Dura-Europos Stories

Katherine Thornton [a,*] Kenneth Seals-Nutt [a] Anne Chen [b]
[a] *Stories Services Collaborative, USA*
E-mail: katherine.thornton@yale.edu
[b] *Bard College, USA*
E-mails: kenneth@seals-nutt.com, achen@bard.edu

Abstract. We introduce Dura-Europos stories, a multimedia application for viewing artifacts and places related to the Dura-Europos archaeological excavation. We describe the process of mapping data to the Wikidata data model and contributing data to Wikidata. We provide an overview of the functionality of an interactive application for viewing images of the artifacts in the context of their metadata. We contextualize this project as an example of using knowledge graphs in research projects in order to leverage technologies of the Semantic Web in such a way that data related to the project can be easily combined with other data on the web. Presenting artifacts in this story-based application allows users to explore these objects visually, and provides pathways for further exploration of related information.

Keywords: Wikidata, Art History, Archaeology, Cultural Heritage, Digital Humanities

1. From Dataset to Multimedia Application

People working in cultural heritage organizations often publish images on the web to make groups of artefacts browsable for people who would like to view the collection remotely. We designed the Dura-Europos Stories project to pull data from a public knowledge graph, Wikidata, that contains information from multiple institutions related to the Dura-Europos archaeological site[1]. Not only does this application provide support for browsing images of artefacts, we also present them in the context of the larger site and connect them to as much supporting information as possible.

Similar projects in the domain of digital humanities include the Biography Sampo project and the other Sampo applications from Finland [1, 2]. Many researchers have explored the use of semantic technologies to support historical inquiry [3, 4]. Other projects have also described the use of Wikidata to support research in the domain of archaeology [5]. Some groups are exploring how to display 3D content related to collections [6]. A diverse range of topics are covered under the term 'digital humanities', and researchers have reused data from Wikidata in many digital humanities projects [7]. Our project contains some features of digital databases described in [8]. We based many of the design decisions for this project on our previous work creating multimedia biographies powered by Wikidata [9, 10]. We describe a collaborative project between

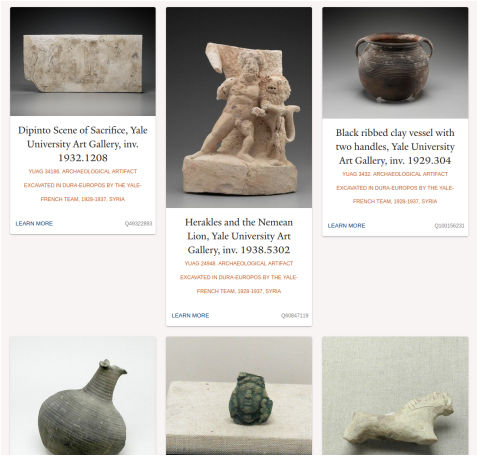

Fig. 1. Detail of the Browse page in the Dura-Europos Stories application.

*Corresponding author. E-mail: katherine.thornton@yale.edu.

[1]The Dura-Europos Stories project is available at: https://dura-europos.stories.k2.services/stories

art historians, information scientists and software engineers to create an interactive application for viewing multimedia content related to the excavations at Dura-Europos.

2. Dura-Europos

The site of Dura-Europos (Syria) was founded around 300 BCE as part of the Seleucid Kingdom, a Hellenistic Greek successor state formed in Western Asia after the death of Alexander the Great. Located on the western bank of the Euphrates river, the city stood on valuable real estate between competing ancient eastern and western powers. As a result, over the course of its history, the city passed from Seleucid control through successive phases of Arsacid (Parthian) and Roman occupation, and ultimately fell to the besiegement efforts of the Sasanian Persians in the 250s CE. Attempting to shore up the city's defenses in advance of the Persian attack, the Roman soldiers garrisoned at Dura constructed a massive earthen embankment to reinforce the city's vulnerable western wall. In a move that would prove fateful for the site's exceptional archaeological preservation, buildings in the vicinity of the west wall–including the oldest excavated Christian church, the most elaborately decorated ancient Jewish synagogue discovered to date, and various pagan temples that attest to the city's ethnic and cultural diversity–were requisitioned and filled with earth and debris.

What was at the time a highly destructive process, in fact, created (together with the hot, dry climate) the conditions for the site's extraordinary degree of preservation. Thanks to these unique circumstances, large sections of rare mural painting, and hundreds of objects made from organic materials survive at Dura-Europos. These rarely-preserved artifact types were found together with sculptures, inscriptions, arms and armor, ceramics, coins, and other objects of everyday life. The site therefore provides unparalleled glimpses into the multicultural, religiously-diverse frontier life, and the running of a military garrison with coexisting soldiers and civilians.

Controlled excavation at the site began in the 1920s, a period when Syria was governed under the French Mandate. In the wake of World War I, British troops stationed at Salihiyeh took notice of well-preserved wall paintings in the fortress' ruins, ultimately setting into motion a few short exploratory campaigns, before the eventual establishment of a long-standing excavation effort sponsored jointly by the French Académie des Inscriptions et Belles-Letres and Yale University. The joint French-American expeditions would continue for ten seasons between 1928 and 1937. Thanks to the practice of partage (division of finds) that was current among archaeological collaborators in the early 20th century, the majority of artifacts from Dura today reside in collections at Yale University and the National Museum of Damascus (established 1919). Due to the importance of the site, however, over time smaller groups of Durene objects have made their way into collections across the world[2].

Following a hiatus of nearly 50 years, systematic archaeological work began once again in 1986 under the auspices of the Mission Franco-Syrienne de Doura-Europos (MFSED). In addition to new excavations, the Franco-Syrian commission undertook re-examination of previously excavated areas and preservation efforts before being forced to terminate in 2011 with the outbreak of armed conflict in Syria[3]. Since 2011, extensive illicit digging and intentional destruction has been documented in the archaeological remains at Dura[4].

3. Dura-Europos Data Model

The core of Dura-Europos data set consists of metadata created by the Yale University Art Gallery. The International Digital Dura-Europos Archive (IDEA) team then created mappings from the metadata to Wikidata properties to align with the Wikidata data model. For each artefact there are pieces of information about the title, location, height, width, length, approximate age, place of excavation, materials used, and type. Currently there are more than fourteen thousand items in the Dura-Europos collection.

After aligning the existing metadata to the Wikidata data model, the IDEA team created thousands of new

[2] https://w.wiki/3MFr

[3] On the history of excavations at Dura-Europos, see Baird, J.A. 2018. Dura-Europos, Bloomsbury, p. 1-16; Brody, Lisa. 2011. "Yale University and Dura-Europos: From Excavation to Installation." In Dura-Europos: Crossroads of Antiquity, 17–32. Chicago: University of Chicago Press.

[4] Wolfinbarger, Susan, Jonathan Drake, Eric Ashcroft, Katharyn Hanson, 2014, "Ancient History, Modern Destruction: Assessing the Status of Syria's Tentative World Heritage Sites Using High-Resolution Satellite Imagery," Accessed November 18, 2022. https://www.aaas.org/resources/ancient-history-modern-destruction-assessing-status-syria-s-tentative-world-heritage-sites-7#12. See also https://duraeuroposarchive.org/digital-accessibility/.

items to represent the artifacts, and contributed tens of thousands of statements about them. The artifact with greatest number of statements is 'Large bowl with rounded sides (twenty-seven rim and side fragments), Yale University Art Gallery, inv. 1938.5999.490' at 66 statements [5]. The bowl consists of twenty-seven fragments. The large number of statements is due to the fact that measurements for the dimensions of each fragment are provided on this item. The artifact with the lowest number of statements is 'Lizard, Yale Peabody Museum of Natural History, YPM ANT 015992' [6]. On average, artifacts from this collection have fifteen statements.

4. Wikidata Data Model

Wikidata is a knowledge base of structured data that allows community members to contribute data [11]. More than 12,000 editors are active in Wikidata each month [12]. The Wikidata data model consists of items and properties. Items have identifiers, called Qids, structured as unique resource identifiers (URIs). Properties are identified with Pids. For example, in Figure 2, we see that the property 'made from material' (P186), is used with three different values 'wood', 'rawhide' and 'paint'. Similarly, property 'depicts' (P180), is used with four values to indicate what figures are represented on the scutum.

The Wikidata community uses this identifier structure in order to be neutral with regard to human languages. Wikidata currently supports more than 300 human languages, so labels for each language are then associated with the numeric identifiers [13]. The multilingual support available in Wikidata is relevant for Dura-Europos because people from many different language backgrounds are interested in the excavation and artifacts. Institutions in several different countries hold materials related to the Dura-Europos excavation, and the excavated objects have been described in several human languages. Due to the fact that the Dura-Europos data is now in Wikidata, people can now add information in many different languages that make this data accessible for more people.

When searching for artifacts, some people will search for the type of object in general. For example, if people use the word 'statue' as a search term, they will find that there are multiple statues that were

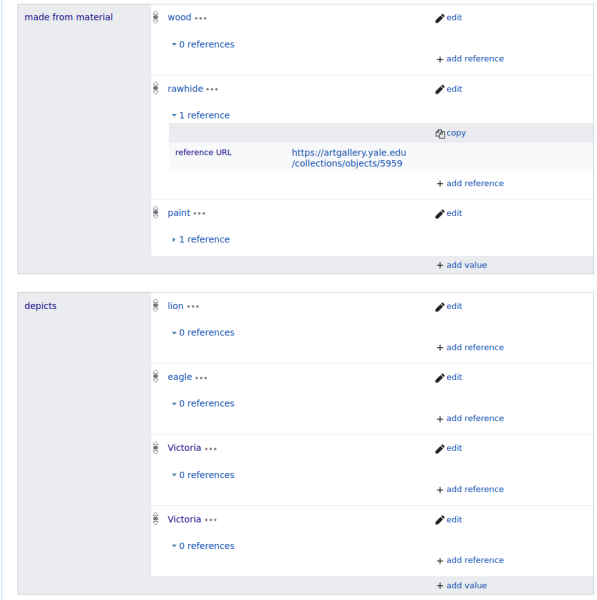

Fig. 2. Properties and values used on the Wikidata item for a scutum.

excavated from the Dura-Europos site. The Wikidata item for 'statue' is Q179700[7]. As of April, 2022 there are more than one-hundred-fourteen labels in different languages on the Wikidata item for 'statue'. In the current version of the Dura-Europos Stories application we have support for search functionality in English. In the future we plan to extend this to additional languages. The multilingual data in Wikidata will allow us to extend search functionality for different languages for words describing types of objects by adjusting the SPARQL query we use to identify these types within Wikidata. If we contrast this with the alternative of commissioning translators to create additional labels in many human languages, leveraging the multilingual data from Wikidata can conserve time and work for people building upon this data.

The Wikidata data model also includes references [14]. Editors can add references to any statement on any Wikidata item. The members of our team who contributed the Dura-Europos data to Wikidata created many references that point back to the Yale University Art Gallery website pages for the artifacts. For researchers interested in Dura-Europos, these references provide pathways to additional information about these materials.

[5] https://www.wikidata.org/wiki/Q100716223
[6] https://www.wikidata.org/wiki/Q98931969
[7] https://www.wikidata.org/wiki/Q179700

Fig. 3. A sample of labels in different languages for the Wikidata item 'statue'.

5. Getting Dura-Europos Data into Wikidata

To bring about the creation of Wikidata items for artifacts excavated from Dura-Europos, the IDEA team worked from spreadsheets of metadata provided by partner institutions. Using their own internal collection management system, each partner institution provided a CSV export of items in their collection related to Dura-Europos. Metadata pertaining to each item was listed in sheet columns. The CSV was then processed using Open Refine[8] to regularize equivalent expressions and correct spelling and spacing errors.

In preparing the initial upload into Wikidata, the aim was to reflect all of the metadata provided by the institution as completely as possible. Some metadata fields recorded by partner institutions in CSV column headings were easily matched with equivalent Wikidata properties with minimal research using the SQID tool[9]. For instance, the object 'medium' provided in institutional metadata mapped neatly to Wikidata Property 'made from material' (P186). Other fields supplied by the institution, like classificatory categories of artifact (ex. coin, lamp, statue), translated to 1instance of' (P31) Wikidata statements.

A few metadata fields, like those expressing uncertainty around the time of an object's date of creation or the object's dimensions required parsing of a single column into multiple columns in preparation for Wikidata upload. In institutional records, object dimensions were expressed with height, width, and depth measurements in the same 'dimensions' column. In Wikidata, however, each dimensional angle corresponds to its own property (ex. height (P2048), width (P2049), horizontal depth (P5524)). To ease the process of item creation, each facet of measurement was parsed into its own column with a new heading corresponding to the relevant Wikidata property. The IDEA team sought the advice of the Wikidata community on how to preserve the uncertainty surrounding an object's precise date of creation. Following the pattern established by more mature projects in the GLAM sector, for objects whose precise date of creation could not be established (ex. via externally datable references in inscripional content) the institution-supplied date column was parsed into three new columns: an 'inception' (P571) column whose value corresponded to 'placeholder for somevalue' (Q53569537), and two additional columns to hold values corresponding to 'earliest date' (P1319) and 'latest date' (P1326) as qualifiers on the 'inception' property. This modeling strategy allows one to query the data according to a span of years rather than forcing the artificial tethering of an object to a single date.

More sustained and specialized human input was required to extract Wikidata 'depicts' (P180) and 'location of discovery' (P189) statements, as well as item labels and descriptions that would allow easy disambiguation and capture metadata content not otherwise reflected in standalone statements. In the Dura datasets, the 'title' values provided by the host institutions contained information that pertained to discovery context (ie. 'location of discovery' (P189)) and/or the scholarly interpretation of content depicted. The institutional labels were thus manually assessed to create new 'location of discovery' (P189) and 'depicts' (P180) columns; values for each field were either matched with an existing item, or a new Wikidata item for the corresponding value was created (prior to running the batch upload)[10]. A summary of the properties used to translate the institutional CSV exports and rationale for specific modeling decisions is available on the Wikiproject IDEA page[12].

The expanded and modified CSV was then used as the basis for a batch upload. Data was written from the

[8]https://openrefine.org/

[9]https://sqid.toolforge.org/#/

[10]For the Dura dataset, all artifacts were given a 'location of discovery' (P189): 'Dura-Europos' (Q464266). In addition, however, the IDEA team is in the process of defining and publishing location identifiers down to the building level of granularity with Pleiades[11], an online gazetteer authority for ancient Mediterranean place-related data. The Dura-Europos urban gazetteer data from Pleiades will ultimately be mirrored in Wikidata to enable on-the-fly visualizations of artifact discovery locations.

[12]https://www.wikidata.org/wiki/Wikidata:WikiProject_IDEA

CSV into Wikidata using QuickStatements[13] and the DuraEuroposBot[14].

5.1. Getting Images into Wikimedia Commons

The IDEA team uploaded thousands of images of artifacts to Wikimedia Commons. Wikimedia Commons is a repository for multimedia content that is used by many projects of the Wikimedia Foundation [15].

Other organizations that hold material related to Dura-Europos may also contribute images or models to Wikimedia Commons. Currently four organizations have uploaded media related to materials from Dura-Europos: Yale University Art Gallery, Beinecke Rare Book & Manuscript Library, Bibliothèque nationale de France, and the Department of Near Eastern Antiquities of the Louvre. As additional organizations make the decision to publish media on Wikimedia Commons, additional images may become available. This means that a more holistic set of images could become available for reuse in the Dura-Europos Stories application. Rather than exclusively reusing images from a single institution, the images can be sourced from multiple collections. All of the metadata for these images is presented alongside them in the IIIF viewer we embed into the stories. In this way we are bringing images together from multiple collections, but also clearly communicating the provenance of the images and crediting the institutions that have made them available online.

5.2. Getting Polygons into Wikimedia Commons

Members of the IDEA team contributed data to Wikimedia Commons in the form of shapefiles, which represent multiple points on a map rather than a single set of coordinates. These shapefiles in Wikimedia Commons can be connected to Wikidata items through the use of Property 3896 ‘geoshape’. In Figure 4, we see a dark gray polygon on a light brown background that represents the footprint of Tower 3 at the Dura-Europos site. By connecting Wikidata items to shapefiles, we can reuse the data in the shapefile within mapping applications to indicate the layout, position, and relative sizes of these features in the Dura-Europos site.

[13] https://quickstatements.toolforge.org/#/

[14] (https://www.wikidata.org/wiki/User:DuraEuroposBot

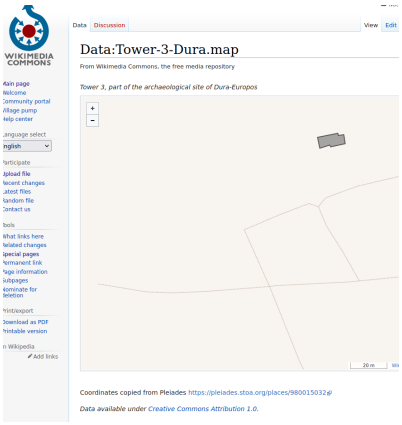

Fig. 4. A shapefile for Tower 3 of Dura-Europos stored in Wikimedia Commons.

6. Semantic Web

After contributing data related to Dura-Europos to Wikidata, we can leverage additional data from the Semantic Web. An example of data that is not in Wikidata itself, but is connected to items from the Dura-Europos dataset in Wikidata is the Pleiades Gazetteer. The Pleiades Gazetteer is a web-based digital gazetteer that describes places in historical contexts [16]. Wikidata has two properties related to the Pleiades Gazetteer project: Property 1584 ‘Pleiades ID’ and Property 2938 ‘Pleiades category identifier’ which can be used to connect Wikidata items to relevant content from the Pleiades Gazetteer. Currently more than thirty Wikidata items representing parts of Dura-Europos have Pleiades identifiers. This means that for each of those items we can follow that external identifier and find out additional information about these places from the Pleiades Gazetteer. For example, the Dura-Europos synagogue has the Pleiades identifier ‘335307374’. in Figure 5, we see the Pleiades page for the Dura-Europos synagogue[15]https://pleiades.stoa.org/places/335307374.

Combining data from many different sources allows us to create a more detailed and complex representation of the people, places, and artifacts related to Dura-Europos.

7. Data Quality

Many different stakeholders contribute data to Wikidata related to Dura-Europos. Multiple institutions

[15] https://pleiades.stoa.org/places/335307374

Fig. 5. Information related to the Dura-Europos synagogue from the Pleiades Gazetteer.

have collection material related to the Dura-Europos site and excavation and independently create items in Wikidata that describe their holdings. Many student researchers are also involved in projects related to data describing Dura-Europos. With a diverse group of contributors, we employ the strategy of sharing relevant data models on-wiki. We share our data models in Wikidata's E namespace which is dedicated to schemas [17]. The Wikidata community uses ShEx to encode schemas within the schema namespace. ShEx is a formal data modeling and data validation language for RDF graphs [18].

We compose our schemas in ShExC, the compact syntax. ShExC is concise, which makes it appropriate for human data modelers. It is also machine-actionable, thus our schemas can be used for validation of entity data from Wikidata. In this way we can share our data models with human editors who can then confidently contribute additional data in alignment with these data models. We can also test entity data for conformance with these schemas to identify items that need review or curation.

People looking for schemas in Wikidata's E namespace use the search bar to find relevant options. By using the pattern E: in combination with a search term, people can search across the labels of the schemas in the E namespace. In Figure 6, when searching for 'E:coin hoard' the schema E366 is returned.

In Figure 7, when searching in Arabic, people will find schema E366. If people who prefer specific languages for interaction add the 'EntitySchemaHighlighter.js' user script to their accounts, they will be able to see labels for items and properties used in the

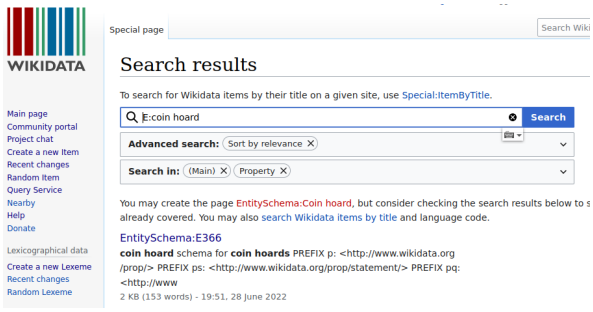

Fig. 6. Searching for 'E:coin_hoard' in the Wikidata search box returns schema E366.

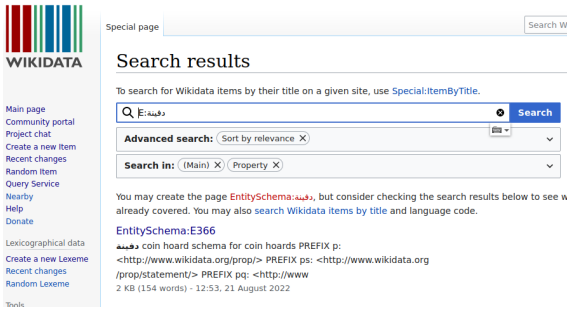

Fig. 7. Searching for 'E:' and Arabic word in the Wikidata search box returns schema E366.

schema when they hover[16]. Through the use of this user script, Wikidata editors have multilingual access to schemas making them an important data modeling tool that reduces barriers to communication about data structuring for people from different language contexts. When collaborating with others in different language communities, sharing data models via schemas in Wikidata's schema namespace allows people to read schemas in the language of their choice.

8. Digital Storytelling with Stories Services

The Stories application supports both browsing and searching for stories. The browse page consists of a set of cards representing individual artifacts. The cards include an image of the artifact, if available, as well as the Qid for the item representing the artifact in Wikidata as well as a 'learn more' button which takes the visitor to the story for the artifact. Visitors can also use the search bar to search for terms, names, places, and more. For example, searching for the word 'amphora' in the search bar of the Dura-Europos Stories applica-

[16] https://www.wikidata.org/wiki/User:Zvpunry/EntitySchemaHighlighter.js

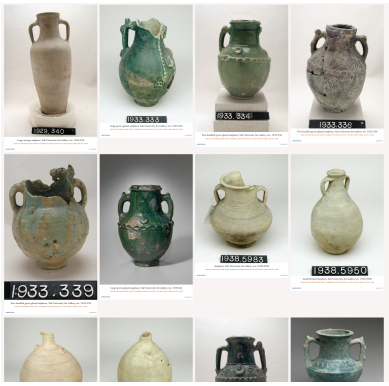

Fig. 8. Results after searching for 'amphora' in the Dura-Europos Stories search bar.

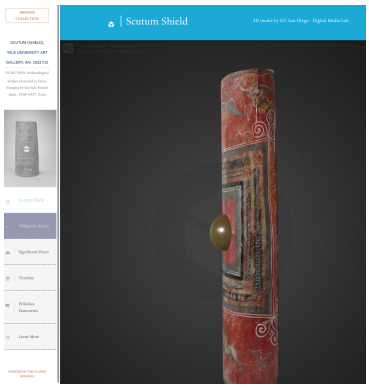

Fig. 9. 3D model of scutum visible in the Sketchfab moment in the Dura-Europos Stories application.

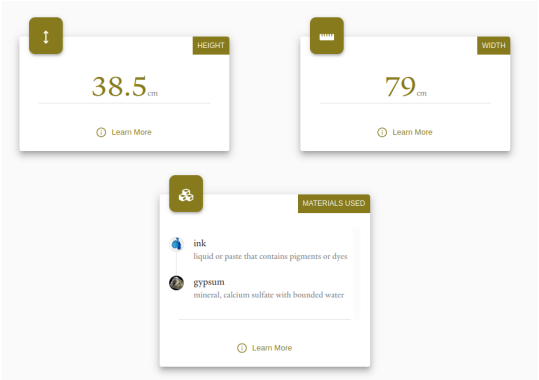

Fig. 10. Detail of the Stats moment in the Dura-Europos Stories application.

tion will return twenty-six results. A subset of the results can be seen in Figure 8.

We decided to highlight the images of these artifacts as the entry point into the stories. For each artifact with an image, the first moment on the story is the image and descriptive name of the object or location. These images are also used on the cards that make up the Browse page in the application as seen in Figure 1. The list of moments is presented on the left-hand side of the screen and serves as the navigational menu for each story. Users click on a moment to select it, and then the content is rendered in the primary window within the application. The moment menu is always available so that users can select the next moment of interest. Moments are interactive, and users can explore the content presented within each moment by hovering, clicking, and scrolling.

The Sketchfab moment provides a viewer for three dimensional models that have been uploaded to Sketchfab. Sketchfab is a website that allows people to upload 3D models and share them on the web [19]. For example, in one story there is a 3D model of a scutum excavated from Dura-Europos, as seen in Figure 9. The ability to view a 3D model of an artifact supports users in getting a sense of what the object looks like from different angles and provides a sense of depth. In future work we would like to integrate with additional sources of 3D content, such as tours of the Dura-Europos site itself, building on the work in [20].

The Artefact moment provides an overview of the dimension of the objects and the materials of which they are composed, as seen in Figure 10. This moment allows users to get a sense of how the object may have been created and an understanding of scale.

The Map moment is a visualization of any geocoordinates related to the object, such as location of excavation and location of the institution in which it is currently held, as seen in Figure 11. Plotting these coordinates provides the user with an understanding of how people have moved the object since the time it was excavated. For certain artifacts, such as the Dura-Europos Route Map, this moment provides detailed information of places mentioned in the artifact itself.

The Library moment contains a set of shelves on which icons are displayed that represent each article or book that describe these artifacts, as seen in 12. This is one strategy for weaving the images of the artifacts themselves into the context of the scholarly literature which describes them.

The Dura-Europos Digital Archive team annotated hundreds of images using the Wikidata Image Positions Tool[17]. This process involved using Wikidata property P2677 'relative position within image' as well

[17]https://www.wikidata.org/wiki/User:Lucas_Werkmeister/Wikidata_Image_Positions

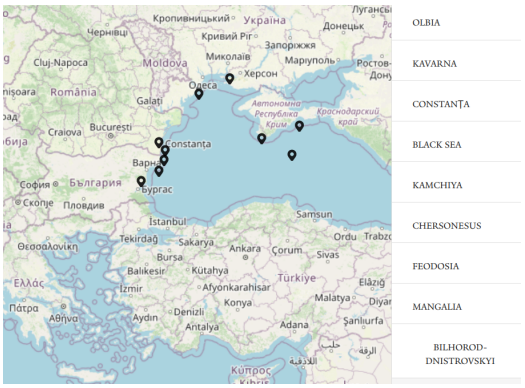

Fig. 11. Detail of the Map moment in the Dura-Europos Stories application.

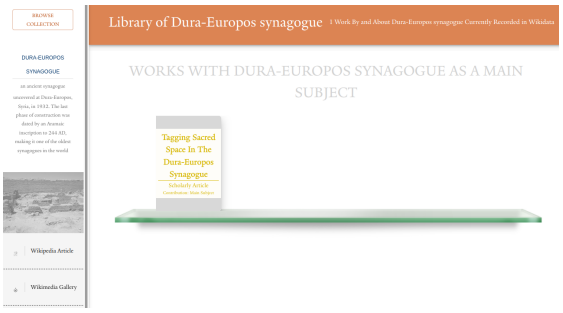

Fig. 12. Detail of the Library moment in the Dura-Europos Stories application.

as property P180 'depicts' to indicate a where something is and then provide an annotation of what is being depicted. This tool uses the International Image Interoperability Framework (IIIF) to display the image as a IIIF canvas with highlighted rectangles around detail areas with tool tips that provide the label for what is depicted. The ability to hover over different parts of the image and learn more about what the artifact represents provides an interactive experience to people exploring the Dura-Europos Stories application.

The Learn More moment displays cards describing information sources for further reading. As seen in 13 these cards include all external identifiers for the item in the story as well as all references used to support statements on the item. External identifiers are the identifiers for items in systems other than Wikidata. People can quickly gather information about where to go for additional information related to the subject of the story.

Many aspects of the stories are customizable. The order of moments, the text, and the colors within moments can all be used with their default settings or

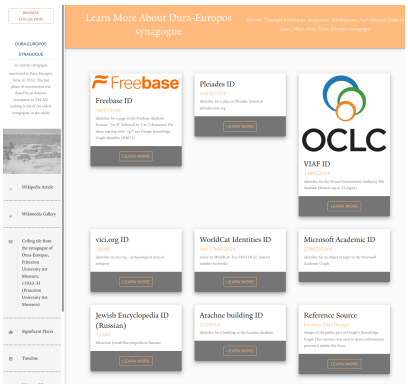

Fig. 13. Detail of the Learn More moment in the Dura-Europos Stories application.

with user-selected configurations. We provide a Publisher Workspace for people to configure their stories and customize their moments.

9. Stories Services

The Stories Services team created the Dura-Europos Stories application based on the framework we created for Science Stories [9]. These applications are both powered by Wikidata as well as the Stories Services API. We chose to use Python as the primary programming language for the backend of the application. We used a PostgreSQL database to store configuration information related to the presentation metadata of each story. We use Redis[18] to cache Wikidata SPARQL query results.

The Stories Services team created a package for working with Wikidata data in a Django application[19]. The Django-wikidata-api library is designed loosely around the core Django object-relational mapping (ORM), modified to interact with Wikibase via SPARQL queries rather than a relational database. The WikidataItemBase python class within the package has an interface for determining which statements are needed to represent a dataset. This class is also used to generate OpenAPI documentation, and to construct serializers compatible with the Django Rest Framework[20] to provide JSON responses used in the Dura-Europos Stories application.

The frontend of Stories Services provides a consistent design for content syndicated from Wikidata.

[18] https://redis.io/
[19] https://github.com/kennethsn/django-wikidata-api
[20] https://www.django-rest-framework.org/

We use the frontend to render and manage multimedia stories. For story rendering, we developed react-stories-api[21], a React.js open source component library. This library controls all visual story elements in the Dura-Europos Stories application. This allows us to decouple infrastructure from presentation so that the Dura-Europos Digital Archive team can host the stories on any domain and data is provided through the Stories Services API. We developed the react-stories-api library to minimize maintenance for people who host their own stories application as a statically-served single-page application (SPA). The library contains rendering components as well as the API client itself. The react-stories-api also provides a component for presenting all stories within a collection as well as search functionality and pagination support. The Dura-Europos Stories website leverages both components to create a full user experience including routing and navigation with no backend server or data store needed. We use Material Design[22] as our design framework and most of the components in the library are built with the Material-UI core library[23].

The Stories Services team also developed a Publisher Workspace to serve as the visual frontend of API operations. The Publisher Workspace can be used to manage the collection and story presentation. The Publisher Workspace can be used by organizations or people who which to customize story structure and styling. While the Stories Services API layer powers the data in the Dura-Europos website, the Publisher Workspace is where admin users can rearrange the ordering of moments, modify the story metadata itself, and most importantly, enhance the stories with curated content such as images, videos, and links found outside of Wikidata. The Publisher Workspace is built using React.js and has react-stories-api as a core dependency. This design allows publishers real-time previews of their story selections.

10. The Case for Knowledge Graphs in Research Projects

The data powering this collection of multimedia stories will grow in complexity and depth over time as members of the Wikidata community add additional information to Wikidata related to these artifacts. Each time a story is rendered, the Stories application takes in new information from Wikidata that matches the data model of a story. In this way, if relevant data is added to Wikidata, it will be included in a story the next time that story is viewed.

The multilingual design of Wikidata offers the potential for access points into this collection via more than three hundred human languages. Creating this project within Wikidata means that, as others learn of the project and decide to contribute additional data related to Dura-Europos, they will be able to interact with Wikidata in any of the languages Wikidata supports. The Wikidata platform enables collaboration between people even if they do not share any human languages in common.

Another type of connection that the Wikidata knowledge base stores is information about external data sources. Wikidata does this through the use of external identifiers. External identifiers are a type of Wikidata property. External identifiers are used to store identifiers for items in information systems, databases, or collections outside of Wikidata. While there are not many external identifiers on the items for the artifacts, there are many external identifiers on the items depicted by the artifacts as well as some of the places. Wikidata serves as a hub for external identifiers on the web [21]. Due to the large number of identifiers stored in Wikidata, it has become a efficient place to find many identifiers for a resource with a single search. Thus by contributing data to Wikidata, a research team will likely gain additional sources of related information. This saves time for research teams, as they no longer need to seek out these other sources, or map their data to each additional source individually.

For research projects that contain multilingual data, or those that attract users from diverse language communities, multilingual knowledge graphs offer a structure that enables the creation of multilingual applications. Many of the artifacts in the Dura-Europos collection contain representations of people, objects, places, or symbols. By using the 'depicts' property to connect artifacts and the items for the entities represented, we open up pathways to additional information about what is represented. For example, multiple gods and goddesses are depicted among the artifacts at Dura-Europos. Wikidata contains many external identifiers for them. These external identifiers are the entry point into other databases and systems that contain additional content about the gods and goddesses. The Wikidata item for Aphrodite has external identifiers pointing out to forty-five different databases or systems. Us-

[21] https://github.com/kennethsn/react-stories-api
[22] https://material.io/
[23] https://material-ui.com/

ing these external identifiers we can quickly locate additional information about Aphrodite from the Oxford Classical Dictionary (P9106), the Getty Iconography Authority File (P5986), the Consortium of European Research Libraries Thesaurus (P1871), and forty-two other systems[24].

Connections between this set of artifact items and other types of items in the knowledge graph provides additional information that supplements the original data set. For example, people who were involved in the Dura-Europos excavation are not described in the set of metadata about the artifacts. The names of these people and information about their involvement are described in archival material and in published documents. As these documents are added to Wikidata, it will be possible to create additional connections between some of the artifacts and the items which represent people. We expect that these connections will be created in the knowledge graph in the future, an example of how the data will grow even after the time of contribution to Wikidata.

Wikidata editors have already connected a number of publications to items related to Dura-Europos. These connections are asserted through the use of 'main subject' (P921) as seen in Figure 14. Connecting the items for scholarly publications to the items that they describe helps people reusing data from Wikidata find groups of publications that are related to topical areas. We anticipate that members of the Wikidata community will add more publications related to Dura-Europos to the knowledge base. We hope that surfacing these publications in the library moments of the Dura-Europos Stories application will allow people interested in this area to find additional information, or to become inspired to add more publications to Wikidata.

11. Conclusion

Powering an application with data from a knowledge base such as Wikidata provides several advantages to research teams. Advantages include the fact that Wikimedians build tools for interacting with Wikidata, that Wikidata itself serves as a platform for collaboration, that it is possible for the data to grow over time, that there are more curators looking at and contributing to the data, that it has built-in support for hundreds of human languages, and that it provides connections to the hub of external identifiers on the web.

[24] https://www.wikidata.org/wiki/Q35500

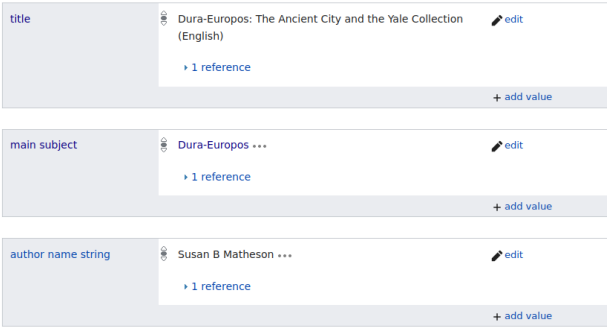

Fig. 14. Detail of an item for a publication that has Dura-Europos as a main subject.

Transforming the metadata for the artifacts related to the Dura-Europos excavation into a set of Wikidata items and statements enabled us to reuse the data within the Stories Services framework. Using Stories Services, we created a custom application for exploring the dataset. This process also enabled new connections between this dataset and other data on the web. The Stories Services framework was built with the Wikidata data model in mind. After mapping the Dura-Europos data to Wikidata and publishing it to the knowledge base, the Stories Services framework pulls this data into the Dura-Europos Stories application for browsing and display of the content. Similarly, the Dura-Europos community can also make use of a wide range of additional tools built by the Wikidata community for working with Wikidata. If the Dura-Europos team had selected a database for storing their data, not only would the data be siloed, they would not be able to leverage the tooling created by Wikimedians.

Wikidata itself serves as a platform for collaboration. Many organizations hold material related to Dura-Europos. If these organizations decide to describe their materials in Wikidata, Wikipedia or Wikimedia Commons, these platforms function as platforms for collaboration. Shared properties, categories, and links between items serve as the structures of organization that bring together different pieces of information. As more people contribute content related to Dura-Europos to Wikidata, the knowledge base itself serves as the infrastructure for collaboration, and this project will be able to reuse the content others contribute.

The Wikidata community of editors actively engage with the knowledge base every day [12]. If we consider the data in Wikidata that we present in the Dura-Europos Stories application to be the Dura-Europos

subset of Wikidata, this subset will grow over time. Contributors to Wikidata may create additional items for other artefacts excavated from Dura-Europos, or add publications related to the excavation or artefacts, or they may add labels for existing items in additional human languages. This model has an advantage over creating a project-specific database that people stop maintaining at the end of a project.

The multilingual design of Wikidata, with support for more than three hundred human languages, has led to the creation of millions of items with labels in many languages. Not only does this mean that more editors from diverse language backgrounds can collaborate and potentially extend data in the domain of your research area, but it also means that applications that reuse data from Wikidata have more multilingual data to show, potentially bringing your research to additional communities of users.

Mapping a new dataset to the Wikidata data model, and contributing the data to the knowledge base replaces the work of structuring the data for a relational database or a set of spreadsheets. We benefit from using Wikidata's graph of external identifiers to unlock pathways to additional information sources with no extra effort for the research team. Having associations between external resources and the original dataset can increase the number of questions the research team can ask about their domain. It can also make it easier to collaborate with a wider range of other researchers because they may be more comfortable thinking about the dataset in terms of the identifiers from a external resource with which they are already familiar.

We enable users of the Dura-Europos stories application to browse multimedia content related to the archaeological site and the excavated artifacts. The application is powered by data from the Wikidata knowledge base. We present images of the artifacts along with geospatial data about their excavation locations, and metadata about the artefacts. This information is presented in the context of information from the web of linked statements about Dura-Europos. The Dura-Europos Stories application provides an interactive interface for presenting this data that users can explore to learn about the site and the excavation, as well as the material culture represented through the artefacts.

12. Acknowledgments

The International Digital Dura-Europos Archive (IDEA) has been made possible in part by a major grant from the National Endowment for the Humanities. We would like to thank the Wikidata community for contributing to the knowledge base that anyone can edit.

References

[1] E. Hyvönen, P. Leskinen, M. Tamper, H. Rantala, E. Ikkala, J. Tuominen and K. Keravuori, BiographySampo–publishing and enriching biographies on the semantic web for digital humanities research, in: *European Semantic Web Conference*, Springer, 2019, pp. 574–589. doi:https://doi.org/10.1007/978-3-030-21348-$0_3$7.

[2] E. Hyvönen et al., "Sampo" Model and Semantic Portals for Digital Humanities on the Semantic Web, in: *Proceedings of the Digital Humanities in the Nordic Countries 5th Conference (DHN 2020)*, CEUR-WS. org, 2020. http://hdl.handle.net/10138/334001.

[3] A. Meroño-Peñuela, A. Ashkpour, M. Van Erp, K. Mandemakers, L. Breure, A. Scharnhorst, S. Schlobach and F. Van Harmelen, Semantic technologies for historical research: A survey, *Semantic Web* **6**(6) (2015), 539–564.

[4] G. Sugimoto, Building linked open date entities for historical research, in: *Metadata and Semantic Research: 14th International Conference, MTSR 2020, Madrid, Spain, December 2-4, 2020, Revised Selected Papers 14*, Springer, 2021, pp. 323–335.

[5] S.C. Schmidt, F. Thiery and M. Trognitz, Practices of Linked Open Data in Archaeology and Their Realisation in Wikidata **2**(3), 333–364. doi:10.3390/DIGITAL2030019.

[6] A. Učakar, A. Sterle, M. Vuga, T. Trček Pečak, D. Trček, J. Ahtik, K. Košak, D. Muck, H. Gabrijelčič Tomc and T.N. Kočevar, 3D Digital Preservation, Presentation, and Interpretation of Wooden Cultural Heritage on the Example of Sculptures of the FormaViva Kostanjevica Na Krki Collection, *Applied Sciences* **12**(17) (2022). doi:10.3390/app12178445. https://www.mdpi.com/2076-3417/12/17/8445.

[7] F. Zhao, A systematic review of Wikidata in Digital Humanities projects, *Digital Scholarship in the Humanities* (2022), fqac083. doi:10.1093/llc/fqac083.

[8] A. Siebold and M. Valleriani, Digital Perspectives in History, *Histories* **2**(2) (2022), 170–177. doi:10.3390/histories2020013. https://www.mdpi.com/2409-9252/2/2/13.

[9] K. Thornton and K. Seals-Nutt, Science Stories: Using IIIF and Wikidata to Create a Linked-Data Application., in: *International Semantic Web Conference (P&D/Industry/BlueSky)*, 2018. http://ceur-ws.org/Vol-2180/paper-68.pdf.

[10] M. Van Remoortel, J.M. Birkholz, P. De Potter, K. Thornton and K. Seals-Nutt, Linking Women Editors of Periodicals to the Wikidata Knowledge Graph (2021). doi:https://doi.org/10.3233/SW-222845.

[11] D. Vrandečić and M. Krötzsch, Wikidata: a free collaborative knowledgebase, *Communications of the ACM* **57**(10) (2014), 78–85. doi:http://dx.doi.org/10.1145/2629489.

[12] Active Editors, 2023. https://stats.wikimedia.org/#/wikidata.org/contributing/active-editors/normal|line|2-year|(page_type)~content*non-content|monthly.

[13] L.-A. Kaffee, A. Piscopo, P. Vougiouklis, E. Simperl, L. Carr and L. Pintscher, A Glimpse into Babel: An Analysis of Multilinguality in Wikidata, in: *Proceedings of the 13th International Symposium on Open Collaboration*, OpenSym '17, ACM, New York, NY, USA, 2017, pp. 14:1–14:5. ISBN 978-1-4503-5187-4. doi:10.1145/3125433.3125465.

[14] A. Piscopo, L.-A. Kaffee, C. Phethean and E. Simperl, Provenance information in a collaborative knowledge graph: an evaluation of Wikidata external references, in: *International semantic web conference*, Springer, 2017, pp. 542–558. doi:https://doi.org/10.1007/978-3-319-68288-4$_3$2.

[15] E.J. Kelly, Reuse of Wikimedia commons cultural heritage images on the wider web, *Evidence Based Library and Information Practice* **14**(3) (2019), 28–51. doi:https://doi.org/10.18438/eblip29575.

[16] E. Barker, R. Simon, L. Isaksen and P. de Soto Cañamares, The Pleiades gazetteer and the Pelagios project (2016). http://oro.open.ac.uk/48328/4/2016_Simon_Barker_etal_Gazetteers.pdf.

[17] K. Thornton, H. Solbrig, G.S. Stupp, J.E. Labra Gayo, D. Mietchen, E. Prud'hommeaux and A. Waagmeester, Using Shape Expressions (ShEx) to Share RDF Data Models and to Guide Curation with Rigorous Validation, in: *The Semantic Web*, Springer International Publishing, 2019, pp. C1–C1. doi:10.1007/978-3-030-21348-0_40.

[18] I. Boneva, J.E. Labra, S. Hym, E.G. Prud'hommeaux, H. Solbrig and S. Staworko, Validating RDF with Shape Expressions, *ArXiv e-prints* (2014). http://adsabs.harvard.edu/abs/2014arXiv1404.1270B.

[19] S. Heath, Digital Creation and Expression in the Context of Teaching Roman Art and Archaeology, *DATAM: Digital Approaches to Teaching the Ancient Mediterranean* (2020), 149–170. https://hcommons.org/deposits/download/hc:29192/CONTENT/datam_heath.pdf/.

[20] Y. Shen, Z. Wang, Q. Sun, A. Chen and H. Rushmeier, Reconstructing Dura-Europos From Sparse Photo Collections Using Deep Contour Extraction, in: *Eurographics Workshop on Graphics and Cultural Heritage*, V. Hulusic and A. Chalmers, eds, The Eurographics Association, 2021. ISSN 2312-6124. ISBN 978-3-03868-141-0. doi:10.2312/gch.20211408.

[21] J. Neubert, Wikidata as a Linking Hub for Knowledge Organization Systems? Integrating an Authority Mapping into Wikidata and Learning Lessons for KOS Mappings., in: *NKOS@ TPDL*, 2017, pp. 14–25. http://ceur-ws.org/Vol-1937/paper2.pdf.

[22] S. Van Hooland and R. Verborgh, *Linked Data for Libraries, Archives and Museums: How to clean, link and publish your metadata*, Facet publishing, 2014.

[23] S. Allison-Cassin, A. Armstrong, P. Ayers, T. Cramer, M. Custer, M. Lemus-Rojas, S. McCallum, M. Proffitt, M. Puente, J. Ruttenberg et al., ARL White Paper on Wikidata: Opportunities and Recommendations (2019). https://scholarworks.iupui.edu/bitstream/handle/1805/18902/ARL-white-paper-on-Wikidata.pdf.

[24] W3C, Semantic Web Activity, 2009. https://www.w3.org/2001/sw/.

[25] D. Abián, G. Candela, J. Birkholz, M. Dolores Saez, P. Escobar, S. Chambers, I. Martinez-Sempere and J. Vicente Berna-Martinez, Wikidata/Wikibase Workshops: Lessons Learned, in: *WikidataCon 2019*, 2019. http://hdl.handle.net/1854/LU-8633763.

[26] A. Waagmeester, G. Stupp, S. Burgstaller-Muehlbacher, B.M. Good, M. Griffith, O.L. Griffith, K. Hanspers, H. Hermjakob, T.S. Hudson, K. Hybiske et al., Science Forum: Wikidata as a knowledge graph for the life sciences, *ELife* **9** (2020), e52614.

[27] M. Klein, H. Gupta, V. Rai, P. Konieczny and H. Zhu, Monitoring the gender gap with wikidata human gender indicators, in: *Proceedings of the 12th International Symposium on Open Collaboration*, 2016, pp. 1–9. doi:https://dl.acm.org/doi/abs/10.1145/2957792.2957798.

[28] P. Konieczny and M. Klein, Gender gap through time and space: A journey through Wikipedia biographies via the Wikidata Human Gender Indicator, *New Media & Society* **20**(12) (2018), 4608–4633. doi:https://journals.sagepub.com/doi/abs/10.1177/1461444818779080.

[29] F. Erxleben, M. Günther, M. Krötzsch, J. Mendez and D. Vrandečić, Introducing Wikidata to the linked data web, in: *International semantic web conference*, Springer, 2014, pp. 50–65. https://doi.org/10.1007/978-3-319-11964-9_4.

[30] F.Å. Nielsen, Ordia: A Web Application for Wikidata Lexemes, in: *European Semantic Web Conference*, Springer, 2019, pp. 141–146. doi:https://doi.org/10.1007/978-3-030-32327-1$_2$8.

[31] T.E. Putman, S. Lelong, S. Burgstaller-Muehlbacher, A. Waagmeester, C. Diesh, N. Dunn, M. Munoz-Torres, G.S. Stupp, C. Wu, A.I. Su et al., WikiGenomes: an open web application for community consumption and curation of gene annotation data in Wikidata, *Database* **2017** (2017). doi:https://doi.org/10.1093/database/bax025.

[32] T. Putman, K. Hybiske, D. Jow, C. Afrasiabi, S. Lelong, M.A. Cano, G.S. Stupp, A. Waagmeester, B.M. Good, C. Wu et al., ChlamBase: a curated model organism database for the Chlamydia research community, *Database* **2019** (2019). doi:https://doi.org/10.1093/database/baz041.

[33] F.Å. Nielsen, D. Mietchen and E. Willighagen, Scholia, scientometrics and wikidata, in: *European Semantic Web Conference*, Springer, 2017, pp. 237–259. doi:https://doi.org/10.1007/978-3-319-70407-4$_3$6.

[34] A. Waagmeester, E.L. Willighagen, A.I. Su, M. Kutmon, J.E.L. Gayo, D. Fernández-Álvarez, Q. Groom, P.J. Schaap, L.M. Verhagen and J.J. Koehorst, A protocol for adding knowledge to Wikidata, a case report, *BioRxiv* (2020). doi:http://dx.doi.org/10.1186/s12915-020-00940-y.

[35] A. Heftberger, J. Höper, C. Müller-Birn and N.-O. Walkowski, Opening up Research Data in Film Studies by Using the Structured Knowledge Base Wikidata, in: *Digital Cultural Heritage*, Springer, 2020, pp. 401–410. doi:https://doi.org/10.1007/978-3-030-15200-0$_2$7.

[36] F.Å. Nielsen, Literature, Geolocation and Wikidata., in: *Wiki@ ICWSM*, 2016. doi:https://doi.org/10.1609/icwsm.v10i2.14833.

[37] M. Miquel-Ribé and D. Laniado, The Wikipedia Diversity Observatory: A Project to Identify and Bridge Content Gaps in Wikipedia, in: *Proceedings of the 16th International Symposium on Open Collaboration*, 2020, pp. 1–4. doi:https://doi.org/10.1145/3412569.3412866.