
Civil War Data 150

About

Civil War Data 150 ("CWD150"), is a collaborative project to share and connect Civil War related data across local, state and federal institutions during the sesquicentennial of the American Civil War, beginning in April of 2011. The project will utilize Linked Open Data to find and create connections between archives and help increase the discovery of these resources by researchers and the general public alike.

[1] Use case: http://www.w3.org/2005/Incubator/lld/wiki/Use_Case_Civil_War_Data_150

Background and Current Practice

• Aggregating diverse data sources:
  – Library
  – Museum
  – Archives
  – Individual researchers
  – Existing books, papers, photographs,
  – a renewed effort on the United States federal and state levels to digitize and release even more photographs, journals and ephemera related to the Civil War.

During the initial phases of this project, CWD150 will work with institutions to publish Civil War related collections metadata as structured data with CC-BY or similar licensing, moving in later phases toward the publishing of metadata in linked data formats.
Goal

• Researchers and the general public alike will be able to discover information about the American Civil War from across multiple institutions, and incorporate that information into their work in new and exciting ways.
• CWD150 will use linked data technology to create connections based on the strong identifiers and taxonomy of the Civil War, particularly the regiments, battles, battlefields, officers, and soldiers and sailors.

Related vocabs/sources

plan to work with the United States National Park Service
• to create an ontology from their Civil War Soldiers and Sailors Database, (dates, places, people, events, etc.)
• with a crosswalk of corresponding identifiers in Freebase, DBPedia, and potentially Library of Congress Subject Headings
  – (though subject headings of granular detail are not at this point available through the id.loc.gov project).
Civil War Data 150

- Will use linked data technology to create connections based on the strong identifiers and taxonomy of the Civil War, particularly the regiments, battles, battlefields, officers, and soldiers and sailors.
- Performing vocabulary alignment to an ontology
- Enabling query information about a particular place, regiment, battle, or officer

Problems and Limitations

- involving no small technology
  - ontology development
  - persistent URIs for all of the key entities
  - possible to utilize dbpedia/Freebase identifiers as a proxy until such time as being able to point to the National Park Service as a source.
- copyright and licensing
  - selecting sources
  - differentiating between the licensing and use of metadata about digital assets, and the assets themselves.