

# Why You Should Care About Linked Data and Open Data Linked Open Data (LOD) in Libraries

Florida Library Webinars  
Novare Library Services  
May 30, 2014

Jason Clark  
Head of Library Informatics & Computing  
Montana State University Library



pinboard.in tag

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

twitter as channel (#hashtag)

@jaclark #lodlib

# LINKED JAZZ



Fixed



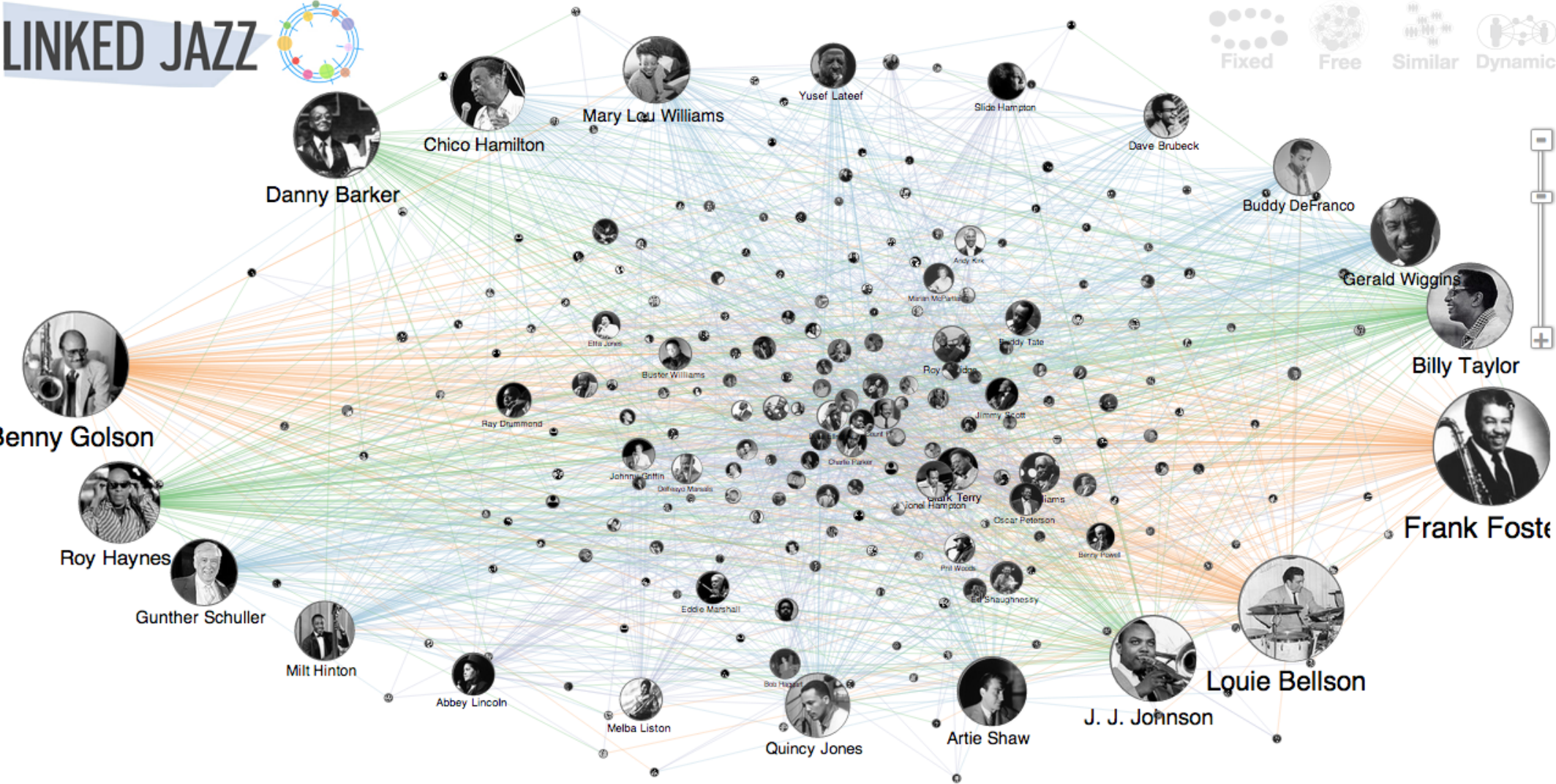
Free



Similar



Dynamic



<http://linkedjazz.org/>

# Overview

- What is Linked Open Data (LOD)?
- LOD in Practice
- Build a sample RDF
- LOD in Libraries
- Questions

# What is Linked Open Data?

- Practice of publishing structured, machine actionable data with an open license for sharing and reuse.

# What is Linked Data?

“a method of publishing structured data so that it can be interlinked and become more useful. It builds upon standard Web technologies such as HTTP, RDF and URIs, but rather than using them to serve web pages for human readers, it extends them to share information in a way that can be read automatically by computers. This enables data from different sources to be connected and queried.”

[http://en.wikipedia.org/wiki/Linked\\_data](http://en.wikipedia.org/wiki/Linked_data)



# What is Open Data?

“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/>

# LOD in Practice

- Design Expectations for LOD
- Expressing and Encoding LOD
- Principles of Publishing LOD

# Design Expectations for LOD

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF\*, SPARQL)
4. Include links to other URIs so that they can discover more things.

- Tim Berners Lee

<http://www.w3.org/DesignIssues/LinkedData.html>

# Expressing and Encoding LOD

## Resource Description Framework (RDF)

A structured data format for expressing relationships (descriptions) between things (resources).

<http://www.xml.com/pub/a/2001/01/24/rdf.html>

# Expressing and Encoding LOD

The data model of RDF is based on the idea of a triple.

A triple is a simple statement with three distinct parts:

1. a subject
2. a predicate
3. an object

	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 <a href="http://en.wikipedia.org">http://en.wikipedia.org</a>
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

(subject)

“has ISBN”

(predicate)

0671776975

(object)

#5 “has author” Norman Maclean  
(subject) (predicate) (object)



1. Use URIs as names for things

<http://www.worldcat.org/oclc/1733412> (subject)

<http://purl.org/dc/terms/creator> (predicate)

[http://dbpedia.org/page/Norman\\_Maclean](http://dbpedia.org/page/Norman_Maclean) (object)

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF PUBLIC "-//DUBLIN CORE//DCMES DTD
2002/07/31//EN"
"http://dublincore.org/documents/2002/07/31/dcmes-xml/dcmes-
xml-dtd.dtd">
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-
ns#"
    xmlns:dc ="http://purl.org/dc/elements/1.1/">
<rdf:Description rdf:about="http://www.worldcat.
org/oclc/1733412">
  <dc:creator rdf:resource="http://dbpedia.
org/page/Norman_Maclean"/>
  <dc:identifier>ISBN 0671776975</dc:identifier>
</rdf:Description>
</rdf:RDF>
```

# LOD Ontologies & Vocabularies

URIs link to:

1. Reputable machine-readable endpoints
  - E.g., Worldcat, VIAF, or DbPedia
2. Open Ontologies & Vocabularies
  - E.g., Dublin Core Terms, Library of Congress Subject Headings

\* See Linked Open Vocabularies (LOV) for a list of possible ontologies and vocabularies. <http://lov.okfn.org/dataset/lov/>

# Principles of Publishing LOD

- Use URIs for description
- Make structured data available on WWW
- Allow for machine-actionability and sharing through common ontologies

\* machine-actionable is a term used by Karen Coyle to mean readable and interpretable by computers or software agents. See her Library Technology Report on *Understanding the Semantic Web* at <http://www.metapress.com/content/g212v1783607/>

# 5 Star Linked Open Data



make your stuff available on the Web (whatever format) under an open license<sup>1</sup>



make it available as structured data (e.g., Excel instead of image scan of a table)<sup>2</sup>



use non-proprietary formats (e.g., CSV instead of Excel)<sup>3</sup>



use URIs to denote things, so that people can point at your stuff<sup>4</sup>



link your data to other data to provide context<sup>5</sup>

# Build a sample RDF

1. Go to Maya Angelou's Wikipedia page
  - a. [http://en.wikipedia.org/wiki/Maya\\_angelou](http://en.wikipedia.org/wiki/Maya_angelou)
2. Come up with 2 statements that can expressed as triples

# Build a sample RDF

What are some triples you identified?

Final step:

Place a triple into RDF and test in W3C  
RDF Validation Service

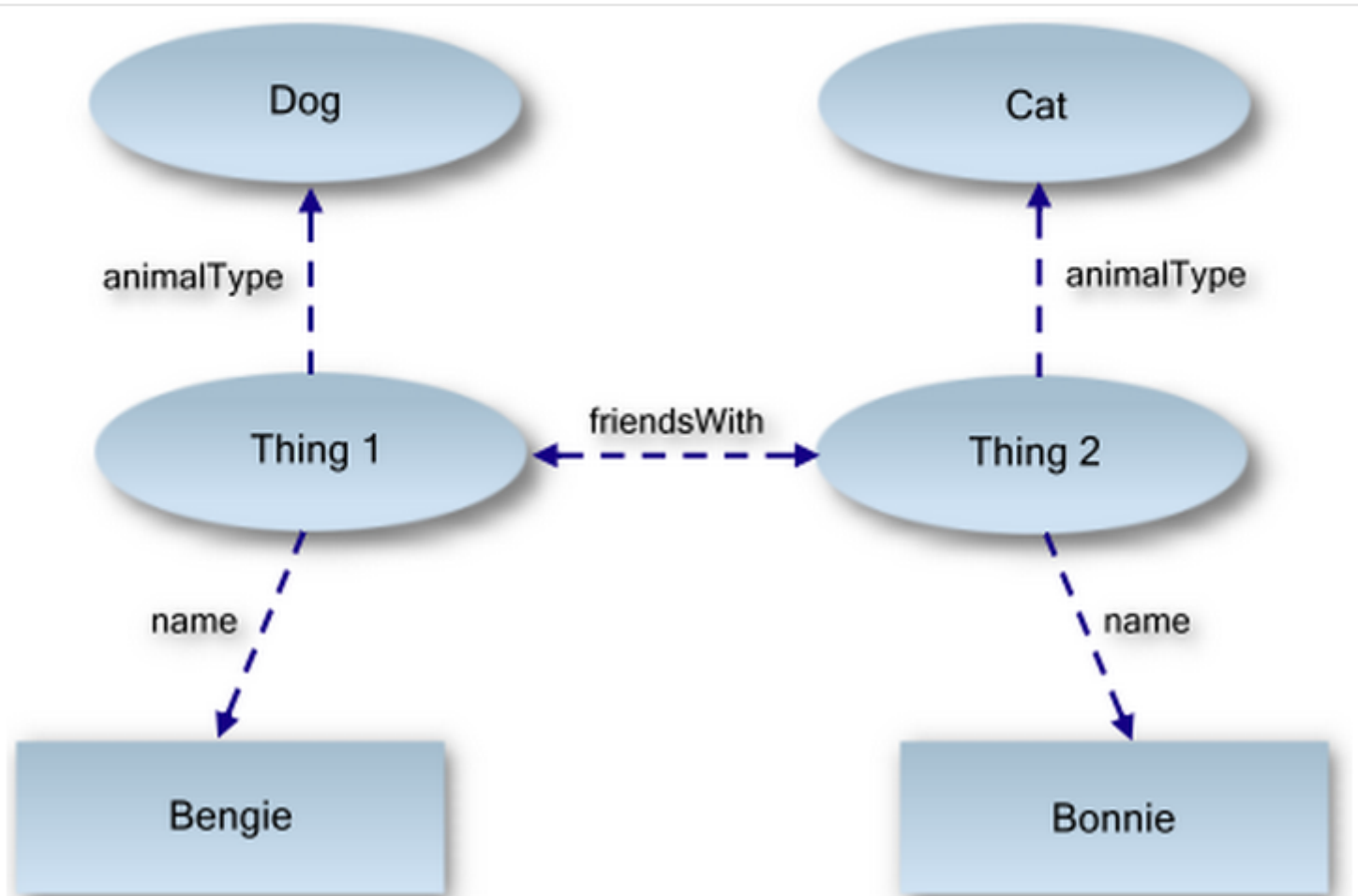


# A Graph Database

- A series of RDF triples that can be queried
- Expressions provide links between triples
- Non-hierarchical, arbitrary object relations

Bengie is a dog.  
Bonnie is a cat.  
Bengie and Bonnie are friends.

Using these three simple statements, let's turn this into a data graph:



# Why LOD?

- Libraries as publishers in Web of Data
- Reuse of our data = Significance
- Can improve findability in search engines (search engine optimization)
- Levels of description in LOD enable entirely new applications



Search for people, places and things



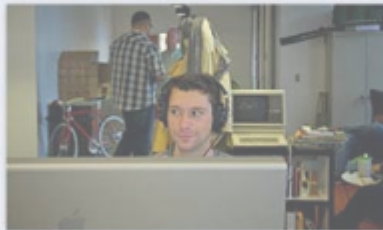
Jason

Home



# Introducing Graph Search

Q Photos I like



<https://www.facebook.com/about/graphsearch>

# LOD in Libraries

- UNLV Linked Data Project
  - <http://www.library.unlv.edu/linked-data>
- Worldcat
  - <http://www.oclc.org/data.en.html>
- Europeana Linked Open Data
  - <http://labs.europeana.eu/api/linked-open-data/introduction/>

# LOD in Libraries - Use Cases

- Web Scale Cataloging
  - Library of Congress BIBFRAME Initiative
  - <http://www.loc.gov/bibframe/>
- Annotated Texts and Books
  - <http://arc.lib.montana.edu/book/home-cooking-history-409/>
- Recommendation Systems
  - <http://linkedjazz.org/network/>

# Introducing Semantic Markup

- HTML5 semantic tags and RDFa that helps classify page types and types of content on the page

“If Google understands the content on your pages, we can create rich snippets—detailed information intended to help users with specific queries.”

[https://support.google.com/webmasters/answer/99170?hl=en&ref\\_topic=1088472](https://support.google.com/webmasters/answer/99170?hl=en&ref_topic=1088472)

Item: **Women posing on college locomotive engine****Title:** Women posing on college locomotive engine**Creator:** unknown**Date:** unknown**Description:** Women posing on a college locomotive engine.**Notes:****Physical Description:** Photo print - Black and White**Subjects:** [Locomotive](#)**Keywords:** locomotive**Photograph ID:** parc-000432 [Higher quality/resolution](#) [Persistent link](#) [Bookmark and Share](#) [Get embed code](#)[<a title="Women posing on college loco](#)[+ Back to MSU Photos Homepage](#)

# BM (Before Microdata)



Item: **Women posing on college locomotive engine****Title:** Women posing on college locomotive engine**Creator:** unknown**Date:** unknown**Description:** Women posing on a college locomotive engine.**Notes:****Physical Description:** Photo print - Black and White**Subjects:** [Locomotive](#)**Keywords:** locomotive**Photograph ID:** parc-000432 [Higher quality/resolution](#) [Persistent link](#) [Bookmark and Share](#) [Get embed code](#)[<a title="Women posing on college loco](#)[+ Back to MSU Photos Homepage](#)

# AM (After Microdata)

```
<dl>
<dt>Title</dt>
<dd>A River Runs Through It and Other
Stories</dd>
<dt>Author</dt>
<dd>Norman Maclean</dd>
<dt>Publication date</dt>
<dd>October 1, 2001</dd>
<dt>ID</dt>
<dd>0226500667</dd>
</dl>
```

```
<dl itemscope itemtype="http://schema.org/Book">
<dt>Title</dt>
<dd itemprop="title">A River Runs Through It and
Other Stories</dd>
<dt>Author</dt>
<dd itemprop="author">Norman Maclean</dd>
<dt>Publication date</dt>
<dd itemprop="pubdate">October 1, 2001</dd>
<dt>ID</dt>
<dd itemprop="isbn">0226500667</dd>
</dl>
```

# Resources

Harper, Corey. Library linked data: tuning library metadata for the semantic web. An ALCTS webcast, March 16. 2011. (open access) <http://www.ala.org/alcts/confevents/upcoming/webinar/cat/031611>



Berners-Lee, Tim. The next web. A TED talk, February 2009. (open access) [http://www.ted.com/talks/tim\\_berners\\_lee\\_on\\_the\\_next\\_web](http://www.ted.com/talks/tim_berners_lee_on_the_next_web)

Koster, Lukas (2011) Brief Introduction to Linked Data (open access) <https://docs.google.com/document/d/1W6UOCLgxTyM0BIPfd5hs58dh4k6CUdLW354AjjtnJfk/edit>

\*See also LODLAM - Linked Open Data in Libraries, Archives & Museums at <http://lodlam.net/>

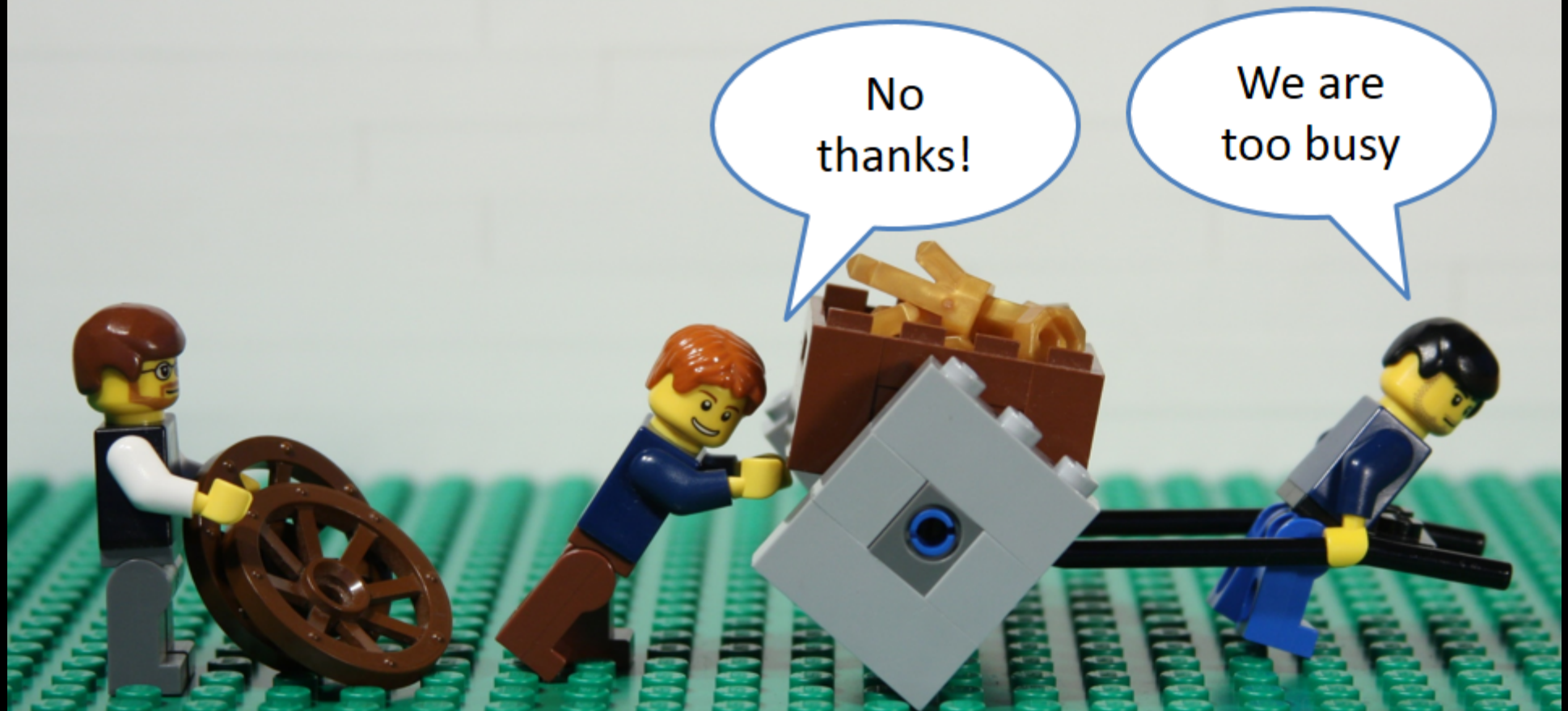
# Acknowledgements

I would like to thank Alison Hitchens for her excellent work in explaining and teaching these linked data concepts and definitions.

<http://www.accessola2.com/olita/insideolita/wordpress/?p=60029>

She was also gracious enough to distribute her work under the Attribution-Noncommercial-Share Alike 3.0 Unported license. And I am doing the same.

# 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/>

# Questions?

[twitter.com/jaclark](https://twitter.com/jaclark)

[www.lib.montana.edu/~jason/talks.php](http://www.lib.montana.edu/~jason/talks.php)