Training the trainers for Linked Data

Seth van Hooland and Ruben Verborgh
September 1st 2015
DC2015 - São Paulo
http://www.google.com/trends/
São Paulo — Wikipédia
São Paulo (prononciation en portugais : /ˈsɐ̃w paw.ˈlu/ ou Sao Paulo (« Saint-Paul » en français) est une ville du sud-est du Brésil. Avec ses onze millions ...
Histoire - Géographie - Population - Criminalité

Images correspondant à sao paulo

Signaler des images inappropriées

Plus d'images pour sao paulo

são paulo | globoesporte.com
globoesporte.globo.com/futebol/times/sao-paulo/ ▶ Traduire cette page
Veja aqui tudo sobre São paulo no globoesporte.com.

São Paulo
Ville au Brésil
São Paulo ou Sao Paulo est une ville du sud-est du Brésil. Avec ses onze millions d'habitants en 2014, elle est la plus grande ville du pays et la deuxième d'Amérique latine. C'est également la ville qui compte le plus de lusophones au monde. Wikipédia
Superficie : 1 523 km²
Météo : 30 °C, vent NO à 11 km/h, 29 % d'humidité
Heure locale : lundi 17:53
Linked Data for Libraries, Archives and Museums
How to clean, link and publish your metadata
Seth van Hooland and Ruben Verborgh
Price: £49.95
Price (to CILIP members): £39.96
ISBN: 978-1-85604-964-1

FORTHCOMING
Pre-order at Bookpoint by calling +44 (0)1235 827702, emailing facet@bookpoint.co.uk or faxing +44 (0)1235 827703

This highly practical handbook teaches you how to unlock the value of your existing metadata through cleaning, reconciliation, enrichment and linking and how to streamline the process of new metadata creation.

Libraries, archives and museums are facing up to the challenge of providing access to fast growing collections whilst managing cuts to budgets. Key to this is the creation, linking and publishing of good quality metadata as Linked Data that will allow their collections to be discovered, accessed and disseminated in a sustainable manner.

Metadata experts Seth van Hooland and Ruben Verborgh introduce the key concepts of metadata standards and Linked Data and how they can be practically applied to existing metadata, giving readers the tools and understanding to achieve maximum results with limited resources.

Readers will learn how to critically assess and use (semi-)automated methods of managing metadata through hands-on exercises within the book and on the accompanying website. Each chapter is built around a case study from institutions around the world, demonstrating how freely available data can be used to advantage.
• Reconciliation : JASIST
• NER : LLC
• REST : JoD
• Links to the three pre-prints available on http://freeyourmetadata.org/publications/
Low hanging fruits

• Clean your metadata
• Reconcile with authoritative sources
• Enrich your metadata
• Link to others who already have the metadata you need
• Publish your added-value metadata on the Web
But wait ... 
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>
</table>
| tabular data | + intuitive approach  
+ very portable  
+ technology agnostic  
– prone to redundancy and leading to inconsistencies  
– inefficient search and retrieval | import and export of data with a simple structure       |
| relational model | + handling of complex data  
+ optimized queries  
+ mature software market  
– binary format  
– schema dependent | management of complex data which require normalization |
| meta-markup  | + platform-independent  
+ both human and machine readable  
– complicated implementation for complex data  
– verbosity | import and export of complex data                     |
| RDF         | + schemaless approach  
+ discovery of new knowledge  
– loss of normalization  
– immature software market | making data available for linking                      |
Don’t be the boy with the hammer!

http://www.flickr.com/photos/48143042@N05/4418727909/in/photolist-7Jt8RX
Low hanging fruits

• Clean your metadata
• Reconcile with authoritative sources
• Enrich your metadata
| guestfc5fbf9 [13,16]  
| jessicapavoramos [14] |
| dns_prefetch: | http://www.slideshare.net/iesjuliocararoba/pablo-picasso-1315023/cdn01.smartliving.com [12]  
| http://www.slideshare.net/iesjuliocararoba/pablo-picasso-1315023/cdn.slidesharecdn.com [12]  
| http://www.slideshare.net/iesjuliocararoba/pablo-picasso-1315023/image.slidesharecdn.com [12]  
| external_id: | Mx4wBVYepwpEbGdrcN5Y29ycA [19] |
| fb_app_id: | 2490221586 [12] |
| gravity: | -2.36094 [2]  
| -2.69741 [5,8]  
| -2.92056 [9]  
| -3.05409 [10]  
| -1.90896 [11] |
| height: | 355 [13,14,16] |
| identifier: | it.bestshopping.com/prezzi/Pablo-picasso.sku=9783791323503%7C.html#product [1,2,4,5,8,9,10,11,12,13,14,16,19]  
| it.bestshopping.com/q/pablo+picasso.html [3,6]  
| waptrik.com/download/Pablo-Picasso-Three-Musicians/?type=W&lang=EN&id=31906 [17]  
| waptrik.com/download/Pablo-Picasso-Seed-Man/?type=W&lang=EN&id=31904 [18]  
| www.fnacspectacles.com/place-spectacle/fr/cheartiste/Pablo-Picasso-b08e10fbc0a8850b017a40f87a76e921.htm [20] |
| label: | Pablo picasso [1,2,4,5,8,9,10,11,12,13,14,16,19]  
| Pablo picasso - Confronta prezzi Pablo picasso su Bestshopping [3,6]  
| Pablo Picasso Three Musicians Photos & Pictures free download [17]  
| Pablo Picasso Seated Man Photos & Pictures free download [18]  
| PabloPicasso-TheArtist [19]  
| Pablo Picasso - Prochains événements [20] |
| nofollow: | show 31 values ➜ |
| og_title: | Pablo Picasso [12] |
| og_type: | slideshare:presentation [12] |
| og_description: | PABLO PICASSO [12]  
| og_url: | http://www.slideshare.net/iesjuliocararoba/pablo-picasso-1315023 [12] |
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 !!!
Your first OR project

• Creating, opening and importing projects

• We’ll create one, based on the SDBM
  http://book.freeyourmetadata.org/chapters/2/schoenberg.csv
Your first OR project

- Creating, opening and importing projects
- We’ll create one, based on the SDBM http://book.freeyourmetadata.org/chapters/2/
- Parsing options:
## Interface

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>ID</th>
<th>Duplicates</th>
<th>Cat./Trans. date</th>
<th>Primary seller</th>
<th>Secondary seller</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>219</td>
<td>219,340,117,536,189,000</td>
<td>19970617</td>
<td>Sotheby's</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>220</td>
<td></td>
<td>19970617</td>
<td>Sotheby's</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>221</td>
<td></td>
<td>19970617</td>
<td>Sotheby's</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>222</td>
<td>22,263,919,885,994,100,000,000,000,000,000,000</td>
<td>19970617</td>
<td>Sotheby's</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>223</td>
<td>11,223</td>
<td>19970617</td>
<td>Sotheby's</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>224</td>
<td>224,877,241,162</td>
<td>19970617</td>
<td>Sotheby's</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>226</td>
<td>226,193,868</td>
<td>19970616</td>
<td>Sotheby's</td>
<td>Beck Collection</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>227</td>
<td>227,384,510,137,147,000,000,000,000,000</td>
<td>19970616</td>
<td>Sotheby's</td>
<td>Beck Collection</td>
</tr>
</tbody>
</table>
Filters

• Ideal functionality to develop a first understanding of the data

• Primary Seller > Text filter > Type « a » or « ab »

• Title > Text filter > Type « bel »

• Remove All or Reset All

• But let’s use Refine to do the guess work...
<table>
<thead>
<tr>
<th>Catalogue ID</th>
<th>Lot or cat.</th>
<th>Price</th>
<th>Currency</th>
<th>Sold</th>
<th>Source</th>
<th>Current location</th>
<th>Author</th>
<th>Author (variant)</th>
<th>Title</th>
<th>Language</th>
<th>Material</th>
<th>Place</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handschriften, B. Acker, Autographen ...</td>
<td>9</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Leonardo Bruni Aretino</td>
<td>Commentary on Bello's Primo Bello Punico</td>
<td>Italian</td>
<td>V</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>No. 51: Bucher, Handschriften, Autographen ...</td>
<td>2</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Josephus Flavius</td>
<td>De Bello Judaico De Judeorum Vetratae Sive Contra Apionem</td>
<td>Latin</td>
<td>P</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>No. 55: Bucher, Handschriften, Autographen ...</td>
<td>5</td>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Josephus Flavius</td>
<td>De Bello Judaico De Judeorum Vetratae Sive Contra Apionem</td>
<td>Latin</td>
<td>P</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>No. 103: Bucher, Autographen</td>
<td>3</td>
<td>3750</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bruni, Leonardo</td>
<td>Commentary on Sallust's Primo Bello Punico</td>
<td>Italian</td>
<td>V</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>No. 103: Bucher, Autographen</td>
<td>21</td>
<td>1250</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sallust</td>
<td>Commentary on Sallust's Primo Bello Punico</td>
<td>Italian</td>
<td>V</td>
<td>Italy, Flore</td>
<td></td>
</tr>
<tr>
<td>No. 124: Bucher, Autographen</td>
<td>18</td>
<td>1750</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sallust</td>
<td>De Conjurante Cistinae, de Bello Jugurthino</td>
<td>Latin</td>
<td>V</td>
<td>Italy, Flore</td>
<td></td>
</tr>
<tr>
<td>Boston Pub</td>
<td>122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Jacopo Palladio Teramo</td>
<td>Consolatio Pecatorum Seu Processus Bellai</td>
<td>German</td>
<td>P</td>
<td>Germany</td>
<td></td>
</tr>
<tr>
<td>Livres précieux et manuscrits avec miniatures composant la bibliothèque de M. E.M.B.</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Camillo Bonadeo</td>
<td>Scelta Dell Piu Belle ...</td>
<td>Italian</td>
<td>P</td>
<td>Italy, Piazo</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Faceting

• One of the core functionalities of Refine, allowing you to discover quickly the true nature of your data

• What’s the difference between Primary Seller and Secondary Seller? Apply a text facet on both

• Message: there are too many options => Click Set choice count limit and raise to 3,000 (but could cause issues with memory allocation)
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 2048M

  • 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
Faceting

- New windows in the left side bar
- By default, ordered alphabetically but click on count
- Apply the same facet on Seller 2, so that we can compare the most popular values of both fields
- Experiment on other fields!
- Also check the outliers!
Faceting

- Identification of fake outliers: apply text facet on *Material* and *Buyer*
- Check the long tail of unique values for misspellings
Field overloading

• Fields such as Author, Author variant, Place and Provenance contain multiple values in the same cell

• Example: Author > Edit Cells > Split multi-valued cells and indicate a vertical bar (|) is used as a separator

• Important: OpenRefine has added extra rows for every unique value from the previously overloaded field (but we have the same amount of records)
<table>
<thead>
<tr>
<th>Catalogue ID</th>
<th>Lot or cat.</th>
<th>Price</th>
<th>Currency</th>
<th>Sold</th>
<th>Source</th>
<th>Current location</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western manuscripts and miniatures (LN7375)</td>
<td>62</td>
<td>71445</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Guillaume de Lorris</td>
</tr>
<tr>
<td>Western manuscripts and miniatures (LN7375)</td>
<td>61</td>
<td>27513.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Jean de Meun</td>
</tr>
<tr>
<td>Western manuscripts and miniatures (LN7375)</td>
<td>60</td>
<td>30360</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Field overloading

• Multiple values referring to other realities (countries, regions, cities)

• Example: Place > Edit columns > Split up into multiple columns

• First value is consistently a country => split up in max 2 columns and keep the original column
Clustering

- Aggregate automatically different values regarding the same reality
- One of the best features of OR
- Example: on the field Artist, apply Edit cells > Cluster and edit
- New window pops up with clustering features and options
Cluster & Edit column "Artist"

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. Find out more ...

<table>
<thead>
<tr>
<th>Cluster Size</th>
<th>Row Count</th>
<th>Values in Cluster</th>
<th>Merge?</th>
<th>New Cell Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>19</td>
<td>Robert Boyvin (12 rows)</td>
<td></td>
<td>Robert Boyvin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Robert Boyvin? (6 rows)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Boyvin, Robert (1 rows)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>Stefano di Tommaso Lunetti? (4 rows)</td>
<td></td>
<td>Stefano di Tommaso Lunetti</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lunetti, Tommaso di Stefano (1 rows)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stefano di Tommaso Lunetti (1 rows)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>Master of Edward IV (21 rows)</td>
<td></td>
<td>Master of Edward IV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Master of Edward IV? (1 rows)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Master of Edward IV (1 rows)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>Girolamo da Cremona (12 rows)</td>
<td></td>
<td>Girolamo da Cremona</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Girolamo, da Cremona (1 rows)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Girolamo da Cremona] (1 rows)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>Memling, Hans, school (4 rows)</td>
<td></td>
<td>Memling, Hans, school</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Memling, Hans, school? (1 rows)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>Mercati, Venturino (2 rows)</td>
<td></td>
<td>Mercati, Venturino</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Venturino Mercati (2 rows)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

91 clusters found

# Choices in Cluster

# Rows in Cluster

Average Length of Choices

Length Variance of Choices
Customized text facets

- Test out Facet > Customized Facets > Text length facet on Comments
- Play around with the handles to influence the distribution spread
- Apply Facet > Customized Facets > Log of text length facet
Numeric facets

- Numbers are colored in green, assuming the Parse cell text into numbers, dates... checkbox was selected at the time of the project creation.

- Example: *Price > Facet > Numeric Facet*
Numeric facets

- Apply Facet > Customized facets > Facet by blank to automatically select only records with actual values
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
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

(Test/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
Case study

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

- Download the data from http://book.freeyourmetadata.org/chapters/3/

- Focus on the Categories field, populated with the Powerhouse museum Object Names Thesaurus (PONT), a locally created vocabulary
<table>
<thead>
<tr>
<th>No.</th>
<th>Description</th>
</tr>
</thead>
</table>

**Untick this box**
<table>
<thead>
<tr>
<th>Record ID</th>
<th>Object Title</th>
<th>Registration Num</th>
<th>Description.</th>
</tr>
</thead>
<tbody>
<tr>
<td>7801</td>
<td>19142 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. Only one point being visible / JEG / D 29.11.88 (SB). 1 to Broken Hill Museum [initials] 14.8.1911 (SB). Cast of meteorite; weight 12 lb; 3.84 ozs found on highest peak, of Battery Mountain junction of Burrowa &amp; Lachlan Rivers (LC).</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 - firmly embedded in the slate, only one point being visible / JEG / R 29.11.88 (SB). 1 to Broken Hill Museum [initials] 14.8.1911 (SB). Cast of meteorite; weight 12 lb; 3.84 ozs found on highest peak, of Battery Mountain junction of Burrowa &amp; Lachlan Rivers (LC).</td>
</tr>
<tr>
<td>31</td>
<td>Hakea macrocarpa (SB). Hakea Macraea (Pol. (LC).</td>
<td>31</td>
<td>One specimen Red Ironbark;</td>
</tr>
<tr>
<td>10</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
</tbody>
</table>
### Categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photographs</td>
<td>10027</td>
</tr>
<tr>
<td>Numismatics</td>
<td>8504</td>
</tr>
<tr>
<td>Clothing and Dress</td>
<td>7635</td>
</tr>
<tr>
<td>Ceramics</td>
<td>7478</td>
</tr>
<tr>
<td>Glass plate negatives</td>
<td>7122</td>
</tr>
<tr>
<td>Pictorials</td>
<td>5142</td>
</tr>
<tr>
<td>Documents</td>
<td>4680</td>
</tr>
<tr>
<td>Photographic prints</td>
<td>3870</td>
</tr>
<tr>
<td>Textiles</td>
<td>2692</td>
</tr>
<tr>
<td>Botanical specimens</td>
<td>2626</td>
</tr>
<tr>
<td>Specimens</td>
<td>2608</td>
</tr>
<tr>
<td>Domestic Equipment-Home</td>
<td>2359</td>
</tr>
</tbody>
</table>
Cluster & Edit column "Categories"

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. Find out more ...

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. Find out more ...

### Cluster & Edit column "Categories"

<table>
<thead>
<tr>
<th>Cluster Size</th>
<th>Row Count</th>
<th>Values in Cluster</th>
<th>Merge?</th>
<th>New Cell Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1026</td>
<td>Audio and Visual Equipment (1021 rows) Audio and visual equipment (5 rows)</td>
<td></td>
<td>Audio and Visual Equipment</td>
</tr>
<tr>
<td>2</td>
<td>683</td>
<td>Photographic Equipment (635 rows) Photographic equipment (48 rows)</td>
<td></td>
<td>Photographic Equipment</td>
</tr>
<tr>
<td>2</td>
<td>615</td>
<td>Food and Drink (614 rows) Food and drink (1 rows)</td>
<td></td>
<td>Food and Drink</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>Office Equipment (83 rows) Office equipment (7 rows)</td>
<td></td>
<td>Office Equipment</td>
</tr>
<tr>
<td>2</td>
<td>4681</td>
<td>Documents (4680 rows) documents (1 rows)</td>
<td></td>
<td>Documents</td>
</tr>
<tr>
<td>2</td>
<td>930</td>
<td>Musical Instruments (925 rows) Musical instruments (5 rows)</td>
<td></td>
<td>Musical Instruments</td>
</tr>
<tr>
<td>2</td>
<td>96</td>
<td>Material technology (89 rows) Material Technology (7 rows)</td>
<td></td>
<td>Material technology</td>
</tr>
<tr>
<td>2</td>
<td>144</td>
<td>Medicines (142 rows) Medicines (7 rows)</td>
<td></td>
<td>Medicines</td>
</tr>
</tbody>
</table>

### Cluster Size | Row Count | Values in Cluster                                                                 | Merge? | New Cell Value       |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>47</td>
<td>Mailbags (43 rows) Mail bags (4 rows)</td>
<td></td>
<td>Mailbags</td>
</tr>
<tr>
<td>2</td>
<td>226</td>
<td>Shirts (128 rows) T-shirts (98 rows)</td>
<td></td>
<td>Shirts</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>Skullcaps (3 rows) Skull caps (1 rows)</td>
<td></td>
<td>Skullcaps</td>
</tr>
<tr>
<td>2</td>
<td>352</td>
<td>Transport-Water (351 rows) Transport - Water (1 rows)</td>
<td></td>
<td>Transport-Water</td>
</tr>
<tr>
<td>2</td>
<td>682</td>
<td>Transport-Air (680 rows) Transport - Air (2 rows)</td>
<td></td>
<td>Transport-Air</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Pillboxes (1 rows) pill boxes (1 rows)</td>
<td></td>
<td>Pillboxes</td>
</tr>
</tbody>
</table>

Select All | Deselect All

### # Rows in Cluster

Graph showing the number of rows in each cluster.

### Average Length of Choices

Graph showing the average length of choices in each cluster.

### Length Variance of Choices

Graph showing the variance of choices in each cluster.
## 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 »

<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>1313194 (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>
Add SPARQL-based reconciliation service

Name: LCSH - PH museum test

Endpoint details
- Endpoint URL: http://sparql.freyourmetadata.org
- Graph URI: http://id.loc.gov/authorities/subjects
  - Leave empty to use the default graph
- Type: Virtuoso
  - This determines the syntax that will be used for search

Label properties
- Select properties that are used to label resources in the endpoint. These properties will be used to match resources:
  - rdfs:label
  - skos:prefLabel
  - dcterms:title
  - dc:title
  - foaf:name
  - Other...
<table>
<thead>
<tr>
<th>Record ID</th>
<th>ID</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>6991</td>
<td>Specimens</td>
</tr>
<tr>
<td>99</td>
<td>6799</td>
<td>Botanical specimens</td>
</tr>
<tr>
<td>99.</td>
<td></td>
<td>Hand fans</td>
</tr>
<tr>
<td>222</td>
<td>369939</td>
<td>Specimens</td>
</tr>
<tr>
<td>222.</td>
<td>369981</td>
<td>Wool specimens</td>
</tr>
<tr>
<td>223</td>
<td>369945</td>
<td>Animal Samples and Products</td>
</tr>
<tr>
<td>226</td>
<td>371995</td>
<td>Textile Technology</td>
</tr>
<tr>
<td>229</td>
<td>371994</td>
<td>Wool specimens</td>
</tr>
<tr>
<td>320</td>
<td>5599</td>
<td>Castings</td>
</tr>
<tr>
<td>314</td>
<td>5499</td>
<td>Sculptures</td>
</tr>
<tr>
<td>373</td>
<td>5299</td>
<td>Sculptures</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Writing equipment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sculpture</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Castings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Draughts sets</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sculpture</td>
</tr>
</tbody>
</table>
Meteorites

From Library of Congress Subject Headings

Details  Visualization  Suggest Terminology

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**

**URI(s):**
- http://id.loc.gov/authorities/subjects/sh87006764
- http://id.loc.gov/authorities/subjects/sh87006764
- http://id.loc.gov/authorities/subjects/sh87006764#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

---

### Specimens

**From: Library of Congress Subject Headings**

**Details**

- 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/sh99001749
- http://id.loc.gov/authorities/subjects/sh99001749#concept

**Instance Of:**
- MADS/RDF GenreForm

**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 - H1096
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 than good
Pre-processed LCSH

• Changes in our version of LCSH:
  • Subdivisions are only present if they do not conflict with an existing main heading with the same 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>Categories</th>
<th>Count</th>
<th>Create new topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meteorites</td>
<td>5664</td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>95183</td>
<td></td>
</tr>
<tr>
<td>Mineral Samples - Geological</td>
<td>103801</td>
<td></td>
</tr>
<tr>
<td>Botanical specimens</td>
<td>3573</td>
<td></td>
</tr>
<tr>
<td>Timber Samples</td>
<td>2279</td>
<td></td>
</tr>
<tr>
<td>Wool specimens</td>
<td>2008</td>
<td></td>
</tr>
<tr>
<td>Specimens</td>
<td>1038</td>
<td></td>
</tr>
<tr>
<td>Animal Samples and Products</td>
<td>4313</td>
<td></td>
</tr>
<tr>
<td>Models (Patents)</td>
<td>557</td>
<td></td>
</tr>
<tr>
<td>Models (Persons) - Vocational guidance</td>
<td>1570</td>
<td></td>
</tr>
<tr>
<td>Models and modelmaking - Radio control systems</td>
<td>1616</td>
<td></td>
</tr>
</tbody>
</table>
Add column based on column Categories (pre-processed LCSH)

New column name: 

On error: set to blank

Expression: cell.recon.match.id

Language: Google Refine Expression Language (GREL)

No syntax error.

Preview:

<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/sh2001000705#concept">http://id.loc.gov/authorities/sh2001000705#concept</a></td>
</tr>
<tr>
<td>6.</td>
<td>Botanical specimens</td>
<td><a href="http://id.loc.gov/authorities/sh2001000705#concept">http://id.loc.gov/authorities/sh2001000705#concept</a></td>
</tr>
<tr>
<td>7.</td>
<td>Botanical specimens</td>
<td><a href="http://id.loc.gov/authorities/sh2001000705#concept">http://id.loc.gov/authorities/sh2001000705#concept</a></td>
</tr>
</tbody>
</table>

Categories (pre-processed L)  Category URL

- Meteorites  http://id.loc.gov/authorities/sh85084315#concept
- Models  http://id.loc.gov/authorities/sh00006390#concept
- Mineral Samples-Geological  Create new topic
- Botanical specimens  http://id.loc.gov/authorities/sh2001000705#concept
- Botanical specimens  http://id.loc.gov/authorities/sh2001000705#concept
- Botanical specimens  http://id.loc.gov/authorities/sh2001000705#concept
- Timber Samples  Create new topic
- Wool specimens  Create new topic
## JASIST results

<table>
<thead>
<tr>
<th></th>
<th>Only prefLabels</th>
<th>Addition of altLabels</th>
<th>Exclusive heading</th>
<th>Omission of qualifiers</th>
<th>Stemming</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCSH</td>
<td>25.5%</td>
<td>56.2%</td>
<td>78.2%</td>
<td>75.9%</td>
<td>81.1%</td>
</tr>
<tr>
<td>AAT</td>
<td>60.1%</td>
<td>64.1%</td>
<td>n/a</td>
<td>77.1%</td>
<td>n/a</td>
</tr>
</tbody>
</table>

**Diagram:**
- **LCSH** + **AAT**: 68.4% of records reconciled
- **LCSH**: 81.1%
- **AAT**: 77.1%
- **PHM Collection**: 89.8% of records reconciled
Frank Lloyd Wright
American, 1867–1959

We have **53 objects** that Frank Lloyd Wright has been involved with. **See all of them.**

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/](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](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, materials 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 cathedrals 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 grids 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 ambition, 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 house 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 homes 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 shared similar interest in the influence of Japanese culture and climate to create contemporary architecture while working with a sense of national identity.

The 1906 Robie House in Chicago was Wright’s most mature expression of the Prairie Style. 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 by moving away from the tight box shape of traditional homes. The low, horizontal form is exaggerated with the use of ribbons of concrete forming the base pilin 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 each level Wright designed a terrace, balcony or porch to break the division between inside and outside. All internal details — including the furnishings, 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 built 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 as screens with glass windows above.

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 home for himself, Mannah Cheney and her two children on land given to him by his mother at Spring Green in Wisconsin. He called the house Taliesin, a Welsh word meaning “shining brow” and the name of a Welsh bard. However, tragedy struck at Taliesin when in 1914 a chef, Julian Carleton, murdered Mamah Borthwick Cheney, her two children and four others and then set fire to the house.

This period of turbulence in Wright’s private life overlapped with the commission for the Imperial Hotel in Tokyo, which consumed him from 1916 to 1922 when he spent a great deal of time in Japan overseeing the work. The hotel’s owner had chosen a western architect to bridge the cultural divide for the western visitors to Tokyo and Wright rose to this challenge. The main feature of the 100-room hotel was the grand three-storey lobby and two-storey dining room, ballroom and auditorium. The use of soft lava block or 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 1968.

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 1917 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 motif 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.
Low hanging fruits

- Clean your metadata
- Reconcile with authoritative sources
- Enrich your metadata
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/page/Washington,_D.C.

• 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.
Named-Entity Recognition extension for Google Refine / OpenRefine

- 90 commits
- 1 branch
- 5 releases
- 3 contributors

Add WikiMeta service.

RubenVerborgh authored 4 days ago

- module
  - implement the "reconstruct" method to allow for history-apply / reapply
  - 3 months ago
- src
  - Add WikiMeta service.
  - 4 days ago
- .classpath
  - Add base class for NER commands with JSON output.
  - a year ago
- .gitignore
  - Initialize service manager with default list if no previous configuration.
  - a year ago
- .project
  - Initialize project with menu placeholders.
  - a year ago
- LICENSE.md
  - Add license.
  - a year ago
- README.md
  - Update paper title.
  - 6 months ago
- build.xml
  - Added missing line to copy DefaultServices.json to classes dir.
  - a year ago

https://github.com/RubenVerborgh/Refine-NER-Extension
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
  - http://www.zemanta.com/
- Click the Named-entity recognition toolbar button and choose Configure API keys
- Add the keys you received and click Update
Case-study

- Experiment with reconciliation operations in a hands-on manner with the metadata of the British Library (CSV conversion from an RDF file available through Europeana)

- Download the data from

- We’re only interested in the description field
  => Choose View > Collapse other columns
<table>
<thead>
<tr>
<th>Rank</th>
<th>Description</th>
<th>Location</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sir David (1758-1825), dancers and musicians, hookahs, portraits, sports</td>
<td>South Asia</td>
<td>South Asian history</td>
</tr>
<tr>
<td></td>
<td>and pastimes, water-colour, Delhi style.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>– History [954]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Gouache painting with gold of village life in Kashmir</td>
<td>Kashmir</td>
<td>Gouache</td>
</tr>
<tr>
<td>3</td>
<td>Watercolour depicting a Durga Puja</td>
<td>Durga Puja</td>
<td>Durga Puja</td>
</tr>
<tr>
<td>4</td>
<td>India, Tamil Nadu, Vellore, clerks, occupations, opaque water-colour,</td>
<td>Tamil Nadu</td>
<td>Outline of South</td>
</tr>
<tr>
<td></td>
<td>Tanjore, Vellore style, South Asia – History [954]</td>
<td></td>
<td>Asian history</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A course of religious instruction for Catholic youth / by John Gerard.

A full course of instructions for the use of catechists : being an explanation of the catechism, entitled "An abridgment of Christian doctrine" / by John Perry.


A new treatise on the duty of a Christian towards God [microform] : 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.

A short life of Our Lord / by P. J. Crean ; with a foreword by e the Cardinal Archbishop of Westminster.

A victim to the seal of confession: a true story / by the Reverend Joseph Spillmann.

Advice to Irish girls in America / by Alton Park; or, Conversations on religious and moral subjects; chiefly designed for the amusement and instruction of young ladies.

An adventure with the Apaches / by Gabriel Ferry [pseud.]

An explanation of the Baltimore catechism of Christian doctrine : for the use of Sunday-school teachers and advanced classes / by Thomas L. Kinkead.

Angelica / from the German of Christopher von Scmid.

Batavia, by Hendrik Conscience. Translated from the original Flemish.

Bobby in Movieland, by Francis J. Finn ...

Breton legends / Translated from the French.

But Thy love and Thy grace / by Francis J. Finn ; with illustrations by Charles C. Svendsen.

http://dh.crc.nd.edu/sandbox/cyl/catalog/
Jean-Jacques Rousseau, Du contrat social

De l'esclavage. Puisque aucun homme n'a une autorité naturelle sur son semblable, et puisque la force ne produit aucun droit, restent donc les conventions pour base de toute autorité légitime parmi les hommes. Si un particulier, dit Grothus, peut alléger sa liberté et se rendre esclave d'un maître, pourquoi tout un peuple ne pourrait-il alléger la sienne et se rendre sujet d'un roi ? Il y a là bien des motifs éloquents qui auraient besoin d'explication ; mais tenons-nous-en à celui d'alléger. Alléger, c'est donner ou vendre. Or, un homme qui se fait esclave de ne se donner pas ; il se vend tout au moins pour sa subsistance ; ou un peuple, pourquoi se vend-il ? Bien loin qu'un roi fournisse à ses sujets leur subsistance, il ne tire la sienne que d'eux ; et, selon Rabelais, un roi ne vit pas de peu. Les sujets donnent donc leur personne, à condition qu'on prenne aussi leur bien. Je ne vois pas ce que les sociétés de la société civile ; soit : mais qu'y gagnent-ils, si les guerres que son ambition leur attire, si son inestimable avidité, si les vexations de son ministère les désolent plus que ne feraient leurs dissensions ? Qu'y gagnent-ils, si cette tranquillité même est une de leurs misères ? On vit tranquille aussi dans les cachots ; en est-ce assez pour s'y trouver bien ? Les Grecs enfermés dans l'antre du Cyclope y vivaient paisiblement; en attendant que leur tour vint d'être les victimes du féroce démon. Dire qu'un homme se donne librement, c'est dire une chose absurde et inconcevable ; un tel acte est illégitime et nul, par cela seul que celui qui le fait n'est pas dans son bon sens. Dire la même chose de tout un peuple, c'est supposer un peuple de fou ; on ne fait pas droit. Quand chacun pourrait s'alléger lui-même, il ne peut alléger ses enfants ; il naisse hommes libres et libres ; leur liberté appartient, nul n'a droit d'en disposer qu'eux. Avant qu'ils soient en âge de se libertaire, le père, en leur nom, stipuler des conditions pour leur conservation, pour leur bien-être, mais non les donner irrévocablement et sans condition ; car un tel don est contraire aux fins de la nature, et passe les droits de la paternité. Il faudrait donc, pour qu'un gouvernement arbitraire fût légitime, qu'à chaque génération le peuple fût le maître de l'admettre ou de le rejeter : mais alors ce gouvernement ne serait plus arbitraire. Renoncer à sa liberté, c'est renoncer à sa qualité d'homme, aux droits de l'humanité, même à ses devoirs. Il n'y a nul dédommagement possible pour quiconque renonce à tout. Une telle renonciation est incompatible avec la nature de l'homme ; et c'est oublier toute moralité à ses actions que d'ôter toute liberté à sa volonté. Enfin c'est une convention vainque et contradictoire de stipuler d'une part une autorité absolue, et de l'autre une obéissance sans bornes. N'est-il pas clair qu'on n'est engagé à rien envers celui dont on a droit de tout exigir ? Et cette seule condition, sans équivoque, sans échange, n'entraîne-t-elle pas la nullité de l'acte ? Car, quel droit mon esclavage aurait-il contre moi, puisque tout ce qu'il a n'appartient et que, son droit étant le mien, ce droit de moi contre moi-même est un mot qui n'a aucun sens ? Grothus et les autres tentent de la guerre une autre origine du prétendu droit d'esclavage. Le vainqueur ayant, selon eux, le droit de tuer le vaincu, celui-ci peut racheter sa vie aux dépens de sa liberté ; convention d'autant plus légitime qu'elle tourne au profit de tous deux. Mais il est clair que ce prétendu droit de tuer le vaincu ne résulte en aucune manière de l'état de guerre. Par cela seul, que les hommes, vivant dans leur primitive indépendance, n'ont point eu de voleur de travail assez constant pour constituer ni l'état de paix ni l'état de guerre, ils ne sont point naturellement ennemis. C'est le rapport des choses et non des hommes qui constitue la guerre ; et l'état de guerre ne pousse nullement les simples relations personnelles, mais seulement des relations réelles, la guerre privée d'homme à homme ne peut épanouir ni dans l'état de nature, où il n'y a point de propriété commune, ni dans l'état social, où tout est sous l'autorité des lois. Les combats particuliers, les duels, les rencontres, sont des actes qui ne constituent point un état ; et à l'égard des guerres privées, autorisées par les États de Louis IX, roi de France, et suspendues par la paix de Dieu, ce sont des abus du gouvernement féodal, système absurde, en qui fut jamais en rien l'autorité du prince naturel et à toute bonne politique. La guerre n'est donc point une action d'homme à un autre, mais une action d'homme à lui-même, ou un acte justiciable par les peuples. Les déclarations de guerre sont moins des avertissements auxquels ils sont attachés qu'à leurs sujets, à l'ennemi, soit vrai, soit particulier, soit peuple, qui vote, tue, ou détient les sujets, sans déclarer la guerre au prince, n'est pas un ennemi, c'est un brigand. Même en paix en guerre, un prince juste s'arrange bien, en pays ennemi, il respecte la personne et les biens de l'ennemi, de tout ce qui appartient au prince, dans le droit de guerre, sans se soucier des arêtes sur lesquels sont fondés les biens. La fin de la guerre étant la destruction de l'ennemi, on a droit d'en tuer eux, les ennemis tant qu'ils ont les armes à la main ; mais siot qu'ils les posent et se rendent, cessant d'être ennemis ou instruments de l'ennemi, ils redeviennent simplement hommes, et l'on n'a plus droit sur leur vie. Quelquesfois, on peut tuer l'ennemi sans un seul de ses membres : or la guerre ne donne aucun droit qui ne soit nécessaire à sa fin. Ces principes fondés sur des autorités de poche ne sont pas ceux de l'État. Ceux de l'État ne dépendent pas de la nature des choses, et sont fondés sur la raison. A l'égard du droit de conquête, il n'a d'autre fondement que la loi du plus fort. Si la guerre ne donne point au vainqueur le droit de massacrer les peuples vaincus, ce droit qu'il n'a pas ne fonderait ceux de les asservir. On n'a le droit de tuer l'ennemi que quand on ne peut le faire sans les laisser vivre ; le droit de le faire esclave ne peut se substituer à celui de le tuer ; c'est donc un échange inégal, au prix de sa liberté vivante, sur laquelle on n'a aucun droit. En établissant le droit de vie et de mort sur le droit d'esclavage, il est le droit d'esclavage sur le droit de vie et de mort, n'est-il pas clair qu'on tombe dans le cercle vicieux ? En supposant même ce terrible droit de tout tuer, je dis qu'un esclave fait à la guerre, ou un peuple conquis, n'est tenu à rien du tout envers son maître, qu'à lui obéir autant qu'il y est force. En prenant un équivalent à sa vie, le vainqueur ne lui en a point fait grâce ; au lieu de le tuer sans fruit, il l'a tué utilement. Loin donc qu'il ait acquis sur lui nulle autorité jointe à la force, l'état de guerre subsiste entre eux comme auparavant, leur relation même en est l'effet ; et l'usage du droit de la guerre ne suppose aucun traité de paix. Ils ont fait une convention ; soit : mais cette convention, loin de détruire l'état de guerre, en suppose la continuité. Ainsi, de quelque sens qu'on envisage les choses, le droit d'esclavage est nul, non seulement parce qu'il est illégitime, mais parce qu'il est absurde et ne signifie rien. Ces mots, esclave et droit, sont contradictoires ; ils s'excluent mutuellement. Soit d'un homme à un homme, soit d'un homme à un peuple, ce discours sera toujours également insensé : "Je fais avec toi une convention toute à ta charge et toute à mon profit, que je t'observerai tant qu'il me plaera, et que tu observeras tant qu'il me plaira."
SINCE no man has a natural authority over his fellow, and force creates no right, we must conclude that conventions form the basis of all legitimate authority among men. If an individual, says Grotius, can alienate his liberty and make himself the slave of a master, why could not a whole people do the same and make itself subject to a king? There are in this passage plenty of ambiguous words which would need explaining; but let us confine ourselves to the word alienate. To alienate is to give or to sell. Now, a man who becomes the slave of another does not give himself; he sells himself, at the least for his subsistence: but for what does a people sell itself? A king is so far from furnishing his subjects with their subsistence that he gets his own only from them; and, according to Rabelais, kings do not live on nothing. Do subjects then give their persons on condition that the king takes their goods also? I fail to see what they have left to preserve. It will be said that the despot assures his subjects civil tranquility. Granted; but what do they gain, if the wars his ambition brings down upon them, his insatiable avidity, and the vexatious conduct of his ministers press harder on them than their own dissensions would have done? What do they gain, if the very tranquility they enjoy is one of their miseries? Tranquility is found also in dungeons; but is that enough to make them desirable places to live in? The Greeks imprisoned in the cave of the Cyclops lived there very tranquilly, while they were awaiting their turn to be devoured. To say that a man gives himself gratuitously, is to say what is absurd and inconceivable; such an act is null and illegitimate, from the mere fact that he who does it is out of his mind. To say the same of a whole people is to suppose a people of madmen; and madness creates no right. Even if each man could alienate himself, he could not alienate his children: they are born men and free; their liberty belongs to them, and no one but they has the right to dispose of it. Before they come to years of discretion, the father can, in their name, lay down conditions for their preservation and well-being, but he cannot give them irrevocably and without conditions: such a gift is contrary to the ends of nature, and exceeds the rights of paternity. It would therefore be necessary, in order to legitimise an arbitrary government, that in every generation the people should be in a position to
Information vs non-information resources

- Information resources are representable in electronic format, non-information resources are not.

- Example: a Wikipedia article about Barack Obama is an information resource, as it can be represented in electronic form. Its URL thus identifies the article, locates it, and allows to send an electronic representation of it. However, Barack Obama is a non-information resource, as we cannot represent him electronically.

- Solution:
  
  - [http://dbpedia.org/resource/Barack_Obama](http://dbpedia.org/resource/Barack_Obama) identifies the person.
  
  - [http://dbpedia.org/page/Barack_Obama](http://dbpedia.org/page/Barack_Obama) identifies the document about the artist.
Get in touch!

- Handbook available from Facet and ALA
- Case-studies, research projects?
- Follow @freemetadata, @RubenVerborgh and @sethvanhooland