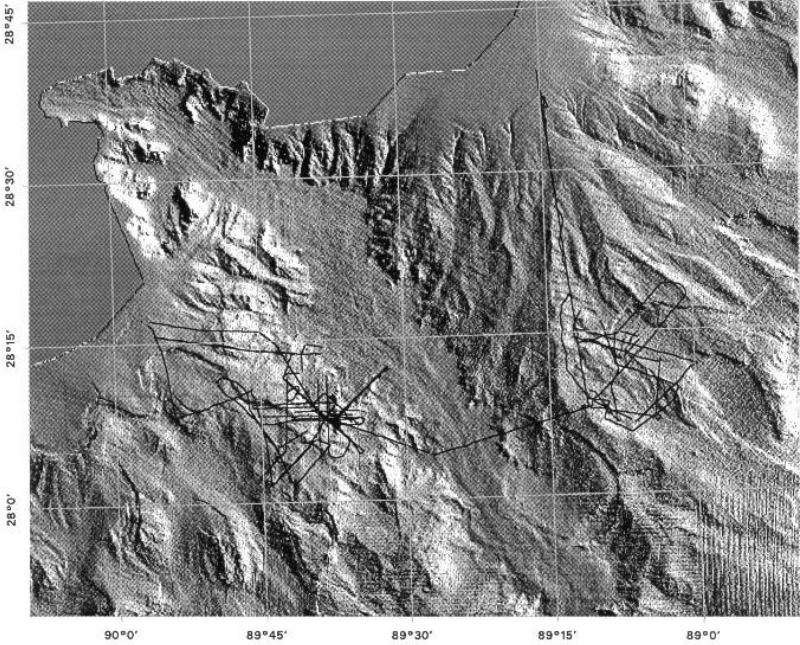


# Developing a Spatial Data Repository using ISO Metadata Standards

STANFORD UNIVERSITY LIBRARIES

Kim Durante  
Metadata Librarian  
kdurante@stanford.edu

# GIS, Data, and Maps at Stanford Libraries



# Open Source Software Development at Stanford Libraries

## Hydra Digital Repository

- Content management
- Preservation
- Access, Use, Re-use

## Blacklight

- Discovery interface for Solr index
- Customizable

### Partners:

DuraSpace

Stanford University

University of Hull

University of Virginia

MediaShelf

University of Notre Dame

Northwestern University

Columbia University

Penn State University

Indiana University

London School of Economics and  
Political Science

Rock and Roll Hall of Fame

Royal Library of Denmark

Data Curation Experts

WGBH

Boston Public Library

Duke University

Yale University

Virginia Tech

University of Cincinnati

Princeton University Library

Cornell University

Case Western Reserve University



# Selection of ISO Metadata Standards and Tools

## Standards

- ISO 19115/19139
- ISO 19110
- ISO 19119

## Features

- Supports unique identification of data and metadata
- *xlink:href* attribute
- Codelists

## Tools

- ArcCatalog
- Python/XSLT

# Develop Organizational Metadata Requirements

## Core Fields

- ISO Core Fields document

## Metadata for Layer-level Discovery

- Keywords (Places, topics, temporal)
- Thumbnail image

## Metadata for Evaluation/Use

- Projection, scale
- Quality reports

## Collection-Level Metadata

- Collection Title, Originator, Publisher, Date

## Re-Use

- Data citation
- Rights
- Transformation & exchange of metadata files

# Persistent URL (PURL)

<http://purl.stanford.edu/zv925hd6723>



## Stanford Digital Repository

### Oil and Gas Wells in the United States, 2011

**Author/Creator:** Hart Energy Publishing (creator)

**Imprint:** 1007 Update – San Diego, California, US : Hart Energy Publishing, 2011

**Type of resource:** Cartographic, Software, multimedia

**Genre:** Geospatial data  
Cartographic dataset

**Format:** Shapefile  
98.011 MB

**Digital origin:** born digital

**Map data:** Scale not given. ; North American Datum (NAD83) (W 124° 38' 46" -- W 73° 23' 14"/N 48° 59' 55" -- N 24° 25' 17")

**Abstract:** This layer shows the locations of oil and gas wells in the United States. Information in this layer is based on data provided by numerous sources in a variety of digital formats, collected at the state or county level.

**Purpose:** The purpose of this file is to combine multiple well data sources into one national file.

**Preferred citation:** Hart Energy Publishing. (2012) Oil and Gas Wells in the United States, 2011. Hart Energy Publishing.

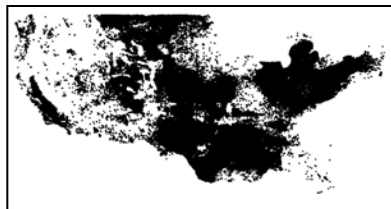
**Supplemental infor...** This release is intended to replace all previous datasets.

**Collection:** Oil and Gas Wells GIS Data: United States

**Subject:** Oil wells  
Gas wells  
Petroleum  
Natural Gas  
United States  
2011  
Utilities and Communication

**Available download formats:**

- [Data \(200 MB\)](#) 



# Unique Identification of Data and Metadata

## **DRUID**

- `zv925hd6723`

## **PURL**

- `http://purl.stanford.edu/zv925hd6723`

## **Metadata Identifier**

- `edu.stanford.purl:zv925hd6723`

## **Feature Catalog**

- `0fd6b47e-ebca-4378-8d86-b2a3a94d9d92`

## **Parent Metadata Identifier**

- `edu.stanford.purl:sv325bx4406`

## **Collection PURL**

- `http://purl.stanford.edu/sv325bx4406.mods`

# Library Catalog Metadata

## **Machine-Readable Cataloging (MARC)**

- Metadata standard for library resources (books, maps, etc.)

## **Metadata Object Description Schema (MODS)**

- Complimentary to MARC metadata standard
- XML Schema for digital objects
- Commonly used digital library standard
- Supports whole/part relationships and other references
- Supports linking through attributes



# Creating ISO Metadata in ArcCatalog

## Institutional template

Distributor Info  
Metadata Info  
Thesauri citation(s)

## Collection template

Purpose  
Parent Metadata ID  
Rights  
Status  
Update Frequency  
Language  
Data Contact(s)  
Larger Work Citation  
Collection PURL  
Originator(s)  
Publisher  
Dates(s)

## Import/Edit

Coordinates  
Projection  
Metadata ID  
Feature Catalog ID

Title  
Abstract  
Keywords  
Lineage  
jpeg  
Attribute info

## Post-process

Data Citation  
Layer Name  
URLs  
xlinks  
19110 Record



Template



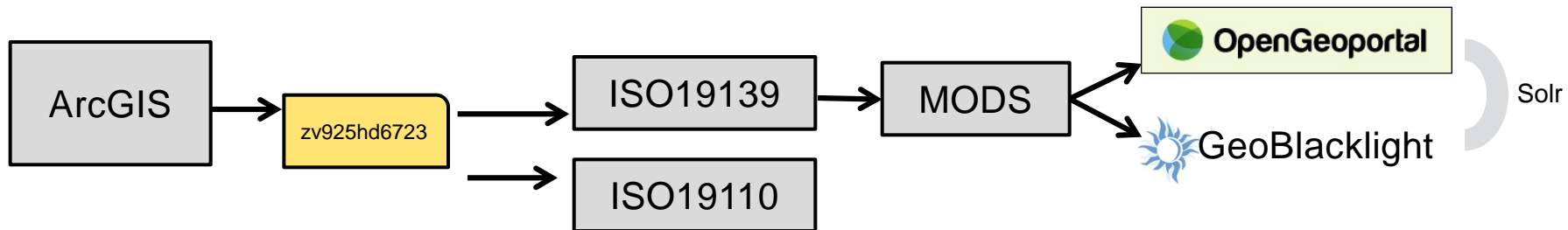
Auto-populate/Auto-generate



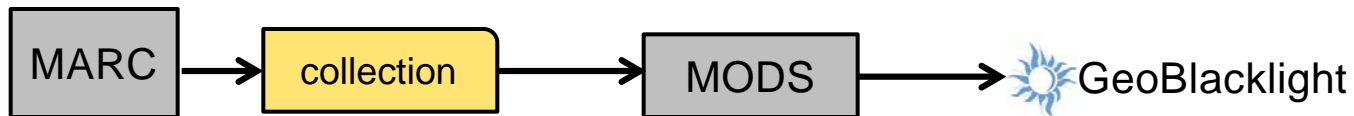
Locally edited

# XML Metadata Transformation

## Layer-level



## Collection-level



# Local ISO Metadata Practices

Element	Field Entry	Example
Title	Use Pattern: What, Where, When	Oil and Gas Wells in the United States, 2011
Projection	EPSG	4269
Originator	Library of Congress Name Authority File (LCNAF)	Geological Survey (U.S.)
Data credit	Data Citation Principles	Hart Energy Publishing. (2012) Oil and Gas Wells in the United States, 2011. Hart Energy Publishing. Available at: <a href="http://purl.stanford.edu/zv925hd6723">http://purl.stanford.edu/zv925hd6723</a>
Theme keywords	Library of Congress Subject Headings (LCSH), and ISO topics	Oil wells Utilities and Communication
Temporal keywords	YYYY or YYYY-YYYY	2011
Place keywords	GeoNames	United States

# Inserting URIs Using gmx:Anchor

## Gazetteer:

place_kw	gn_name	gn_id	lc_name	lc_id
United States	United States	625200 1	United States	n78095330

## ISO:


```
<gmd:keyword>  
  <gmx:Anchor xlink:href=http://sws.geonames.org/625001">United States</gmx:Anchor>  
</gmd:keyword>
```

## MODS:

```
<subject>  
<geographic authority="geonames" valueURI="http://id.loc.gov/authorities/names/n78095330">  
  United States</geographic>  
</subject>
```

# GeoBlacklight


Search..

Search 


## Limit your search

Institution 


Publisher 

Subject 

Place 


Collection 

Year 

Access 

Data type 

Format 

Language 

## Welcome!

Find GIS data!

Discover Maps!

## Featured Datasets

- [Amazing Data 1](#)
- [Amazing Data 2](#)
- [Amazing Data 3](#)



# GeoBlacklight

oil wells

Search

## Limit your search

Institution

Publisher

Subject

Place

Collection

Year

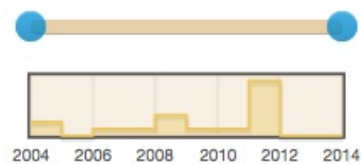
2004 to 2013

17

2004

2013

Limit



Access

Data type

Point

17

You searched for:

Keyword > oil wells

Data type > Point

Year > 2004 to 2013

Start Over

« Previous | 1 - 10 of 17 | Next »

Sort by relevance

10 per page

1. Oil and Gas Wells in the United States, 2011



2. World (Wells, 2004)



3. Oil Stores Colombia 2006



4. World (Oil Refineries, Dec. 2008)



5. Oil and Gas Fields in the United States, 2011



6. Oil Production Facilities Points 50k Scale Ecu...



7. Liquid or Gas Extraction Wells 50k Scale Ecu...



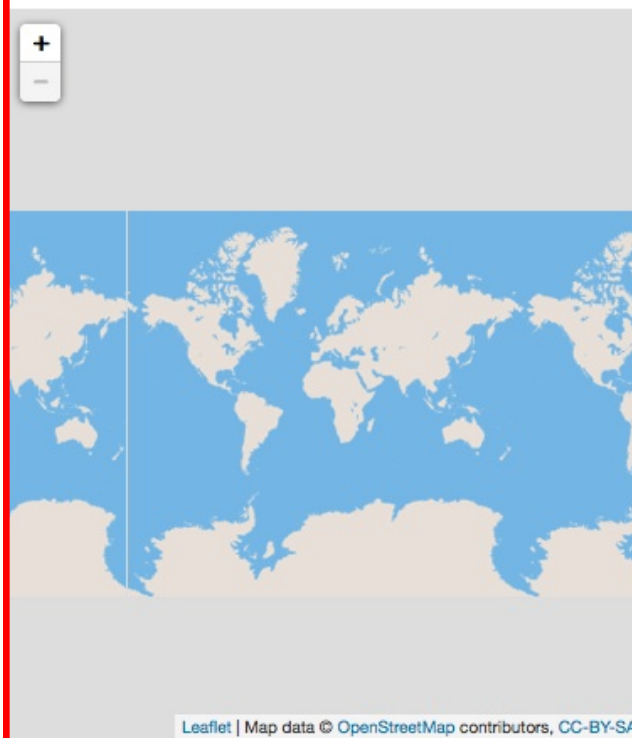
8. World (Mineral Extraction point structures, 20...



9. Planimetric Feature Points, Amsterdam (East)...




10. Planimetric Feature Points, Amsterdam (We...



# GeoBlacklight

oil wells

Search 

« Previous | 1 of 17 | Next »

[Back to Search](#)

[Start Over](#)

## Oil and Gas Wells in the United States, 2011

**Publisher** Hart Energy Publishing

**Year** 2011

**Abstract** This point dataset shows the locations of oil and gas wells in the United States. Information in this layer is based on data provided by numerous sour...[more](#)

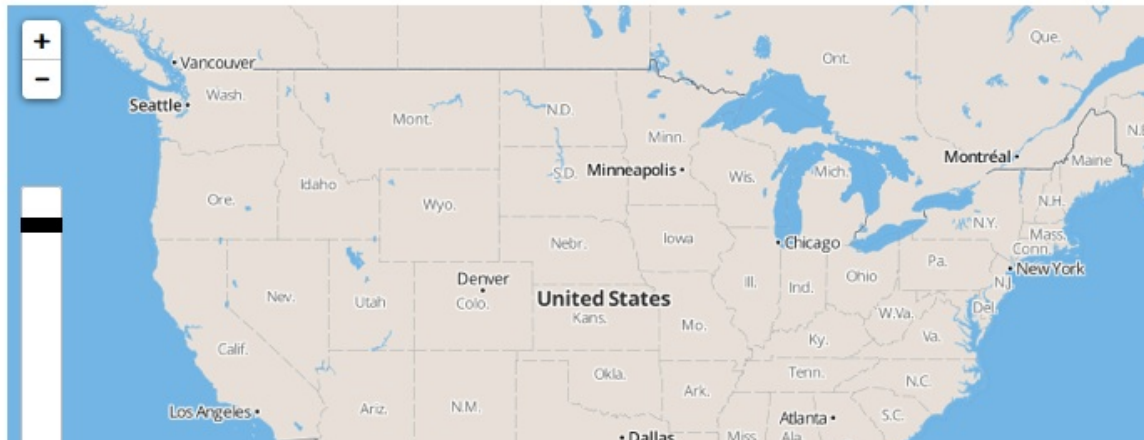
**Citation** <http://purl.stanford.edu/zv925hd6723>

**Places** United States

**Subjects** Gas wells, Natural Gas, Oil wells, Petroleum, Utilities and Communication

### Tools

- Bookmark
- Email
- SMS This
- Cite This
  
-  Metadata
-  KML
-  Shapefile
  
-  WMS Link
-  WFS Link





## GeoBlacklight – GIS Data Discovery

- Solr metadata schema derived from MODS
- Aggregated metadata portal for discovery of GIS data
- Faceted browsing
- Indexes collection-level metadata
- Persistent and citeable URLs for spatial data
- Layer availability



# OpenGeoPortal



## Home

[OpenGeoportal.org](http://OpenGeoportal.org) is a new site that brings together geospatial professionals, developers, metadata specialists, and librarians to coordinate the [Open Geoportal \(OGP\)](http://Open Geoportal (OGP) project) project. The [Open Geoportal](http://Open Geoportal) is a collaboratively developed, open source, federated web application to rapidly discover, preview, and retrieve geospatial data from multiple repositories. [OpenGeoportal.org](http://OpenGeoportal.org) is also a collaborative effort to share resources and best practices in the areas of application development, metadata, data sharing, data licensing, and data sources in support of geospatial data repositories. Please contact [Patrick.Florance@Tufts.edu](mailto:Patrick.Florance@Tufts.edu) for more information. The Open Geoportal is now receiving additional support from the [Alfred P. Sloan Foundation](http://Alfred P. Sloan Foundation).

A public beta instance of Open Geoportal 2.0 can be found [here](http://here).



The project's [code repository](http://code repository) is hosted on github. Documentation can be found [here](http://here). However, some technical documentation is still on the old google code site: [technical documentation/wiki](http://technical documentation/wiki).

## New Developments

### OGP 2.0 Release

Open Geoportal 2.0 in beta! The Open Geoportal 2.0 is currently in beta. It includes new enhanced interface, interoperability for several web mapping platforms, etc. The

### UNH Receives IMLS Award

Empowering the University of New Hampshire User Community with the Power of PLACE. The University of New Hampshire Library and

### Alfred P. Sloan Foundation Grant Awarded to Tufts

Tufts University has received an award from the Alfred P. Sloan Foundation for the development of The Open

## Upcoming Events

### NYC Open Geoportal Meeting – December 13th

LINK TO SUMMARY OF GREATER NYC OPEN GEOPORTAL MEETING  
NYC Open Geoportal Meeting Friday, December 13th Location: New York University (NYU) Map of Meeting Location (PDF) Interactive Map Morning Sessions (8:30am-12:30pm) NYU Kimmel Center for University Life – Room 914, 60 Washington Square S, New York Afternoon Working Meeting (1:30-4:45pm) NYU Bobst Library - Room 619, 70 Washington Square S, New York, NY Through the support of the Alfred P. Sloan Foundation and NYU, we [Continue Reading](http://Continue Reading)→

### Tufts Presents OGP at UN Annual Plenary

Tufts University has been invited to present the Open Geoportal at the United Nations Geographic Information Working Group (UNGIWG) 2013 Annual Plenary in Istanbul, Turkey on February 27th - March 1st. [Continue Reading](http://Continue Reading)→



## Partners:

- Tufts University
- Harvard University
- Massachusetts Institute of Technology
- Massachusetts State GIS Office
- Baruch College – CUNY
- California State University – Fresno
- CIESIN
- Columbia University
- Cornell University
- Emory University
- GeoCat, GeoNetwork
- New York University (NYU)
- Purdue University
- Stanford University
- State of Wisconsin
- State of New Hampshire
- United Nations
- University of California Berkeley
- University of California Los Angeles
- University of California Santa Barbara
- University of Connecticut
- University of Iowa
- University of Maryland
- University of Minnesota
- University of New Hampshire
- University of Virginia
- University of Washington
- University of Wisconsin Madison
- USAID – GIST
- Yale University

# OpenGeoPortal

## GeoData@Stanford

BETA

Getting Started

Search

Cart (0)

SEARCH

rajasthan demographics

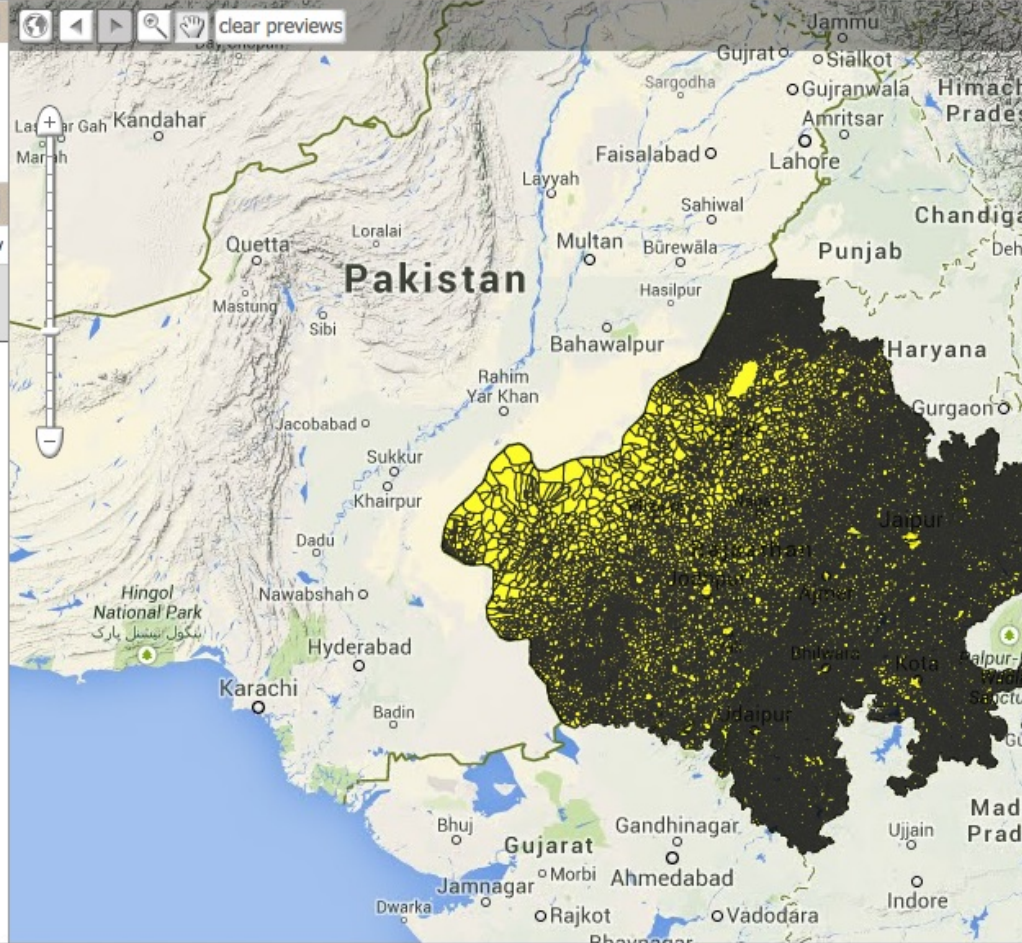
Search

[advanced options](#)

66 RESULTS by Relevancy

Columns

Type	Name	Rep	i	Preview
<input type="checkbox"/>	Rajasthan, India: Village Socio-Demographic and Economic C	S	i	<input checked="" type="checkbox"/>
opacity: 80% border: 1px				
<input type="checkbox"/>	Rajasthan, India: Pre Delimitation State Assembly Constituen	S	i	<input type="checkbox"/>
<input type="checkbox"/>	Rajasthan, India: Post Delimitation State Assembly Constitue	S	i	<input type="checkbox"/>
<input type="checkbox"/>	Madhya Pradesh, India: Village Socio-Demographic and Econ	S	i	<input type="checkbox"/>
<input type="checkbox"/>	Gujarat, India: Village Socio-Demographic and Economic Cer	S	i	<input type="checkbox"/>
<input type="checkbox"/>	Uttar Pradesh West, India: Village Socio-Demographic and E	S	i	<input type="checkbox"/>
<input type="checkbox"/>	Uttarakhand, India: Village Boundaries with Socio-Demograp	S	i	<input type="checkbox"/>
<input type="checkbox"/>	Haryana, India: Village Socio-Demographic and Economic Ce	S	i	<input type="checkbox"/>
<input type="checkbox"/>	Punjab, India: Village Socio-Demographic and Economic Cen	S	i	<input type="checkbox"/>
<input type="checkbox"/>	Daman and Diu, India: Village Socio-Demographic and Econ	S	i	<input type="checkbox"/>
<input type="checkbox"/>	Uttarakhand, India: Village Points with Socio-Demographic ar	S	i	<input type="checkbox"/>
<input type="checkbox"/>	Himachal Pradesh, India: Village Points with Socio-Demograp	S	i	<input type="checkbox"/>
<input type="checkbox"/>	District Boundaries of India, 1991	S	i	<input type="checkbox"/>
<input type="checkbox"/>	State Boundaries of India, 1971	S	i	<input type="checkbox"/>
<input type="checkbox"/>	State Boundaries of India, 2001	S	i	<input type="checkbox"/>
<input type="checkbox"/>	State Boundaries of India, 1961	S	i	<input type="checkbox"/>
<input type="checkbox"/>	State Boundaries of India, 1981	S	i	<input type="checkbox"/>



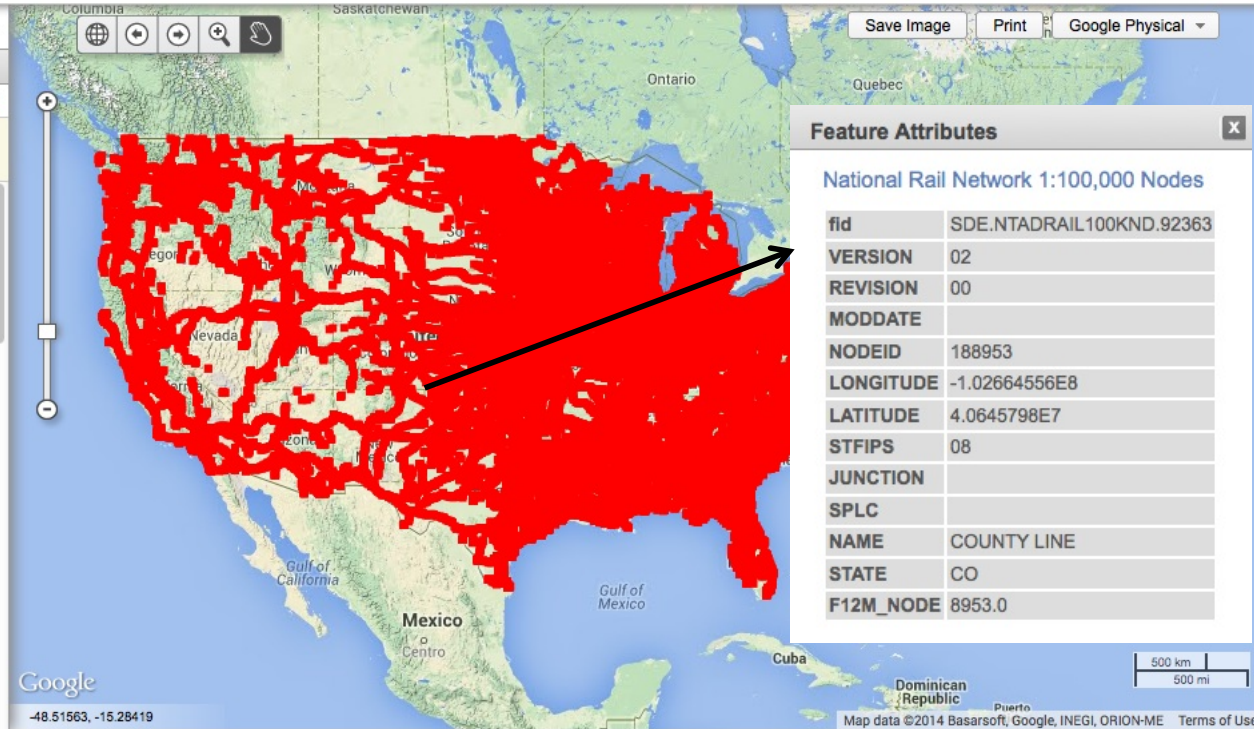
# GeoData at Tufts

Where:  What:  [Search](#) [Advanced Search](#)

636 Results  Cart (0)

Sort Results Columns Clear Views

Type	Name	Originator	Rep	Meta	View
<input checked="" type="checkbox"/>	National Rail Network 1:100,000 Nodes	Bureau of Trans			<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/>	National Rail Network 1:100,000	Bureau of Trans			<input type="checkbox"/>
<input checked="" type="checkbox"/>	National Rail Network 1:2,000,000	Bureau of Trans			<input type="checkbox"/>
<input checked="" type="checkbox"/>	National Rail Network 1:2,000,000 N	Bureau of Trans			<input type="checkbox"/>
<input checked="" type="checkbox"/>	USA (Amtrak Stations, 2009)	Federal Railroad			<input type="checkbox"/>
<input checked="" type="checkbox"/>	Amtrak Stations	Federal Railroad			<input type="checkbox"/>
<input checked="" type="checkbox"/>	Rail Lines USA 2011	United States. F			<input type="checkbox"/>
<input checked="" type="checkbox"/>	Rail Stations USA 2011	United States. F			<input type="checkbox"/>
<input checked="" type="checkbox"/>	Amtrak Stations USA 2009	United States. F			<input type="checkbox"/>
<input checked="" type="checkbox"/>	USA (Amtrak Stations, 2008)	Federal Railroad			<input type="checkbox"/>
<input checked="" type="checkbox"/>	Amtrak Lines USA 2011	United States. F			<input type="checkbox"/>
<input checked="" type="checkbox"/>	USA (Fixed-Guideway Transit Networ	Federal Transit			<input type="checkbox"/>
<input checked="" type="checkbox"/>	USA (Fixed-Guideway Transit Stator	Federal Transit			<input type="checkbox"/>
<input checked="" type="checkbox"/>	MassGIS 2000 Rail Trails Datalayer	Massachusetts.			<input type="checkbox"/>
<input checked="" type="checkbox"/>	Playgrounds Cambridge MA 2003	Cambridge (Mas			<input type="checkbox"/>
<input checked="" type="checkbox"/>	Playgrounds Nantucket MA 2004	Nantucket (Mas			<input type="checkbox"/>
<input checked="" type="checkbox"/>	Railroad Trails Concord MA 2006	Concord (Mass.			<input type="checkbox"/>



### Feature Attributes

National Rail Network 1:100,000 Nodes

fid	SDE.NTADRAIL100KND.92363
VERSION	02
REVISION	00
MODDATE	
NODEID	188953
LONGITUDE	-1.02664556E8
LATITUDE	4.0645798E7
STFIPS	08
JUNCTION	
SPLC	
NAME	COUNTY LINE
STATE	CO
F12M_NODE	8953.0

# Multi-Institutional XML Metadata Harvesting

Harvard, Tufts, MIT, etc.



# Metadata Transformations Using XSLT

ArcGIS to ISO 19139

ArcGIS to ISO 19110

ISO 19139 to MODS

FGDC to MODS

MARC to MODS

MODS to GeoBlacklight

MODS to OGP Solr



# Issues

- Volume of data
- Evolving nature of standards
- Collection-level metadata
- Rights
- Frequently updated data, other types of versioning
- Duplication of work across institutions

## More Information

### GeoBlacklight on GitHub

- [github.com/geoblacklight](https://github.com/geoblacklight)

### Hydra

- [projecthydra.org](https://projecthydra.org)

### OpenGeoPortal

- [opengeoportal.org](https://opengeoportal.org)

GIS Team: Patricia Carbajales, Darren Hardy, Amy Hodge, David Mederios, Salim Mohammed, Jack Reed, Bess Sadler, Julie Sweetkind-Singer