Training the trainers for Linked Data
A workshop
Seth van Hooland and Ruben Verborgh
October 11th 2014
DCMI 2014 conference – Austin
Agenda

• Introduction: understanding data models

• Four step process
  • Cleaning
  • Linking
  • Enriching
  • Publishing

• Discussion: current bottlenecks
Outcomes

- Give answers to the following questions in a pedagogical setting:
  - Possibilities and limits of RDF?
  - Importance of data quality?
  - How to clean, link and enrich metadata in a simple way?
  - What are the bottlenecks?
- Use the methods, tools and data of today’s workshop in your own teaching
Our approach

• Set the right expectations
• Theory, models and frameworks are means, not ends
• Tools that work for everyone
• Real-life data with real-life issues
• Methods to assess quality of results
Context

- Journal papers:
  - Cleaning - reconciliation: JASIST
  - NER: Literary and Linguistics Computing
  - REST: Journal of Documentation

- Handbook: ALA - Facet


- Courses at ULB and UW iSchool
Low hanging fruits

- Clean your metadata
- Reconcile with authoritative sources
- Enrich your metadata
- Publish your added-value metadata on the Web
http://www.google.com/trends/
Lana Del Rey rocks lime green-striped vintage dress for her sold-out concert in Austin
Daily Mail - 1 day ago
Lana Del Rey rocked a lime green-striped vintage dress for her sold-out concert at the Austin City Limits Music Festival last Saturday.

Meet Chuck Grant. Lana Del Rey's Equally Gorgeous And Talented Sister
Huffington Post - 3 days ago
Jessica Lange Got Freaky on 'American Horror Story' With David Bowie's 'Life On Mars'
SPIN - 2 days ago

More news for lana del rey

Lana Del Rey: Homepage
lanadelrey.com/
LANA DEL REY • HOME NEWS LIVE MUSIC GALLERY VIDEOS STORE. SIGNUP.
Facebook • Twitter • Youtube • Tumblr • Soundcloud ...

Lana Del Rey - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Lana_Del_Rey • Wikipedia
Elizabeth Woolridge Grant (born June 21, 1985), known by her stage name Lana Del Rey, is an American singer-songwriter. Del Rey started songwriting at the age of 15 and signed her first recording contract with 5 Points ...
Born: June 21, 1985 (age 36), New York City, NY
Full name: Elizabeth Woolridge Grant
Parents: Pat Grant, Rob Grant
Siblings: Caroline Grant, Charlie Grant

Lana Del Rey (@LanaDelRey) | Twitter
https://twitter.com/LanaDelRey
The latest Tweets from Lana Del Rey (@LanaDelRey). The roses had the look of flowers that are looked at. New album Ultraviolence out now. Download on ...

Lana Del Rey | Facebook
https://www.facebook.com/lanadelrey
Lana Del Rey. 109,412,657 likes · 8,690 talking about this. New album Ultraviolence out now. Buy on iTunes: http://lanadelreyiTunesSRfb...

18 Things You Learn After Two Long Days With Lana Del Rey
Jul 24, 2014 - There's only one person, it seems, who doesn't find Lana Del Rey a confounding mystery. "I know everything about myself," Del Rey says in her ...
Hold on ... Why do you actually want/need LD?
Data models
<table>
<thead>
<tr>
<th>Data Model</th>
<th>(Dis-)Advantages</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabular Data</td>
<td>+ intuitive approach</td>
<td>Import and export of data with a simple structure</td>
</tr>
<tr>
<td></td>
<td>+ very portable</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ technology agnostic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– prone to redundancy and leading to inconsistencies</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– inefficient search and retrieval</td>
<td></td>
</tr>
<tr>
<td>Relational Model</td>
<td>+ handling of complex data</td>
<td>Management of complex data which require normalization</td>
</tr>
<tr>
<td></td>
<td>+ optimized queries</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ mature software market</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– binary format</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– schema dependent</td>
<td></td>
</tr>
<tr>
<td>Meta-Markup</td>
<td>+ platform-independent</td>
<td>Import and export of complex data</td>
</tr>
<tr>
<td></td>
<td>+ both human and machine readable</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– complicated implementation for complex data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– verbosity</td>
<td></td>
</tr>
<tr>
<td>RDF</td>
<td>+ schemaless approach</td>
<td>Making data available for linking</td>
</tr>
<tr>
<td></td>
<td>+ discovery of new knowledge</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– loss of normalization</td>
<td></td>
</tr>
<tr>
<td></td>
<td>– immature software market</td>
<td></td>
</tr>
</tbody>
</table>
Don’t be the boy with the hammer!
Low hanging fruits

- Clean your metadata
- Reconcile with authoritative sources
- Enrich your metadata
- Publish your added-value metadata on the Web
# Pablo Picasso

**Picture:**
- [Image](#)
- [Image](#)
- [Image](#)
- [Image](#)
- [Image](#)
- [Image](#)

**Comment:**
- Pablo picasso [7,8,12,14]
- fotos, biografía y cuadros de pablo picasso [16]
- Voir toute l'actuálité de Pablo Picasso [20]

**Admins:**
- 729353303 [20]

**App ID:**
- 2490221586 [12,14,16]
- 210301922379857 [20]

**Alternate:**
- [http://www.slideshare.net/rss/latest](http://www.slideshare.net/rss/latest) [12,14,16]
- [http://www.slideshare.net/api/oembed/1?format=xml&url=http://www.slideshare.net/guest1c5fb9/pablo-picasso-2491014](http://www.slideshare.net/api/oembed/1?format=xml&url=http://www.slideshare.net/guest1c5fb9/pablo-picasso-2491014) [16]
- [http://www.slideshare.net/api/oembed/1?format=json&url=http://www.slideshare.net/guest1c5fb9/pablo-picasso-2491014](http://www.slideshare.net/api/oembed/1?format=json&url=http://www.slideshare.net/guest1c5fb9/pablo-picasso-2491014) [16]

[show 46 more values →](#)
<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>external id</td>
<td>Mx4rwBVYewpwEbGdrcN5Y29ycA [19]</td>
</tr>
<tr>
<td>fb app id</td>
<td>2490221586 [12]</td>
</tr>
<tr>
<td>height</td>
<td>355 [13,14,16]</td>
</tr>
<tr>
<td>noreferrer</td>
<td>show 31 values +</td>
</tr>
<tr>
<td>og title</td>
<td>Pablo Picasso [12]</td>
</tr>
<tr>
<td>og type</td>
<td>slideshare:presentation [12]</td>
</tr>
<tr>
<td>og description</td>
<td>PABLO PICASSO [12]</td>
</tr>
<tr>
<td>og url</td>
<td><a href="http://www.slideshare.net/iesjuliocarobaroga/pablo-picasso-1315023">http://www.slideshare.net/iesjuliocarobaroga/pablo-picasso-1315023</a> [12]</td>
</tr>
</tbody>
</table>

http://sig.ma/search?q=Pablo+Picasso&templateName=
The ROI of Data Quality seminar on June 10th, read more on the event page.

Data Quality Association Belgium

Activities
- DQA will be organizing courses, seminars and a yearly congress about Data Quality.
- Are you looking for information? A knowledge center will bring together all the generalizations and sector information you read about Data Quality.
- Membership is free. Just subscribe to our newsletter to stay posted.

Newsletter
- Sign up for the DQA newsletter to stay informed of all our current and upcoming activities.

http://www.dqa.be/

The MIT Total Data Quality Management Program

MIT TDQM Program Highlights
- MIT Information Quality Program
- The International Conference on Information Quality
- The ACM Journal of Data and Information Quality
- Information Quality Certificate Program

Corporate Householding Research
- CHH White Paper (Size 359KB)
- MIT-CHH Introduction (Size 35KB)
- MIT-CHH Description (Size 163KB)
- MIT-CHH Password required area

http://web.mit.edu/tdqm/
Evolution through time ...
Data cleaning tools

- Different communities:
  - Academia: Wrangler (http://vis.stanford.edu/wrangler) and Potters Wheel ABC (http://control.cs.berkeley.edu/abc/)
  - Hackers: Data pipes (http://datapipes.okfnlabs.org) and Nomenklatura (http://nomenklatura.okfnlabs.org)
  - General purpose tool: OpenRefine (http://openrefine.org)
OpenRefine

• Background and history :
  • Formerly Freebase Gridworks
  • GoogleRefine
  • OpenRefine

• Spreadsheet on steroids: simple but powerful
• Runs in a browser, but locally!
• Active development community (extensions) and discussion list
Installing OR

• Available for different platforms:
  
  • Windows: download the ZIP archive file and extract its contents to a folder of your choice. The software can be started by double-clicking the refine file.

  • Mac: download the DMG disk file and open it. You will see an OpenRefine icon, which you can drag into your Applications folder. Then double-click the icon to start it.

  • Linux: download the archive file and extract it. OpenRefine can be started by the ./refine command.

• Install version 2.5 !!!
Hands-on

• Working OpenRefine install?
• Download Powerhouse Museum data
• Create a new project
• Import the data
Untick this box
Memory allocation

• Under http://127.0.0.1:3333/preferences you can define the number of facet choices

• Allocate more memory to Refine:
  
  • Windows: open openrefine.l4j.ini file, find the line that starts with -Xmx and override the default allocated memory of 1024M with for example 2048 M
  
  • Mac: close Refine, hold control and click on its icon, selecting Show package contents from the pop-up menu. Open the info.plist file from the Contents folder. Navigate to the Java settings and edit the value of VMOptions. Look for the part that starts with -Xmx and change its default value of 1024 M to the desired amount of memory
  
  • Linux: instead of starting OpenRefine with ./refine as you usually would do, just type in ./refine -m 2048M
Afraid to mess up your data?

- You are working on a copy of your data
- OpenRefine keeps a history of your actions, so you can always roll back (see the Undo/Redo tab)
- Possibility to export your history
Low hanging fruits

- Clean your metadata
- Reconcile with authoritative sources
- Enrich your metadata
- Publish your added-value metadata on the Web
RDF Refine - a Google Refine extension for exporting RDF

Main features

Reconcile & interlink
- Reconcile against SPARQL endpoints, RDF dumps
- Search the Web for related RDF datasets

Export RDF
- GUI for defining the shape of the RDF graph
- Use your own vocabulary or import existing ones
- Autocomplete for property and class names

Download extension 0.8.0

(Stable/ALPHA) Download ORefine extension 0.9.0

Getting started...

1. Make sure "extensions" folder exists in your Google Refine workspace
2. Download the extension
3. Extract the downloaded zip file to the "extensions" folder
4. Restart Google Refine

Visit the installation guide for more detailed instructions. For tutorials-like material see our showcases. For detailed and technical documentation see the documentation section.

News

2014–03–06: a test version compatible with Open Refine 2.6 beta is released!
2012–07–27: version 0.8.0 released
Support for reconciling using Apache Stanbol is added. This feature has been initially implemented for LMF and partially funded by the Republic of Austria within the COMET project Salzburg NewMediaLab.
2012–01–18: version 0.7.5 released

http://refine.deri.ie/
Case study

- Experiment with reconciliation operations in a hands-on manner with the metadata of the Powerhouse museum and the LCSH

- Focus on the Categories field, populated with the Powerhouse museum Object Names Thesaurus (PONT), a locally created vocabulary
<table>
<thead>
<tr>
<th>Record ID</th>
<th>Object Title</th>
<th>Registration Num</th>
<th>Description</th>
<th>Marks</th>
<th>Production Date</th>
<th>Provenance (Prc)</th>
<th>Provenance (His)</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>7801</td>
<td>19142</td>
<td>Meteorite (cast), Weight 12 lbs; 3.84 oz (av.); Sp. Gr. 7.802; Battery Mountain (SB). Permission given by Mines Dept. to make the cast (SB). Found by miner named O'Shaughnessy - now living at Cowra on the highest peak of the Battery Mountain at the junction of the Burrowa and Lachlan Rivers.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Facet, Text filter, Edit cells, Edit column, Transpose, Sort...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19142</td>
<td>Meteors (cast), Weight 12 lbs; 3.64 oz (av.); Sp. Gr. 7.802; Battery Mountain (SB). Permission given by Mines Dept. to make the cast (SB). Found by miner named O'Shaughnessy - now living at Cowra on the highest peak of the Battery Mountain at the junction of the Burrowa and Lachlan Rivers.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>View</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using facets and filters

Use facets and filters to select subsets of your data to act on. Choose facet and filter methods from the menus at the top of each data column.

Not sure how to get started? Watch these screencasts.
This feature helps you find groups of different cell values that might be alternative representations of the same thing. For example, the two strings "New York" and "new york" are very likely to refer to the same concept and just have capitalization differences, and "Gödel" and "Godel" probably refer to the same person.

### Cluster & Edit column "Categories"

**Method:** key collision

**Keying Function:** fingerprint

**Ngram Size:** 2

#### Cluster Size 2

<table>
<thead>
<tr>
<th>Row Count</th>
<th>Values in Cluster</th>
<th>Merge?</th>
<th>New Cell Value</th>
</tr>
</thead>
</table>
| 1026      | Audio and Visual Equipment (1021 rows)  
            Audio and visual equipment (5 rows) |         | Audio and Visual Equipment |
| 683       | Photographic Equipment (635 rows)  
            Photographic equipment (48 rows) |         | Photographic Equipment |
| 615       | Food and Drink (614 rows)  
            food and drink (1 rows) |         | Food and Drink |
| 90        | Office Equipment (83 rows)  
            Office equipment (7 rows) |         | Office Equipment |
| 4661      | Documents (4680 rows)  
            documents (1 rows) |         | Documents |
| 930       | Musical Instruments (925 rows)  
            Musical instruments (5 rows) |         | Musical Instruments |
| 96        | Material Technology (89 rows)  
            Material Technology (7 rows) |         | Material Technology |
| 144       | Medicines (142 rows)  
            medicines (5 rows) |         | Medicines |

#### Cluster Size 2

<table>
<thead>
<tr>
<th>Row Count</th>
<th>Values in Cluster</th>
<th>Merge?</th>
<th>New Cell Value</th>
</tr>
</thead>
</table>
| 47        | Mailbags (43 rows)  
            Mail bags (4 rows) |         | Mailbags |
| 226       | Shirts (128 rows)  
            T-shirts (98 rows) |         | Shirts |
| 4         | Skullcaps (3 rows)  
            Skull caps (1 rows) |         | Skullcaps |
| 352       | Transport-Water (351 rows)  
            Transport - Water (1 rows) |         | Transport-Water |
| 662       | Transport-Air (680 rows)  
            Transport - Air (2 rows) |         | Transport-Air |
| 2         | Pillboxes (1 rows)  
            pill boxes (1 rows) |         | Pillboxes |

**Select All**  **Deselect All**  **Merge Selected & Re-Cluster**  **Merge Selected & Close**  **Close**
## Download

### Bulk Downloads

Bulk downloads, serialized as RDF/XML, Turtle, or N-triples, are available for the authorities and vocabularies. These may be downloaded as compressed files (ZIP format), with the exception of the LC Name Authority File, which is GZipped. Unless otherwise noted, each download contains MADS/RDF and SKOS/RDF representations of the data.

[Read more »](http://id.loc.gov/download/)

<table>
<thead>
<tr>
<th>Vocabulary/Authority</th>
<th>Date</th>
<th>Size</th>
<th>Download</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC Children's Subject Headings</td>
<td>11 Aug 2011</td>
<td>350 (kb)</td>
<td>nt</td>
</tr>
<tr>
<td>LC Children's Subject Headings</td>
<td>11 Aug 2011</td>
<td>147 (kb)</td>
<td>rdfxml</td>
</tr>
<tr>
<td>LC Children's Subject Headings</td>
<td>11 Aug 2011</td>
<td>116 (kb)</td>
<td>ttl</td>
</tr>
<tr>
<td>LC Genre/Form Terms</td>
<td>11 Aug 2011</td>
<td>493 (kb)</td>
<td>nt</td>
</tr>
<tr>
<td>LC Genre/Form Terms</td>
<td>11 Aug 2011</td>
<td>279 (kb)</td>
<td>rdfxml</td>
</tr>
<tr>
<td>LC Genre/Form Terms</td>
<td>11 Aug 2011</td>
<td>232 (kb)</td>
<td>ttl</td>
</tr>
<tr>
<td>LC Name Authority File (MADS/RDF only)</td>
<td>10 Mar 2014</td>
<td>3076988 (kb)</td>
<td>nt</td>
</tr>
<tr>
<td>LC Name Authority File (SKOS/RDF only)</td>
<td>10 Mar 2014</td>
<td>1313104 (kb)</td>
<td>nt</td>
</tr>
<tr>
<td>LC Name Authority File (MADS/RDF only)</td>
<td>10 Mar 2014</td>
<td>3296103 (kb)</td>
<td>rdfxml</td>
</tr>
<tr>
<td>LC Name Authority File (SKOS/RDF only)</td>
<td>10 Mar 2014</td>
<td>1426830 (kb)</td>
<td>rdfxml</td>
</tr>
<tr>
<td>LC Subject Headings (MADS/RDF only)</td>
<td>10 Mar 2014</td>
<td>182214 (kb)</td>
<td>nt</td>
</tr>
</tbody>
</table>

[http://id.loc.gov/download/](http://id.loc.gov/download/)
Reconcile column "Categories"

Reconcile each cell to an entity of one of these types:

- skos:Concept
  http://www.w3.org/2004/02/skos/core#Concept

Also use relevant details from other columns:

<table>
<thead>
<tr>
<th>Column</th>
<th>Include? As Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Record ID</td>
<td></td>
</tr>
<tr>
<td>Object Title</td>
<td></td>
</tr>
<tr>
<td>Registration</td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td></td>
</tr>
<tr>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>Marks</td>
<td></td>
</tr>
<tr>
<td>Production Date</td>
<td></td>
</tr>
<tr>
<td>Provenance</td>
<td></td>
</tr>
<tr>
<td>(Production)</td>
<td></td>
</tr>
<tr>
<td>Provenance</td>
<td></td>
</tr>
<tr>
<td>(History)</td>
<td></td>
</tr>
<tr>
<td>Persistent Link</td>
<td></td>
</tr>
<tr>
<td>Height</td>
<td></td>
</tr>
</tbody>
</table>

- Reconcile against type:
- Reconcile against no particular type
- Auto-match candidates with high confidence

Add Standard Service...  Add Namespaced Service...  Start Reconciling  Cancel
Meteorites

From Library of Congress Subject Headings

Meteorites

URI(s)

- http://id.loc.gov/authorities/subjects/sh85084315
- info:lc/authorities/sh85084315
- http://id.loc.gov/authorities/sh85084315#concept

Instance Of

- MADS/RDF Topic
- MADS/RDF Authority
- SKOS Concept

Scheme Membership(s)

- Library of Congress Subject Headings

Collection Membership(s)

- LCSH Collection - Authorized Headings
- LCSH Collection - General Collection
- LCSH Collection - May Subdivide Geographically

Broader Terms

- Meteors

Narrower Terms

- Achondrites
### Specimens

From Library of Congress Subject Headings

<table>
<thead>
<tr>
<th>Details</th>
<th>Visualization</th>
<th>Suggest Terminology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specimens</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Specimens**

- **URI(s)**
  - http://id.loc.gov/authorities/subjects/sh8706764
  - info:lc/authorities/sh8706764
  - http://id.loc.gov/authorities/sh8706764#concept

**Instance Of**
- MADS/RDF Topic
- MADS/RDF Authority
- SKOS Concept

**Scheme Membership(s)**
- Library of Congress Subject Headings

**Collection Membership(s)**
- LCSH Collection - General Collection

---

### Mineral Samples - Geological

<table>
<thead>
<tr>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Models (Patents) (0.375)</td>
</tr>
<tr>
<td>Models (Persons) - Vocational guidance (0.162)</td>
</tr>
<tr>
<td>Models and modelmaking - Radio control systems (0.133)</td>
</tr>
<tr>
<td>Create new topic</td>
</tr>
</tbody>
</table>

---

### Botanical Specimens

#### 2.

- **Mineral Samples - Geological**
  - Create new topic

#### 3.

- **Botanical Specimens**
  - Choose new match

#### 4.

- **Timber Samples**
  - Create new topic

- **Specimens**
  - Create new topic

---

### Wool Specimens

#### Create new topic

---

### Use as a form subdivision under types of publications, printed matter, etc., for actual specimens of the material.

**URI(s)**
- http://id.loc.gov/authorities/subjects/sh96001749
- info:lc/authorities/sh96001749
- http://id.loc.gov/authorities/sh96001749#concept

**Instance Of**
- MADS/RDF GenreForm
- MADS/RDF Authority
- SKOS Concept

**Scheme Membership(s)**
- Library of Congress Subject Headings

**Collection Membership(s)**
- LCSH Collection - Subdivisions
- LCSH Collection - GenreForm Subdivisions
- LCSH Collection - General Collection
- Pattern Heading Collection - H1095
Problems?

- OpenRefine does not differentiate between different types of labels.
- If one concept uses a label as its preferred label, and another uses the same label as alternate label, OpenRefine will not be able to choose between the two.
Example

- Skating is an alternative label of the term with preferred label Ice skating (sj96005713)
- But a separate term with preferred label Skating (sh85123105) already exists
- => Refine can not choose between both
- As long as this issue stands, alternate labels can do more harm then good
Pre-processed LCSH

• Changes in our version of LCSH:
  • Subdivisions are only present if they do not conflict with an existing main heading
  • Alternate labels were added, to the extent they did not cause clashes with other labels
Adding pre-processed LCSH

- **Name**: LCSH (pre-processed)
- **Endpoint URL**: http://sparql.freeyourmetadata.org
- **Graph URI**: http://sparql.freeyourmetadata.org/authorities-processed
- **Type**: Virtuoso
- **Label properties**: tick only skos:prefLabel, then tick other and enter the alternate label URI: http://www.w3.org/2004/02/skos/core#altLabel
Pre-processed LCSH

• Duplicate the column values of Categories in order to compare the results: Edit column > Add column based on this column...

• Give the new column a name like Categories (LCSH pre-processed)

• Remove the pre-existing reconciliation facets by clicking Remove all
<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meteorites</td>
<td>5664</td>
</tr>
<tr>
<td>Models</td>
<td>96183</td>
</tr>
<tr>
<td>Models (Patents)</td>
<td>103801</td>
</tr>
<tr>
<td>Models (Persons)--Vocational guidance</td>
<td></td>
</tr>
<tr>
<td>Models and modellmaking--Radio control</td>
<td></td>
</tr>
<tr>
<td>systems</td>
<td></td>
</tr>
<tr>
<td>Mineral Samples--Geological</td>
<td></td>
</tr>
<tr>
<td>Botanical specimens</td>
<td></td>
</tr>
<tr>
<td>Timber Samples</td>
<td></td>
</tr>
<tr>
<td>Wool specimens</td>
<td></td>
</tr>
<tr>
<td>Animal Samples and Products</td>
<td></td>
</tr>
</tbody>
</table>

Facet by choice counts:
- (blank) 5664
- matched 96183
- none 103801
Add column based on column Categories (pre-processed LCSH)

New column name

On error
- set to blank
- store error
- copy value from original column

Expression

`cell.recon.match.id`

Language
- Google Refine Expression Language (GREL)

No syntax error.

<table>
<thead>
<tr>
<th>row</th>
<th>value</th>
<th>cell.recon.match.id</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Meteorites</td>
<td><a href="http://id.loc.gov/authorities/sh85084315#concept">http://id.loc.gov/authorities/sh85084315#concept</a></td>
</tr>
<tr>
<td>2.</td>
<td>Models</td>
<td><a href="http://id.loc.gov/authorities/sh00006390#concept">http://id.loc.gov/authorities/sh00006390#concept</a></td>
</tr>
<tr>
<td>3.</td>
<td>Models</td>
<td><a href="http://id.loc.gov/authorities/sh00006390#concept">http://id.loc.gov/authorities/sh00006390#concept</a></td>
</tr>
<tr>
<td>4.</td>
<td>Mineral Samples-Geological</td>
<td>Error: Cannot retrieve field from null</td>
</tr>
<tr>
<td>5.</td>
<td>Botanical specimens</td>
<td><a href="http://id.loc.gov/authorities/sh20010000705#concept">http://id.loc.gov/authorities/sh20010000705#concept</a></td>
</tr>
<tr>
<td>6.</td>
<td>Botanical specimens</td>
<td><a href="http://id.loc.gov/authorities/sh20010000705#concept">http://id.loc.gov/authorities/sh20010000705#concept</a></td>
</tr>
<tr>
<td>7.</td>
<td>Botanical specimens</td>
<td><a href="http://id.loc.gov/authorities/sh20010000705#concept">http://id.loc.gov/authorities/sh20010000705#concept</a></td>
</tr>
</tbody>
</table>

OK  Cancel
JASIST results

LCSH + AAT
68.4%

LCSH
81.1%

AAT
77.1%

PHM Collection
89.8% of records reconciled
Frank Lloyd Wright has had a hand in 53 objects that are part of our collection. Specifically: 39 objects as Architect and 13 objects as Designer and one object as Style of.

Frank Lloyd Wright has been involved in work collected by the following departments: 39 objects from Drawings, Prints, and Graphic Design and 5 objects from Textiles and 5 objects from Product Design and Decorative Arts and 4 objects from Wallcoverings.

This person has collaborated with Kelmscott Gallery and Henry J. Allen and Karin Zonis and George Mann Niedecken and Imperial

Short URL: http://cprhw.tt/p/2AtgP/
Person ID: 18042049
Tag as: ch:person=18042049

We're also pretty confident we know who Frank Lloyd Wright is at Design Museum (London) and Freebase and Indianapolis Museum of Art and MoMA and Victoria and Albert Museum and Virtual International Authority File and Wikipedia.

http://collection.cooperhewitt.org
Believing that "the space within that building is the reality of that building", FRANK LLOYD WRIGHT (1867-1959) was one of the most prolific and influential architects of the 20th century. From his early Prairie Style homes, to the sculptural curves of the Guggenheim Museum in New York he defined a North American style of architecture which was rich in emotion and sensitive to its surroundings.

One of the founders of modern architecture in North America, Frank Lloyd Wright embraced the use of new technology, media and engineering to create some of the 20th century's most influential and iconic buildings. During a long and productive career spanning 70 years he designed over 1,000 buildings of which over 400 were built.

Wright developed a language of architecture that did not look to Europe but was unique to the United States. As well as creating buildings which were radical in appearance, Wright had a rare ability to integrate them with the landscape - stemming from his deep love and knowledge of nature. It was this gift that marked him out from contemporary pioneers of modern architecture, such as Le Corbusier and Mies van der Rohe, and make his buildings seem in tune with our environmentally conscious era.

Born in 1867, Wright was the eldest child of William Russell Cary Wright, a Unitarian minister and music teacher, and Anna Lloyd Jones Wright. His father gave him a love of music, but it was his mother who encouraged him to become an architect. As well as hanging prints of landscape on his bedroom wall, she bought him a Frederich Froebel Kindergarten system on a visit to the Philadelphia Centennial in 1876. This system consisted of a set of coloured strips of paper, two-dimensional geometric prisms and a set of wooden bricks comprising cubes, spheres and pyramids. Later Wright wrote “the maple wood blocks... all are in my fingers to this day”. An infinite and playful combination of these geometric shapes gave Wright the core forms of his architecture.

At 18, Wright enrolled to study engineering at the University of Wisconsin, Madison, but, desperate to pursue a career in architecture, he dropped out and moved to Chicago where he quickly found work with the architectural firm of Joseph Lyman Silsbee. Wright's ambitions, however, soon took him to Adler and Sullivan, Chicago's most progressive architects. Louis Sullivan was an important influence on Wright and put him in charge of the firm's residential building work. He also gave him a loan in 1889 to purchase land to build a home for himself and his new wife, Catherine Lee Tobin, in the Oak Park district of Chicago. In 1893 Wright was asked to leave the firm for pursuing too much private work and at the age of 26 he started his own practice.

During the next 16 years Wright developed the Prairie Style of architecture in a large number of commissions for private houses in Chicago, in particular, in Oak Park. It is to his credit that most of his clients were extremely pleased with the homes Wright built. One of his less published achievements was his mastery of the internal environment, with great attention paid to lighting, heating and climate control. The Prairie Style aimed to create a truly North American architecture, but Wright also drew inspiration from Europe: from the French rationalist writings of Eugene Viollet-le-Duc and the British Arts and Crafts movement. He also had great knowledge of the art and architecture of Japan and the culture of pre-Columbian America. Although radical, Wright can be viewed within the context of a group of US architects and designers, who included Gustav Stickley and the brothers Charles and Henry Greene. They had similar ideas about how to incorporate traditional elements and climate into their work. These ideas would soon influence Wright's own work and eventually he developed his own unique style.

The 1900 Robie House in Chicago was Wright's most mature expression of the Prairie Style of architecture. Frederick Robie, an engineer and industrialist, wanted a house full of light with views of the street, but without his neighbours looking in. Using brick, concrete, steel and glass, Wright constructed a massive cantilever on the west side of the house that gave the living room privacy and shelter from the sun. It also opened out the house from moving away from the tight box shape of traditional homes. The low, horizontal form is exaggerated with the use of ribbon of cream stone for the base plinth and corbelling and red brick for the walls. A central fireplace open above the mantel gave greater unity of space to the large living and dining rooms, which Wright saw as the centre of family life. Although there was no external garden, the use of massive planters and urns softened the hard edges of the building and at the main entrance Wright designed a terrace, balcony or porch to break the division between inside and outside. All internal details - including the furnishing, light fittings, rugs and the essential art glass - were also designed by Wright.

Wright was also asked to build the 1905 Unity Temple, a place of worship for the Universalist Church in Oak Park. Coming from a long tradition of Universalists, he accepted the commission on a very slim budget of $45,000. Due to these financial constraints Wright bought for the first time with poured concrete. A square two-storey space housed the temple of worship and behind it was a rectangular parish meeting house for socialising. The temple of worship had to seat 400 people yet Wright still managed to create an intimate space. To enhance the visual drama, these two structures were connected by a modest entrance with low ceiling. The roof of the building was supported by the four square masses in the room, the poured concrete walls therefore became screens with glass windows above windows.

Wright was now a popular and established architect, but he entered a phase of emotional turmoil. In 1909 after falling in love with the wife of a client and neighbour, Mamah Borthwick Cheney. Leaving Wright's wife and six children and closing his studio, the couple fled to Berlin. During this time Wright worked on a book of his work for the German publisher Ernst Wasmuth as well as travelling to Austria, Italy and France. He returned to the US in 1911 and managed to secure enough money to build a house for himself, the Robie House. He and his wife and two children moved into the new house in February 1912. The Robie House was designed in collaboration with the Imperial Hotel, the hotel which consumed him from 1912 to 1915 and his design was finally completed in 1922. The main feature of the 100-room hotel was the central three-storey lobby and two-storey dining room, ballroom and auditorium. The use of soft, sandy brick and Oya stone enabled Wright to use extensive carving and decoration. When the Great Kantō earthquake struck Tokyo in 1923, the floating foundations and reinforced steel construction ensured that the Imperial was one of the few buildings to survive, although most of it was demolished in 1969.

While in Japan, Wright received a commission from the oil heiress and theatrical producer, Aline Barnsdall to build a house, shops and theatre complex for her in Los Angeles. Only the main house, the 1926 Hollyhock House, and residences A and B were constructed. Inspired by his experiences in Japan, Wright had a new sense of freedom with decoration and applied the abstracted motifs of a hollyhock, a favourite flower of the client, in cast concrete to parapets, pinnacles and planters. In form the Hollyhock House is the link between Wright's early Prairie Style and his later textile block concrete houses. It also reflected his newfound interest in Mayan temple design.

http://designmuseum.org
Hold on ...
So we can’t connect objects who refer to the same link?
Re-decentralisation

By design the web has no centre. Anybody can create a new website. When one site fails, the rest of the web continues unabated. Individual links are allowed to break so the entire web does not. This architecture enabled the web to scale and produced the long-tail distribution of sites so conducive to innovation and an open market. However, some popular and successful services (search, social networking, email) have achieved near-monopoly status. Although industry leaders often spur positive change, we must remain wary of concentrations of power as they can make the web brittle.

By continually "re-decentralising" the web, we will unleash the next generation of technology, business and social innovators. In particular, I look forward to new approaches to video, photos, music and game distribution. We have seen some progress (such as DRM-free music) but there are still hard technical, business and legal problems to solve. Some solutions may disrupt people's lives and livelihoods, an important reason to pursue social inclusion via the web.

http://www.wired.co.uk/magazine/archive/2014/03/web-at-25/tim-berners-lee
« I continue Doug’s work by keeping the links outside the document »
Low hanging fruits

• Clean your metadata
• Reconcile with authoritative sources
• Enrich your metadata
• Publish your added-value metadata on the Web
What is NER?

• Consider the sentence « On 25 September 2006, we visited Washington to see the White House »

• First step => identification
  • 25 September 2006
  • Washington
  • White House

• Second step => disambiguation
What is NER?

- Each entity is associated with a meaning:
  - [http://dbpedia.org/resource/White_House](http://dbpedia.org/resource/White_House)

- NE extraction workflow consists of analyzing input content for detecting named entities, assigning them a type weighted by a confidence score and by providing a list of URIs for disambiguation.
Free Your Metadata
Learn how to get more value out of metadata easily

Cleanup  Reconciliation  Extraction  Access  Reaching out  About

Named entity extraction
The techniques we discussed in the Cleanup and Reconciliation parts come in very handy when your data is already in a structured format. However, many fields (notoriously description) contain unstructured text, yet they usually convey a high amount of interesting information. To capture this in machine-processable format, named entity recognition can be used.

Download our Refine extension
Named-entity recognition has never been easier: thanks to our brand new OpenRefine extension, you can enrich your description fields right from your workspace.

Refine extension
Our free OpenRefine extension makes named-entity recognition easier than ever before.
Download

Publication
Read the pre-print of our forthcoming publication on named-entity recognition for cultural heritage collections.

http://freeyourmetadat.org/named-entity-extraction/
Adding extra services

- You need to request an API key to make use of the services of Alchemy and Zemanta:
  - http://www.alchemyapi.com/api/register.html
- Click the Named-entity recognition toolbar button and choose Configure API keys
- Add the keys you received and click Update
LLC results

DBpedia Spotlight

AlchemyAPI

Zemanta

1.1%

0.0%

18.3%

11.3%

0.0%

2.2%

11.3%

2.2%

4.3%

19.4%
<table>
<thead>
<tr>
<th>Title</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>A course of religious instruction for Catholic youth</td>
<td>by John Gerard</td>
</tr>
<tr>
<td>A full course of instructions for the use of catechists</td>
<td>being an explanation of the catechism, entitled “An abridgement of Christian doctrine” by John Perry</td>
</tr>
<tr>
<td>A garden of roses. Stories and sketches</td>
<td>by Maurice Francis Egan</td>
</tr>
<tr>
<td>A new treatise on the duty of a Christian towards God [microform]</td>
<td>being an enlarged and improved version of the original treatise / written by J.B. de La Salle ; translated from the French by Mrs. J. Sadlier</td>
</tr>
<tr>
<td>A short life of Our Lord</td>
<td>by P. J. Crean ; with a foreword by e the Cardinal Archbishop of Westminster</td>
</tr>
<tr>
<td>A victim to the seal of confession a true story</td>
<td>by the Reverend Joseph Spillmann</td>
</tr>
<tr>
<td>Advice to Irish girls in America</td>
<td>by Alton Park; or, Conversations on religious and moral subjects; chiefly designed for the amusement and instruction of young ladies</td>
</tr>
<tr>
<td>An adventure with the Apaches</td>
<td>by Gabriel Ferry [pseud.]</td>
</tr>
<tr>
<td>An explanation of the Baltimore catechism of Christian doctrine</td>
<td>for the use of Sunday-school teachers and advanced classes / by Thomas L. Kinkead</td>
</tr>
<tr>
<td>Angelica / from the German of Christopher von Scmid</td>
<td></td>
</tr>
<tr>
<td>Batavia, by Hendrik Conscience</td>
<td>Translated from the original Flemish</td>
</tr>
<tr>
<td>Bobby in Movieland, by Francis J. Finn</td>
<td></td>
</tr>
<tr>
<td>Breton legends / Translated from the French</td>
<td></td>
</tr>
<tr>
<td>But Thy love and Thy grace</td>
<td>by Francis J. Finn ; with illustrations by Charles C. Svendsen</td>
</tr>
</tbody>
</table>

[http://dh.crc.nd.edu/sandbox/cyl/catalog/]
Batavia, by Hendrik Conscience. Translated from the original Flemish.

http://dh.crc.nd.edu/sandbox/cyl/catalog/
Low hanging fruits

• Clean your metadata
• Reconcile with authoritative sources
• Enrich your metadata
• Publish your added-value metadata on the Web
REST
How to expose data in a sustainable way?
How to expose data in a sustainable way?

The “final” answer has evolved over time.

1997  HTML 3.2  You need a website!
2000  XML
2005  JSON  You need an API!
2014  RDF  ??
2020  ???????  You need two APIs?
Introducing REST

Publishing data on the Web

Core principles of REST

Fallacy of the multi-API culture
Introducing REST

Publishing data on the Web

Core principles of REST

Fallacy of the multi-API culture
You have a website.
Now you want something more.

Museum: “Could you build me an API?”
IT company: “Of course!”

If you ask for an API, you’ll get one.
Even if you didn’t need it.
(And you don’t.)
What the customer asked for.
How the project leader understood it.
How the developers implemented it.
How the customer was billed.
How the marketeers sold it.
What the customer really needed.
So what does our “customer” actually need?

A way to publish data on the Web for humans and machines now and in the future.
What does “sustainable” actually mean?

The method to access information should not change over time.

The format of the information should adapt to technological change.
Do current Web publishing methods provide sustainable access?

http://museum.org/api/v3/showObject.php?id=34&format=json&api_key=GYVS311dgjTJhkdgIHBDLKKP

Will PHP still exist in the future? JSON?

Can you share this URL on Twitter?

Will this URL work with RDF clients?
Introducing REST

Publishing data on the Web

Core principles of REST

Fallacy of the multi-API culture
REST (Representational State Transfer): what it is *not*. Nice URLs, HTTP, non-XML, a format, a tool.
REST (Representational State Transfer): is an architectural style for distributed hypermedia systems.

a set of constraints
If you follow them,
you inherit their benefits.

created by analyzing the Web

Why does the Web work?
The REST architectural style consists of several constraints:

- client-server constraint
- statelessness constraint
- cacheability constraint
- uniform interface constraints
- layered system constraint

The uniform interface constraints are the most characteristic ones.

client-server constraint

statelessness constraint

cacheability constraint

uniform interface constraints

layered system constraint

The **uniform interface constraints** are the most characteristic ones.

- identification of resources
- manipulation through representations
- self-describing messages
- hypermedia as the engine of application state
A resource relates an identifier to a concept.

**Don’t.**

http: //example.org/collection/showObject.php

**Do.**

http://example.org/objects/18353113/

What is this?
Can I bookmark this?
Can I share this?

What is this?
Can I bookmark this?
Can I share this?
Resources are **conceptual** relationships.

A museum Web page shows a certain object of its collection at the URL `http://example.org/objects/18353113/`. Another object is accessible at `/objects/35460799/`, and similarly, each object has its own URL, which we can share or bookmark for later usage.

The important difference with for instance database systems, is that the URL relation is conceptual. This has been illustrated in Figure 2. The entities of our application domain are indicated on top in white; they are objects in a collection and exist independently of the Web application we built on top of it. On the bottom row, resources and their corresponding URL can be seen. For instance, the conceptual mapping "Toy Theater" will always correspond to the object with that name in the collection through the URL `/objects/18353113/`. Interestingly, the conceptual mapping "latest addition" will always point to the most recently acquired object in the collection, but the identity of that concrete object will of course change over time. This indicates the conceptual nature of resources.

### 3.3.2 Resource manipulation through representations

Now that a shared concept of identification between client and server has been established, the next question is how information can be transferred between them. An essential property of REST architectures is that resources themselves are **not** transferred; instead, client and server exchange a representation of a resource. This contrasts with file systems, wherein an identifier (file name) always corresponds to a specific physical representation (the file). On the Web, an identifier (a URL) corresponds to a conceptual entity (a resource), which can have different representations, depending on the capabilities of the client. For instance, a resource can be represented in HTML for human viewers, and in JSON for consumption by software such as JavaScript applications in the browser.

**Non-REST compliant**

A museum provides access to the HTML version of an object in its collection at `http://example.org/objects/18353113/`. However, the JSON version must be accessed through `http://api.example.org/getObjectJson.php?id=18353113`, and an API key is necessary for all requests.

**REST compliant**

The object is accessible through `http://example.org/objects/18353113/` and, depending on the request, the server replies with HTML or JSON. In the future, RDF might be supported through this same URL.

Note how in the non-compliant example, the identification happens on the technical level instead of the conceptual level. The URL identifies "the HTML representation of object 18353113" instead of "object 18353113". This makes the exchange of URLs between different systems difficult, as the choice for a specific representation is tied to the URL. Furthermore, the addition of new representation formats (such as RDF) would imply that new URLs have to be assigned again, not to mention the complexity of removing support for old representation formats. Without...
Each resource can have multiple representations.

Don’t.

http://example.org/objects/18353113 gives HTML

http://api.example.org/getObjectJson.php?id=18353113 gives JSON

Can I bookmark this?
Can I share this?

Do.

http://example.org/objects/18353113 gives HTML.

http://example.org/objects/18353113 gives JSON.
gives RDF.

Can I bookmark this?
Can I share this?
Identifiers can be exchanged between current and future clients.

---

**Collection Item**

*"Toy Theater"

- JSON
- RDF
- HTML

(objects/18353113/)

**Photograph**

*"Toy Theater"

- JPEG
- PNG
- PDF

(objects/18353113/photos/1/)
Servers and clients send self-descriptive messages.

Don’t.
/objects?filter=toy
/?page=2

Do.
/objects?filter=toy
/objects?filter=toy&page=2

What is this?
Can I bookmark this?
Can I share this?

What is this?
Can I bookmark this?
Can I share this?
The interaction should be driven by hypermedia.

Don’t.

```
{
  "title": "Spun Chair",
  "producer": {
    "id": 1804
  }
}
```

Can I act on this?

Do.

```
{
  "title": "Spun Chair",
  "producer": {
    "url": "/producers/1804"
  }
}
```

Can I act on this?
The uniform interface constraints are the most characteristic ones.

- identification of resources
- manipulation through representations
- self-describing messages
- hypermedia as the engine of application state
Why is hypermedia so important for the Web?

http://www.google.com/
Why is hypermedia so important for the Web?

To use this site, enter your search term as follows:
http://www.google.com/search?q=search+term
Don’t create an API.
Your website is the API.

The REST architectural style treats all clients the same way.

It uses information as the corner stones of your API.

Your information remains in place while technologies evolve.
Introducing REST

Publishing data on the Web

Core principles of REST

Fallacy of the multi-API culture
Europeana and DPLA chose a non-REST API.

Consequences:

They had to actually build an API, making development and maintenance much more expensive.

Their APIs are not sustainably designed, so much more money will need to be spent.
How do I use the Europeana website as a human?

Dispute between an eight-year old boy and Confucius (Hasa no ko to Koshi no mondo) ()

Title: Nitiren in gesprek met een kind bij een waterval (andere titel)

Title: Uji shui Monogatari (Supplementary tales from Uji), nr. ? (?) (serietitel)

Description: De Chinese filosof Confucius (Koshi) in gesprek met een kleine jongen die voor hem staat. Achter hen een waterval en opkommende zon. Links rotsen en takken van een pijnboom.

Creator: : Gakutei, Yashima
Contributor: Legaat van J.A. Bierens de Haan, Amsterdam
How do I use the Europeana website as a machine?

You don’t...

You have to use the API.
How do I use the Europeana API as a machine?

In this section you will find all information necessary to jump-start using Europeana API.

**Request**

Every Europeana Search API call is an HTTP request in a specified format that is sent to the Europeana API service. The API root URL is located at:

```
http://www.europeana.eu/api/v2
```

For obligatory request parameters look into the documentation of specific calls. The authentication section provides information on the obligatory authentication parameter `wskey`.

**Response**

A response to an API call will always contain a number of standard fields that precede the fields specific for the call. The standard part contains the following fields:

<table>
<thead>
<tr>
<th>Field</th>
<th>Datatype</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>the authentication parameter <code>wskey</code></td>
</tr>
</tbody>
</table>
How do I use the Europeana API as a machine?
Europeana’s design choices to publish data have consequences.

You cannot bookmark a URL.
You cannot share a URL.
You cannot share a message’s body.
You cannot use it without documentation.
You cannot use this API in the future.
API keys to shield off representations are a fallacy.

Protect the information, not the representation.

Blocking access to JSON is meaningless if the HTML of the same resource is accessible.

An access mechanism is not part of a resource.
Multiple APIs are a fallacy, their maintenance is a nightmare.

Your website is your API.

Offer access for humans and machines.

Identify resources, extend representations.
What the customer really needed?

Machine-readable representations.
The bigger picture

• How could we describe the current state of LD? What bottlenecks need to be resolved?

• Let’s step back from the practical work we’ve done and think about the bigger picture:
  • Role of statistics and probability to make sense out of LD
  • Impact of market forces
  • Current ambiguities regarding the usage of URLs
Statistics and probability

- Confronted with the volume of triples, decisions have to be made in regards to how many and in what order triples are shown.

- Try to ask DBPedia for example:

  ```sql
  SELECT COUNT(DISTINCT ?s) WHERE { ?s ?p ?o }
  ```
There is a tension between:

- the search for generally applicable laws: Chomsky (linguistics), Willey and Philips (archeology)
- the hermeneutical tradition as defined by Dilthey, focusing on the subjective and the singular
We use tools from disciplines whose epistemological foundations are at odds with, or even hostile to, the humanities. Positivistic, quantitative and reductive, these techniques preclude humanistic methods because of the very assumptions on which they are designed: that objects of knowledge can be understood as ahistorical and autonomous... Probability is not the same as ambiguity or multivalent possibility within the field of humanistic inquiry. The task of calculating norms, medians, means and averages will never be the same as the task of engaging with anomalies and taking their details as the basis of an argument.

— Johanna Drucker,
Debates in the DH
Henri IV, né Henri de Bourbon le 13 décembre 1553 à Pau et assassiné le 14 mai 1610 à Paris, fut roi de Navarre (Henri III de Navarre, 1572-1610) puis roi de France (1589-1610), premier souverain français de la branche dite de Bourbon de la dynastie capétienne. Il était le fils de Jeanne III, de son nom patronymique Jeanne d'Albret, reine de Navarre, et d'Antoine de Bourbon, chef de la maison de Bourbon, descendant du roi Louis IX et premier prince du sang.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
</table>
- Henry IV (13 December 1553 – 14 May 1610), Henri-Quatre, was King of Navarre (as Henry II) from 1572 to 1610 and King of France from 1589 to 1610. He was the first French monarch of the House of Bourbon. Baptised a Catholic, he converted to Protestantism along with his mother Jeanne d'Albret, Queen of Navarre. He inherited the throne of Navarre in 1572 on the death of his mother. As a Huguenot, Henry was involved in the French Wars of Religion; he barely escaped assassination at the time of the St. Bartholomew's Day massacre, and he later led Protestant forces against the royal army. As a French Prince of the Blood by reason of his descent from King Louis IX, he ascended the throne of France upon the death of his childless cousin Henry III in 1589. In accepting the throne, he found it prudent to abjure his Calvinist faith. Regardless, his coronation was followed by a four-year war against the Catholic League to establish his legitimacy. One of the most popular French kings, both during and after his reign, Henry showed great care for the welfare of his subjects. As a pragmatic politician, he displayed an unusual religious tolerance for the time. Notably, he enacted the Edict of Nantes in 1598, which guaranteed religious liberties to Protestants, thereby effectively ending the Wars of Religion. He was assassinated by François Ravallié, a fanatical Catholic.  
- Enrique de Borbón, fue rey de Navarracon el nombre de Enrique III entre 1572 y 1610 y rey de Francia como Enrique IV entre 1589 y 1610, primero de la Casa de Borbón en este país, conocido como Enrique el Grande (Henri le Grand) o el Buen Rey (Le bon roi Henri) y coprinicé de Andorra (1562-1610). A menudo es considerado por los franceses como el mejor monarca que ha gobernado su país, siempre intentando mejorar las condiciones de vida de sus súbditos. Se le atribuye la frase: «Un polo en las ollas de todos los campesinos, todos los domingos», que simboliza perfectamente su política de hacer feliz a su pueblo, no sólo con poder y conquistas, sino también con paz y prosperidad. Es el referente de los monárquicos franceses, los cuales realizan todos los años un homenaje frente a su estatua del Pont Neuf (Puente Nuevo) de París el día de su entrada a la ciudad. Le fue dedicada la marcha Vive Henri IV (Viva Enrique IV), que llegó a ser el himno de facto del reino y el de luce durante la restauración, y es conocida aún hoy día por los franceses. |
Henry IV of France

Henry IV, Henri-Quatre, also known by the epithet "Good King Henry", was King of Navarre from 1572 to 1610 and King of France from 1589 to 1610. He was the first French monarch of the House of Bourbon. Baptised as a Catholic but raised in the Protestant faith by his mother Jeanne d'Albret, Queen of Navarre, he inherited the throne of Navarre in 1572 on the death of his mother. As a Huguenot, Henry was involved in the French Wars of Religion, he barely escaped assassination at the time of the St. Bartholomew's Day massacre, and he later led Protestant forces against the royal army. As a French "Prince of the Blood" by reason of his descent from King Louis IX, he ascended the throne of France upon the death of his childless brother-in-law Henry III in 1589. In accepting the throne, he found it prudent to abjure his Calvinist faith. Regardless, his coronation was followed by a four-year war against the Catholic League to establish his legitimacy. As a pragmatic politician, he displayed an unusual religious tolerance for the time. Notably, he promulgated the Edict of Nantes in 1598, which guaranteed religious liberties to Protestants, thereby effectively ending the Wars of Religion. He was assassinated by François Ravaillac, a fanatic Catholic, and was succeeded by his son Louis XIII. Wikipedia [-]

<table>
<thead>
<tr>
<th>Properties</th>
<th>18n</th>
<th>Keys</th>
<th>Links</th>
</tr>
</thead>
</table>

View and edit specific domains, types, or properties.

Filter options: Show all domains and properties

### Metaweb System Types

- Object
- Common
- Topic
- Art
- Art Subject
- Freebase
- Object profile
- People
- Person
- Deceased Person
- avh's types
- Elterdale Topic
- Organization

http://www.freebase.com/m/0fw6r
About: **Henri IV de France**  
An Entity of Type: [http://yago-knowledge.org/resource/wordnet_whole_100003553](http://yago-knowledge.org/resource/wordnet_whole_100003553), within Data Space: [lod.openlinksw.com](http://lod.openlinksw.com) associated with source dataset(s).

Type: whole

New Facets Session with This Class

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>entity</td>
</tr>
<tr>
<td></td>
<td><a href="http://yago-knowledge.org/resource/wikicategory_Dauphins_of_France">http://yago-knowledge.org/resource/wikicategory_Dauphins_of_France</a></td>
</tr>
<tr>
<td></td>
<td><a href="http://yago-knowledge.org/resource/wikicategory_Dauphins_of_Viennois">http://yago-knowledge.org/resource/wikicategory_Dauphins_of_Viennois</a></td>
</tr>
<tr>
<td></td>
<td>convert</td>
</tr>
<tr>
<td></td>
<td>dauphin</td>
</tr>
<tr>
<td></td>
<td>»more»</td>
</tr>
</tbody>
</table>

is [http://yago-knowledge.org/resource/linksTo](http://yago-knowledge.org/resource/linksTo) of

André de Brancas  
André Dulaurens  
Arnoul de Lisle  
Anthony Mildmay  
Antoine de Chandieu  
»more»

Faceted Search & Find service v1.13.53  
Alternative Linked Data Views: **iSPARQL** | **ODE**  
Raw Data in: **CXML** | **CSV** | **RDF** (N-Triples N3/Turtle JSON XML) | **OData** (Atom JSON) | **Microdata** (JSON HTML) | **JSON-LD** | **About**

http://www.yago-knowledge.org/resource/Henry_IV_of_France
Difference?

- **DBPedia**:
  ```
  dbpedia:Henry_IV_of_France
  dbpprop:religion dbpedia:Catholic_Church,
  "previously Huguenot";
  dcterms:subject category:Roman_Catholic_monarchs,
  category:Converts_to_Roman_Catholicism_from_Calvinism,
  People_excommunicated_by_the_Roman_Catholic_Church;
  rdf:type yago:Convert109962414.
  ```

- **Freebase**:
  ```
  ns:m.0fw6r ns:people.person.religion ns:m.0c8wxp.
  ```

- **Yago**:
  ```
  yago:Henry_IV_of_France rdf:type
  yago:wikicategory_Converts_to_Roman_Catholicism_from_Calvinism.
  ```
Market forces

• Small number of players are imposing their way of rendering semantics on the Web explicit

• Impact on the long term consequences of this tendency?
OpenGraph

<html prefix="og: http://ogp.me/ns#">
<head>
<title>The Rock (1996)</title>
<meta property="og:title" content="The Rock" />
<meta property="og:type" content="video.movie" />
<meta property="og:url" content="http://www.imdb.com/title/tt0117500/" />
<meta property="og:image" content="http://ia.media-imdb.com/images/rock.jpg" />
</head>
<div itemtype="http://schema.org/Movie">
<h1 itemprop="name">The Rock</h1>
  <div itemprop="director" itemscope itemtype="http://schema.org/Person">
    <span itemprop="name">Rob Marshall</span>
  </div>
</div>

<div itemprop="actor" itemscope itemtype="http://schema.org/Person">
</div>
Usage of URLs

• Cornerstone of your Linked Data

• TBL in regards to cool URIs : « Do what I say, not what I do »

• Reflect about the affordance of URLs : what do they exactly identify ? A resource or a document about that resource, or a document about that document about that resource, etc
Get in touch!

- Host a workshop - tutorial?
- Handbook - a review copy anyone?
- Follow @freemetadatadata, @RubenVerborgh and @sethvanhooland