

Report on and demonstration of the PeriodO period gazetteer

Ryan Shaw · Adam Rabinowitz · Patrick Golden · Eric Kansa

The PeriodO period gazetteer¹ documents definitions of historical period names. Each entry of the gazetteer identifies the definition of a single period. To be included in the gazetteer, a definition must a) give the period a name, b) impose some temporal bounds on the period, c) have some implicit or explicit association with a geographical region, and d) have been formally or informally published in some citable source. Much care has been put into giving period definitions stable identifiers that can be resolved to RDF representations of period definitions. Anyone can propose additions of new definitions to PeriodO, and we have implemented an open source web service and browser-based client for distributed versioning and collaborative maintenance of the gazetteer.

The primary goal of PeriodO was to provide stable identifiers that data curators could use to unambiguously identify period *definitions* associated with the period *names* they use to label and organize records. This has a number of advantages over simple text-searching labels containing period names, as structured approximations of temporal and spatial extent can be compared and not simply lexical similarity. A secondary goal was to provide a resource that students and researchers could use to understand patterns of periodization by historical scholars and scientists.

PeriodO models period definitions as SKOS concepts. These are grouped into concept schemes sharing the same bibliographic source. Wherever possible bibliographic sources are identified with Worldcat URIs or Crossref DOIs, and creators are identified with VIAF URIs. Temporal extent is expressed via a direct textual quotation from the source, as well as via a structured approximation of this expression modeled using the OWL-Time ontology. Similarly spatial extent is represented both by a textual quote (where one was given) and a set of identifiers referring to spatial entities in external resources such as DBpedia.

Although the PeriodO dataset is modeled in RDF and can be worked with using any serialization, we chose to focus on JSON-LD as the canonical serialization. This allows the large pool of browser-based web developers to more easily work with PeriodO data without relying on triple-based libraries. It also provided a way for us to manage distributed version control of the PeriodO dataset via JSON Patch documents. Each change is associated with ORCID identifiers identifying the submitter and acceptor of the patch. The entire editorial history of the PeriodO gazetteer is published as a separate dataset using the PROV ontology.

Depending on the interests of workshop attendees, Ryan Shaw could simply demonstrate the functionality of the browser-based client, or he could present a “deep dive” into the PeriodO data model, the ins and outs of working with JSON-LD, or the distributed version control architecture.

References

Shaw R, Rabinowitz A, Golden P, Kansa E (forthcoming). A Sharing-Oriented Design Strategy for Networked Knowledge Organization Systems. *International Journal on Digital Libraries*.

¹ <http://perio.do>