Appendix E: Data Sharing Platforms
Introduction and Examples

The following introduction of data sharing platforms is intended to provdie the Regional Data Advisory Committee and staff with an overview of data sharing technologies that have been noted by committee members, identified by HUD technical assistance staff, and others examples that are notable within the local or greater region.

Geoportal

Leading the open source conversation is Geoportal. This is because Geoportal is already being used by agencies within the region and is something that the Regional GIS Partnership is already investigating. Geoportal is a type of web portal used to find and access geographic information (geospatial information) and associated geographic services (display, editing, analysis, etc.) via the Internet. Geoportals are important for effective use of geographic information systems (GIS) and a key element of Spatial Data Infrastructure (SDI).

Geographic information providers, including government agencies and commercial sources, use geoportals to publish descriptions (geospatial metadata) of their geographic information. Geographic information consumers, professional or casual, use geoportals to search and access the information they need. Thus geoportals serve an increasingly important role in the sharing of geographic information and can avoid duplicated efforts, inconsistencies, delays, confusion, and wasted resources

Modern web-based geoportals include direct access to raw data in multiple formats, complete metadata, online visualization tools so users can create maps with data in the portal, automated provenance linkages across users, datasets and created maps, commenting mechanisms to discuss data quality and interpretation, and sharing or exporting created maps in various formats. Open portals allow the possibility for users to contribute datasets as well.

Figure 1: Lane County's Geoportal website

Lane County has been working with Geoportal in recent years. The **Lane County Geoportal** is a web-based application running on an intranet server using LDAP (Lightweight Directory Access Protocol) user authentication (which establishes security protocols) connected to a Microsoft SQL server database.

The Lane County Geoportal has been customized to have a similar look and feel as other standard county web applications and uses a familiar data directory organization scheme as well as a local map service as a geographic reference. There are presently about 110 metadata records in the Lane County Geoportal and the growth of the metadata is an ongoing effort. For
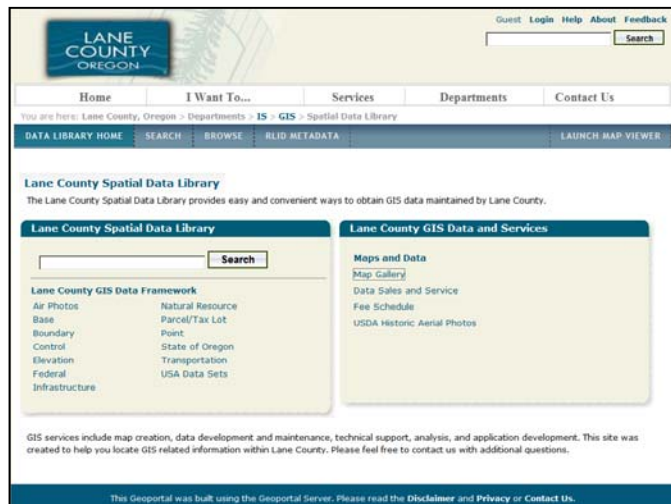
regional staff the Lane County Geoportal application can be viewed at:
http://lcapp11:8080/geoportal

Upon display, the user will be presented with the home page that offers the ability to perform metadata record searches or access other County GIS information. Geoportal is also offered as a means to connect to other "published" metadata resources so that users can discover and access external records. There are also features in the Geoportal that can be implemented to provide data downloads using the metadata application as an initial data discovery point. Although Geoportal is designed for spatial data, non-spatial data can also be added and made available for download.

Lane County staff prepared a more in-depth summary of Geoportal, including an evaluation of strengths and weaknesses and potential use in addressing qualitative and non-spatial data. That summary is included as Attachment A.

The **Oregon Spatial Data Library** was built on the ESRI Geoportal Server, a joint effort between the Department of Administrative Services Geospatial Enterprise Office and Oregon State University. Currently, more than 200 spatial datasets are accessible from this online library. These datasets serve as "base data" for a variety of Geographic Information System (GIS) applications that support research, business and public services. The targeted primary users for the Oregon Spatial Data Library are staff of agencies, but secondary users include the general public, consultants, educators, and students. This could be a very useful model for future data sharing efforts using Geoportal.

Figure 2: Cropped screenshot from the Oregon Spatial Data Library

World Wind
World Wind is a free, open source application programming interface (API) for a virtual globe. The application was written so as to be cross-platform, intuitive, precise and flexible in supporting many needs. World Wind is used worldwide for a number of purposes including visualizing cities and terrain, analyzing spatial data and simply helping people understand earth systems. Unlike Google Earth, World Wind is a software development kit (SDK) with a powerful geographic rendering engine well suited for large remote sensing data sets, complex 3D renderings and large area regional networks (e.g. hydrography across the Pacific Northwest, transportation systems across the state and global imagery).

Figure 3: Image below displays hi resolution imagery and over 16 million points LiDAR (Light Detection and Ranging), two forms of remote sensing data, in World Wind on a standard laptop
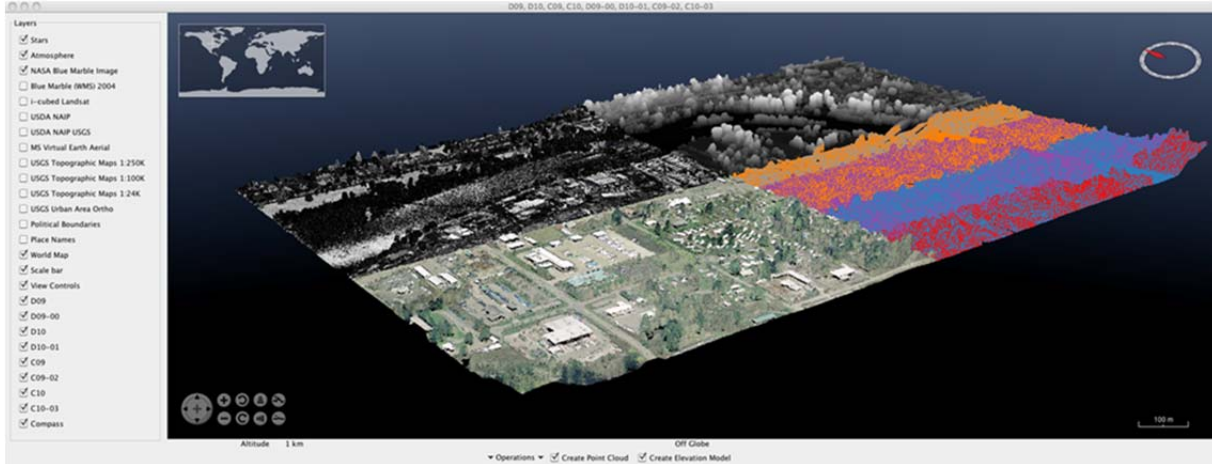
If the region develops modern web-based portals, World Wind presents an opportunity to support raw data and metadata into a single on-line visual virtual globe. Web browser balloons could display HTML, JavaScript, and flash content from metadata repositories. Viewers and layer managers can support loading, displaying and interacting with various mapping formats such as KML, SHP, GML, GeoJSON, GeoTIFF, MrSID, etc. as well as live streaming information (such as WMS) from servers across the region. Collaborative development with World Wind could be a natural next step to investigate if initial efforts with Geoportal succeed.

OpenColorado (Colorado) *http://data.opencolorado.org/group/drcog*
OpenColorado is hosted by the Denver Regional Council of Governments. The vision of Open Colorado is to support a transformation that will lead to a "simple, beautiful, and easy-to-use government." The primary resource is the OpenColorado data catalog. The data catalog allows any municipality, county, government agency, nonprofit, or individual to share open data with the public. The Catalog allows governments and community organizations to provide open access to data on their own websites through a centralized online data catalog (clearinghouse). One can search for data based on groups, data formats or "tags" (subjects). OpenColorado is powered by CKAN an Open Source provider http://ckan.org/.

Greater Portland Pulse *http://portlandpulse.org/*
Greater Portland Pulse uses data and dialogue to encourage coordinated action for better outcomes across the region. With Weave as the main analytical platform, the site uses indicator data to show where the region is successful and where it's lagging behind in the areas of economy, education, health, safety, the arts, civic engagement, environment, housing and transportation. The indicators often reflect who's being left behind and how communities and the region are impacted as a result. Greater Portland Pulse makes a significant effort to tell data stories and even has a theme based drop-down menu specifically dedicated to "Data Stories."

Figure 4: Cropped screenshot from Greater Portland Pulse Website

Open Application Programming Interfaces (APIs)

There is a rich background of application development within our region. RLID.org provides a good example of such applications. However, providing data to support openly developed APIs can be an effective way to utilize untapped social and commercial value in developing applications. When access to data is provided openly to the masses (in bulk or otherwise), an obstacle to data delivery innovation is arguably removed. More individuals and entities working to identify and address data issues can only result in increased innovation. The standard "Open Data" model suggests that data must be free, but open data services (such as an API) can be charged for. This provides one of the most immediate opportunities around open data by incentivizing innovation. Figure 27 presents examples of applications developed as a result of the Open Data approach taken by governments in the Denver Area through OpenColorado.

Figure 5: Examples of applications (APIs) developed through OpenColorado
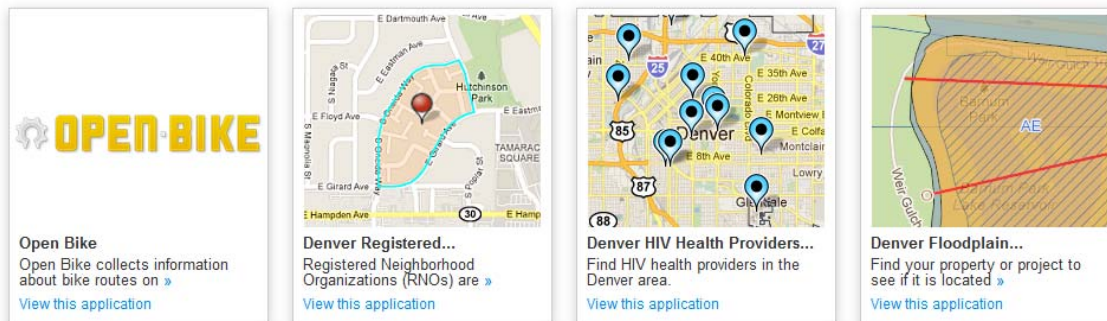
Data Dashboard Approaches

One data sharing concept which was well received by the Regional Data Advisory Committee was that of "Data Dashboards." Generally speaking, the idea of Data Dashboards is to provide at-a-glance views of key performance indicators relevant to a particular objective. This approach is becoming increasingly commonplace in numerous arenas, including business, education, public health, and other public administration. The term "dashboard" originates from the automobile dashboard, where drivers monitor the major functions at a glance via the

dashboard. Just as in a car, dashboards allow users to quickly know if something is wrong or if something is right. Following are a few examples of the dashboard concept in application.

Kansas City, KCStat https://kcstat.kcmo.org/

KCStat focuses on the City's six strategic priority areas: Public Infrastructure, Economic Development, Public Safety, Healthy Communities, Neighborhood Livability, and Governance. One additional priority cuts across all outcome areas and concentrates on Customer Service and Communication. The KCStat Dashboard provides an "at-a-glance" view of each priority's current status. KCStat is powered by Socrata, and Open Source provider http://www.socrata.com/

Figure 6: Example Dashboard figures for Priority Areas on KCStat

MetroBoston Data Common www.metrobostondatacommon.org/snapshots/

The MetroBoston DataCommon provides accessible information about the Boston Metropolitan Region's people and communities through a variety of topics including arts, education, environment and transportation. It is a resource for residents, stakeholders, planners, city and town officials, educators and journalists to explore data and make informed decisions. The site is very sophisticated and multi-faceted, but a "snapshots" component of the site provides information in a "dashboard" format. Users can isolate data for specific geographies within the region and can view user friendly and current summary statistics (and maps) for a number of topics. DataCommon does not go so far as to explicitly suggest whether key indicators are "on task" or "need improvement" as KCStat does (see Figure X for an example).

Figure 7: Examples of MetroBoston DataCommon "Snapshot" functionality
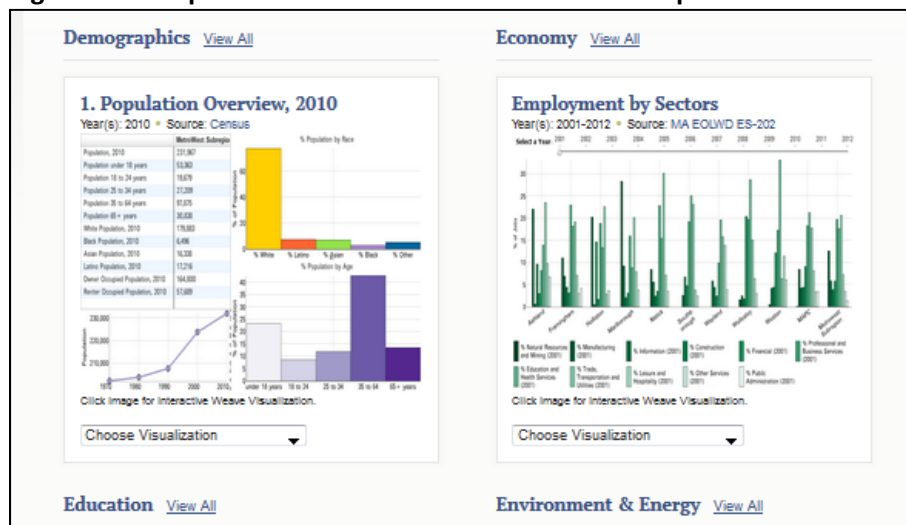
The MetroBoston DataCommon supports a much larger region (in size and population). It does, however, provide some good examples for potential dashboard elements and approaches for a future data sharing strategy.

## Data Mapping Approaches (GIS Portals)

There is significant local interest in providing more interactive mapping resources. This addresses a frequently noted concern that data can be difficult to weave into a poignant "story." Maps can be an effective to convey or discover such data "stories." There is significant existing web mapping capacity within GIS divisions in the region, as well as existing tools (e.g. RLID Maps, LaneCountyMaps, MapSpring). And there are a number of innovative and relatively accessible tools and examples to expand data mapping in the region. Mapping-only platforms are interactive, though they tend not to have a user-friendly interface. GIS Portals typically do not provide space for collaborative analysis or interpretation of data. Following are a few examples.

Regional Equity Atlas 2.0, Portland (Coalition for a Livable Future)
https://gis.oregonmetro.gov/equityAtlas/

The Regional Equity Atlas is a tool that was designed to enables understanding of how well different neighborhoods and populations across the Portland region are able to access resources and opportunities to meet their basic needs. Through the use of maps and data, the Atlas provides a visual depiction of disparities and illuminates how the benefits and burdens of growth and change are distributed across the region. One advance utility of the Regional Equity Atlas 2.0 is that the web-based Mapping Tool enables users to make and edit their own maps (in addition to providing finished thematic maps and preliminary analysis of many key issues.)

**Figure 8: Screen shot of "Proximity to Community Spaces"
within Regional Equity Atlas 2.0**
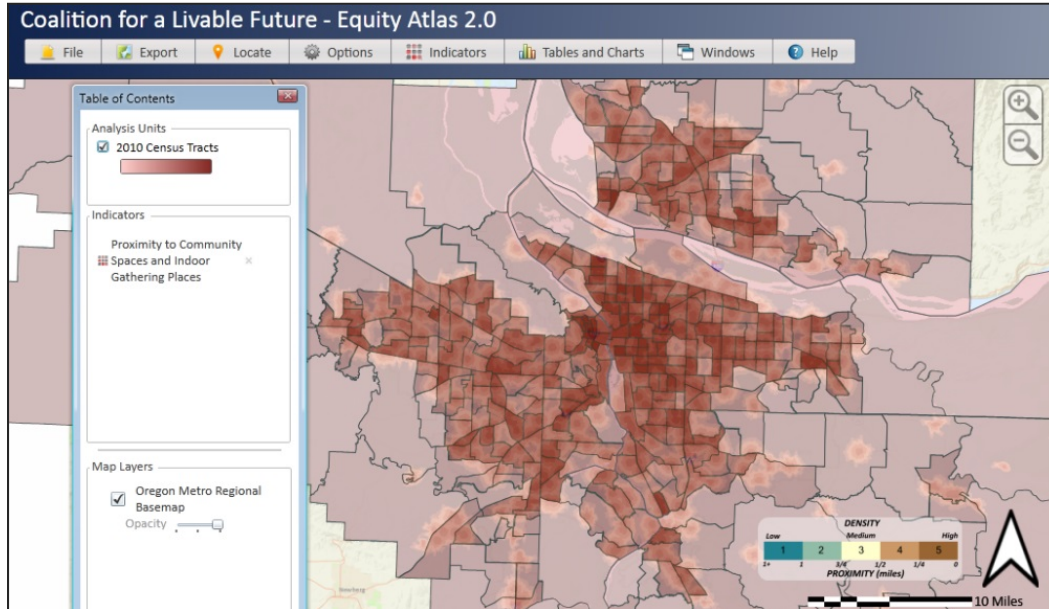
InstantAtlas (E.S.R.I.)

InstantAtlas™ enables information analysts and researchers to create highly-interactive dynamic and profile reports that combine statistics and map data to improve data visualization, enhance communication, and engage people in more informed decision making. InstantAtlas enables numerous dynamic reports with built in templates, including time series alternatives. The service is available under license (desktop or server).

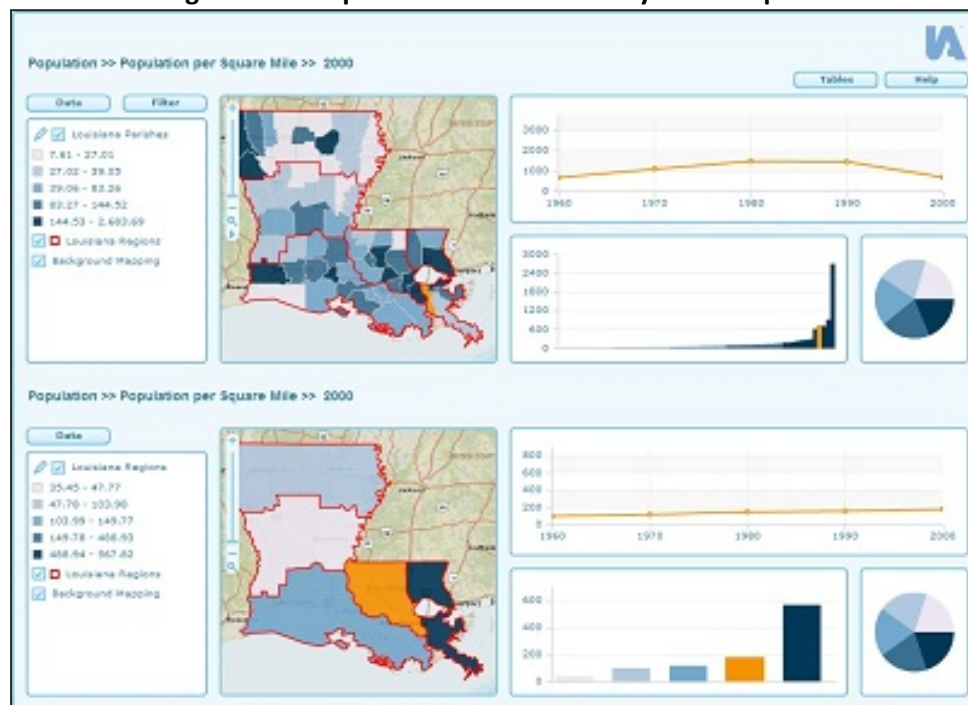**Figure 9: Example of the InstantAtlas Dynamic Reports**

Table 1: Summary of Sharing Frameworks outlined in Section 4.2

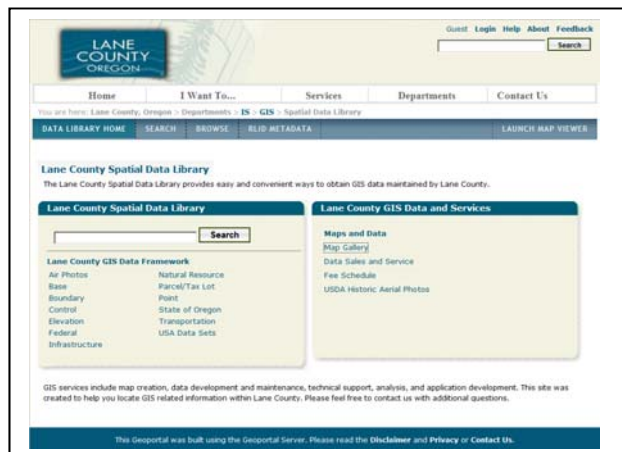| Samples	Framework Type	Open vs Vendor	Powered By
Technologies			
Geoportal Server	Inventory/Catalog or Warehouse	Open (Built on ESRI software)	ESRI
World Wind	Data Commons	Open	World Wind (NASA)
InstantAtlas	Data Commons(Mapping focus)	Vendor	ESRI
Applications			
OpenColorado	Inventory/Catalog or Warehouse	Open	CKAN
Greater Portland Pulse	Data Commons	Open	Weave (in part)
KCStat	Data Commons (Dashboard)	Open	Socrata
MetroBoston DataCommon	Data Commons	Open	Weave (in part)
Regional Equity Atlas 2.0	Data Common (Mapping)	Unknown	Portland Metro
Key Existing Local Examples (Mapping)			
RLID Maps	Data Commons	Vendor	ESRI (Geocortex Essentials)
MapSpring	Data Commons	Vendor	MetroPlanning Inc.
LaneCountyMaps	Data Commons	Unknown	Unknown

Geoportal and Regional Data Sharing

Background

In the fall of 2011, Lane County GIS staff began investigating the possible use of the Geoportal software as a tool to develop metadata for GIS data holdings. Geoportal is Open Source software developed by ESRI (Environmental System Research Institute) that offers a metadata web-enabled product capable of connecting to several varieties of relational databases. Open source software is "software that can be freely used, changed, and shared … by anyone. Open Source software is made by many people and distributed under licenses that comply with the Open Source Definition." ESRI chose to develop the Geoportal as an Open Source offering in order to provide improved integration with other systems, promote collaboration, and meet government and individual agency requirements. Geoportal is highly customizable and initial investigations indicated that it could be a viable tool for developing a web-enabled platform for creating and maintaining metadata in a shared environment.

An initial test of Geoportal was implemented on a standalone server running Apache Tomcat 6.x as the Java Servlet (web container) and PostgreSQL 8.x as the database engine. Simple user authentication (no user login/password or security) was used during the test. And in keeping with our local decision to develop metadata using the FGDC (Federal Geographic Data Committee) standard, the test Geoportal was implemented using a slightly modified FGDC standard; although Geoportal does support other metadata standards such as the International Organization for Standardization (ISO) and the Infrastructure for Spatial Information in Europe (INSPIRE). After about 10 months of testing, Lane County staff determined that Geoportal was a viable platform for creating a local metadata application.

Lane County has since ported the Geoportal application to an intranet server using LDAP (Lightweight Directory Access Protocol) user authentication (establishes security protocols) connected to a Microsoft SQL server database. The Lane County Geoportal has been customized to have a similar look and feel as other standard county web applications and uses a familiar data directory organization scheme as well as a local map service as a geographic reference. There are presently about 110 metadata records in the Lane County Geoportal and the growth of the metadata is an ongoing effort.

For regional staff the Lane County Geoportal application can be viewed at:
http://lcapp11:8080/geoportal

Upon display the user will be presented with the home page that offers the ability to perform metadata record searches or access other County GIS information.

Geoportal is also offered as a means to connect to other "published" metadata resources so that users can discover and access external records. There are also features in the Geoportal that can be implemented to provide data downloads using the metadata application as an initial data discovery point.

Other Metadata Tools

A regional metadata application currently exists that grew out of an effort to build an Access-based application to document regional GIS data holdings. The Access application was also organized around the FGDC standard and was well supported by Lane Council of Governments and championed by the City of Eugene. This application has since been "web-enabled" and contains almost 400 metadata records; most of them related to Eugene data. The "RLID Metadata" application is accessible from the Lane County Geoportal via the "RLID Metadata" tab:

http://lcapp11:8080/geoportal/catalog/RLIDmdlink/RLIDmdlink.page

Lane County never wholly embraced the RLID Metadata application since initial efforts to add records through the Access application were thwarted due to application incompatibilities between different versions of Access (County on one version of Access, application running on a more current version) and regional network problems gaining access to that application.

ESRI's ArcCatalog product has a built-in metadata editor and was tested by Lane County as part of a work flow process to build initial metadata records in ArcCatalog; export to an .xml FGDC standard file; and then import to the Geoportal. Since the Geoportal already has a built-in metadata editor it was determined that using ArcCatalog as an initial data entry point only added additional (and unnecessary) steps to the work flow process. (*It should be noted that this work flow process is still under review by Lane County GIS since Geoportal also provides an "add-on" tool called the Publisher Client which offers the ability to create metadata in ArcCatalog and then, using the Publish Client, publish the metadata record to Geoportal. Metadata validation is provided by the Publish Client which can catch errors before the record is committed to the Geoportal database.)*

ESRI has also provided an update to the FGDC metadata editor "plug-in" to ArcCatalog 10.x. This was also tested and while it offers the advantage of conforming to the FGDC standard it met with the same work flow issues mentioned above.

Lane County GIS staff also tested the Environmental Protection Agency (EPA) metadata editor tool (EME). This tool had a steeper learning curve and presented the same work flow problems mentioned above.

For Lane County the Geoportal application offers an opportunity to document data holdings in a standardized environment, capable of being easily shared and presented back to the user in an easy-to-read format. It also offers greater security, a peer review process, and an administration tool for managing the metadata records. In addition it keeps the metadata entry work flow process clean and simple. Geoportal is essentially being used by Lane County GIS staff as a "one stop" solution for entering, updating, and managing metadata records.

Geoportal – Strengths and Weaknesses

The Lane County experience with Geoportal implementation has indicated a number of concerns and also highlighted many advantages over existing local metadata applications.

WEAKNESSES	
Ease of implementation	ESRI has provided a nice user guide for getting a basic geoportal up and running quickly. However, implementing some of the more advanced features (such as security protocols) and even customizing the user interface, can be daunting.
Technical help	Initially ESRI provided geoportal technical help as part of their software maintenance package. ESRI has dropped that and now offers technical help for geoportal on a contract basis. There is a user forum for finding answers to questions but this can be time-consuming.
FGDC Standard deficiency	ESRI did not provide the dedicated metadata editor for geoportal with the function for documenting an FGDC detailed description of the "entity and attribute" data; only an overview description. Lane County was able to obtain a patch for this function but it took some time for ESRI to come through with the patch and still does not appear to be part of the core product.
Security Protocol	Lane County GIS worked diligently with IS staff to implement security features using the LDAP standard. This was an uphill climb that took a lot of time to troubleshoot and get working properly. Even ESRI staff were unable to help on this.

Despite the difficult learning curve for fully implementing the Geoportal, many nice advantages are apparent.

STRENGTHS	
Very Customizable	The Geoportal is built using .xml files and java script files that provide the user with a lot of flexibility to apply custom functions and a custom look and feel to the application.
Easy to share data	The geoportal is a web-enabled application that can be shared across the intra/internet with other users. Lane County has plans to create an internet version of their geoportal in the near future. It is also possible to link to other geoportals through a "federated search" mechanism so that external records can be viewed as part of the geoportal application.
Administration/Security available	Geoportal offers a means to assign administrator privileges so that metadata records can be reviewed before being posted to the database. Access can be designated as users (view only); contributors (add but not changes), editors (add and change) and full administrator.
Better Integration with other applications	Geoportal offers an opportunity to integrate metadata access into other GIS applications: • Link from web mapping apps to the Geoportal metadata • Link from ArcMap to Geoportal
FGDC Compatible	We are using a fairly "bare bones" approach to creating metadata records. This means that an FGDC-compliant record can be added to the data base with a minimal amount of information and still pass the validation tests. However, we are also filling in as much information about the entity and attribute data as possible to give the user a better understanding of the data.
"Fun" to use	Adding metadata is fairly simple and done through a series of tabs and fill-in menus. Creating a browse graphic adds a map element to the Geoportal records making it visually more appealing and understandable.
Has strong data search features	The Geoportal can be used to search for metadata records using keywords, by user defined data categories, by partial words, and even by geographic extent or place names.
Customizable Metadata Report	The Geoportal stores standard .xml files to the database which can be displayed back to the user. However, these are not in a very user-friendly format. A "Details" view of the report is offered that uses an xml stylesheet to create a custom report in a much nicer format capable of being printed.

Geoportal and Project Documentation

The original intent of developing the Geoportal application was (and still is) to provide a means to document our GIS data holdings and offer users easy access to finding metadata. However, Geoportal can also be used to document not only GIS data layers but other project-related information as well.

The Lane Livability project is consuming, and even generating, GIS data. Some of this data is already, (or soon will be) documented in Geoportal. The Geoportal can be organized so that metadata records can be grouped by various categories – such as base data, boundary data, airphotos, transportation data, and so on. In addition, the Geoportal can be used as a dynamic tool for organizing other information about a project for which documentation would be useful to users, stakeholders, and project participants. This might include the following:

- Creating a series of "Project Groups" in the Browse Tree that presents information about past and ongoing projects; such as Lane Livability.
- Including metadata records that document not only the GIS data; but also other components of a project; such as documents, published plans, meeting notes, public comment, and so forth. These types of metadata records deal more with qualitative descriptions of a project and the project data. But users still benefit by being able to easily find this type of information which is one of the key strengths of the Geoportal.
- Creating a series of related websites that are easily accessed through the Geoportal. In this sense the Geoprotal becomes a "jumping off" spot for locating other web-related applications and resources.

The dedicated metadata editor in Lane County's Geoportal is using the FGDC standard. However, records can be entered into this standard that describe other "data" beside GIS layers. The Geoportal editor has been implemented such that only a minimal amount of FGDC standard information is required. For example, if someone wanted to add a metadata record that documents the content, description and location of a project document (such as a plan); then the Geoportal could be used in that fashion.

Geoportal User Site Examples

Oregon Spatial Data Library:
http://spatialdata.oregonexplorer.info/geoportal/catalog/main/home.page

Open Raleigh – City of Raleigh GIS Data Portal:
http://maps.raleighnc.gov/geoportal/catalog/main/home.page

Match – Metadata Access Tool for Climate and Health – US Global Change Research Program:
http://match.globalchange.gov/geoportal/catalog/main/home.page

EPA Environmental Dataset Gateway (EDG):
https://edg.epa.gov/metadata/catalog/main/home.page

				Total Annual Staffing Cost:	$	25,611

Meeting Coordination and Facilitation   Annual FTE Cost: $ 8,938

Coordination role will be responsible for preparing and facilitating approximately (8) meetings per year at approximately (4) hours per meeting. Preparation of each meeting will require approximately (4) hours to schedule, develop meeting notices and address questions from participants. Meeting follow up will require an additional (4) hours to summarize meeting findings, update impacted documents and respond to questions from participates. Estimate includes some additional time for non-meeting activities.

Hours Annually (8 meetings)*((4 hrs. preparation)+(4 hrs. meeting)+(4 hrs. follow up)) =   96

Anticipated Hourly Cost:	Rate	Indirect	Total Hourly		
	$ 49.00	90% $	93.10		

Anticipated FTE Cost:	FTE	Day	Week	Month	Year
	0.05 $	36 $	178 $	712 $	8,938

Technical Support   Annual FTE Cost: $ 16,013

Technical role will be responsible for updating data elements of the Lane Livability Tool Kit and coordinating regional data catalog metadata into GeoPortal (GP) pilot site, as well as any unique research necessary to support investigation and evaluation of other data platforms. Updates to the Lane Livability Tool Kit (LLTK) will occur monthly or (10) times per year at approximately (4) hours per update. Updating the GeoPortal site will be ongoing, but is estimated to occur at an average of (18) times per year and require (4) hours per update. Estimates above include time to organize and prepare data and materials prior to making updates to Committee and respective sites.

Hours Annually ((10 LLTK updates)* (4 hrs.)) + ((18 GP updates)* (4 hrs.)) + 60 hrs evaluation support =   172

Anticipated Hourly Cost:	Rate	Indirect	Total Hourly		
	$ 49.00	90% $	93.10		

Anticipated FTE Cost:	FTE	Day	Week	Month	Year
	0.09 $	64 $	319 $	1,276 $	16,013

Work Hour Assumptions	100% FTE	
Work hours per Day	8	
Work hours per Week	40	
Work hours per Month	160	
Work Hours per Year	2008	--> Source: http://www.workingdays.us/#aepd
		(July 1, 2014 - June 30, 2015 = 251 work days)

Meeting and Travel Expense   Annual Cost: $ 660

Meeting and Travel Expense will include approximately (4) out of town trips at approximately (200) miles per trip and approximately (3) meals per day or (1) daily per diem per trip to discuss data advisory issues, methods and proposed solutions with comparable metro regions such as Portland, Seattle and San Francisco. For purposes of estimating costs, the number of trips and mileage have been normalized. Actual costs will be contained by conducting conference calls, exchanging emails and using various forms of remote communication.

Travel Miles Annually (4 trips) * (200 miles per trip) * (56 cents per mile) =	$	448
Travel Meals Annually (4 trips) * ($53 daily per diem) =	$	212

IRS Standard Rates	Rate	
2014 Standard Mileage Rates for business miles driven	$ 0.56	per mile
GSA Meals and Incidental Expenses (M&IE) Breakdown	$ 53	per day average