Linked Open Greek Pottery

Ethan Gruber, American Numismatic Society, gruber@numismatics.org
Tyler Jo Smith, McIntire Department of Art, University of Virginia, tjs6e@virginia.edu

Keywords: linked data, pottery, xforms, information architecture

Abstract

Linked open data methodologies have tremendous potential in facilitating data interoperability, aggregation, and analysis. With lessons learned in the development of Nomisma.org, a collaborative project dedicated to the definition of numismatic concepts in the semantic web, the authors have endeavored to apply these methodologies to the field of Greek pottery, through Kerameikos.org. While several thesauri that contain pottery concepts have been published online, not all employ linked data, and none of them are “five star” linked data, as defined by Berners-Lee. This paper discusses the development of a discipline-specific thesaurus which serves as a bridge between existing vocabulary systems, the open source XForms/REST/SPARQL framework for its publication, and the development of web-based tools to analyze and visualize pottery data aggregated from the Getty Museum and British Museum as a proof of concept of the utility of Kerameikos.org as a research tool for Greek pottery.

1. Accessing Greek Vases: A Potted History

The study of figure-decorated pottery (or ‘vases’) from archaic and classical Greece has undergone tremendous change over the past century. Open and closed vessels, varying greatly in their form, scale, function, subject, and quality of embellishment, not only line the shelves of many Classical art collections in the major museums of Europe and the United States, but also continue to be discovered in abundant quantities from excavated contexts. A cursory glance at publications devoted to the locally produced vases excavated from the Athenian Agora, such as those in Volume 23 for black-figure (Moore and Philippides 1986), offers a sense of the range of shapes and styles, as well as their standard method of presentation as archaeological material - i.e. the data is presented in catalogue form according to shape categories rather than according to specific context, findspot, or assemblage. On closer inspection one would notice that, in addition to the expected information about condition and size, the individual pots, cups, and fragments have been associated with or attributed to specific painters or workshops. The depth of detail and the number of artistic personalities to choose from would seem staggering to archaeologists unfamiliar with Classical (and esp. Greek) archaeology and its unique system of pottery classification.

Greek vases were produced in great quantities in the potter’s quarters of ancient Corinth and Athens, as well as in other regions of Greece, among them Laconia, Boeotia, the Cyclades, Crete, and various cities in the East (Boardman 2008; Cook 1997). Although a large number of vases have been discovered in Etruscan tombs, and their Etruscan reception and context has begun to interest many scholars, it is important to recall that vast amounts of figure-decorated pottery have been uncovered in both Greece itself and from Greek colonial sites across the Mediterranean and around the Black Sea. The mechanisms of production for these vases, regardless of place of manufacture or final spot of deposition, are not fully understood. Ancient authors, who have a great deal to tell us about sculptors, architects, and painters, are virtually silent on the topic pottery, which may indicate something significant about its standing as a craft as opposed to ‘high’ art. That being said, many would argue that Greek vases, especially those produced in the city of Athens from the middle years of the 6th
century onward, have much to offer aesthetically, and that figural vases of all dates, styles, and locations are the best surviving visual source for myths about the gods and heroes, religious life and practices, and many other everyday activities, such as dancing, drinking, and dining. Iconography, namely the imagery (or ‘vase-painting’) and its interpretation, has also been a dominant trend in the modern analysis of Greek vases.

While the early history of Greek vase scholarship cannot be recounted in detail here, it is useful to remember that their scholarly presentation as archaeological material derives from the combination of their position in the early history of antiquities collecting and their high esteem in that regard, their large survival numbers, and the classification established for Athenian (or ‘Attic’) examples during the last century by Sir John Beazley of Oxford University (Nøskov 2002; Rouet 2001; Kurtz 1985). At present there are several major resources used by archaeologists to access, identify, and classify Greek vases, and each is applicable regardless of the nature (i.e. whole pots vs sherds) or origin (i.e. Athenian vs non-Athenian) of the evidence they showcase. Volumes, such as one listed above for the Agora in Athens, are among the first places vase specialist consult in order to identify forms and their decoration (both subsidiary ornamentation and human/animal figures), particularly if one is making a study of unpublished excavation pottery. The large quantities of well-studied figure-decorated pottery from Athens, Corinth, and Samos, among other places – complete with drawings, photographs, and detailed descriptions - can provide either immediate comparanda or at the very least serve as a vital starting point for further exploration. A second go-to place, and one where many would turn for Athenian vases, is the Beazley Archive Pottery Database (BAPD) maintained by the Beazley Archive at Oxford University (www.beazley.ox.ac.uk). The BAPD was based initially on the famous ‘lists’ of Attic black- and red-figure vase-painters compiled by Sir John Beazley and published in several volumes. Since Beazley’s death in 1970, the paper archive comprised of photographs, notes, and drawings, has been digitized and made available as an open access searchable database (Smith 2005). A large number of entries and accompanying photographs have been added since the death of Beazley, and the Beazley Archive was able to provide the following statistics as of 21 July 2014:

- Greek vases all fabrics: 107,610
- Greek vases with images: 60,289
- All Athenian vases: 80,183
- Athenian black-figure: 42,664
- Athenian red-figure: 37,519
- All vases in Beazley’s lists: 34,000

Although on average the BAPD adds 5000 entries per year based on newly published material, it is most readily accessible to specialists and assumes a certain level of technical expertise. For non-Athenian vases, other scholars, such as A.D. Trendall for South Italian wares, have applied similar methods of attribution to establish their own classification lists complete with artists’ names and a legacy archive (http://www.latrobe.edu.au/trendall/). A third important resource for vase scholars is the Corpus Vasorum Antiquorum [CVA] an ongoing international project founded in 1922 with the aim of publishing in a more or less standardised form all vases from all collections around the world (http://www.cvaonline.org/cva/). Digitisation of the existing CVA volumes, many of which are now out of print, was undertaken by the Beazley Archive beginning in 2000. The contents of the CVA volumes were scanned as PDFs and are now searchable in their own right or as part of a BAPD search for an individual object. As with the excavation reports and the BAPD itself, the CVA volumes are written by specialists for experts in the field of Greek vase scholarship; and despite their production cost, more are being put into print each year. Each of these resources, furthermore, should be
understood as a starting point for research and, one that leads on to the next step, rather than as an end in itself. A fourth resource, and one which has become a bit less fashionable (and even slightly controversial in the wake of Beazley), is the scholarly monograph devoted to a single vase-painter (e.g. Oakley 1990). Unlike other areas of archaeology, Greek vase scholars have shied away from statistical analysis and are well aware of the shortcomings of using any of the resources listed above (including the BAPD) for such a purpose. Although attempts in this direction can be met with harsh criticism, the tide may slowly be turning towards a slightly more scientific approach to Greek pottery presentation (cf. Walsh 2014; Smith 2007).

The place of Greek figural pottery within the larger context of both online and print resources for the study of Classical art and archaeology more generally, should be emphasized. Because of their rich and informative iconography, Greek vases from all known centres, have found their place in major print reference works such as the Lexicon Iconographicum Mythologiae Classicae (LIMC) and its companion publication Thesaurus Cultus et Rituum Antiquorum (ThesCRA). Digital resources for Classical studies, most notably the Perseus Digital Library (http://www.perseus.tufts.edu/hopper/), includes Greek vases in its Art and Archaeology Artifact Browser, where one is able to search according to: collection, context, painter, period, potter, region, shape, ware, and keyword. The catalogue of vases in Perseus is by no means comprehensive, but the keyword function makes it more user-friendly than the more focused options mentioned above. It is also of critical importance to stress that “most online citation databases subscribed to by academic and research libraries do not cover the scope of ancient art” (Stylianopoulos 2012: 714). And, although an increasing number of museums are putting their collections on the Web, their searchable databases are of most use to Greek vase scholars in search of publication or teaching illustrations (Stylianopoulos 2012: 720-721), and are only of limited value either as research tools or as reliable introductions for non-specialists.

Finally, due to the wide variation in the native languages of these databases (and lack of consistent, controlled vocabulary even among databases of the same languages), it is difficult to conduct research across multiple systems. Inspired by the recent successes of Nomisma.org, a collaborative project dedicated to the definition of numismatic concepts with Linked Open Data (LOD) methodologies, we are endeavoring to create a similar project, Kerameikos.org, which would define the intellectual concepts of Greek pottery according to similar standards. Rather than relying on textual strings as controlled vocabulary (the painter, Exekias, in most Western scholarship is transliterated to 埃克塞基亚斯 in Chinese), concepts are represented by HTTP URIs, following the definition of Linked Open Data by Tim Berners-Lee (Berners-Lee 2006). These URIs, when dereferenced, may offer both human-readable HTML or machine-readable RDF or JSON-LD, following standard ontologies. Lastly, by definition these concepts, encoded in RDF, link to concepts in other information systems. Kerameikos.org may provide a pathway to normalisation for many pottery databases, enabling the large scale data aggregation and subsequent analyses that are not currently possible within the discipline.

2. Kerameikos.org: Linked Data Applied to Ceramics

Kerameikos.org is primarily a thesaurus of ceramic concepts, but the user interface includes a handful of visualisation and analysis features that will be detailed further in this paper. In its initial development phase, more than 60 URIs were minted to define concepts in a small handful of categories: production place, person or organization responsible for some aspect of production (i.e. a painter or potter), shape, material, style, period, ware, and technique. These URIs contain a list of preferred labels in a variety of languages, which would enable multilingual user interfaces, and a list
of URIs of identical or related concepts defined in other linked data systems. The latter component is an especially important one in the development of large-scale interoperable systems, as we will demonstrate.

Many pottery concepts defined on Kerameikos.org have also been defined in the LOD thesauri published by the British Museum (http://collection.britishmuseum.org/) and the Getty (http://vocab.getty.edu). Unfortunately, neither thesaurus is “five star” linked data (yet), as defined by Berners-Lee—that is, they do not link to related external resources. Kerameikos.org is intended to serve as the intermediary bridge between disparate systems, enabling the aggregation of content from collections that have adopted Kerameikos.org URIs natively, as well as collections that provide data that conform to British Museum or Getty identifiers. Therefore, the concept of the Black Figure technique, defined by http://kerameikos.org/id/black_figure, is the same as http://collection.britishmuseum.org/id/thesauri/x14736 in the British Museum thesaurus and http://vocab.getty.edu/aat/300387209 in the Getty Art and Architecture (AAT) thesaurus. Matches are also made with dbPedia, when possible. Painters and potters defined in Kerameikos.org are linked to URIs in the Virtual International Authority File (http://viaf.org) and the Lexicon of Greek Personal Names (LGPN: http://lgpn.ox.ac.uk/). Ancient production places are linked to URIs defined by the Pleiades Gazetteer of Ancient Places (http://pleiades.stoa.org). Kerameikos.org concepts will incorporate links to other systems as they become available.

Pottery concepts defined in this project have a varied audience, predominately vase scholars and their students. However, the benefit of Kerameikos.org to the museum community cannot be understated, particularly those institutions that own ceramics, but whose databases do not already incorporate controlled vocabulary that conform to existing thesaurus systems (e.g., the Getty). Even for those institutions that use Getty or British Museum identifiers to categorize their objects, Kerameikos.org could provide a pathway to wider publication, access, and scholarly query through its SPARQL endpoint, as semantic reasoning makes it possible to associate similar objects together regardless of whether they are categorised by Kerameikos.org, Getty, or British Museum URIs.

Archaeological databases have tremendous potential in shaping scholarly research in the discipline. Kerameikos.org could represent a stable set of identifiers, curated and maintained by pottery specialists, that excavations might incorporate into their own databases. This reduces the need for individual projects to maintain internal typologies, while making it easier for them to publish content into the larger LOD cloud. Even among those projects that implement semantic web methodologies, like CLAROS (http://www.clarosnet.org/XDB/ASP/clarosHome/), URIs for concepts (whether ceramic ones from Kerameikos.org or numismatic ones published by Nomisma.org) can aid in normalisation and enhancement of the user experience.

With the goals of Kerameikos.org in mind, the authors developed specifications for the first phase of the project:

- The ability for pottery specialists to create, edit, and publish identifiers (with training for the semantic meaning of properties, but without need to understand RDF).
- Human readable pages for each identifier.
- Geographic and quantitative visualisation features that specialists are accustomed to create in their research, delivered nearly instantaneously.
- Machine readable serialisations in relevant models (e.g., RDF/XML, Turtle, JSON-LD, KML, geoJSON, etc.).
• REST APIs that deliver data efficiently.
• A published ontology, with accompanying documentation.
• Receive, validate, and post RDF of vase data into an RDF triplestore to facilitate queries and visualisations mentioned above.

The prototype interface (both the front end and administrative back end) was developed upon a small subset of several dozen identifiers that represent the essential categories by which Greek vases are classified. The architecture of this system will be discussed in greater detail below, but we must first detail the thought processes that informed the development of an ontology and model that maps the organization of specialised ceramics knowledge into machine terms.

2.1 Ontologies and Models

2.1.1 Ontology and Model of Kerameikos.org Concepts

Before delving into the technical specifics relating to RDF models and ontologies, it was imperative to approach the challenge of knowledge representation by evaluating the traditional classification methods of Greek vases. Firstly, how does a pottery specialist categorize objects? As indicated above, Greek vases are classified and presented by the following fairly limited set of typologies: artist/creator (painter and/or potter: could be an individual or organization), decoration/iconography, material, period, place of production, shape, style, technique, and ware. Although these typologies (or classes, in ontological terms) apply specifically to Greek pottery, they are generally applicable to ancient ceramics of other periods and cultures.

In order to promote reusability, we opted to incorporate classes and properties from other RDF ontologies when possible. Several classes from CIDOC-CRM were adopted: E4_Period, E53_Place, and E57_Material. While the CRM provides an Actor class, we adopted Person and Organization from Friend of a Friend (FOAF) to designate painters, potters, and workshops. Most identifiers are simultaneously designated as concepts, as defined by Simple Knowledge Organization System (SKOS). Accordingly, since a production place, e.g. Athens, is reckoned to be a concept, but also has a geographic component, we have chosen to implement SpatialThing from the World Wide Web Consortium (W3C) geographic vocabulary (World Wide Web Consortium 2004) to encapsulate coordinates in the form of simple points (with geo:lat and geo:long) or more complex polygons that represent regions, encoded in geoJSON (Figure 1).

<ecrm:E53_Place rdf:about="http://kerameikos.org/id/athens">
  <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
  <skos:prefLabel xml:lang="en">Athens</skos:prefLabel>
  <skos:prefLabel xml:lang="ja">アテネ</skos:prefLabel>
  <skos:prefLabel xml:lang="es">Atenas</skos:prefLabel>
  <skos:prefLabel xml:lang="fr">Athènes</skos:prefLabel>
  <skos:definition xml:lang="en">Athens dominates the Attica region and is one of the world’s oldest cities, with its recorded history spanning around 3,400 years.</skos:definition>
  <skos:exactMatch rdf:resource="http://collection.britishmuseum.org/id/place/x22744"/>
  <skos:relatedMatch rdf:resource="http://pleiades.stoa.org/places/579885"/>
  <skos:exactMatch rdf:resource="http://dbpedia.org/resource/Athens"/>
  <skos:broad rdf:resource="http://kerameikos.org/id/attica"/>
  <geo:location rdf:resource="http://kerameikos.org/id/athens#this"/>
</ecrm:E53_Place>
<geo:SpatialThing rdf:about="http://kerameikos.org/id/athens#this">
  <geo:lat rdf:datatype="xsd:float">37.972</geo:lat>
  <geo:long rdf:datatype="xsd:float">23.726443</geo:long>
  <ecrm:P88i_forms_part_of rdf:resource="http://kerameikos.org/id/attica"/>
</geo:SpatialThing>
Some classes—such as shape, ware, or technique—are particular to the discipline and have no equivalent in other ontologies. The British Museum and Getty thesauri are intended to be applied to a diverse array of object types found in museums, and so their applications are more general. For example, in the British Museum thesauri, vessel shapes such as lekythos and amphora are considered object types (rather than forms or shapes). As a result, we endeavored to create an ontology, identified by http://kerameikos.org/ontology# (prefix: ‘kon’), for pottery-specific classes and properties that would fill in these gaps. Therefore, http://kerameikos.org/id/lekythos and http://kerameikos.org/id/amphora are defined by the kon:Shape class; http://kerameikos.org/id/black_figure is a kon:Technique; http://kerameikos.org/id/boeotian is a kon:Ware.

As mentioned previously, the most vital aspects of any concept defined in Kerameikos.org are the multilingual labels and links to related concepts in other LOD systems. SKOS provides a clear avenue for encoding labels with skos:prefLabel (preferred label) and skos:altLabel (alternative label) and linking external concepts with skos:exactMatch and skos:relatedMatch. The Exact Match property is used when linking to the Getty AAT, British Museum, VIAF, LGPN, and most other systems. The Pleiades Gazetteer of Ancient Places is the exception; skos:relatedMatch is used because the distinction between places as concepts and spatial features is not clearly defined. All IDs require an English definition (denoted by the skos:definition property), but definitions in other languages may be used. SKOS scope notes may be included to define rules governing the semantic application of a concept.

The other essential linking property incorporated from the SKOS ontology is skos:broadere, used to connect concepts hierarchically. Logically, this would apply to production places: Athens is in Attica, which is in Greece. Techniques, shapes, periods, and styles may also be linked together with skos:broadere, following the traditional organizational structure of Greek pottery. Bell kraters and volute kraters are specific renditions of the krater shape. The Black Figure technique, by definition, comprises both silhouette and incision techniques. Early Helladic III is the third phase of the Helladic period, which is one of a number of eras of the Aegean Bronze Age. Applying the SKOS ontology to these concepts makes it possible for Kerameikos.org identifiers to reflect the current methods of pottery classification.

2.1.2 Using CIDOC-CRM with Kerameikos.org Properties to Encode Vases
In order to demonstrate the utility of Kerameikos.org as both a LOD thesaurus and research portal, we ingested vase data into the project’s underlying triplestore to facilitate analysis and visualisation. Following examples of applying CIDOC-CRM to Greek vases established by the British Museum, we developed two small datasets of vases: one derived from a CSV file provided by the Getty Museum and processed with PHP into RDF/XML conforming to the British Museum CIDOC-CRM model and the other extracted programmatically from the British Museum SPARQL endpoint. In all, data for more than 20 vases in the Getty and several hundred from the British Museum were processed into CIDOC-CRM and published into Kerameikos.org’s RDF triplestore. An example of one vase is shown in Figure 2.

  <dcterms:title xml:lang="en">Oil Jar with a Departing Warrior</dcterms:title>
  <dcterms:identifier>71.AE.442</dcterms:identifier>
  <ecrm:P50_has_current_keeper rdfs:resource="http://kerameikos.org/id/getty_museum"/>
In this example, although many CRM properties were used, several properties from other ontologies were incorporated to simplify the model and improve processing efficiency: title from Dublin Core Terms and FOAF properties for encoding image URLs. In addition to a handful of pottery-specific classes, the Kerameikos.org ontology also includes a number of properties that more closely resemble the real-world organization of ceramic knowledge, e.g., kon:hasShape. With the exception of crm:P50_has_current_keeper, which links to a Kerameikos.org URI defining a legal body that curates the object, British Museum or Getty vases are defined by URIs in their respective vocabulary systems. Semantic reasoning inherent to Kerameikos.org’s linked data architecture makes it possible to query objects defined in different thesauri.

2.2 Architecture
Kerameikos.org’s software architecture is nearly identical to that of Nomisma.org and similar to other American Numismatic Society projects, due to Gruber’s software development role at the American Numismatic Society. Kerameikos.org is an open source collection of scripts that handle the front-end user interface and back-end administrative functions. The framework is available on Github at https://github.com/kerameikos/framework. This collection of scripts—predominantly XSLT, XForms (and supporting standards), and Javascript—binds together open source server applications that run in Java, specifically in Apache Tomcat.

The web application architecture is based on XRX (XForms, REST, and XQuery) (McCreary 2008), though XQuery has been substituted for the combination of SPARQL and Apache Lucene/Solr. Kerameikos.org identifiers are stored in the filesystem as RDF/XML files, maintained in a Github repository (https://github.com/kerameikos/ceramic-ids), which enhances access in addition to providing version control. Apache Solr is the de facto standard search index in the cultural heritage sector, prevalent in repository applications, like Fedora, or content aggregators, like Europeana. XForms, a W3C standard schema for advanced web form functionality, is the backbone of Kerameikos.org, handling the editing and publication process of ceramic concepts in addition to processing RDF data dumps from contributors (World Wide Web Consortium 2009). XForms enables the publication of data into the project’s internal Solr index and RDF triplestore (Apache Fuseki is currently implemented in Kerameikos.org), but also supports REST interactions with a variety of external Linked Open Data services, which will be discussed in greater detail below. The XForms processor used by Kerameikos.org is Orbeon (http://orbeon.com). Orbeon provides an array of processors that serialise the canonical RDF/XML concepts, Solr and SPARQL query results, and
various APIs into all of the content types required to power the user interface—from human-readable HTML to machine-readable geoJSON-LD, Turtle/RDF, and KML for mapping.

Kerameikos.org’s architecture is modularised. With minor modifications, the Github/filesystem-based storage could be replaced with any number of off-the-shelf NoSQL database solutions. Fuseki could be substituted with any SPARQL 1.1-compliant RDF triplestore. Solr could be replaced with any search index software that supports REST, and Orbeon could be swapped with other XForms processors, such as XSLTForms or betterFORM. While the Kerameikos.org framework does include code fragments which are specific to this project, we aim to make this application more generalisable and adaptable for publishing other types of LOD thesauri.

The system is flexible and scalable. This framework differs from the LOD thesauri published by the British Museum, Library of Congress, and the Getty Museum. Kerameikos.org concepts are made available through linked data methodologies, but the project’s user interface includes various query APIs and visualisation tools; it is a research portal for Greek pottery in addition to a vocabulary system. Furthermore, the back-end system is enterprise-ready: scalable to millions of ceramic concepts and hundreds of millions of triples, managed through stable information systems workflows. These applications scale well beyond the ceiling we ever anticipate encountering, even if we expand beyond Greek pottery into the typologies of ceramics from other periods and cultures.

2.2.1 Administrative Editing and Publishing Interface

The most important aspect of the administrative interface for Kerameikos.org is the editing interface that pottery scholars may use to create well-defined concepts without the need to edit RDF directly or have any specific technical knowledge about the RDF publication process (although knowledge of what a property means semantically, e.g., when to use skos:exactMatch rather than skos:relatedMatch, is required). Kerameikos.org presents editors with an intuitive web-based interface, with controls to expedite the lookup and linking mechanisms with other LOD thesaurus systems (Figure 3).

To reiterate, XForms drives the editing interface. While a comprehensive discussion of the standard is beyond the scope of this paper, a brief synopsis will be provided:

In recent years, support for the standard has moved from native browser support to standalone processors, whether client or server-side. XForms’ XML controls are embedded within XHTML files, and the processor is responsible for transforming this document into a functional HTML-based web form, including all of the necessary CSS for styling and Javascript to handle sophisticated validation, REST interactions, and other controls on the web page (e.g., to add a new skos:prefLabel field into the document model when a user clicks a button in the user interface). XForms adheres to the Model-View-Controller (MVC) architectural framework. As an XForms developer, one needs only to focus on the development of the schematic functionality of the web form. The XForms code can be ported from processor to processor with the expectation of continued performance.

XForms applications support the editing and validation of complex XML models based on XPath and XML Schema. Below are some of the more common validation scenarios:

1 For further information about XForms applied to cultural heritage informatics, see Gruber et al. 2010.
A concept must have an English preferred label and definition, and there may be no duplicate languages for these properties.

- Latitude and longitude must be decimal values between -180 and 180.
- URIs inserted into linking properties, like skos:exactMatch, must be valid.
- A geo:SpatialThing must contain either geo:lat and geo:long or a geoJSON-encoded polygon.

If all validation scenarios are met while editing an ID, the Save button will become active, and upon saving, the XForms application will execute events to serialise the RDF/XML model into an XML file and store it on the disk (for pushing to Github later). The RDF triples will be refreshed in the triplestore through SPARQL/Update, and the RDF/XML will be transformed into another XML model conforming to the Solr schema and indexed for keyword and faceted search in the public user interface.

This RDF editor allows a user to insert all of the aforementioned properties used in Kerameikos.org’s data model, but the application includes a number of features that expedite the editing process. When creating a Place concept, the user does not manually enter geographic coordinates, but rather uses an OpenLayers-based map to either create a point or draw a polygon. With a Javascript extension of the application, Orbeon interacts with OpenLayers to update the RDF with these features (Figure 4).

The XForms application also interacts with REST APIs provided by external thesaurus systems. The editor may wish to link a Greek vase painter defined in Kerameikos.org with identities defined on VIAF. VIAF provides query results in an RSS feed. A search for “Exekias” may yield numerous results, as VIAF still has much work to do in the disambiguation of entities (Figure 5). The editor may select applicable boxes, defining them as Exact Matches, and the URIs will be imported in the RDF model. Furthermore, the XForms application will request the RDF for each of these entities from VIAF, extracting owl:sameAs URIs for matches in other systems, such as dbPedia or the Gemeinsame Normdatei (GND), maintained by the Deutsche Nationalbibliothek. If the Kerameikos.org identifier contains a skos:exactMatch for a dbPedia entry, preferred labels may be extracted directly from dbPedia’s RDF. Similar lookups are available for other types of concepts—like shapes and techniques—in the Getty AAT, which are made available through their SPARQL endpoint. This mechanism will be extended in the Kerameikos.org editing interface when the other Getty vocabularies become available as LOD. Essentially, specialists can hit the ground running, rapidly making links to other vocabulary systems in the creation of Kerameikos.org concepts, with labels harvested programmatically from dbPedia.

2.2.2 Public User Interface

Presently, the public user interface of Kerameikos.org is generated primarily through Orbeon and the XML Pipeline Language (World Wide Web Consortium 2002). It includes a browse page generated from Solr query results (http://kerameikos.org/id/), a page which serves as a placeholder for documentation of the ontology (http://kerameikos.org/ontology), a SPARQL endpoint (http://kerameikos.org/sparql), and pages for each individual concept. We intend to introduce a variety of REST APIs which either interface with SPARQL to deliver common queries more efficiently or deliver machine readable serializations more expediently. The ontology page is generated dynamically from the OWL/RDF file and available in RDF/XML and Turtle. There is much work to be done throughout the public user interface, but the pages for individual concepts provide the clearest picture of the potential for Kerameikos.org as both a dynamic thesaurus of ceramic concepts and a research portal for students and scholars of the discipline.
The pages for each entity typically contain the following in the body: the RDF predicates and objects rendered in HTML5 (with underlying RDFa attributes in the HTML, enabling extraction of triples with an RDFa parser), a map and timeline (with the Timemap Javascript library [https://code.google.com/p/timemap/]) showing the distribution of a particular typology over time and space (Figure 6), a list of thumbnails representing the typology, and a very simple form which allows a user to perform basic quantitative analysis of the typology, e.g., to show the numeric distribution of shapes created by a Greek potter (Figure 7). The maps and quantitative analyses are generated by using XSLT to serialise underlying SPARQL query results into the JSON models required by the map and chart Javascript libraries. The sidebar includes links to alternate serialisations of the concept in RDF/XML, Turtle, JSON-LD, and KML, as well as contextual information extracted from other LOD sources (e.g., birth and death dates extracted from Text Encoding Initiative (TEI) XML provided by LGPN). Furthermore, Orbeon supports semantic HTTP 303 redirects and content negotiation, enabling technical users to extract JSON-LD by requesting the content type of “application/json” from http://kerameikos.org/id/black_figure. We support both content negotiation and REST to maximise access and reuse of data.

We will eventually enhance these quantitative analysis features in order to deliver the types of visualisations that scholars are accustomed to generating in their own research. With more typologies defined in Kerameikos and more vase data in the system, we might alter research possibilities in the discipline. Scholars could, in theory, spend more time on analysis and interpretation rather than data gathering, with museum and archaeological materials made available for query through the same system. Although the BAPD achieves this to the same extent at present, its primary function is to serve as a repository of images and bibliography (built on the format of Beazley’s ‘lists’), its nucleus is Athenian pottery specifically, and its search terms are not linked externally. With Kerameikos.org’s analysis and visualisation tools, it might be possible to conduct sophisticated queries based on geography and statistics, paving the way for studies of the ancient economy (akin to research methodologies in numismatics). These tools may even extend the boundaries of the discipline, making it possible to formulate new types of questions that may never have been considered without large-scale aggregation and visualisation of ceramics data.

3. Future

In this paper, we have effectively demonstrated the potential of Kerameikos.org and provided a glimpse of the sophisticated query and visualisation mechanisms made possible by applying linked open data methodologies to Greek vases. On a technical level, this is not experimental digital work. We are applying stable, tried-and-true technologies to a field in need of a re-envisioning of the role that digital tools play enhancing the research and synthesis process. Kerameikos.org will only succeed with community buy-in, and this interest can take years of hard work to generate. With this in mind, we are forming a steering committee that consists both of Greek pottery and cultural heritage informatics specialists. Our goal is to develop a system that effectively represents the intellectual organisation of ceramics knowledge, based upon current Linked Open Data standards.

Our initial phase has focused on the creation of a thesaurus of the simplest typologies to model, and we will progress toward more complex challenges in later iterations of the Kerameikos.org project: a system for maintaining data dumps of CIDOC-CRM describing vases, iconographic topic modeling, enhanced models for representing and linking painters, potters, and workshops (which, in turn, would enable social network analysis of the ancient pottery industry), and further interaction with other LOD
projects, such as the Pelagios Project and Standards for Networking Ancient Prosopographies. We will continue to build this project and adapt it to the needs of the Classical archaeology and museum communities.

4. Acknowledgments
The authors would like to acknowledge Natasha Dakouri-Hild at the University of Virginia, Thomas Mannack of the Beazley Archive at Oxford University, Sebastian Rahtz, Director (Research Support) of Academic IT Services at Oxford, Vladimir Alexiev at Ontotext, for insights into ontology development, and the linked data team at the Getty Museum (specifically, Joan Cobb, Patricia Harpring, and Brenda Podemski), for providing vase data.

5. Bibliography


