

What to do

When Google ignores your Fedora objects

Robin Dean

robin@coalliance.org

Digital Repository Director

Colorado Alliance of Research Libraries

Jonathan Green

jonathan@discoverygarden.ca

Chief Technology Officer

discoverygarden inc.

Open Repositories 2013

Charlottetown, Prince Edward Island, Canada

Fedora Users Group – Thursday July 11, 5:30 PM

Alliance Digital Repository



COLORADO
ALLIANCE
OF RESEARCH LIBRARIES

Fedora 3.4.2

Islandora 11.3

+ Drupal 6

+ Core Islandora Modules

+ Custom Modules/Viewer


<http://adr.coalliance.org>

Custom Compound Object Viewer in Islandora 6

Home > Digital Repository Top > Archives and Special Collections > Richard H. Truly U.S. Space Program Collection > Richard H. Truly Speech Collection >

1990 Recap

Overview Viewer

Description		Files (Total: 1)
Title	1990 Recap	 RUART00167.pdf RUART00167.pdf <input type="button" value="View"/> <input type="button" value="Download"/> Page 1 of 1
Creator	Truly, Richard H.	
Type of Resource	text	
Date Created	1990-12-18	
Digital Origin	digitized other analog	
Description	Richard H. Truly's handwritten notes recapping NASA highlights for 1990.	
Topics	NASA Space Shuttle Program NASA budget Augustine Report Space Station Freedom human space flight and exploration	
People	Truly, Richard H.	
Organizations	NASA	
Handle	http://hdl.handle.net/10176/codr:2098	
Rights Statement	The Richard H. Truly U.S. Space Program Collection is the physical property of Regis University Libraries. Permission to publish unpublished materials must be obtained in writing from an officer of the Regis University Libraries.	
Genre	speech	
Place of Publication	Denver, Colo.	
Publisher	Regis University, Dayton Memorial Library	
Language	English	

Improving repository search results in Google

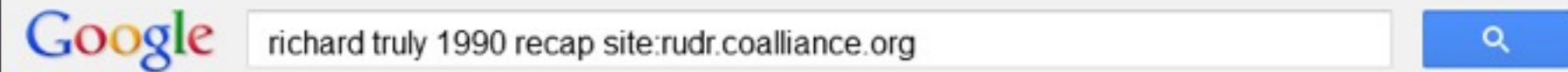
1. Create RSS sitemaps of Fedora objects for search engines
2. Improve search results snippets by re-using MODS metadata in HTML <meta> tags
3. Add meaning for search engines “behind the scenes” with RDFa

Without Sitemaps (Bing)



No results found for **richard truly 1990 recap site:rudr.coalliance.org**.

With Sitemaps (Google)



Web Images Maps Shopping More ▾ Search tools

4 results (0.28 seconds)

[1990 Recap | Regis University Digital Repository](#)

rudr.coalliance.org/fedora/repository/codr.2098 ▾

1990 Recap. Title Information:Alternative. Name:Personal Role :Text(marcrelator) creator. Truly, Richard H. Name:Personal Role :Text(marcrelator) contributor ...

How did we make RSS sitemaps in Islandora?

1. Index all Fedora objects in Solr
2. Add public objects in Solr to an RSS feed using Islandora
3. Submit the RSS feed as a sitemap in [Google Webmaster Tools](#)

Sitemap Details from Google Webmaster Tools

Sitemaps

By me (1)

All (1)

Sitemaps content

All content types

Submitted

Indexed

Web pages

28,944 Submitted

27,282 Indexed



Sitemaps (All content types)

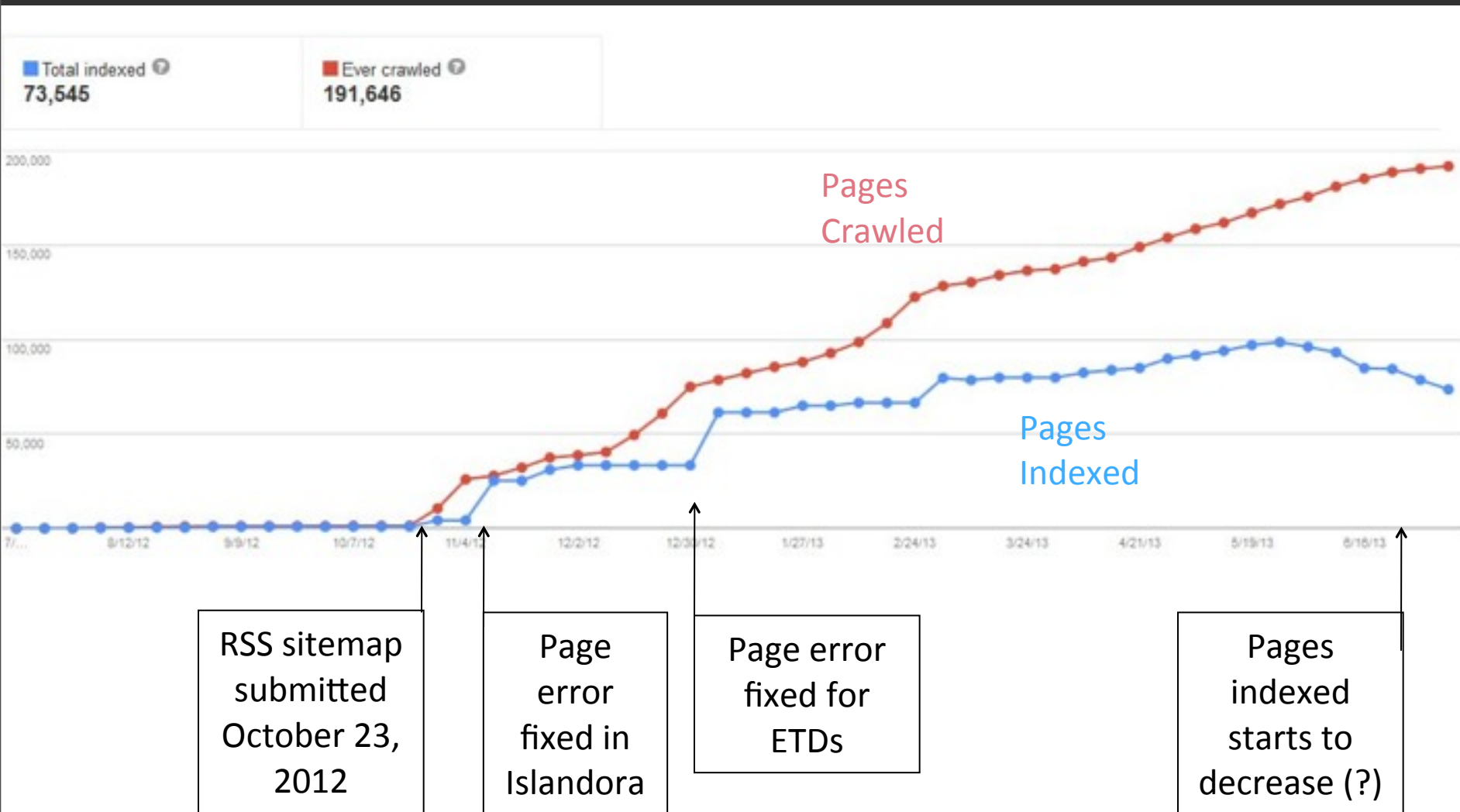
Download All

Resubmit

Delete

<input type="checkbox"/>	#	Sitemap ▲	Type	Processed
<input type="checkbox"/>	1	/islandora/solr/search?solr_profile=rss&limit=-1	RSS	Jul 10, 2013

Trend in Pages Crawled and Pages Indexed



What about Google Scholar?

1. Google scholar is trickier.
Uses a different web crawler
than normal Google.
2. Doesn't accept RSS sitemaps.
3. In a large repository, indexer
tends to request a lot of
documents very quickly.

Google Scholar Sitemaps

1. Create simple HTML pages.
2. Create a list of recently added papers.
3. Ensure that Google can quickly crawl your pages.

Re-use descriptive metadata in HTML <meta> description

Islandora Solr Results



Herbert A. Howe 1904 Diary

Professor Howe's diary for 1904 recounts his brief weather and astronomical observations, administrative and teaching activities at the University of Denver, entertainment of visitors to the Chamberlin Observatory, participation in events in University Park and elsewhere in Denver, as well as his family and religious life. Special note is made of: electrical service in the University Park neighborhood being raised to 108 volts; and in September, Julian and Hubert Howe left to attend the Culver Military Academy in Indiana.

Created by: Howe, Herbert A.--(Herbert Alonzo),--1858-1926;

Date Created: 1904

Google Search Results

[Herbert A. Howe 1904 Diary | Digital DU](#)

digitaldu.coalliance.org/fedora/repository/codu%3A64948 ▼

Professor Howe's diary for 1904 recounts his brief weather and astronomical observations, administrative and teaching activities at the University of Denver, ...

Google Scholar Meta Tags

1. Scholar expects meta tags in its own format.
2. Google Scholar supports
 - Highwire Press tags
 - Eprints tags
 - PRISM tags.
3. Limited support for Dublin Core tags.

Google Scholar <meta> Tag Examples

```
<meta name="citation_title" content="12-Acetoxy pseudopteroide" />
```

```
<meta name="citation_author" content="Ata, A." />
```

```
<meta name="citation_author" content="Kerr, R. G." />
```

```
<meta name="citation_author" content="," />
```

```
<meta name="citation_publication_date" content="2000" />
```

```
<meta name="citation_journal_title" content="" />
```

```
<meta name="citation_volume" content="53" />
```

```
<meta name="citation_issue" content="3" />
```

```
<meta name="citation_firstpage" content="717" />
```

Rich results with schema.org microdata

[Lobster Rolls Recipe - Allrecipes.com](#)



allrecipes.com/recipe/lobster-rolls/ ▾

★★★★★ Rating: 4.6 - 62 reviews - 15 mins - 369 cal

I grew up in Nova Scotia where seafood was cheap. My Mom and grandmothers had some pretty inventive ways of making lobster 'not-boring!' Now, lobster is a ...

[Lobster Rolls Recipe Photos](#) - [Recipes Like](#) - [Reviewer](#)

<http://schema.org/Recipe>

[Lobster Roll Recipe : Ellie Krieger : Recipes : Food Network](#)



www.foodnetwork.com ▸ [Recipes](#) ▸ [Seafood](#) ▾

★★★★★ Rating: 5 - 24 reviews - 33 mins

Get this all-star, easy-to-follow Food Network Lobster Roll recipe from Ellie Krieger.

[The Lobster Roll](#)

www.lobsterroll.com/ ▾

northside.

[Island Favorites "PEI's BEST Lobster Roll" - TripAdvisor](#)

www.tripadvisor.ca ▸ ... ▸ [Island Favorites "PEI's BEST Lobster Roll"](#) ▾

★★★★★ Rating: 4.5 - 38 reviews

Island Favorites "PEI's BEST Lobster Roll": Great tasting Lobster Rolls on PEI - See 38 traveler reviews, 12 candid photos, and great deals for Cavendish, ...

<http://schema.org/Review>

Before RDFa

[Beauty and elegance in the world around us: | Regis University ...](#)

rudr.coalliance.org/fedora/repository/codr:2652 - United States

Title, **Beauty and elegance in the world around us:** titleInfo: : Alternative. Title. Title,

Elucidating the higher order structure of the B12 Riboswitch in Thermatoga ...

After RDFa

[Beauty and elegance in the world around us: | Regis University ...](#)

rudr.coalliance.org/fedora/repository/codr:2652 ▼

Beauty and elegance in the world around us: Title Information:Alternative Elucidating the higher order structure of the B12 Riboswitch in Thermatoga maratima ...

[Regis College Honors Theses | Regis University Digital Repository](#)

rudr.coalliance.org/fedora/repository/codr:32 ▼

Beauty and elegance in the world around us: Over the years, science has shattered and reshaped our understanding of the universe, revealing and illuminating ...

RDFa fixes and improvements

Use an object type that exists in an ontology we can link to

Try the [MODS RDF](#) elements instead of our homegrown MODS crosswalk

Empty elements should not be marked up with RDFa

Next Steps for Google Scholar Indexing

Improve crawling

Look to others in the community
for advice

Create caches that Google
Scholar can crawl faster

Questions and Discussion

Robin Dean

robin@coalliance.org

Digital Repository Director
Colorado Alliance of Research Libraries

Jonathan Green

jonathan@discoverygarden.ca

Chief Technology Officer
discoverygarden inc.