Next Generation Data Discovery

Fusing Structured and Unstructured Content from Multiple Repositories

Chris Meredith – UDOT
Dan Quinn – PTFS
Questions

Socrata

Shared Drives

UPlan

UDOT Map Center

electronic PROGRAM MANAGEMENT

DATA PORTAL

LTDOT

Keeping Utah Moving
Vision: Enterprise-wide Data Discovery

- Information Transparency
- UDOT information should be discoverable to the entire department and its partners
- Sometimes, information is desired but the question is unknown
What Have We Done?

• You don’t need to know where the information resides
• You don’t need to know what information exists
What Stays the Same?

The Index...

- References data from the source system. Nothing is copied!
- Utilizes source system credentials for document access.
  - If you’re required to log into the source system, the index does not provide a way around that.
- Allows data owners to remain data owners and continue to collect and maintain their information.

So you can more easily use the information already provided by the department without duplication!
Basic Knowvation operation

**Data Sources**
- Web Service
- Database
- SharePoint
- Esri
- Shared Drives
- ProjectWise

**Knowvation Services**
- Crawl Service
- Index/Ingest Service
- Search Service
  - OGC REST API
  - OAI-PMH
  - Z39.50

**Knowvation Clients**
- OGC-Compliant Apps
- ArcGIS Sites
- OGC Web Services (CSW, WCS, WFS)
- Knowvation Javascript-based UI
- OAI-PMH Client
- Z39.50 Client

[Diagram showing the integration of data sources, services, and clients in the Knowvation system]
Basic Search screen has several options

The UDOT Index provides access to a number of documents, file types, and data layers from multiple UDOT sources.

Search by File Type
- Text Search
- Video Search
- Image Search
- Geo Search

Search by File Source
- ProjectWise
- Dashboards
- UPlan
- Region 2 Share Drive
Knowvation supports a browse hierarchy
Browse enables navigating a folder structure
Knowvation can support optimal UDOT hierarchy
There are many search options from this interface.
The full text search box can start easy searches
Items presented as dots on base map
Various layers can be easily turned on/off
Now mileposts are off
Base map options can be switched with a click.
Select Open Street Map
Open Street Map used to start search, display data
A geospatial search is a common starting point
The search can now be limited to Route 48
And further limited with a PIN
And further limited with a Document Type
Data can be presented in different views: Grid View
Data can be presented in different views: List View
Data can be presented in different views: Thumbnail View
An Esri widget enables Knowvation searches in ArcGIS
An Esri widget enables Knowvation searches in ArcGIS.
A drop down makes selecting target route easy
A Search simply on route returns 4,282,556 files
Adding full text search on “ramp” narrows to 146,763 files.
Adding PIN 10711 narrows list to five files
Selecting “all” brings back all records = 685 records
Pattern Search is a fuzzy text search
Correct/incorrect spellings are highlighted.
Start A Search

- Enter a query
- Use the search criteria to narrow down possible records
- Manually enter an Area Of Interest
- Draw an Area Of Interest using the Map
SPECIAL PROVISION

PROJECT # S-0018(49)2
PIN # 10783

SECTION 03055M

PORTLAND CEMENT CONCRETE

July 5, 2017

Delete Article 2.2, subparagraph A.1 and replace with the following:

1. Use Type V Portland cement, or equivalent according to Table 4, or blended hydraulic cement unless otherwise specified.
What Else?

- Improving data attributes/metadata improves searchability
- Aligning data standards across the Department
- The index reflects how well data governance functions within the department. With data governance improvements, the index improves.
Moving Forward

• Findability Study using machine learning to help make documents more findable
• Power user testing – Region Designers
• Training
• Incorporate additional data sources
Chris Meredith
Utah Department of Transportation
Central Right of Way GIS Administrator
cmeredith@utah.gov

Dan Quinn
PTFS
VP, Sales & Marketing
dquinn@ptfs.gov
How Can You Do That?

You can search by...

Location on a map
How Can You Do That?

You can search by...
- Address
- Route and Milepost
- Source system
- Project information (PIN)
  - PIN
  - Route
  - Name
How Can You Do That?

You can search by...

Metadata categories

<table>
<thead>
<tr>
<th>Year</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>3</td>
</tr>
<tr>
<td>2011</td>
<td>6</td>
</tr>
<tr>
<td>2012</td>
<td>5</td>
</tr>
<tr>
<td>2013</td>
<td>20</td>
</tr>
<tr>
<td>2014</td>
<td>15</td>
</tr>
<tr>
<td>2015</td>
<td>19</td>
</tr>
<tr>
<td>2016</td>
<td>12</td>
</tr>
<tr>
<td>2017</td>
<td>19</td>
</tr>
<tr>
<td>2018</td>
<td>9</td>
</tr>
<tr>
<td>2020</td>
<td>18</td>
</tr>
</tbody>
</table>

Document Type

- ARCGIS feature[11125]
- TIFF[19]
- Shapefile[9]
- KML[7]
- JPEG[1]

<table>
<thead>
<tr>
<th>PIN</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>5469</td>
<td>95</td>
</tr>
<tr>
<td>14566</td>
<td>84</td>
</tr>
<tr>
<td>11864</td>
<td>81</td>
</tr>
<tr>
<td>8289</td>
<td>60</td>
</tr>
<tr>
<td>14827</td>
<td>48</td>
</tr>
<tr>
<td>15124</td>
<td>32</td>
</tr>
<tr>
<td>15239</td>
<td>30</td>
</tr>
<tr>
<td>14730</td>
<td>27</td>
</tr>
<tr>
<td>16527</td>
<td>25</td>
</tr>
<tr>
<td>13489</td>
<td>23</td>
</tr>
<tr>
<td>7928</td>
<td>22</td>
</tr>
<tr>
<td>16533</td>
<td>21</td>
</tr>
</tbody>
</table>

Collection

- 1411 - region 1 maint admin[5560]
- Region 1[2144]
- 1411 - region 1 maint ...[310]
How Can You Do That?

You can search by...

Full text across metadata and text in files using Boolean, Exact, Concept and Pattern search techniques