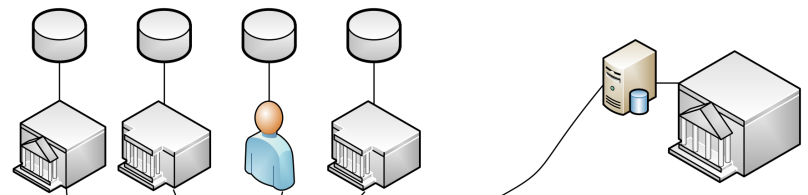




www.archive.org

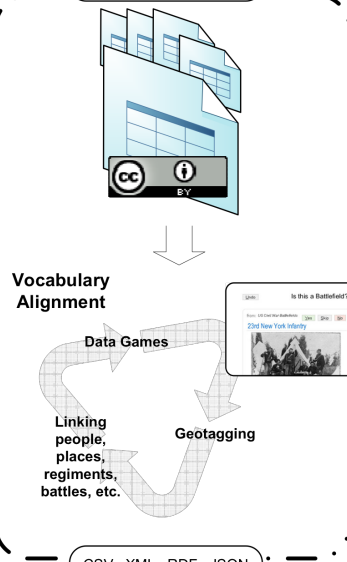
Archives, Libraries, Individual Researchers provide structured data to CWD150

Institutions and companies that provide structured data via various means, such as API's, RDF, etc.



CSV XML RDF APIs Z39.50

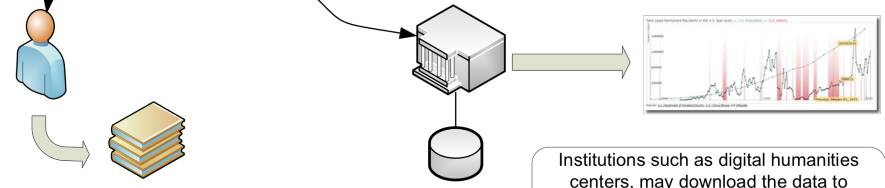
Civil War Data 150



The data will be published to Freebase and other Semantic Web publishing platforms and repositories. This enables the general public to create further mashups and visualizations.

Individuals can download published data and may use it in reports, publications, visualizations, citing source material.

Institutions such as digital humanities centers, may download the data to incorporate into existing research or provide further analysis.



What Inhibits Adoption/Traction?

- Lack of Awareness
- Organizational Inertia
- Lack of shared ontologies/taxonomies – what is the standard? Who is the authority?
- Dirty Data/Diversity of data formats, e.g. disambiguation of place names over time, people, etc.
- Legacy systems and/or endless technology debates
- Lack of Clarity of Rights to Use/Reuse data sets and/or limitations on use that inhibit creative integration & expression
- Concerns about persistence, ongoing support

What can we do about this?

- Get Involved: Expose or Consume a collection or two
 - even if only under the auspices of an “experiment”
 - Start small
- Embrace Best Efforts/Be Flexible
 - an imperfect approach is better than none at all
- Embrace Shared Services/Ontologies
 - why reinvent the wheel?
- Promote Maximum Reuse
- Partner



100,000+ Hours of TV

400,000+ Moving Images

700,000+ Audio Recordings

2,500,000+ Books

150,000,000,000+ Web Pages