PHILCO
DIVISION OF GEOSPECIAL SCIENCE

REPORT

ON THE QUALITY OF GOVERNMENT INSITUTION'S GIS DATA SHARING SERVICES

FORWARD

This report will discuss sources of free GIS data shared by State and Federal Institutions within the United States. The will specifically outline good data sharing practices from bad data sharing practices. At Philco, We have three pillars of good data presentation that these institutions will be judged by. These pillars are:

ORGANIZATION
People need to be able to actually find the data, Is you data organized properly, is the website findable, and is there a way to search for specific datasets?

UNDERSTANDABILITY
People need to understand what they are looking at when they get the data. Is there documentation for what your fields in tables mean? Can the fields make sense on their own?

ACCESSIBILITY
People need to be able to access the data. Are the links on your webpage current, do paths actually go to where data is being stored? Is any of the data locked behind some kind of specific login or paywall?

In order to judge these institutions, we will be scoring them using a rubric we use to score our own data before presenting them to the public.

SCORE	ORGANIZATION	UNDERSTANDABILITY	ACCESSIBILITY
5	The data is shared using well designed ArcGIS Open Data website or independent well designed website that as an effective search feature	The data is well documented and has the tables and fields are self-explanatory	It is completely accessible, finding the data was easy, grabbing the data was simple and available in multiple formats and free
4	The data was shared using ArcGIS Open Data or an equivalent data sharing tool. Or hosted on a website with a working search engine	The tables, file names and databases were selfexplanatory or could be understood with little effort. The databases and tables had some kind of documentation	The data could be found with and download in at least 2 different formats
3	Data was shared on a website with a search engine that works effectively	Table, files names or fields were self-explanatory but had little or no documentation	Data could be freely and easily accessed but only had one format or the website had some broken links
2	The data was shared in a giant web table with no way to search	Table and file names look significant effort to understand had no documentation	Data could be accessed but many links were broken
1	The data was shared on some webservers directory browsers no effort put into webpage construction or organization	The file names are proprietary or do not make sense and have no documentation	Data could not be access and was locked behind a login or paywall (Usually Universities lock files, This is also on Automatic 0)

As examples we will examine the following state institutions:

Using all of these sources, we great a simple map in order to illustrate the useability of these services.

UTAH GEOSPATIAL RESOURCE CENTER

The UGRC is the definition of well-designed data presentation. It has a website that has data effectively organized into different sections (Using these very cute little Periodic table looking buttons).

It has an effective search feature, which I found was a blessing on my search. Every piece of data contained elaborate documentation about every dataset along with multiple file types as well as a preview/download using ArcGIS open data. All the data was freely accessible and finding it was easy, all it took was a simple google search and I was able to get all the data I needed. I give the UGRC Library a perfect **15/16**, scoring a prefect **5** in **Organization, Understandability and Availability**.
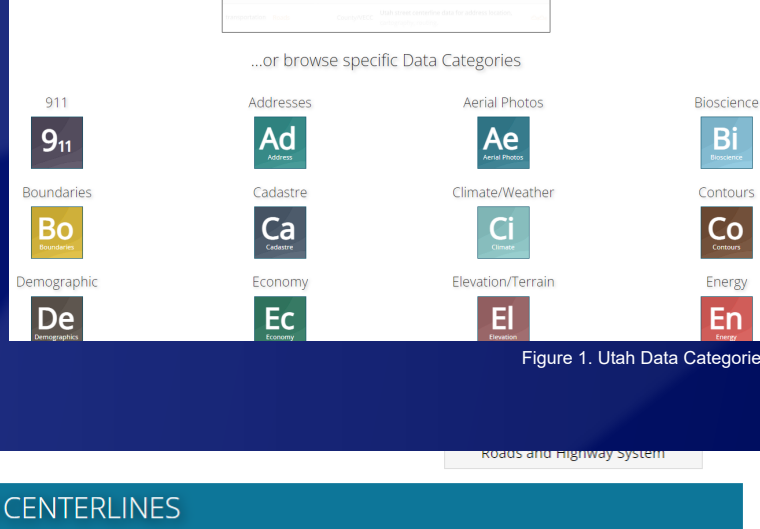

Figure 1: Utah Data Categories

Figure 2: Utah Data Documentation

Using this data, we created a basic map of North Utah(Note: Both States share GIS Data)
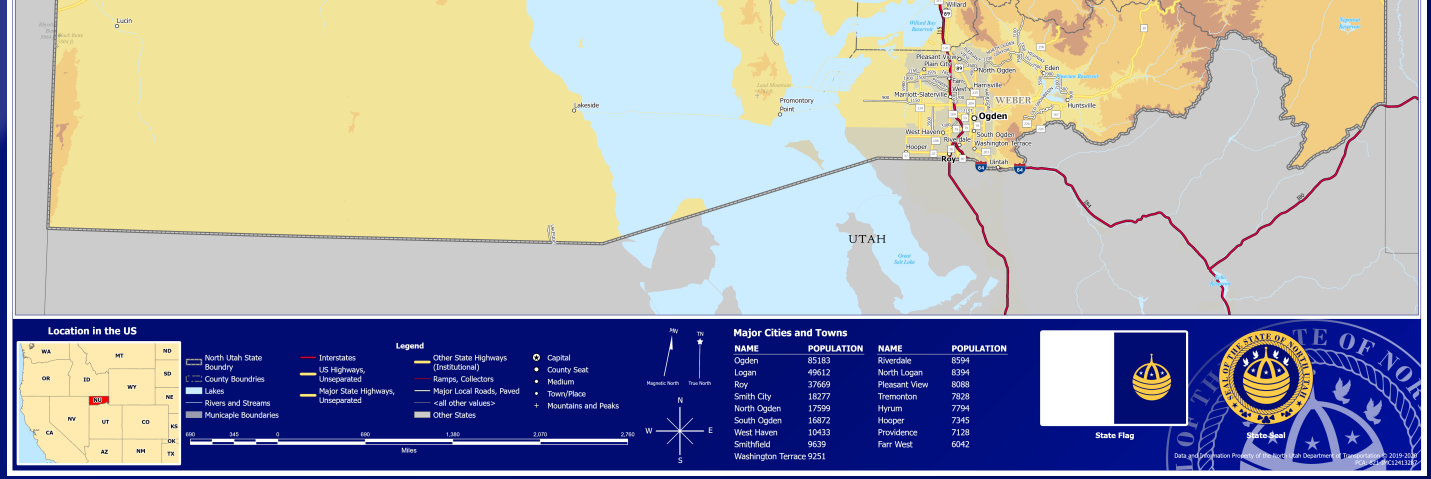
MARYLAND AND DELAWARE GIS CATALOGS

To create the following me I used two data sources, both the Maryland and Delaware State GIS Libraries. I will judge both independently.

The Maryland State GIS Library was well organized, although it required a bit of digging to find all the data I needed. While it had a search engine, it sometimes didn't find what I needed based on my query even though the data would be found based on a more general slightly different query. For this I give it a 4 as it was built on ArcGIS hub and was well modified to its purposes.

Where the Datasets fall short is the database documentation. It took me a good amount of time to document the codes that were used for the Maryland files, I had to apply a different symbol to each and correlate with google maps to understand what the roads were. For this, I give it a score of 2. I give it a score of 5 for Availability and Accessibility, as all the data was freely downloadable and accessible.

Giving it a total score of 11/15

Figure 3: Maryland Road Prefixes

Figure 4: Delaware Road Types

The Delaware datasets had a similar search problem as the Maryland Datasets, so I give it a 4 there was well. Where they really surpass the Maryland Datasets is the readability of the datasets. Most if not all the fields were readable, except for fields that related to other tables. But it lacked any kind of documentation, because of this I have to give it a 3 in Understandability. Like the Maryland GIS library it was easily accessible and gets a score of 5.

Giving it a total score of 12/15

Using this data, we created a map of Saint Tersa County in Maryland
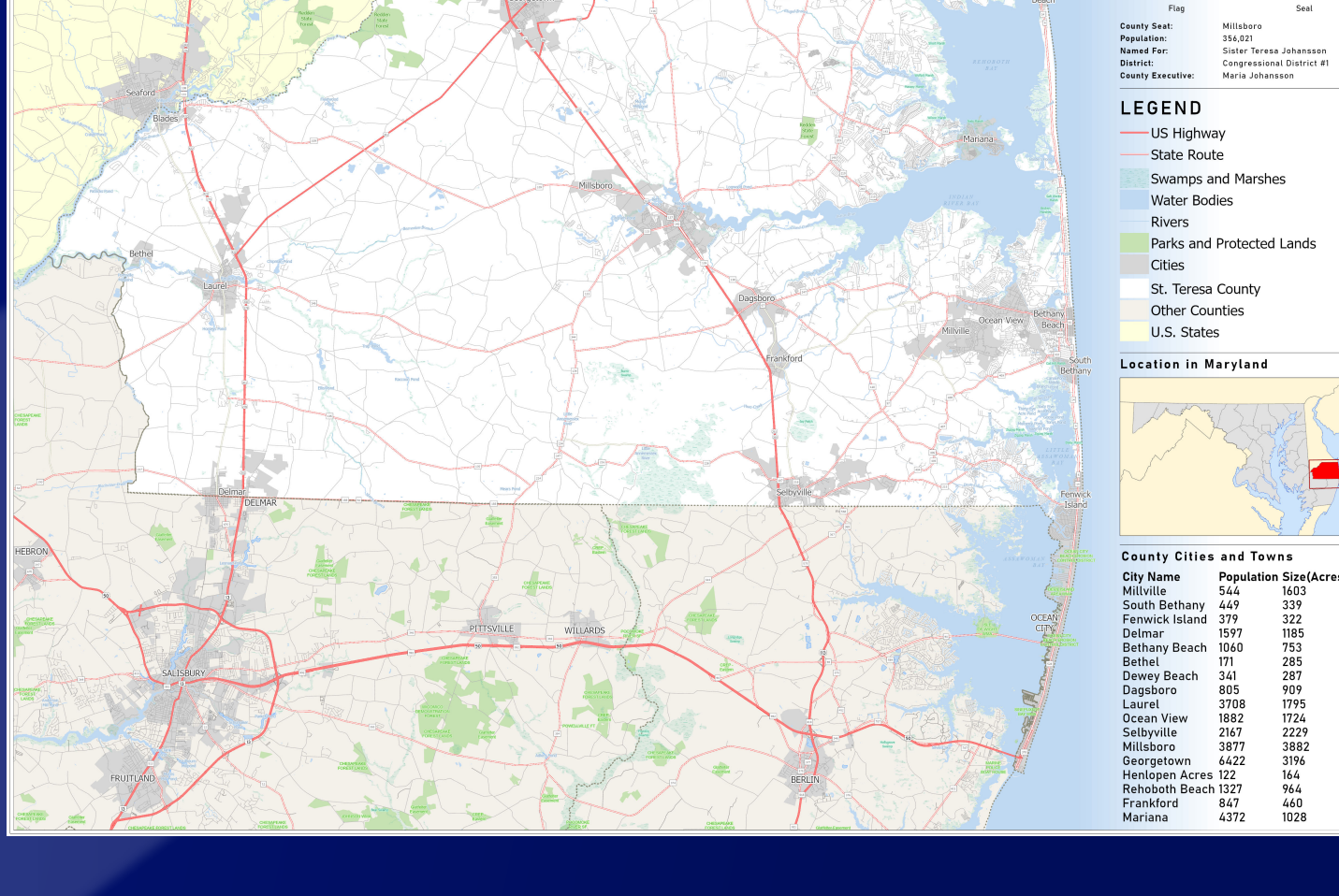
OREGON GEOSPATIAL DATA LIBRARY

While I have the utmost respect for my state and the work that people that work in the Oregon Geospacial Enterprise Office. I really dislike the Oregon Geospatial Data Library. While on the surface it may look clean, some of its fundamental properties don't work as well as they could or should. The biggest reason for this is that the website is mostly designed and managed by Oregon State University and does not use ArcGIS Open Data. In my experience, the school is infamous for its lack of web construction skills and maintenance.
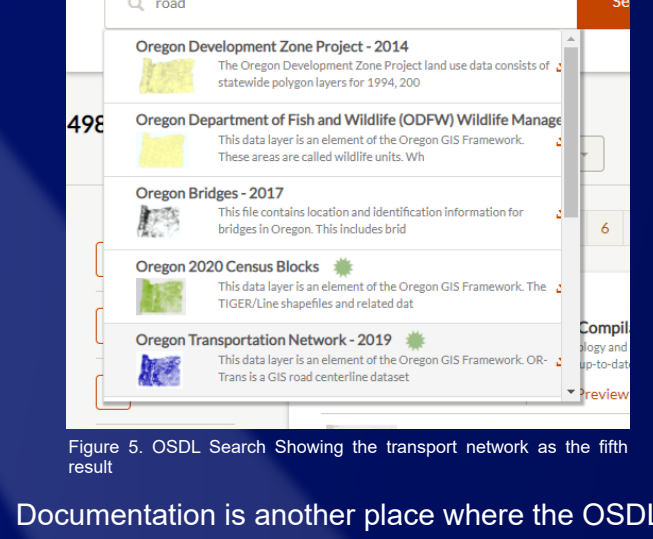

Figure 5: OGDL Search Showing the transport network as the 5th result

First Organization, While the data is fairly well organized, trying to get to that data can be tedious. The search engine does not work as well as it could, for example, searching up roads, the first thing that might come up is Land usage or fish and wildlife boundaries. A search for Transportation might yield better results since that is exactly what the geodatabase is called. There don't seem to be any tags that allow you to find items that are close enough to the search field. Another notable mention is the many of the previews don't work which is more of a nitpick than anything. I give the site a 3 only because the search engine doesn't work as well as it could, which I think is a major component of a good public data sharing site.

Documentation is another place where the OSDL falls short. Most feature classes contain very short descriptions of what the data is and what it could be used for. Or they are the complete opposite and have a massive wall of text explaining licensing and a few bits of what each field means in the table. It doesn't help that the fields are also mostly acronyms or can't make sense without documentation. For this I give it a score of 2.

For Accessibility I give it a score of 3. While the data was freely available most of it was only in one format of a .shp file.

Giving it a total score of 8/15

Description

This data layer is an element of the Oregon GIS Framework. OR-Trans is a GIS road centerline dataset compiled from numerous sources of data throughout the state. Each dataset is from the road authority responsible for (or assigned data maintenance for) the road data each dataset contains. Data from each dataset is compiled into a

Figure 6: Description for the Oregon Transportation Network

Using this data, we created a map of Saranoma County, Oregon

One that that you might have notice is that most of these data libraries lack of thing, and that's documentation. No matter what platform they use whether it be ArcGIS Open Data or something else, they all do not document their database. Documentation is extremely important when sharing data to the public. If people don't know how to work with your data it slows down workflow because they either must take time to understand or contact the developers to get documentation. So, remember:

ALWAYS DOCUMENT YOUR DATABASE

Or your database could end up here!