the use of MSDI.
- Also, include a standardized feedback mechanism for both the data and delivery system.

Purpose: One definitive download location for all MSDI layers will greatly reduce confusion among users for the most up to date download source. This is also an excellent avenue to identify users of the data, a theme’s download rate and use, and who is utilizing the data. This would also provide a platform for theme stewards and leads to communicate to users important information about their respective data including updates to schema, symbology, geometry, etc. Finally, theme leads and stewards can include disclaimers of use and readme files to ensure correct usage of the data and guide data users.

Ideally, a user would navigate to the MSDI homepage layers would be listed out, and a user could select the download option, which would include the specified FGDB or open source GIS layer, the metadata .xml, defined symbology for the layer, .lyr for a web mapping service, and other pertinent information. Figure 1, below, is an example of what the user might see when trying to download MSDI Structures.

8. Each MSDI download package includes ONLY its respective MSDI theme, and no MSDI theme download package includes any other MSDI theme or data.

- Including only one MSDI theme per download greatly reduces end user confusion on what data/datasets available in the download are MSDI data.

Purpose: At present, some of the available MSDI download packages include other MSDI themes or otherwise related datasets. This can create confusion from the data user’s perspective, since it is somewhat unclear what the MSDI data source actually is. Also, if other MSDI themes and data are inside other download packages, it is unclear if the additional datasets represent the most current state of the data, and whether those data states coincide with each download package.

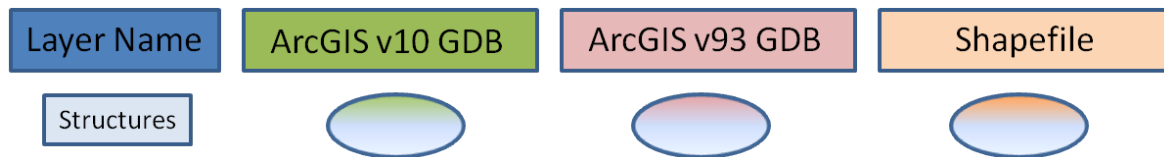


Figure 1. Example of data download page for MSDI Structures.

Current State of MSDI Framework Layers:

Below are brief descriptions of the current state of each MSDI layer as they pertain to the proposed best practices listed above. These descriptions are meant to identify possible solutions to ensure all MSDI framework layers meet the criteria listed above.

Best Practice 1:

- See Table 1 for a list of the currently available data formats.
- No theme is served in all three data formats.

Best Practice 2:

- All MSDI layers contain embedded metadata upon download, though only a handful include a standalone .xml file per download package.
- Table 1 lists the metadata availability for each MSDI layer.

Best Practice 3:

- For vector data, Geology and SSURGO Soils are not currently available at a statewide extent.

- Geology is served through Montana Bureau of Mines and Geology (MBMG) [here](#). can As an interim solution they could be downloaded as a list of 100K, 250K, and 500K tile extents and merged together, extract the downloaded zip files, run “Import from Interchange” geoprocessing tool to convert the .e00 interchange files to coverages, then run “Feature Class to Feature Class” geoprocessing tool, then merge all the pieces together and apply defined symbology. We have been apprised that MBMG is interested in pursuing a geodatabase solution that could serve as a better long-term solutions
- SSURGO Soils is also a series of separate downloads typically available as county polygons (available for download [here](#)). Several geoprocessing scripts are available [here](#) to rename the feature classes and tables inside each PGDB to enable a statewide distributed layer, so this approach can most likely be automated.
- Landuse_Landcover is currently being served up at a statewide Extent.
- Elevation is approximately 55 GB in total in 10 m National Elevation Dataset (NED). Elevation can probably be served up in buffered tiles or as a geoprocessing service for most applications and analyses.
- Orthoimagery is of course too large (over one terabyte) to make available as a statewide internet download. However, orthoimagery is available [here](#) for 2009 imagery, though 2011 NAIP is not yet available. A potential solution is to use the NAIP index shapefile (available [here](#)) and the NAIP Image Service to select small enough areas for AOIs to select and buffer the NAIP index tile, clip out that section from the NAIP raster, and store to a local drive, very similar to how NRIS currently serves the downloadable 2009 imagery.
- Table 1 lists the MSDI datasets available for download at a statewide extent.

Best Practice 4:

- At the time of this report, arcpy code development and automation to describe available web mapping services and specific parameters is still being developed. This code can be shared with MSL to assist in implementation.
- Most BMSC maintained MSDI datasets have either geoprocessing or web mapping services.
- Table 2 lists MSDI datasets with web mapping services (WMS), WMS metadata, and WMS service properties availability.

Best Practice 5:

- Only LandUse_LandCover contains a .lyr file for a service and a standalone .xml file for metadata purposes.
- Table 2 lists the MSDI datasets that contain .lyr files with the distribution package.

Best Practice 6:

- Currently, no known or specific update schedule has been identified for any MSDI dataset.

Best Practice 7:

- Currently, the closest thing to an organized MSDI data download page is the MSDI Coordination site located [here](#).
- Many BMSC MSDI GIS layers are available for download in varying formats from the MSDI Coordination site, but many links are pointers to other locations, where a user must continue to navigate through another site to download sought after data in various formats.

Best Practice 8:

- Currently, some duplication of data exists for some MSDI data downloads.
- Within Administrative Boundaries, several datasets are duplicated in the State and Counties download packages.
- The Counties feature class is also in the Cadastral data download.
- Hydrography and Hydrologic units contain the identical datasets per download.
 - The concern with duplication per download is whether the duplicated data is of the same version/state for each download.
 - End user confusion will be reduced greatly if one download package exists for each MSDI theme.

Conclusion:

Montana has long been blessed with framework efforts that have been widely held up as shining examples of how to build a state's spatial data infrastructure. However, MSDI framework layers can benefit greatly from a better organized and more easily accessible data distribution platform. MSDI data that is easily accessed from one central location encourages users of these data to continually obtain the data, services, and symbology from one source. This will create an environment for theme stewards and leads to communicate with users of their data, thereby creating transparency between users and creators/maintainers of the data. At the same time, implementing data distribution best practices for MSDI will serve as a model for all GIS professionals in Montana that publish their data.

Lastly, it is recommended to the MAGIP Board that after careful review and potential revision of the proposed MSDI best practices, that the proposed best practices and any accompanying methodology be put up to the MAGIP Board for a vote of approval and be presented to the Montana Land Information Advisory Council for their consideration and support as general GIS data distribution best practices for the state.

References:

[Administrative Boundaries Page](#)
[ArcGIS Layer Method Help](#)
[BMSC Data Download Page](#)
[Implementing European Metadata Using ArcCatalog](#)
[GDAL](#)
[MAGIP Best Practices & Standards Page](#)
[MAGIP MSDI Page](#)
[MBMG GIS Digital Data Links Page](#)
[MSDI Home Page](#)
[NAIP Download Page](#)
[NRCS FTP Directory for Scripts, Tools, Etc.](#)
[Open Geospatial Consortium \(OGC\)](#)
[SSURGO Soils Data Access Page](#)

Figures:

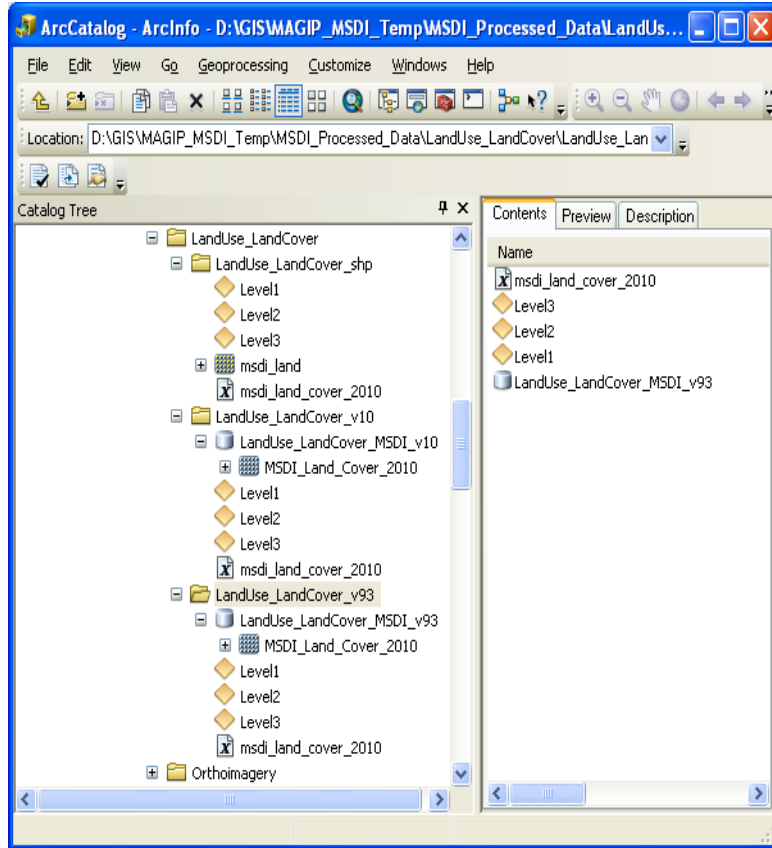


Figure 2. Example of Land Use/Land Cover data formats after download, extraction, and geoprocessing.

Tables:

Table 1. Data download format breakdown for MSDI layers, metadata availability, and if the current download package. Note: Data download availability is based off the original download packages supplied by the theme steward. 'NA' means not applicable' to that theme.

Layer Name	Opensource format	Current ArcGIS Version	Current ArcGIS Version - 1	Metadata	Statewide Format
Boundaries_Counties	Yes	No	No	Yes	Yes
Boundaries_State	No	No	No	Yes	Yes
Boundaries_FireDistricts	Yes	No	Yes	Yes	Yes
Boundaries_Reservations	Yes	No	No	Yes	Yes
Boundaries_Cities&Towns	Yes	No	Yes	Yes	Yes
Boundaries_TIFD	Yes	No	Yes	Yes	Yes
Boundaries_SchoolDistricts	Yes	Yes	No	Yes	Yes
Cadastral	Yes	No	Yes	Yes	Yes
Elevation	NA	NA	NA	Yes	No
Geodetic Control	No	No	No	Yes	Yes
Geographic Names	Yes	No	No	Yes	Yes
Geology	No	No	No	No	No
Hydrography	No	No	No	Yes	Yes
Hydrologic Units	No	No	No	Yes	Yes
Land Use/Land Cover	NA	No	Yes	Yes	Yes
Orthoimagery	NA	NA	NA	Yes	No
Soils	Yes	No	No	No	No
Structures	No	Yes	No	Yes	Yes
Transportation	No	Yes	Yes	No	Yes
Wetlands	Yes	No	No	Yes	Yes

Table 2. Web mapping service (WMS), WMS metadata, defined symbology, and layer file (.lyr) availability with current download package. If an MSDI dataset has a service, then Defined Symbology availability is 'Yes' since that .lyr file will be created from the WMS service and can be applied to the data download package.

Layer Name	Service Available?	Service Metadata Available?	Defined Symbology?	Symbology Available w/Download?
Boundaries_Counties	Yes	No	Yes	No
Boundaries_State	Yes	No	Yes	No
Boundaries_FireDistricts	Yes	No	Yes	No
Boundaries_Reservations	Yes	No	Yes	No
Boundaries_Cities&Towns	Yes	No	Yes	No
Boundaries_TIFD	Yes	No	Yes	No
Boundaries_SchoolDistricts	Yes	No	Yes	No
Cadastral	Yes	No	Yes	No
Elevation	No	No	NA	No
Geodetic Control	Yes	No	Yes	No
Geographic Names	No	No	No	No
Geology	No	No	No	No
Hydrography	Yes	No	Yes	No
Hydrologic Units	Yes	No	Yes	No
Land Use/Land Cover	Yes	No	Yes	Yes
Orthoimagery	Yes	No	Yes	No
Soils	No	No	Yes	No
Structures	Yes	Yes	Yes	No
Transportation	Yes	No	Yes	No
Wetlands	No	No	No	No