Open Data and Open Libraries:
Supporting the Open Data Project in Library Services

Jason A. Clark
Associate Professor
Head of Library Informatics & Computing
Montana State University (MSU) Library

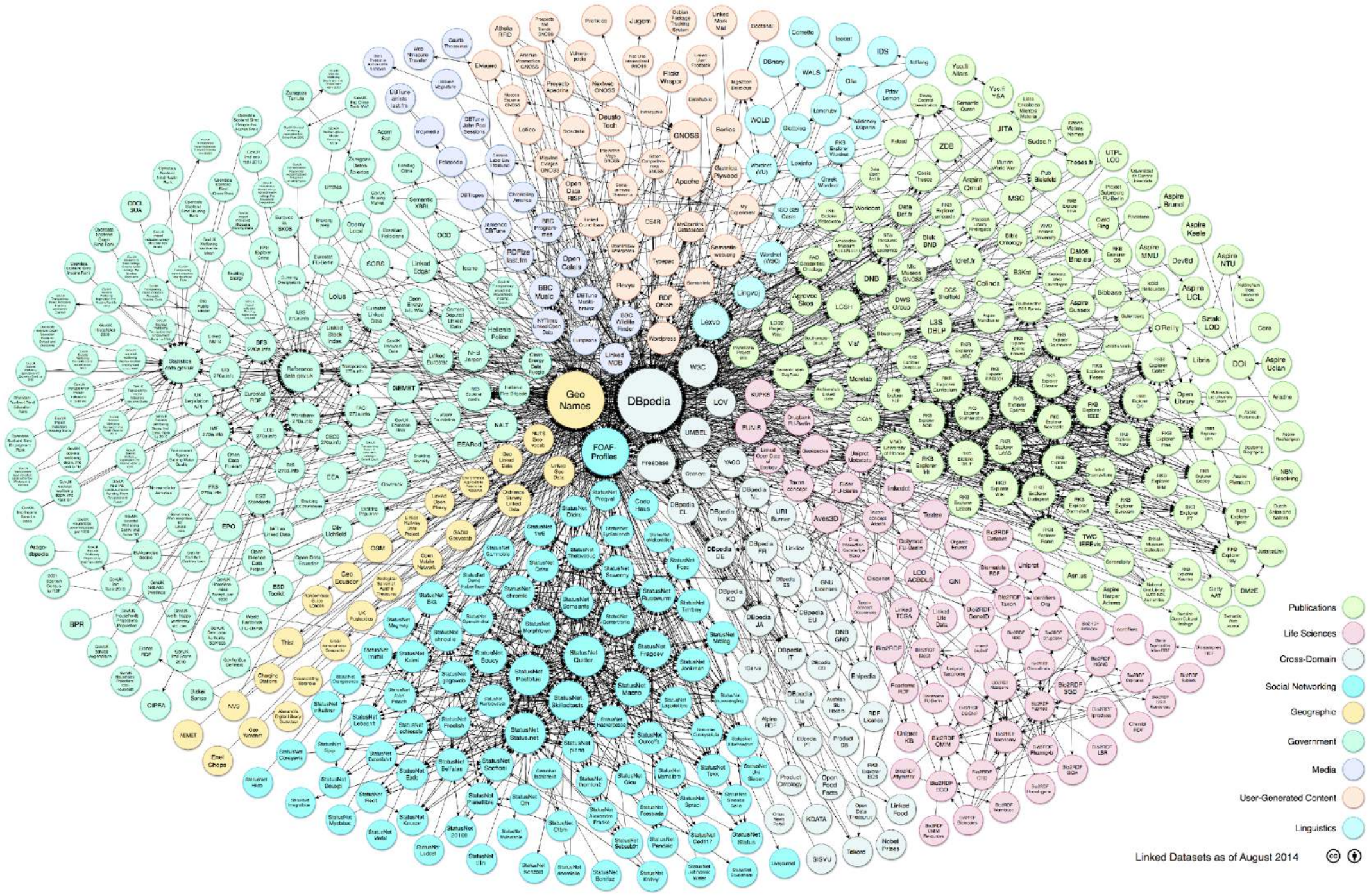
Florida Library Webinars
July 22, 2015

pinboard.in tag

pinboard.in/u:jasonclark/t:open/

twitter as channel (#hashtag)

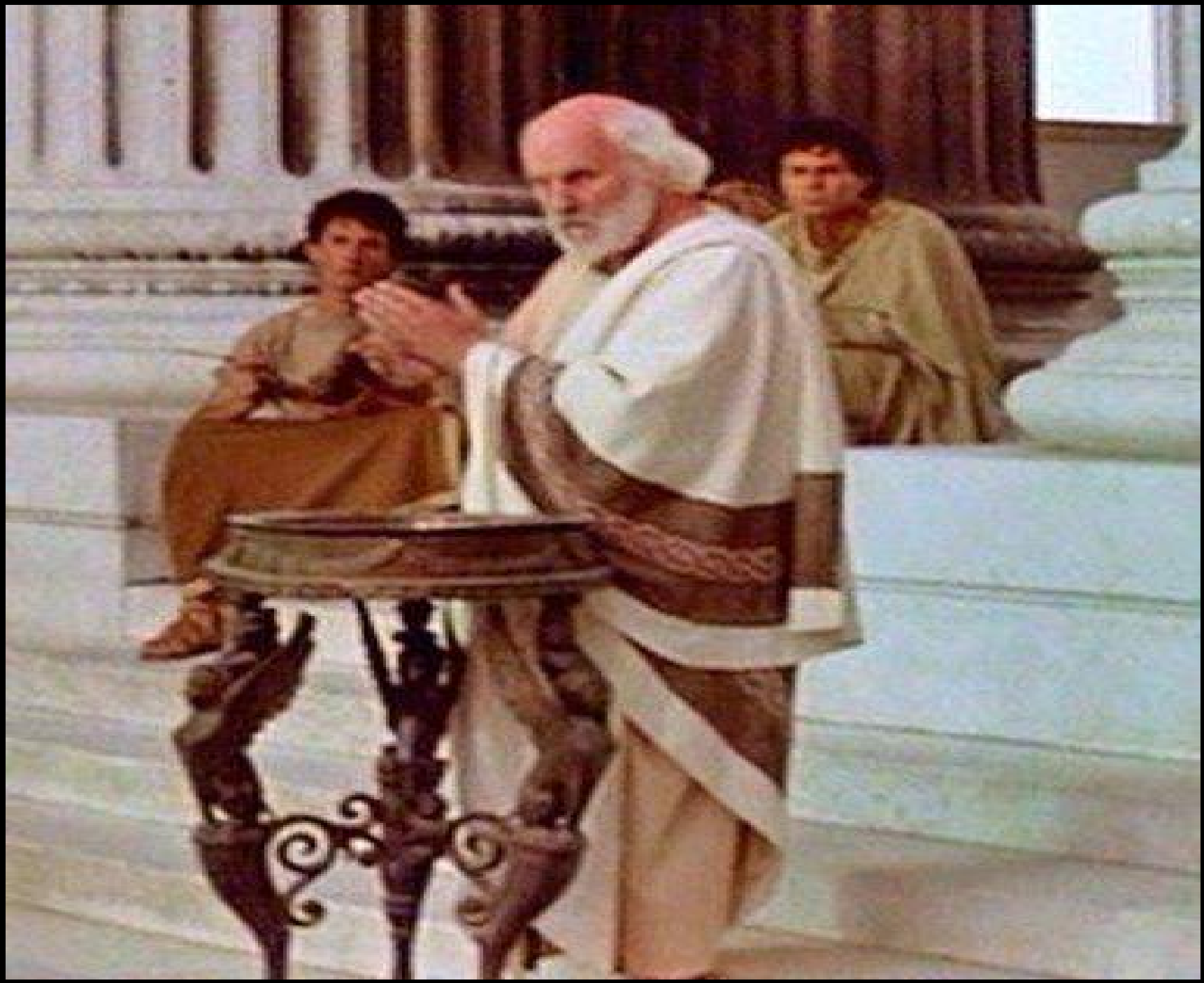
@jaclark #opendata



Linking Open Data cloud diagram 2014, by Max Schmachtenberg, Christian Bizer, Anja Jentzsch and Richard Cyganiak. http://lod-cloud.net/

Overview

- Why Open Data?
- Open Data in Practice
- Open Data + Libraries
- Open Data Processes + Tools
- Questions



Defining Terms and Truths

What do you mean by “open”?

- An open work must satisfy requirements around open licenses, open access, and open formats.

It can be summed up in the statement that:

> *"Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)."*

Put most succinctly:

> *"Open data and content can be freely used, modified, and shared by anyone for any purpose"*

Read the full Open Definition »

http://opendefinition.org/od/

“A piece of data or content is open if anyone is free to use, reuse, and redistribute it - subject only, at most, to the requirement to attribute and/or share-alike.”

http://opendefinition.org/

Why Open Data?

Part of a Library Continuum

Sharing

https://twitter.com/opencityapps

Attribution

FIND, ACCESS, AND REUSE DATA

Member Login | Contact | FAQ Search...

DataCite

ABOUT DATACITE ~ DATACITE FOR YOU ~ PARTICIPATE ~ EVENTS ~ NEWS JOIN DATACITE

Home › Services › Cite your data

Cite your data

Why is it so important to cite data? Books and journal articles have long benefited from an infrastructure that makes them easy to cite, a key element in the process of research and academic discourse. We believe that you should cite data in just the same way that you can cite other sources of information, such as articles and books. Data citation can help by:

- enabling easy reuse and verification of data
- allowing the impact of data to be tracked
- creating a scholarly structure that recognises and rewards data producers

Examples of data citation

We recognise that the challenges associated with data publication vary across disciplines, and we encourage research communities to develop citation systems that work well for them. Our recommended format for data citation is as follows:

Events

Keep an eye out for the upcoming future events

News

Announcing the THOR Project

We are pleased to announce DataCite's participation in THOR (Technical and Human...

DataCite Event -- Persistent Identifiers: Enabling Services for Data Intensive Research

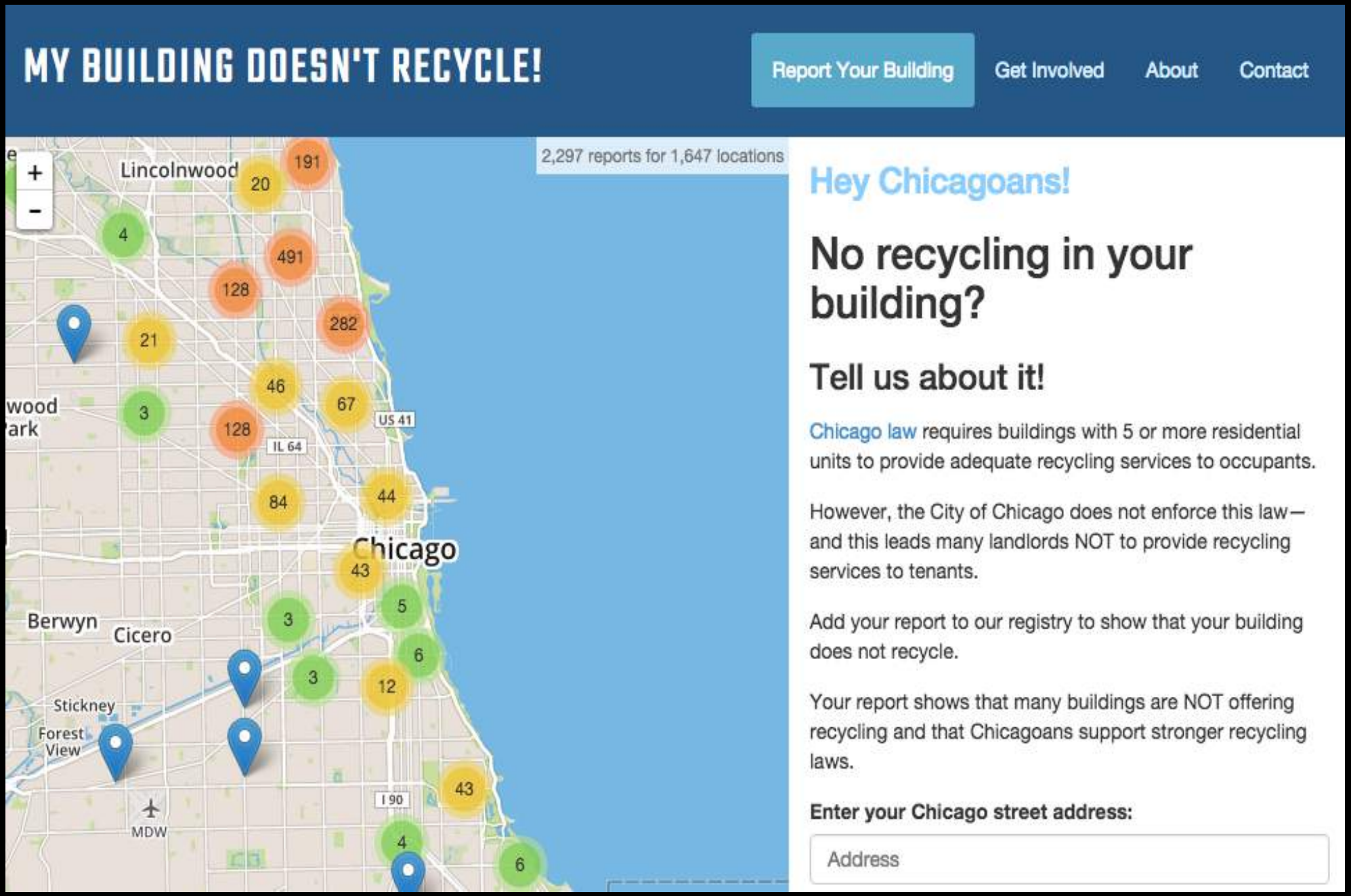
Save the date and join...

https://www.datacite.org/services/cite-your-data.html

Access

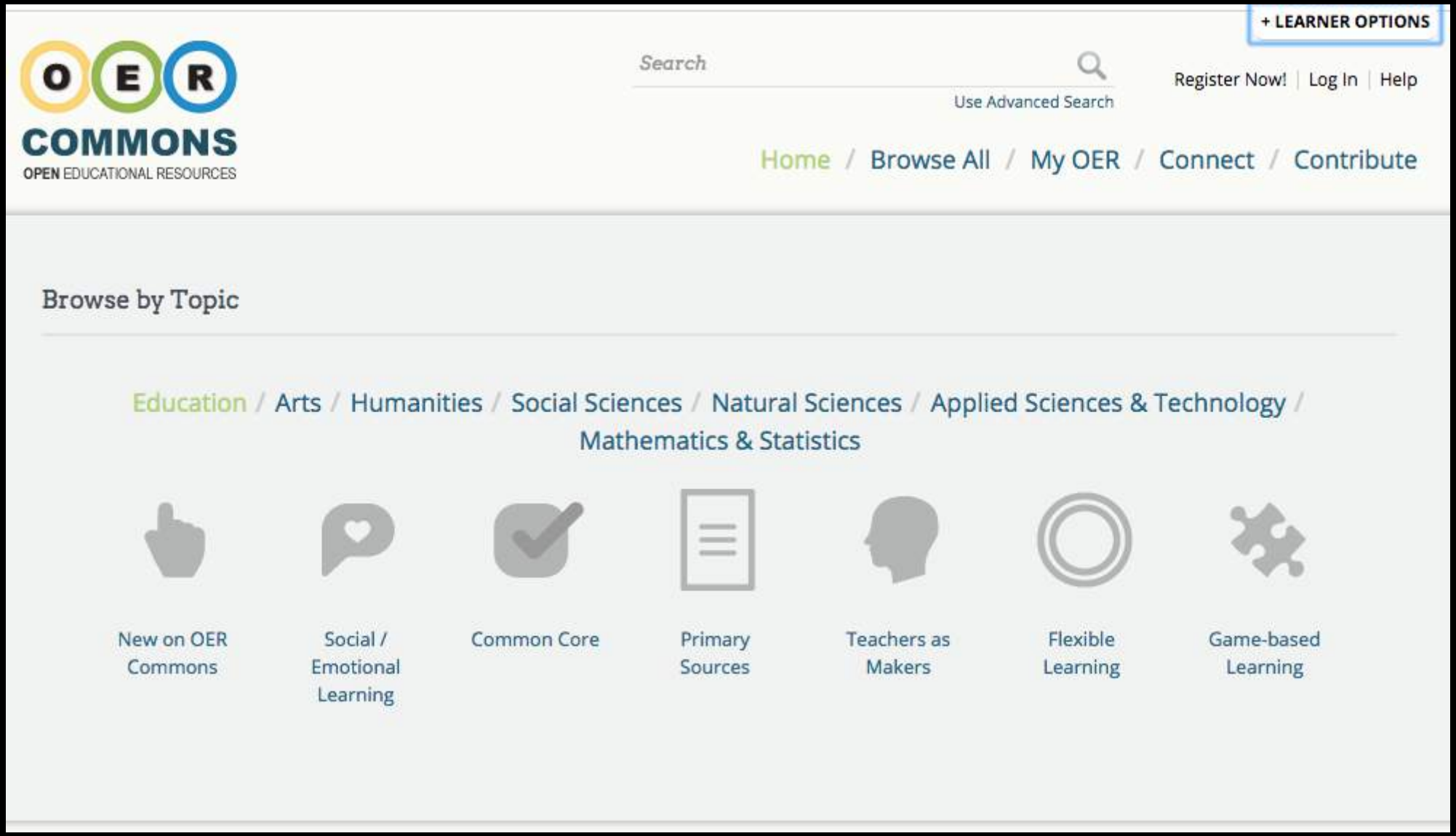
http://www.data.gov/

Reuse and Repurpose



http://mybuildingdoesntrecycle.com/

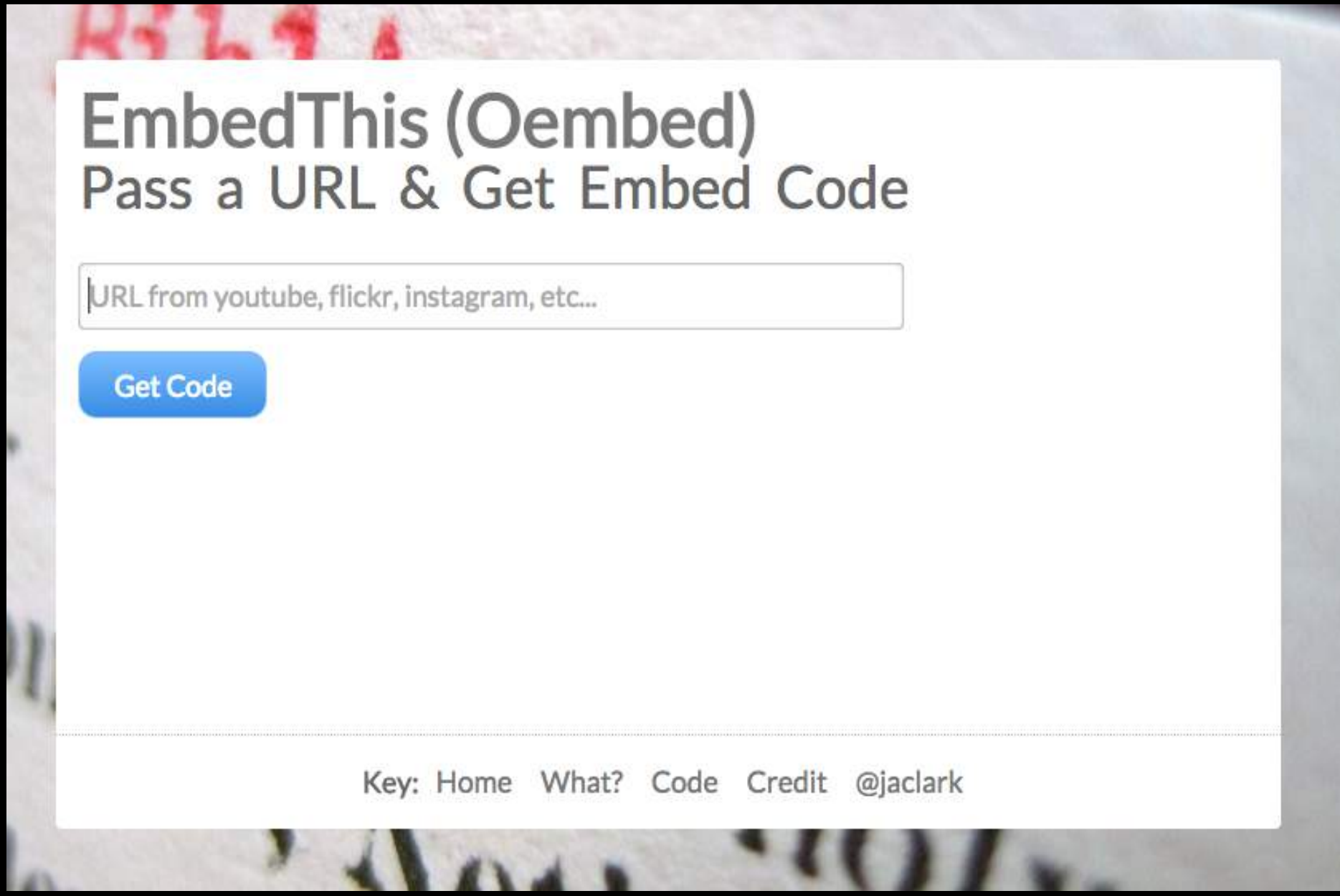
Archive



https://www.oercommons.org/

Open Data in Practice

So what?



http://www.lib.montana.edu/~jason/files/oembed-this/

EmbedThis (Oembed)
Pass a URL & Get Embed Code

Hmmm... Thanks for trying out the URL at https://www.youtube.com/watch?v=H7pkpSVbyc8

Here's what I found:

MSU Library - Everywhere You Go

embed source code:

```
<iframe width="320" height="180" src="https://www.youtube.com/embed/H7pkpSVb
yc8?feature=oembed" frameborder="0" allowfullscreen></iframe>
```

Key: Home What? Code Credit @jaclark

https://github.com/jasonclark/oembed-this/

Application Programming Interfaces (API)

```
<?xml version="1.0"?>
<items>
    <item>
        <id>1</id>
        <name>Angel Food Cake</name>
        <creator>Laura Cameron</creator>
        <url>http://arc.lib.montana.edu/book/home-cooking-history-409/article-laura-cameron.html</url>
        <description>Angel food cake recipe written by Laura Cameron for Montana State University's upper division history c
        <about>baking cake</about>
        <image>http://upload.wikimedia.org/wikipedia/commons/thumb/6/64/AngelFoodCake.jpg/200px-AngelFoodCake.jpg</image>
        <cookingMethod>Baking</cookingMethod>
        <ingredients>18 egg whites, 2 t. cream of tartar, 1 pinch salt, 1 1/2 cups white sugar, 1 cup cake flour, 1/2 cup co
        <recipeInstructions>&lt;li&gt;Preheat oven to 350 degrees&lt;/li&gt;
        &lt;li&gt;Sift cake flour and confectioners sugar together 5 times. Set aside.&lt;/li&gt;
        &lt;li&gt;In a large bowl whip egg whites with salt until foamy. Add cream of tartar and continue to beat until soft
        &lt;li&gt;Quickly fold in flour mixutre.&lt;/li&gt;
        &lt;li&gt;Pour into a 10 inch tube pan.&lt;/li&gt;
        &lt;li&gt;Bake for 45 minutes.&lt;/li&gt;</recipeInstructions>
        <cookTime>PT45M</cookTime>
        <prepTime>PT15M</prepTime>
        <totalTime>PT60M</totalTime>
        <recipeCategory>dessert</recipeCategory>
        <recipeCuisine>American</recipeCuisine>
        <recipeYield>8 to 10 servings</recipeYield>
        <sameAs>http://dbpedia.org/page/Angel_food_cake</sameAs>
        <keywords>cream of tartar, confectioners sugar, vanilla extract, egg whites, cup cake flour, cake flour</keywords>
        <publisher>Montana State University Library</publisher>
        <dateCreated>2012-12-01T18:34Z</dateCreated>
        <dateModified>2012-12-01T18:34Z</dateModified>
        <datePublished>2012-12-01T18:34Z</datePublished>
        <inLanguage>en</inLanguage>
    </item>
</items>
```


...api?id=1&format=xml

...api?type=search&q=poetry&format=json

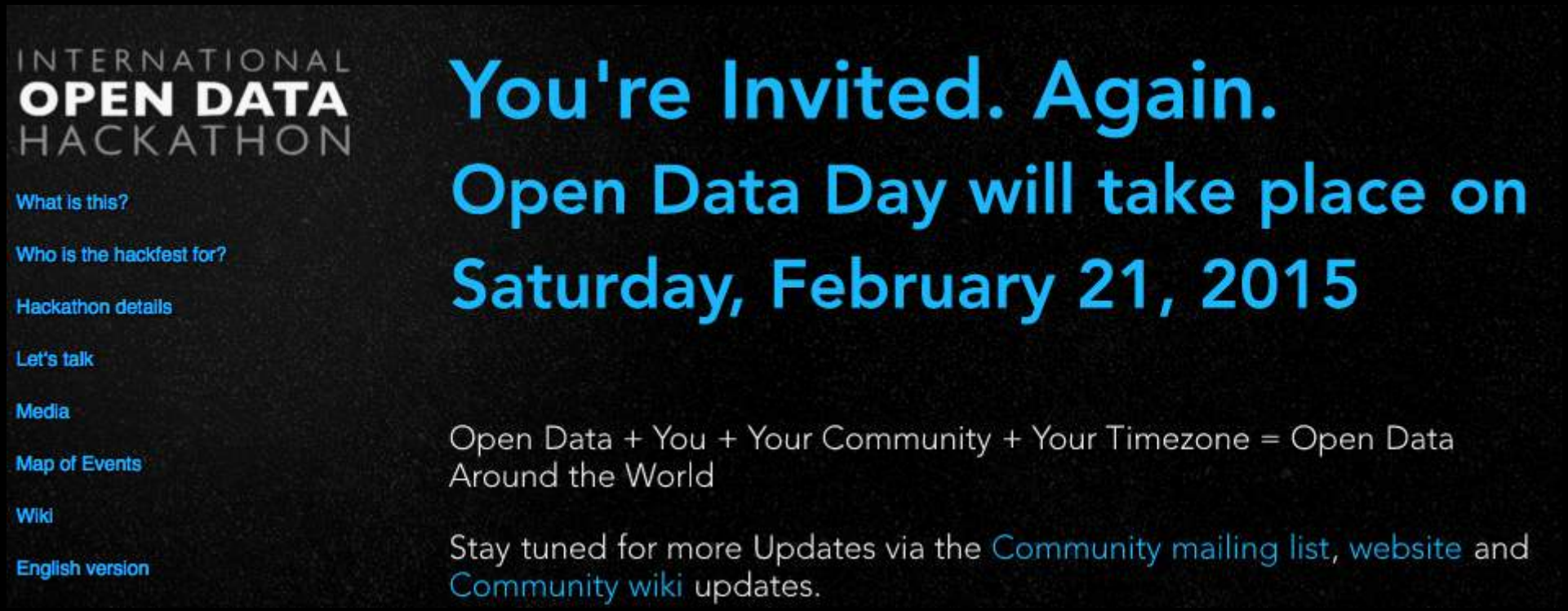
Open Data + Libraries

Where does it fit?

Teaching Data Science or Data Literacy Skills

Archiving and Packaging Data for Discovery and Reuse

Sponsoring Open Data Creation Moments



http://opendataday.org/

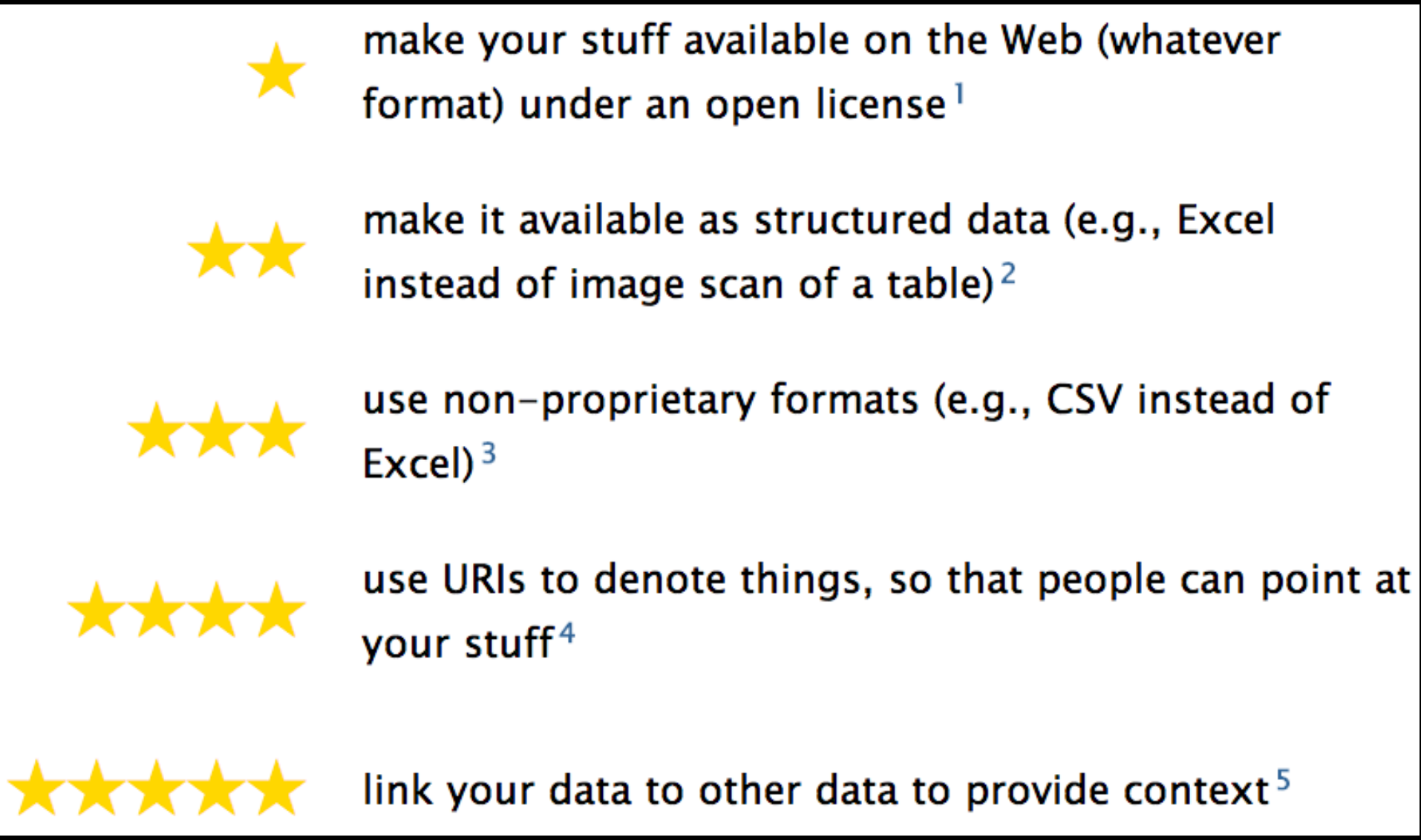
Open Data Processes + Tools

Getting Started

Structured Data

	A	B	C	D	E	F
1	Title	Author	Publisher	OCLC #	ISBN	Notes
2	This house of sky : landscapes of a Western mind	Ivan Doig	New York: Harcourt Brace Jovanovich, 1992.	25629631	0151900558	
3	To skin a cat : stories	Thomas McGuane	New York: Dutton/S. Lawrence, 1987.	15549047	0394755219	
4	Young men & fire	Norman Maclean	Chicago : University of Chicago Press, 1992.	283736539	0226500624	
5	A river runs through it, and other stories	Norman Maclean	Chicago: University of Chicago Press, 1976.	442266245	0671776975	Also a film version from Robert Redford http://en.wikipedia.org
6	The horse whisperer	Nicholas Evans	New York: Delacorte Press, 1995.	35722048	0440222656	
7	Hiking Yellowstone National Park	Bill Schneider	Helena, Mont: Falcon Press, 2003.	52500539	0762725397	2nd edition (Hiking Guide Series)

5 Star Open Data



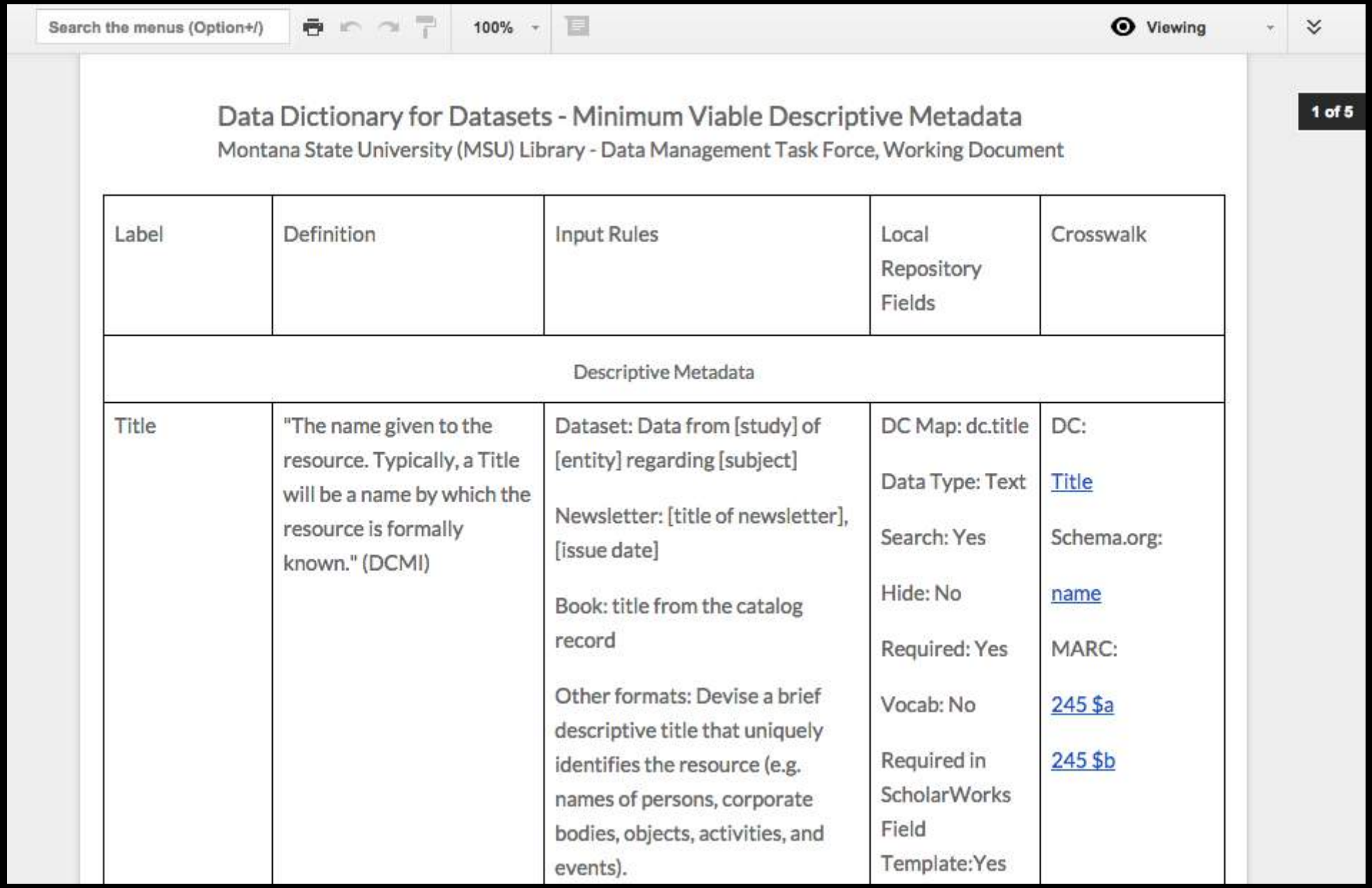
http://5stardata.info/

```
<?xml version="1.0"?>
<items>
    <item>
        <id>1</id>
        <name>Angel Food Cake</name>
        <creator>Laura Cameron</creator>
        <url>http://arc.lib.montana.edu/book/home-cooking-history-409/article-laura-cameron.html</url>
        <description>Angel food cake recipe written by Laura Cameron for Montana State University's upper division history c
        <about>baking cake</about>
        <image>http://upload.wikimedia.org/wikipedia/commons/thumb/6/64/AngelFoodCake.jpg/200px-AngelFoodCake.jpg</image>
        <cookingMethod>Baking</cookingMethod>
        <ingredients>18 egg whites, 2 t. cream of tartar, 1 pinch salt, 1 1/2 cups white sugar, 1 cup cake flour, 1/2 cup co
        <recipeInstructions>&lt;li&gt;Preheat oven to 350 degrees&lt;/li&gt;
        &lt;li&gt;Sift cake flour and confectioners sugar together 5 times. Set aside.&lt;/li&gt;
        &lt;li&gt;In a large bowl whip egg whites with salt until foamy. Add cream of tartar and continue to beat until soft
        &lt;li&gt;Quickly fold in flour mixutre.&lt;/li&gt;
        &lt;li&gt;Pour into a 10 inch tube pan.&lt;/li&gt;
        &lt;li&gt;Bake for 45 minutes.&lt;/li&gt;</recipeInstructions>
        <cookTime>PT45M</cookTime>
        <prepTime>PT15M</prepTime>
        <totalTime>PT60M</totalTime>
        <recipeCategory>dessert</recipeCategory>
        <recipeCuisine>American</recipeCuisine>
        <recipeYield>8 to 10 servings</recipeYield>
        <sameAs>http://dbpedia.org/page/Angel_food_cake</sameAs>
        <keywords>cream of tartar, confectioners sugar, vanilla extract, egg whites, cup cake flour, cake flour</keywords>
        <publisher>Montana State University Library</publisher>
        <dateCreated>2012-12-01T18:34Z</dateCreated>
        <dateModified>2012-12-01T18:34Z</dateModified>
        <datePublished>2012-12-01T18:34Z</datePublished>
        <inLanguage>en</inLanguage>
    </item>
</items>
```


...api?id=1&format=xml

...api?type=search&q=poetry&format=json

Minimum Viable Metadata



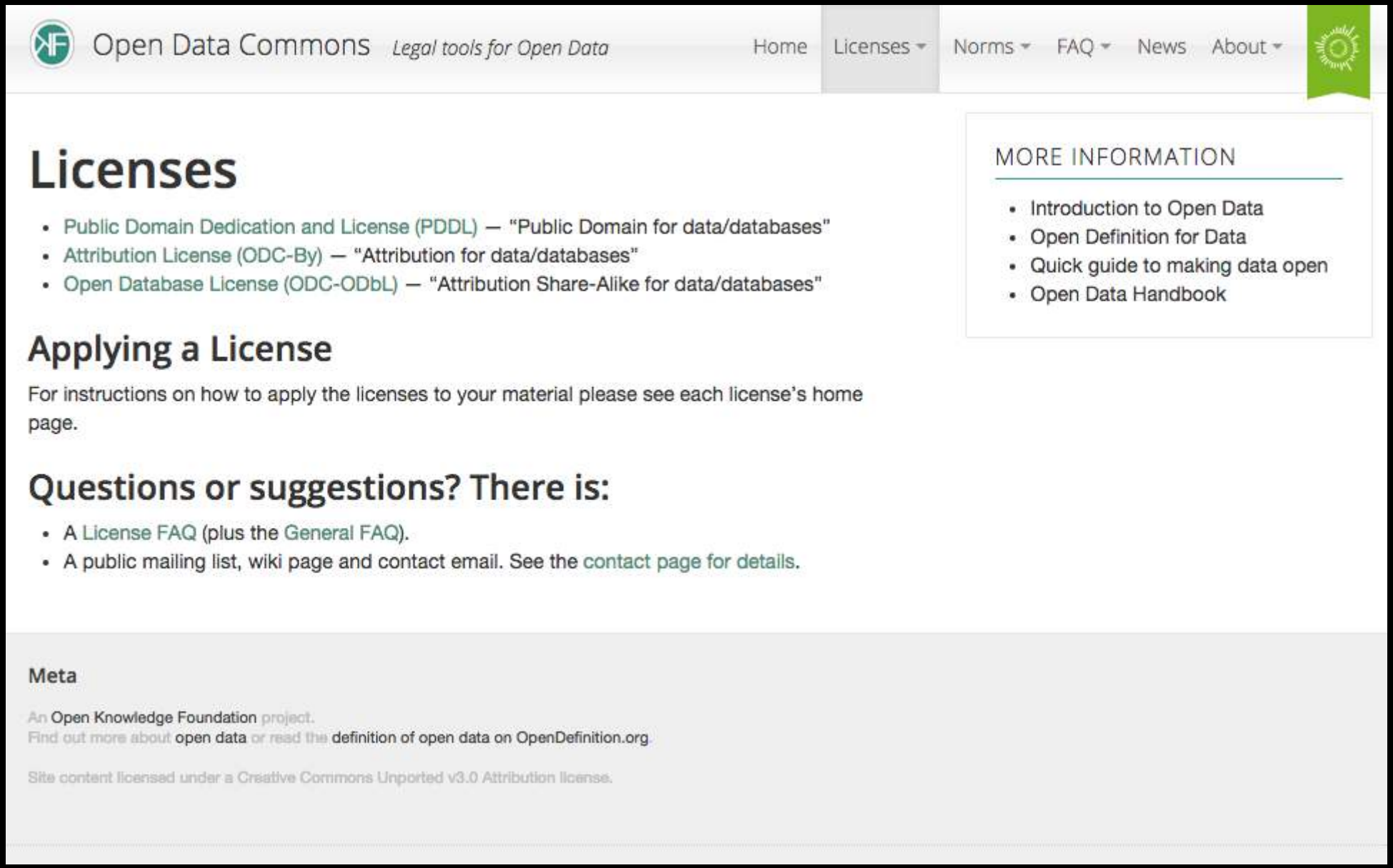
Data Dictionary for Datasets - Minimum Viable Descriptive Metadata

Montana State University (MSU) Library - Data Management Task Force, Working Document

Label	Definition	Input Rules	Local Repository Fields	Crosswalk
Descriptive Metadata				
Title	"The name given to the resource. Typically, a Title will be a name by which the resource is formally known." (DCMI)	Dataset: Data from [study] of [entity] regarding [subject] Newsletter: [title of newsletter], [issue date] Book: title from the catalog record Other formats: Devise a brief descriptive title that uniquely identifies the resource (e.g. names of persons, corporate bodies, objects, activities, and events).	DC Map: dc.title Data Type: Text Search: Yes Hide: No Required: Yes Vocab: No Required in ScholarWorks Field Template:Yes	DC: Title Schema.org: name MARC: 245 $a 245 $b

Data Dictionary for Datasets - Montana State University (MSU) Library

Licensing



http://opendatacommons.org/licenses/

Tools

OPEN DATA HANDBOOK

OPEN DATA HANDBOOK

Guides, case studies and resources for government & civil society on the "what, why & how" of open data.

Open Data Guide

This guide discusses the legal, social and technical aspects of open data. It can be used by anyone but is especially designed for those seeking to open up data. It discusses why to go open, what open is, and the how to 'open' data.

Start Reading

Value Stories

Use cases, stories and case studies highlighting the social and economic value, the impact and the varied applications of open data from cities and countries across the globe.

Value Stories

Resource Library

A curated collection of open data resources, including articles, longer publications, how to guides, presentations and videos, produced by the global open data community.

Open Data Resources

OPEN KNOWLEDGE SUPPORTED BY: SHUTTLEWORTH FUNDED THE WORLD BANK

Contribute Terms of use Privacy policy OPEN CONTENT

Related Projects: OpenGovernmentData.org | The DataHub | DataCatalogs.org | OpenSpending.org | SchoolOfData.org | OpenDefinition.org | Introducing Open Data

http://opendatahandbook.org/

Open Data Tools

Tools to Explore, Publish and Share Public Datasets

Home | Visualization | Data | Source Code | Information

EXMERG

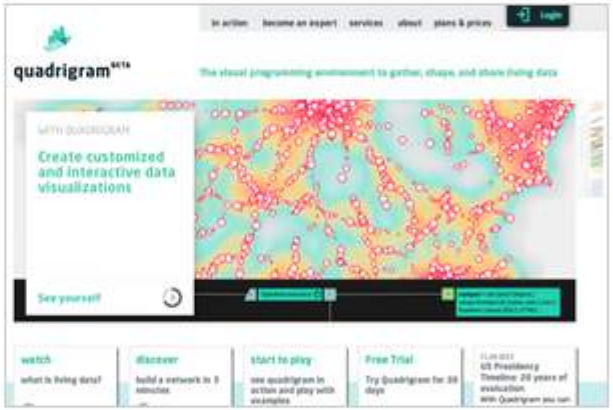
Tool for merging data from different spreadsheet or CSV files and creating dynamic charts.

http://www.exmerg.com

Wikidata

Large free knowledge base and repository for structured data for the Wikipedia.

http://www.wikidata.org



Quadrigram

Professional platform with paid plans for creating customized data visualizations.

http://www.quadrigram.com

http://opendata-tools.org/en/

OpenRefine

A free, open source, powerful tool for working with messy data

- Home
- Download
- Documentation
- Community
- Post archive

Mapping OpenRefine Ecosystem

2014 survey results

A Governance Model for OpenRefine

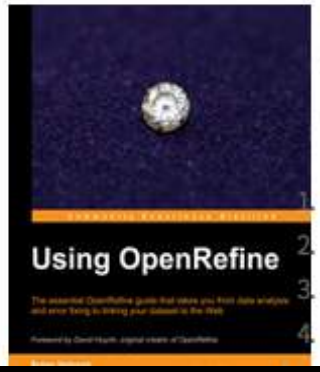
Using OpenRefine: a

Welcome!

OpenRefine (formerly Google Refine) is a powerful tool for working with messy data: cleaning it; transforming it from one format into another; extending it with web services; and linking it to databases like Freebase.

Please note that since October 2nd, 2012, Google is not actively supporting this project, which has now been rebranded to OpenRefine. Project development, documentation and promotion is now fully supported by volunteers. Find out more about the history of OpenRefine and how you can help the community.

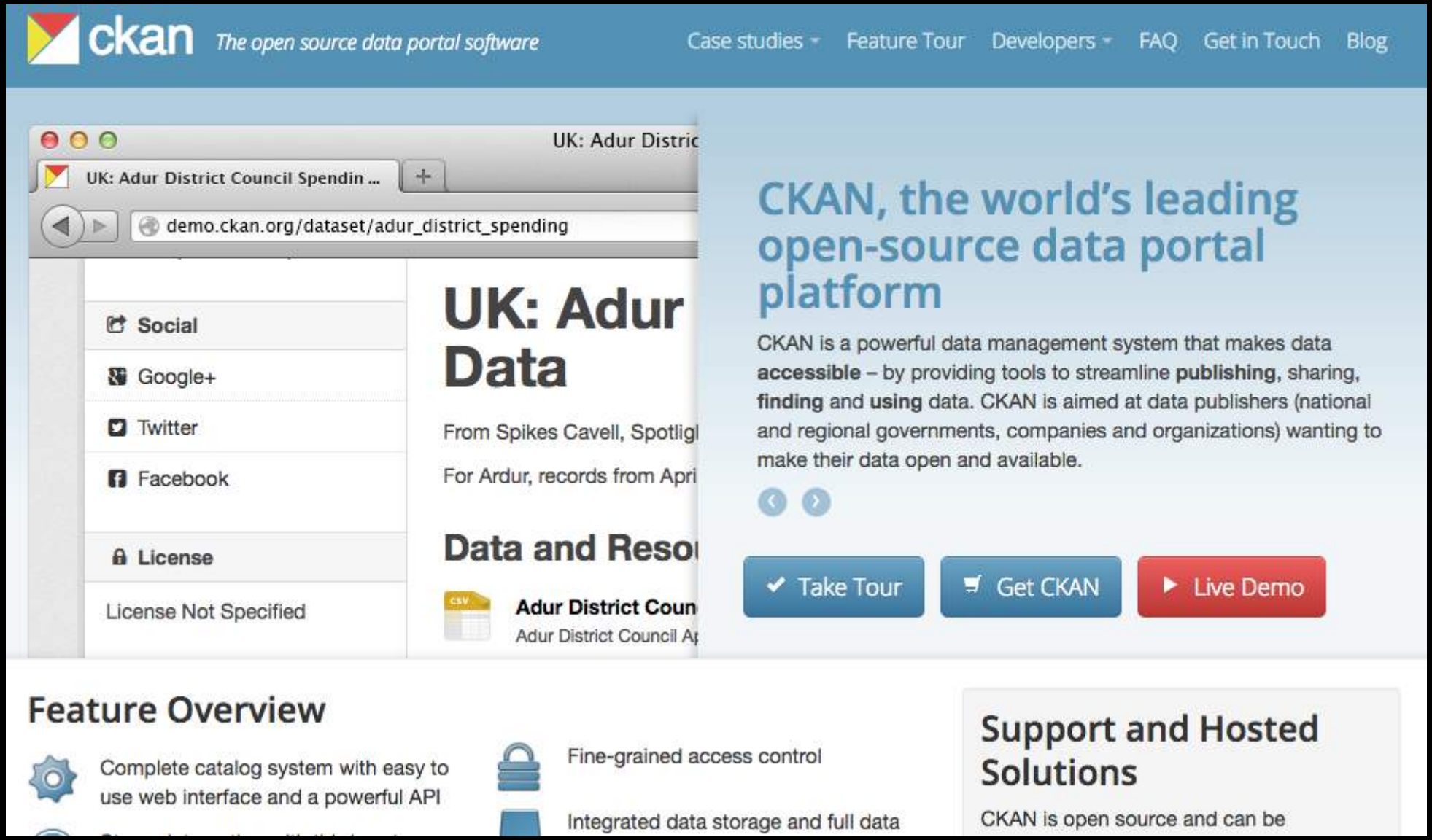
Using OpenRefine - The Book



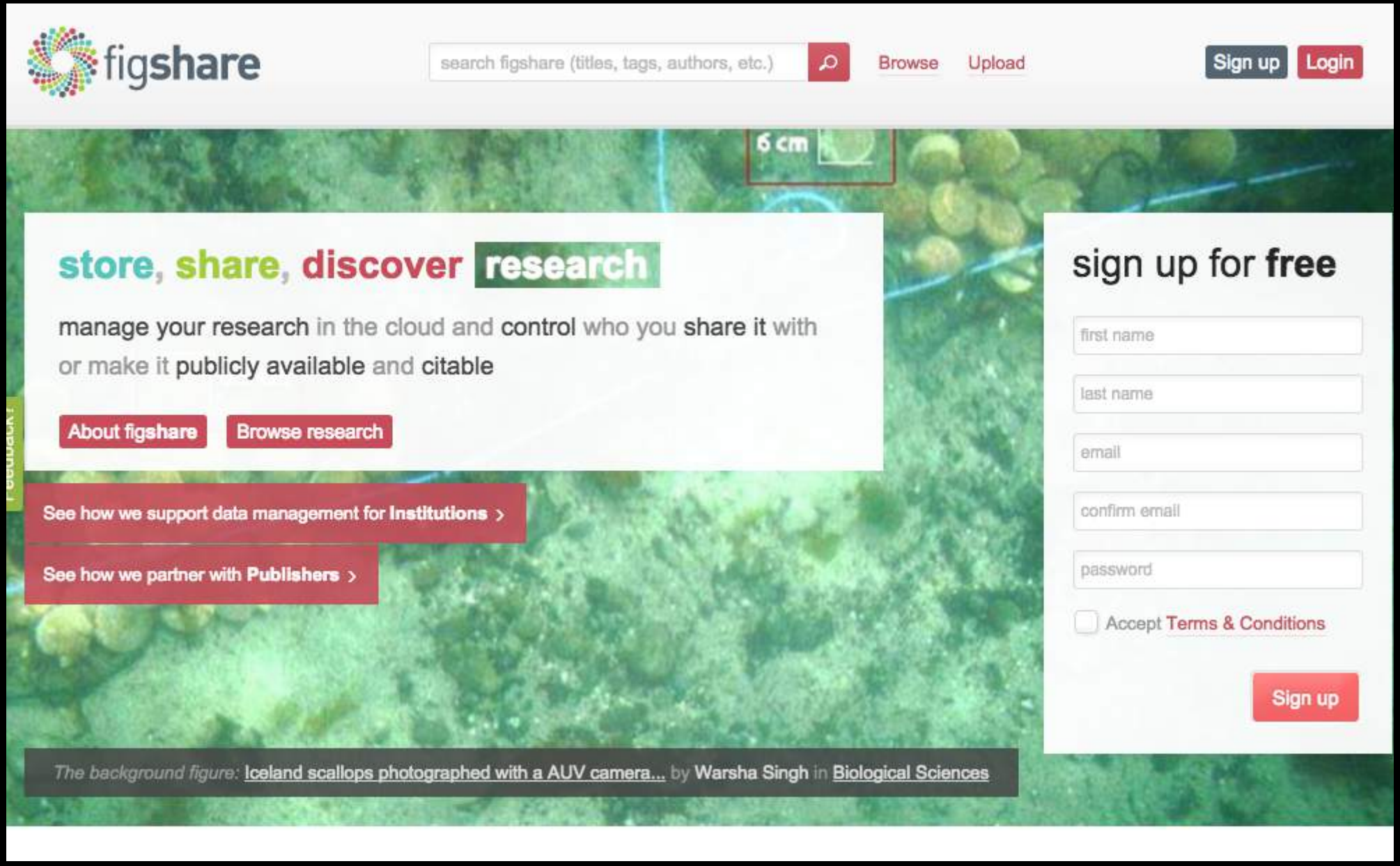
Using OpenRefine, by Ruben Verborgh and Max De Wilde, offers a great introduction to OpenRefine. Organized by recipes with hands on examples, the book covers the following topics:

1. Import data in various formats
2. Explore datasets in a matter of seconds
3. Apply basic and advanced cell transformations
4. Deal with cells that contain multiple values

http://openrefine.org/



http://ckan.org/



http://figshare.com/

Questions?

twitter.com/jaclark
www.lib.montana.edu/~jason/talks.php

Are you too busy to improve?

Håkan Forss @hakanforss http://hakanforss.wordpress.com

This illustration is inspired by and in part derived from the work by Scott Simmerman, "The Square Wheels Guy" http://www.performancemanagementcompany.com/