




## Article

# Practices of Linked Open Data in Archaeology and Their Realisation in Wikidata

Sophie C. Schmidt <sup>1,\*</sup> , Florian Thiery <sup>2</sup>  and Martina Trognitz <sup>3</sup> 

- <sup>1</sup> Institut für Prähistorische Archäologie and Berlin Graduate School of Ancient Studies, Freie Universität Berlin, 14195 Berlin, Germany
- <sup>2</sup> Römisch-Germanisches Zentralmuseum, Department of Scientific IT, Digital Platforms and Research Tools, 55116 Mainz, Germany; florian.thiery@rgzm.de
- <sup>3</sup> Austrian Centre for Digital Humanities and Cultural Heritage, Österreichische Akademie der Wissenschaften, 1010 Vienna, Austria; martina.trognitz@oeaw.ac.at
- \* Correspondence: s.c.schmidt@fu-berlin.de

**Abstract:** In this paper, we introduce Linked Open Data (LOD) in the archaeological domain as a means to connect dispersed data sources and enable cross-querying. The technology behind the design principles and how LOD can be created and published is described to enable less-familiar researchers to understand the presented benefits and drawbacks of LOD. Wikidata is introduced as an open knowledge hub for the creation and dissemination of LOD. Different actors within archaeology have implemented LOD, and we present which challenges have been and are being addressed. A selection of projects showcases how Wikidata is being used by archaeologists to enrich and open their databases to the general public. With this paper, we aim to encourage the creation and re-use of LOD in archaeology, as we believe it offers an improvement on current data publishing practices.

**Keywords:** linked open data; Wikidata; archaeology



**Citation:** Schmidt, S.C.; Thiery, F.; Trognitz, M. Practices of Linked Open Data in Archaeology and Their Realisation in Wikidata. *Digital* **2022**, *2*, 333–364. <https://doi.org/10.3390/digital2030019>

Academic Editors: Markos Katsianis, Tuna Kalayci and Apostolos Sarris

Received: 20 April 2022

Accepted: 17 June 2022

Published: 22 June 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Legacy data is ubiquitous in Archaeology because excavations are not reproducible, and therefore their original records are always relevant. It exists in analogue and digital form. The latter can be born-digital or digitised and comes in a variety of formats [1]. Surveys show that after completion of a project, data is often saved on local networks or servers provided by the employer, and rarely in professional repositories [2–4], which is also due to national legislation not always requiring deposition of data with professional repositories [5]. Data and databases that have been published online are often hard to find and/or inaccessible ([6] pp. 133–136). They are not always managed by IT professionals, complicating their integration into follow-up research. As the special issue of *Internet Archaeology* (58, 2021) [7] shows, the state of development of digital repositories is diverse, with some regions of the world having a well established infrastructure, while others are just beginning a transformation. Richards et al. conclude:

*“To be successful, archaeology needs not only better policies for data curation, but also the harmonising of the processes of data creation and its deposition for archiving.”* [1]

Here, we propose one building block towards the solution to these problems: Linked Open Data (LOD). LOD is part of a vision for the World Wide Web [8,9] that aims at connecting openly available datasets and facilitating machine analysis by applying a standardised technology stack. LOD enables the combination and joint processing of existing data from various digital sources to address archaeological research questions. Since LOD by its definition (cf. Section 2) is interoperable, we believe that archaeology data and metadata adhering to LOD standards can effectively be combined and integrated into follow-up research and increase the findability and interoperability of published data on the Web.

With this paper, we would like to encourage the creation and use of LOD by offering an accessible entry point to the use of LOD and a bridge between Digital Humanities practitioners and less technically proficient researchers. To this end, the use and creation of LOD within recent projects that re-use and add data to Wikidata [10] is described. Wikidata acts as a linked data hub and readily provides extensive multidisciplinary LOD and respective user-friendly tools. The presented projects demonstrate the potential of LOD together with Wikidata and exemplify common challenges in turning legacy data into LOD. They cover a variety of sub-disciplines, dealing with Early Neolithic European ceramics, Roman terra sigillata, Irish Early Medieval inscribed stones, and bibliographic data about Aegean Bronze Age seals and sealings. We argue that the Wikidata ecosystem greatly simplifies the provision of highly connected LOD, regardless of its provenance and legacy status.

The remainder of this contribution is structured as follows: The next section provides an introduction to LOD and related concepts, general pointers to LOD creation and publication, and a more in detail description of Wikidata and the creation of LOD with Wikidata. The section concludes with an overview of the benefits and challenges of LOD. Next, we describe in brief the adoption of LOD in archaeology by building upon already published summaries [11,12] before presenting and discussing five Wikidata projects. The inline links given in the text have been accessed last 02 June 2022.

## 2. Understanding Linked Open Data (LOD)

The overall aim of LOD is to provide openly available online data sets that are inter-linked and allow for cross-querying to, e.g., address archaeological research questions. To fully understand what LOD is, we will briefly present some of its core building blocks: Semantic Web, Linked Data (LD), and Open Data. A more in-depth coverage of the topic from a humanities perspective is given in, e.g., [13,14].

Before moving on, it has to be understood that LOD is not a data model, but a design principle for providing data with a standardised technology stack in its background. Data modelling (semantic and conceptual) ideally precedes exposing data as LOD. Comparable alternatives to LOD include publishing data in structured and standardised formats or allowing access via programming interfaces (Web APIs), which often lack comprehensibility of the data and their underlying schema, accessibility as well as share-ability, and hinder data integration [15]. Other alternative developments include Microformats and Dataspaces ([16] pp. 13–15). The former adds structured data directly in HTML but only has a restricted vocabulary and leaves integrated data without identifiers ([16] p. 13). The latter is a small-scale LOD approach for databases that might eventually contribute to LOD ([16] p. 15).

The Semantic Web is a term introduced by Tim Berners-Lee in 1998 ([8] p. 157) [17,18], that refers to a World Wide Web not only accessible to humans but also processable by machines, i.e., a web with machine-readable information. For information to become machine-readable, it needs to be annotated (also: marked up or tagged) with metadata. This metadata adds the semantic meaning to data and links, which makes them processable by computers without human intervention ([19] p. 89) [20]. For this purpose, the World Wide Web Consortium (W3C) provides standardised technologies [21], e.g. the Resource Description Framework (RDF) [22] as a data model for the representation of data and the SPARQL Protocol and RDF Query Language (SPARQL) [23] as a language to query data represented in RDF.

Let us take the Wikipedia article about the Phaistos Disc [24] as an example (Figure 1), which renders an informative text with various sections, tables, and images to the human user. While a human can easily read where the disc was found, a machine cannot do this because the underlying facts and information such as the statement *‘The Phaistos Disc was found at the site of Phaistos.’* are buried in sentences like *“The disc was discovered in 1908 by the Italian archaeologist Luigi Pernier in the Minoan palace-site of Phaistos, [...]”* [24]. In the



1. Resources (or things) being described should be named with URIs, a universal form of URLs.
2. The URIs used, should be HTTP URIs, to enable looking up the URI names.
3. Useful information should be provided using standards, such as RDF\* (i.e., the entire family of RDF standards [26]) and SPARQL.
4. Links to other URIs should be included.

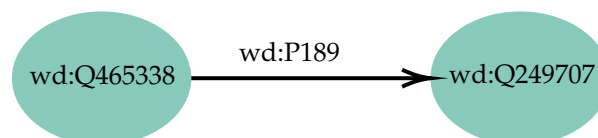
The last rule is what allows humans and machines alike to explore the ‘Giant Global Graph’ also called the ‘Linked Open Data Cloud’ [27] or ‘Web of Data’, which are other terms used for the Semantic Web (also known as ‘Web 3.0’) or ‘Linked Data’.

The application of the first two of the four LD rules to Example 1 leads to the graph in Figure 3. The individual parts of the statement are now represented with HTTP URIs coming from Wikidata. The subject (‘The Phaistos Disc’) is indicated with <https://www.wikidata.org/entity/Q465338>, the object (‘the site of Phaistos’) with <https://www.wikidata.org/entity/Q249707>, and the predicate (‘was found at’) with <https://www.wikidata.org/entity/P189>.



**Figure 3.** The first two rules of Linked Data applied to the statement ‘The Phaistos Disc was found at the site of Phaistos.’ with HTTP URIs coming from Wikidata.

Before applying LD rules three and four, we introduce the concept of namespaces [28], which are used to shorten the URI notation in Figure 3 by replacing the URI part ‘<https://www.wikidata.org/entity/>’ with a short prefix like ‘*wd:*’. This allows for a simpler visualisation of Example 1 as shown in Figure 4.



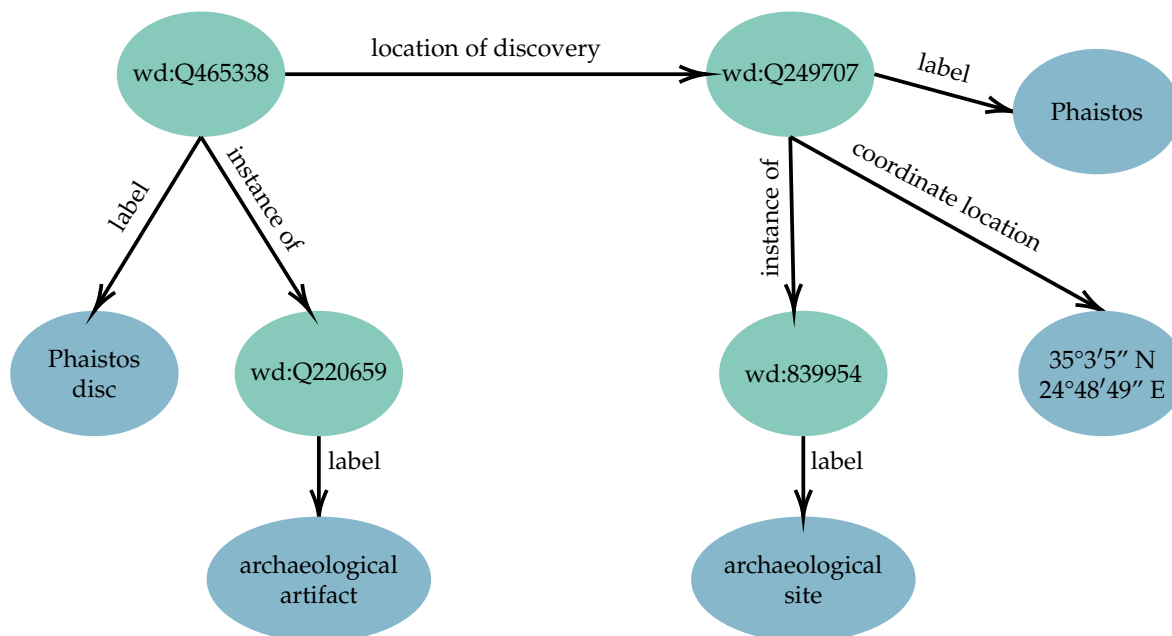
**Figure 4.** The first two rules of Linked Data applied to the statement ‘The Phaistos Disc was found at the site of Phaistos.’ with the namespace ‘*wd:*’ in use to replace the URI part ‘<https://www.wikidata.org/entity/>’.

LD rule number 3 requires the application of standards, such as RDF, and the provision of useful information [9]. While our example is already represented in RDF, the very little information given by the statement is not very useful yet. This can be improved by adding properties to further characterise our resources, that is the subject, the object and also the predicate from Example 1. We could, e.g., add a label and a type. Depending on the type, we could add a property indicating the material or the geographic coordinates. Which properties can be used for distinct types of resources and how they can be related is formalised with a set of rules called an ontology [29]. When needed, a list of allowed terms for the description of resources can be defined in a controlled vocabulary [30]. For both ontologies and controlled vocabularies, widely used standards exist, including dedicated ones for cultural heritage and archaeology mentioned in Section 3.

Adding further information to the statement of Example 1 results in the larger graph in Figure 5. The initial triple is now embedded in a network of triples and we can see that the resource representing our initial subject ‘*wd:Q465338*’ has the label ‘Phaistos Disc’ and is an instance of the class ‘*wd:Q220659*’. The latter represents another resource, which itself is further described with properties, such as the label ‘archaeological artifact’. This means our subject ‘Phaistos Disc’ is classified as an ‘archaeological artifact’. The object in



an RDF triple can either be another resource or a plain literal. As can be seen for the case of *'wd:P189'* (location of discovery) in Figure 4, properties are also resources that can be further described with, e.g., their labels used in Figure 5.

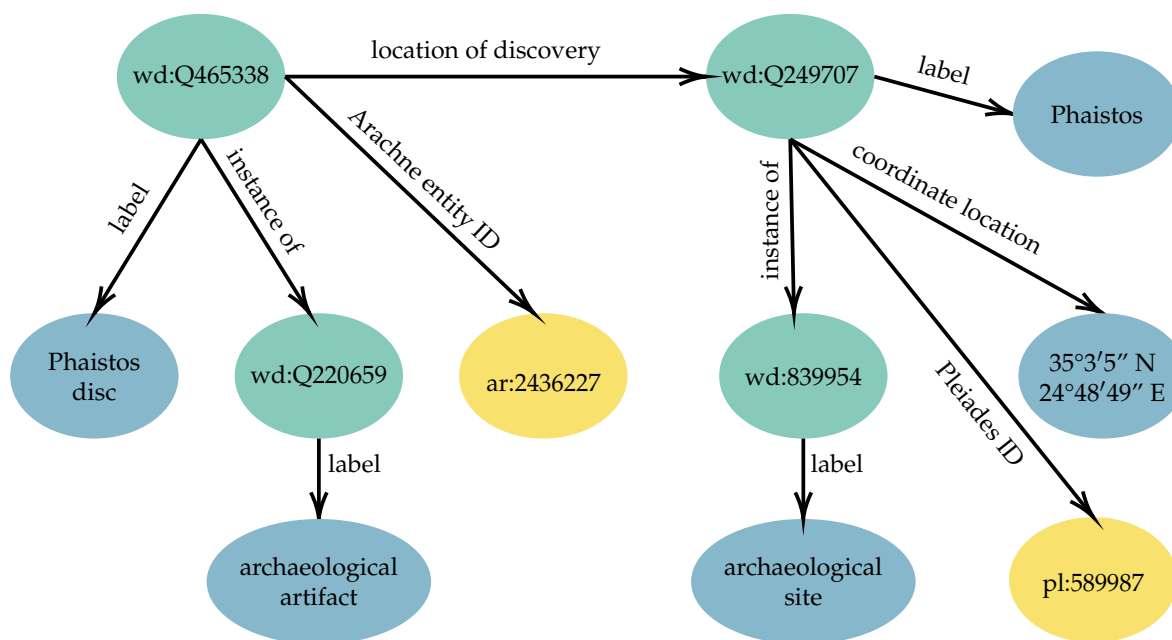


**Figure 5.** An RDF graph with the first three rules for Linked Data applied. The graph represents the statement *'The Phaistos Disc was found at the site of Phaistos.'* with additional information for the resources *'wd:Q465338'* and *'wd:Q249707'* (green). Additional information can either be another linked resource (green) or a plain literal (blue). For legibility, the property labels were directly included in the graph.

To complete our LD example, we shall now apply the fourth rule, which calls for links to other URIs. While links are already present, they all point to internal resources, i.e., all resources linked to are stored in the same namespace (*wd*), representing Wikidata. However, as Berners-Lee states, links to external sources are *"necessary to connect the data we have into a web, a serious, unbounded web in which one can find all[] kinds of things"* [9].

A record for the Phaistos Disc can also be found in the object database iDAI.objects arachne (Arachne) [31]. The URI for the record is <https://arachne.dainst.org/entity/2436227> and with the namespace *'ar:'* replacing *'https://arachne.dainst.org/entity/'* we get a link to the resource *'ar:2436227'* in Figure 6. Similarly, we can proceed with an external link for the place Phaistos, which has the URI <https://pleiades.stoa.org/places/589987> in Pleiades [32,33] and is now represented as *'pl:589987'* in Figure 6. With these additional external links, records from different data bases are now connected and can be analysed jointly.

Finally, we come to what constitutes Linked Open Data (LOD). The prerequisites include LD, as explained above, and the concept of Open Data. The Open Definition [34] defines *'open'* as: *"anyone is free to access, use, modify, and share"* the data, limited only by measures to preserve provenance and openness. This definition implies that data as a whole is provided openly and with free access in an open format that can be processed with at least one free/libre/open-source software tool (e.g., CSV [35]). Furthermore, the definition stresses that data should be machine-readable, i.e., it *"must be provided in a form readily processable by a computer and where the individual elements of the work can be easily accessed and modified"*, and—if the data is not already in the public domain—it must be provided under an open license, such as the Creative Commons licenses [36] CC BY 4.0 or CC0.



**Figure 6.** An RDF graph with all four rules for Linked Data applied. The graph represents the statement ‘The Phaistos Disc was found at the site of Phaistos’. with additional information for the resources ‘wd:Q465338’ and ‘wd:Q249707’ (green) in the form of connected resources (green) or plain literals (blue). Links to external resources are included as well (yellow and namespace other than ‘wd’). For legibility, the property labels were directly included in the graph.

In short, LD becomes LOD when it is free for anybody to access, use, modify, and share it. Differently put, LOD is LD released under an open license. LOD that complies with these requirements is also referred to as five star LOD [9].

### 2.1. Creating and Publishing LOD

This section provides general pointers on how to create and publish LOD before describing in more detail what Wikidata is and how LOD publication can be achieved with it. Dedicated practical guides from a humanities perspective with recommendations for specific tools include [13,37,38].

The W3C [39] has identified ten steps for the creation and provision of LOD, which can be summarised as having four activity areas: planning, choosing and curating the dataset, data modelling, and publishing and maintaining LOD. Planning for LOD includes defining the overall aim of the project, assessing the overall needs and required expertise, as well as identifying suitable tools and resources to be used ([38] p. 5). When choosing a dataset, licensing, re-use potential, and quality of data should be considered. The latter can be improved by curating, cleaning up, and refining the dataset ([38] pp. 6–7).

Data modelling is a process consisting of determining the external resources to link to, choosing and applying a model for knowledge representation, and converting the data to the chosen schema as well as to a suitable linked data format ([38] pp. 7–15). Modelling for LOD also involves defining how data represented in one model, such as a local spreadsheet or a relational database model, are going to be transformed into the chosen knowledge representation model ([39] Step #3). It is best practice to use a well-known and standard knowledge representation model (i.e., an ontology) ([38] p. 11) ([39] Step #6) for LOD instead of creating a new bespoke ontology, as the latter can lead to data silos and limit the re-usability of LOD [40].

Finally, the prepared dataset can be served to the public in various ways ([37] Thing 8) ([38] p. 15), e.g., as a simple downloadable file (a so-called ‘data dump’) or a full-fledged RDF database (a so-called ‘triplestore’, a database system for graph-data represented in RDF) with a dedicated SPARQL endpoint (e.g., a website for receiving and

processing SPARQL queries) for querying the data. Offering an RDF triplestore means that the 'self-hosted' URLs used as identifiers in the triples have to be provided in a persistent way.

After publication maintenance is required, which not only includes maintenance of the used hard- and software framework but also of the data itself. Data maintenance includes tasks like taking care of broken links and applying corrections and updates to the dataset itself.

#### LOD with Wikidata

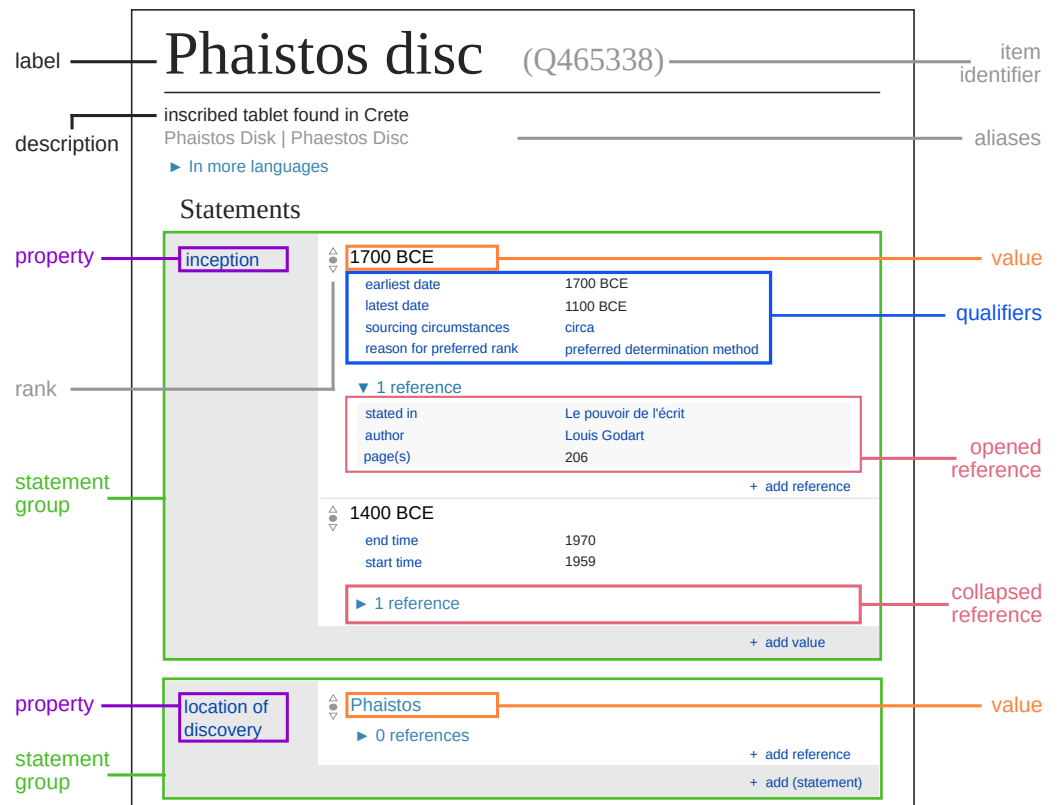
Providing LOD requires significant effort when done with a custom framework and database. Given the time and cost constraints many humanities and especially archaeology projects face, using ready-made solutions can immensely help in providing standardised LOD without much additional effort. One of these solutions is the use of Wikidata, which is already used for science and research [41].

The open knowledge base Wikidata [10] was established in 2012. It acts as a central storage for structured data of other Wikimedia projects including Wikipedia, Wikimedia Commons, or Wiktionary. Wikidata is a secondary database, which does not only record statements but also their sources, as well as connections to other databases. Data within Wikidata can be edited and added by anybody, is multilingual and available under the free licence CC0. The data is accessible to humans and machines alike and can be exported in standardised formats like e.g. JSON [42], RDF [22], or CSV [35]. Researchers can immensely benefit from using Wikidata because of the wide range of disciplines, topics, and research areas already represented there. This facilitates the combination of different knowledge domains and allows for interdisciplinary queries on the data.

While Wikidata's data model represented in Figure 7 is not defined in terms of RDF [43], it is still close to Semantic Web triples. The main entities in Wikidata are items and properties [44]. Items can be uniquely identified by the Wikidata Q identifier (QID), which consists of a 'Q' followed by a number, e.g., 'Q465338' from Figure 7, which identifies the item representing the Phaistos Disc. Similarly, the properties are identified by a 'P' followed by a number, e.g., the property 'location of discovery' in Figure 7 is identified by P189. Items in Wikidata are further described with statements. Each statement consists of a property and a value which can be supplied with qualifiers and references to indicate where the statement comes from and narrow the scope of a statement, e.g., state when a population number was valid. Several properties in Wikidata are external identifiers that point to authority control files or external databases [45,46]. Examples include the properties P496 to add ORCID iDs for researchers [47], P356 to append a DOI to publications [48], or P1566 for adding GeoNames IDs [49] to places. The external identifiers help in integrating the data within Wikidata with other data sources, enabling Wikidata to act as a linking hub [50].

Wikidata is part of the LOD Cloud [51–53], meaning that all data included in Wikidata is already LOD and can be accessed as such [54]. Multiple options are available for accessing the data in an LOD format: querying via the Wikidata Query Service (further described in Section 4.2), downloading a full or partial data dump, or per item access via the item's URI [54].

Contributing data to Wikidata is simply achieved by just editing existing items or adding new ones. This is possible even without an account, which is a must for batch editing (cf. Section 4.2). Due to Wikidata's multidisciplinary and flexibility, no specific data model is prescribed for items. Nonetheless, a project that aims at adding data to Wikidata, should take into account the used set of statements for existing items of a specific class and follow relevant recommendations to ensure that the newly added data is consistent with existing data.



**Figure 7.** The data model in Wikidata with two statement groups and an opened reference for the item ‘Phaistos Disc’ with the identifier Q465338. Mtrognitz, CC0, via Wikimedia Commons.

2.2. Benefits and Drawbacks of LOD

LOD comes with a number of benefits, as well as some shortcomings which we will expand upon here.

Decentralised and distributed data storage is at the centre of LD and its core benefit lies in the linking of otherwise discrete data silos and the possibility to query across them, e.g., [15,38] ([55] p. 269) [56,57], regardless of the discipline [58]. This linking in particular enhances and accelerates research processes [59,60] ([61] p. 84) and aids in distributing and reducing costs, e.g., by enriching a database with little effort by referring to and re-using external data sources [15,38] ([59] p. 3). With LOD, links between free-text reports and associated data can be established, which facilitates the use of data-heavy publications such as site reports and increase their usefulness. This also enables tracking the process of interpretation and re-interpretation in different research stages [55] and to better represent complex relationships between different entities ([62] p. 5).

The required use of formal ontologies for LOD leads to self-describing and self-contained data because definitions for elements of an unknown ontology can be looked up via the provided dereferenceable URI ([16] pp. 3–4) [63] (see rules 2 and 3 in Berner-Lees 5 Star LOD, explained in Section 2.1). The use of ontologies also leads to disambiguation [63] and the formal logic of ontologies allows for computational reasoning and inferencing ([8] pp. 177–194), which enables information validation [64] or may even lead to drawing new conclusions [59,63].

LOD is represented in RDF format, which in turn represents a graph—something well suited for visualisation and exploration [57]. Graphs can easily be extended with new relations and nodes, which means that LOD is more open to further enrichment than other systems might be [15,38]. The distribution of LOD has the benefit of allowing for bespoke ontologies that fit specific needs [65]. The underlying technologies of LOD enable this by allowing a single ontology to contain links (‘mappings’) to other ontologies. Furthermore, one graph can contain elements from multiple differing ontologies.



It has been argued that digitisation without reflection and subsequent classification of data may lead to the digital resources being cemented as self-evident in later archaeological discourse and that this effect might be intensified by the use of LOD ([62] p. 5). We believe, though, that the distribution and flexibility of LOD are capable of capturing the academic discourse with varying scholarly opinions and attributions, as exemplified by the projects presented in Sections 4.1.4 and 4.1.5.

Applying LOD is considered to enhance the discoverability of data and foster community engagement [59], but tracking actual re-use of LOD is difficult ([38] p. 3), which is why no reliable data on this topic within archaeology is available [5]. Nonetheless, as an archaeological example, the numismatic project Nomisma.org [66] has been mentioned as a success story in this regard ([61] p. 46) ([62] p. 8).

The features that make up the advantages of LOD are also behind some of the disadvantages. Due to data in Archaeology often being MEAN (Miscellaneous, Exceptional, Arbitrary, Nonconformist) [67] disadvantages are often exacerbated. Nonetheless, most drawbacks can be mitigated with adequate measures we present here.

The decentralised and distributed nature of LD leads to issues with data quality because external data is under the control of autonomous data providers with differing intentions. Thus, linked data and its metadata might be inconsistent, inaccurate, incomplete, or out of date ([61] pp. 78–79) [68]. Linking to trusted sources can lessen quality issues. For LD to be sustainable, persistent URIs are vital. Applications have failed, because a data provider changed their infrastructure, leading to thousands of dead links [69] and abandoned SPARQL endpoints ([61] p. 37) [70]. Sustainable LD infrastructures are required for the provision of stable LOD resources, which implies that short-term financed LD/LOD projects will need institutional partners with a stable infrastructure or should use well-established open alternatives, such as Wikidata (cf. Section 2.1).

Another challenge related to the distribution of LOD is that the same concept has several URIs originating from different sources, e.g., the place ‘Phaistos’ can be identified with the Wikidata URI <https://www.wikidata.org/entity/Q249707>, the GeoNames URI <https://sws.geonames.org/262531/> or the Pleiades URI <https://pleiades.stoa.org/places/589987>. Data hubs, such as Wikidata, that collect multiple identifiers for their records, can help in pulling the resources together and ultimately bring LOD a step forward by connecting more sources. In addition, different sources often use different standards and ontologies to represent their resources, which makes it difficult to map resources to each other and leads to a lack of interoperability. These problems are further exacerbated when the provenance of the data has not been properly declared and vocabularies fail to adequately label and classify their containing concepts. In order to interconnect such sources and resources, a cross-disciplinary semantic alignment can be implemented, which is, e.g., an endeavour of TRAIL 4.2 of the NFDI4Objects research data infrastructure [71]. Wikidata enables the interlinking of dispersed resources with dedicated properties for external sources, thereby acting as a linking hub.

It can be challenging to find suitable ontologies or an adequate level of resolution for metadata to link differing concepts and understandings of terminology ([62] p. 6) without losing information. This holds especially true when adhering to the best practice recommendation of re-using existing ontologies (cf. Section 2.1). If a new ontology has to be created, additional effort is required to ensure interoperability by mapping it to standard ontologies.

LOD served via a triplestore can lead to performance issues when dealing with a large number of triples ([61] pp. 9–10) and this can only be mitigated with the development of other solutions or by removing a subset of the triples [72].

Creating and using LOD often requires technical skills, which not every archaeologist has or will have, as learning the technical concepts and needed programming languages is a high investment that does not yet lead to academic recognition. This is applicable to most Digital Humanities skills within the traditional humanities, see e.g., ([73] p. 358) and

has to be addressed from two sides: on one side, easy-to-use tools should be provided, on another, scholars need to be trained in basic digital technologies.

When it comes to the creation and publication of LOD, the most significant investment is that of human labour ([38] p. 3), since individual steps require careful planning and might have to be re-iterated. Furthermore, the maintenance of published LOD is a continuous effort in order to adhere to the latest standards ([38] p. 4). Analysing the cost-benefit of creating LOD is difficult, as there are various factors to calculate for costs (development and mapping to ontologies, output of RDF, hosting of triplestores, etc.) as well as benefits (see above), which may be valued differently by individual stakeholders. The development of new tools, which facilitate and simplify the workflows can be expected to reduce the costs of providing LOD ([61] pp. 56-63) and we describe some of those in Section 4.2.

### 3. LOD in Archaeology

Drawing from different data sets, mixing and merging information from diverse sources, deriving own ideas and conclusions as well as forming new arguments based on already published material is a core practice of archaeological and historical research. Data was and often still is published in form of printed catalogues, figures, and tables. Instead of copying and publishing this kind of data anew, it is standard practice to reference the published table or figure. This practice can be expressed as an LD triple:

#### Example 2.

|                    |                      |   |                     |
|--------------------|----------------------|---|---------------------|
| <i>(The table)</i> | – was published in – | > | <i>(literature)</i> |
| <i>subject</i>     | <i>predicate</i>     |   | <i>object</i>       |

It is, though, only a small part of what has always been linked. Citations of articles and monographs are links to the information given in them. Concordance lists, bibliographies, and index volumes of journals, where entries are sorted by keywords are other solutions to a problem researchers have always been faced with: finding information as the very first step in any scholarly undertaking.

With the growth of the Open Access movement, the awareness of the need for finding aids such as standardised and consistent tagging systems and indexes increases. In contrast, the ideas of Open Data and LOD are still little known in archaeology, though they can be seen as nothing but a further consequent development of good citation practices for a digital and interconnected research environment. LOD allows to entrust computers with the tedious task of information tracking and enables computational reasoning approaches ([8] pp. 177–194) [63].

LOD was first adopted by Cultural Heritage institutions, libraries, archives, and museums, before being integrated in archaeological research processes ([61] pp. 42–43). The uptake and use of LOD in archaeology have been extensively described by Leif Isaksen in 2011 [11]. He pointed out that contributions to Semantic Web technologies at the international Computer Applications and Quantitative Methods in Archaeology (CAA) conferences increased since the early 2000s ([11] p. 41 Table 2.5). An update on LOD in archaeology including a vision for its future has been provided by Guntram Geser in 2016 [12] and the Ariadne LOD SIG in 2017 [61]. This section builds on existing reviews and contributes towards completing the picture until the present day.

The first step towards LOD in archaeology was done by publishing data and metadata on online platforms, of which some have already been developed in the 1990s. Examples include the German Arachne [31], born in 1995 with a focus on providing object data, now part of the iDAI.world [74], or the Archaeology Data Service (ADS) [75] founded in 1996 in the UK for the long-term preservation of digital archaeological data [76].

In the early 2000s, a need for dedicated discipline-specific standards to combine different data collections was identified, e.g., ([11] pp. 39–40) ([12] p. 7) [77–79]. Dedicated ontologies, e.g., [80–82] as well as vocabularies, e.g., [78] were developed to close this gap.

Relevant ontologies include Lightweight Information Describing Objects (LIDO) [83], Cultural Heritage Abstract Reference Model (CHARM) and the CIDOC Conceptual Reference Model (CIDOC CRM) (ISO 21127:2014) [84]. The latter is widely used in archaeology ([12] pp. 78–79) [62]. Vocabularies relevant for archaeology include the Getty Art & Architecture Thesaurus (Getty AAT) [85] and the Forum on Information Standards in Heritage (FISH Vocabularies) [86] that include more general concepts, as well as the community-built gazetteer for ancient places Pleiades [32,33], and the public gazetteer for historical, art-historical, and archaeological periods PeriodO [87,88].

The now completed STAR (Semantic Technologies for Archaeological Resources, 2010) and STELLAR (Semantic Technologies Enhancing Links and Linked data for Archaeological Resources, 2010–2011) projects of the ADS are forerunners for the uptake of Linked Data in archaeology. Projects like these fuelled the emergence of interest groups, such as the Linked Ancient World Data Initiative (LAWDI) or the Special Interest Group for Semantics in Archaeology at CAA, which was active from 2010 until 2017. The group re-formed in 2020 as Semantics and LOUD in Archaeology (SIG Data-Dragon). Further groups aiming for increasing the visibility and usage of LOD in archaeology exist: the Pelagios Network established in 2011, has partnered with many initiatives and communities, such as Pleiades [32] or Nomisma.org [66,89], to develop linking between different resources based on the common references to places [90]. The Linked Pasts Network (and symposium) emerged from Pelagios in 2015 and contributes to the spread of LOD in the wider humanities.

In the last decade, the Semantic Web has still been a popular topic at international CAA conferences (Table 1), with e.g., contributions in the domain of numismatics [91,92], ceramics [57,93,94], or periods [95]. Notably, the popularity of LD and LOD has significantly increased since 2018, which might be explained by the introduction of the FAIR Data Principles in 2016 [96]. The catchy acronym is used to describe a set of principles for making data and its metadata Findable, Accessible, Interoperable, and Re-usable. While not directly connected to LOD, the rules behind these principles partially overlap with LOD principles. In 2018, Linked Art introduced the concept of Linked Open Usable Data (LOUD) [97], to further include usability for human users—with an emphasis on human-readability—and thus lowering barriers to using LOD.

**Table 1.** Number of papers presented at CAA on semantic technologies or linked data (2012–2021). The table is a continuation of [11] (p. 41 Table 2.5) (years 2001–2011).

| Year               | Number of Papers |
|--------------------|------------------|
| 2012 (proceedings) | 13               |
| 2013 (proceedings) | 6                |
| 2014 (proceedings) | 8                |
| 2015 (proceedings) | 11               |
| 2016 (proceedings) | 8                |
| 2017 (proceedings) | 2                |
| 2018 (conference)  | 24               |
| 2019 (conference)  | 30               |
| 2020 (cancelled)   |                  |
| 2021 (conference)  | 23               |

Important and reliable providers of LOD in archaeology are repositories dedicated to long-term preservation, such as the aforementioned ADS. A recent overview of such repositories on a national and regional level is provided in a special issue of Internet Archaeology (58, 2021) [7]. Only a few countries “have repositories with the required specialist knowledge” [1] and “most do not have persistently available data in interoperable formats” [1], which renders reliable linking impossible. The overview in [7] is skewed towards the ‘global West’ because it is an outcome of the European COST Action SEADDA

(CA 18128), which is a European network (including Turkey and Israel) with international partners from Argentina, Canada, the United States, and Japan, leaving an obvious gap with missing or underrepresented partners from Africa, Asia, and South America. Steps towards data integration are being made by aggregation platforms functioning as a single entry point for searching across multiple data providers. Examples include the European Europeana (since 2008) [98] and ARIADNE Portal (since 2013) [99] as well as the US-American Digital Index of North American Archaeology (DINAA) (since 2012) [100], which links i. a. to the [Paleoindian Database of the Americas \(PIDBA\)](#) and the [Digital Archaeological Record \(tDAR\)](#) [101]. These aggregators combine different subject-specific data sources via standardised Metadata.

Since the [Transatlantic Archaeology Gateway](#) project, which created links between ADS and tDAR [102], there have been, to our knowledge, no significant attempts to connect data on a worldwide level. Therefore the national scope and the lack of interoperable data (including LOD) preclude a true worldwide cloud of archaeological knowledge. Data integration is further complicated by the employment of different workflows [103] and heterogeneity of data ([61] p. 45).

The vision developed by Geser for LOD in archaeology ([12] p. 7), in which archaeological data is interlinked with e.g. biological or geoscientific information, has not been realised yet. A survey among repository managers by ARIADNEplus and SEADDA about data management policies and practices of digital archaeological repositories shows that LD is considered an important technology: The answer *“Use Linked Data to interlink own and other (meta)data”* to the question *“What would help the repository most for improving data access?”* was selected by 43% of the respondents [5].

To conclude, we can observe a long-standing tradition of linking information in archaeology, that is being transferred into the digital space since the early 2000s. Since the 2010s there has been an EU-wide political effort towards enhancing LOD and several networks, including ARIADNE and Europeana, got funded by the European Union. Although a number of national repositories are not interlinked there are a number of successful LOD implementations. Geser’s assertion from 2016, that *“To meet expectations such as automatic reasoning over a large web of archaeological data many more (consistent) conceptual mappings of databases to the CIDOC CRM would be necessary.”* ([12] p. 16), still holds true today, although one might argue that any other agreed-upon standard would work as well.

#### 4. Wikidata for Archaeology

A growing hub for researchers and volunteers interested in archaeology and LOD is Wikidata (cf. Section 2.1). Archaeological data has already been integrated and interlinked in Wikidata to a certain degree. Besides actual items, such as the period ‘Bronze Age’ (Q11761), the place ‘Olympia’ (Q38888), the object ‘Dipylon amphora’ (Q331805), the publication ‘Architecture and Consumption in the Terrace House 2 in Ephesos’ (Q102078120), or the person ‘Gertrude Bell’ (Q231360), more general items for classifying items exist. Examples include the discipline ‘archaeology’ (Q23498) or the Wikidata class ‘archaeological site’ (Q839954). Examples of Wikidata properties related to archaeology include the ‘Art & Architecture Thesaurus ID’ (P1014) to link concepts with the Getty AAT, the ‘Pleiades ID’ (P1584) to connect ancient places in Wikidata with Pleiades, or the ‘Nomisma ID’ (P2950), which is used to link coin-related items to Nomisma.org.

The driving force behind getting archaeology (and any other) data into Wikidata is the community. Teams can get organised via [WikiProjects](#) to bundle items, properties, tasks and activities concentrating on a specific topic. This section can only highlight a very small example set from the vast amount of archaeology-related data in Wikidata. A more complete overview is being collected in the [WikiProject Archaeology](#), which includes, among others, archaeological datasets and relevant queries.

#### 4.1. Practical Approaches to Realising LOD in Wikidata

This section presents six Wikidata-related projects pursued by the authors. They showcase how Wikidata can be used to enrich and provide digital archaeological data as LOD. First, the Linked Open Samian Ware project (Section 4.1.1) exemplifies a workflow for turning legacy data from printed collections into a relational database and into FAIR resources. Second, the Linked Open Ogham project (Section 4.1.2) showcases how to use Wikidata as an information hub to map and align different cataloguing systems of archaeological artefacts. Third, the establishment of an up-to-date bibliography on a certain topic within Wikidata is presented (Section 4.1.3). Fourth, how to reference a bibliographical entry as a source for a Wikidata statement is explained in Section 4.1.4. Fifth, the ARS3D project demonstrates how Wikidata can be used as a hub for scientific data – in this case iconographic items – referring to modelled concepts defined in external resources. At the end, a citizen science project exemplifies an easy workflow for collaborative data collection. All projects link to, contribute to, or use data from Wikidata in various ways and make use of several free Wikidata-related tools, which are presented in Section 4.2.

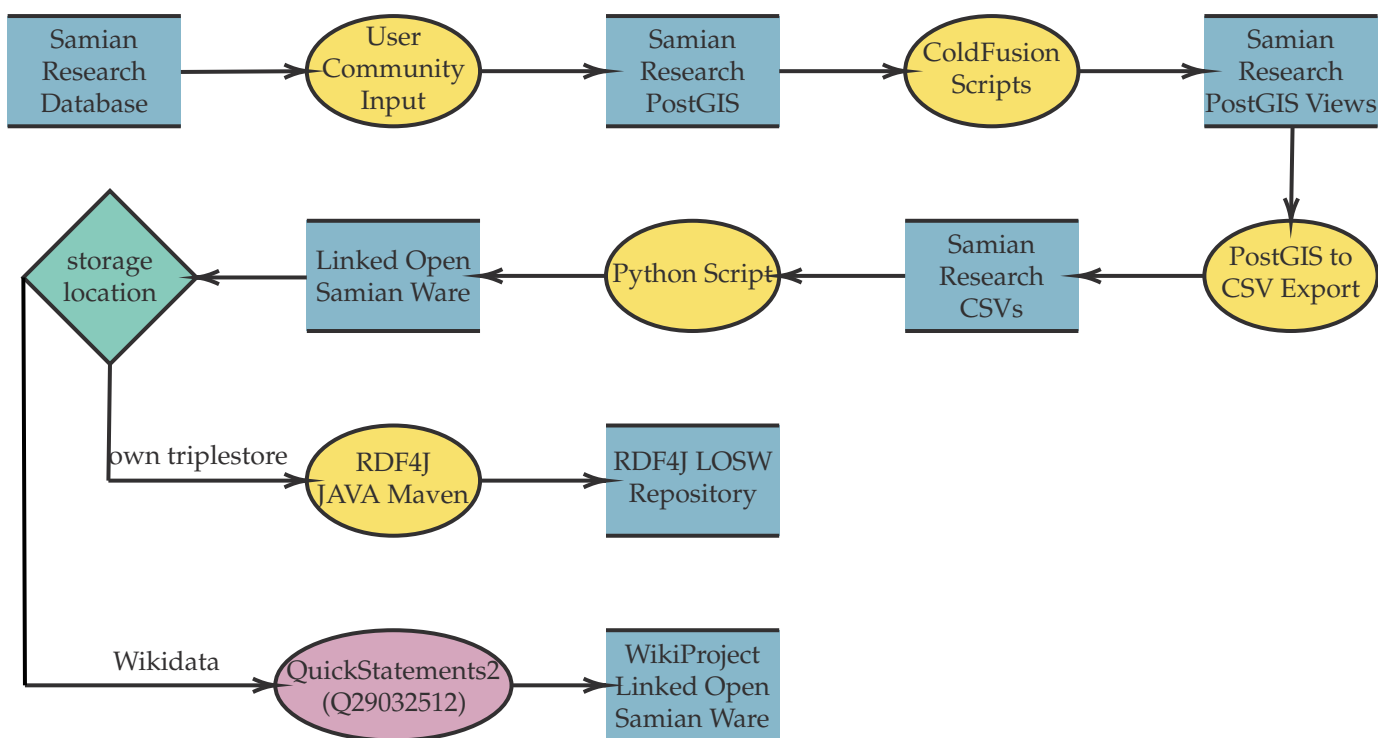
##### 4.1.1. A Workflow from Legacy Data to Wikidata: Linked Open Samian Ware

The Department of Scientific IT (WissIT) at the Römisch-Germanisches Zentralmuseum (RGZM) currently focuses on the transformation of Roman ceramics data into LOD to provide FAIR data [96] and to connect distributed databases. The project is also part of the NFDI4Objects initiative [104] and describes digital resources, e.g., about ceramic objects, with authority files, gazetteers, and vocabularies, which are linked to each other via community-driven vocabulary hubs, such as Wikidata, DANTE (Eng. data hub for authority files and terminologies) [71,105,106], or domain-specific data hubs, such as [archaeology.link](#). The RGZM WissIT works on the publication of expert data in Wikidata and on establishing bidirectional links between existing (re)sources. Geographic, iconographic and typologic items from expert databases can now be used by the research community and citizen scientists to create a common understanding of these entities. This way the dissemination of data can be maximised and new input integrated into the expert database.

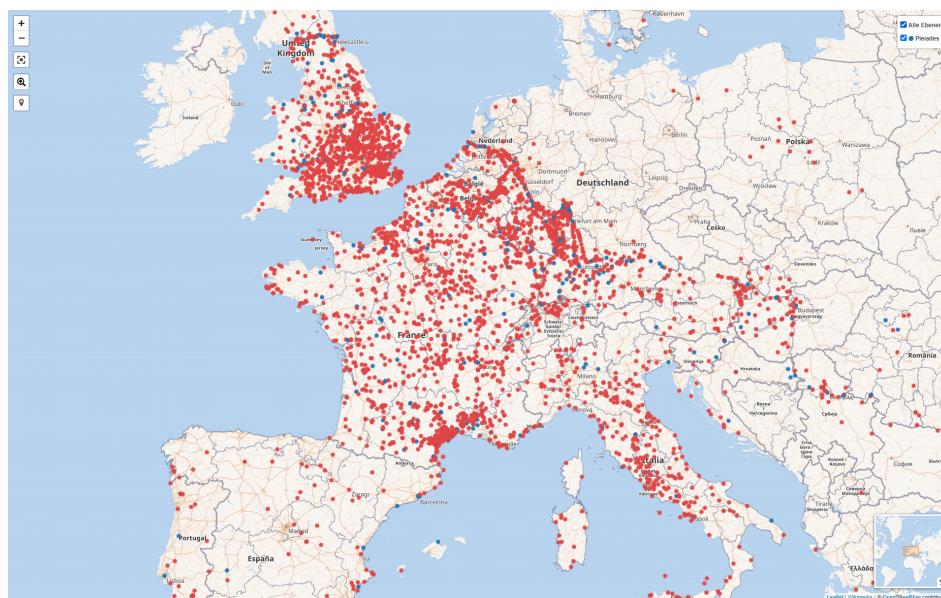
The project Linked Open Samian Ware (LOSW) is based on the online database Samian Research [107]. It relies on the printed corpus ‘Names on Terra Sigillata’ [108], which documents over 200.000 potters’ stamps and signatures. The Samian Research database was created as a statistical exploration tool for data interpretation and to provide up-to-date information including new material and revisions of existing information. The database comprises a quarter of a million potters’ stamps on terra sigillata from all over the Roman Empire. The data is actively curated by an international research community. A series of machine-readable interfaces and interoperable data formats such as a [RESTful API](#) and [RDF](#) were implemented by the RGZM WissIT [93,109,110]. The database contains (bidirectional) external links to, e.g., Pleiades, and Wikidata [111].

A reproducible workflow, presented in Figure 8, to transform Samian Ware data into LOD [97] and FAIR data was developed to enable researchers to re-use the data [112]. The Samian Research Database grows through the input by the user community and several steps are necessary in order to normalise this source data and add it to the LOD Cloud. Entries like potters’ stamps are curated in a web application and transferred into a PostgreSQL database, from which database views are created via ColdFusion/SQL scripts. Next, these database views are exported as CSV files, which are [transformed to RDF](#) with the help of [Python scripts](#) according to the [Samian Ontology](#). The Linked Open Samian Ware project then uses two separate sub-workflows, (i) one leading to a self hosted triplestore, (ii) the other one to the WikiProject in Wikidata. In the first sub-workflow (i) with the help of the Java [Maven](#) application, the RDF data [113] is imported into a RDF4J triplestore and links to the Linked Open Data Cloud, e.g., Pleiades and Wikidata are added. The second sub-workflow (ii) employs the QuickStatements [114] tool (see Section 4.2) to transform the CSV data to Wikidata entries.





**Figure 8.** Linked Pipe: Linked Open Samian Ware, as Data Flow Diagram using the Yourdon And/Or De Marco notation [115] and the Linked Pipes style ([116] slide 41), CC BY 4.0, Florian Thiery, Timo Homburg, Martina Trognitz [117].



**Figure 9.** Linked Open Samian Ware Discovery Sites (red dots) and discovery sites with references to Pleiades (blue dots) on Wikidata queried via <https://w.wiki/4pWG>, Wikidata Community, CC0 (Public Domain).

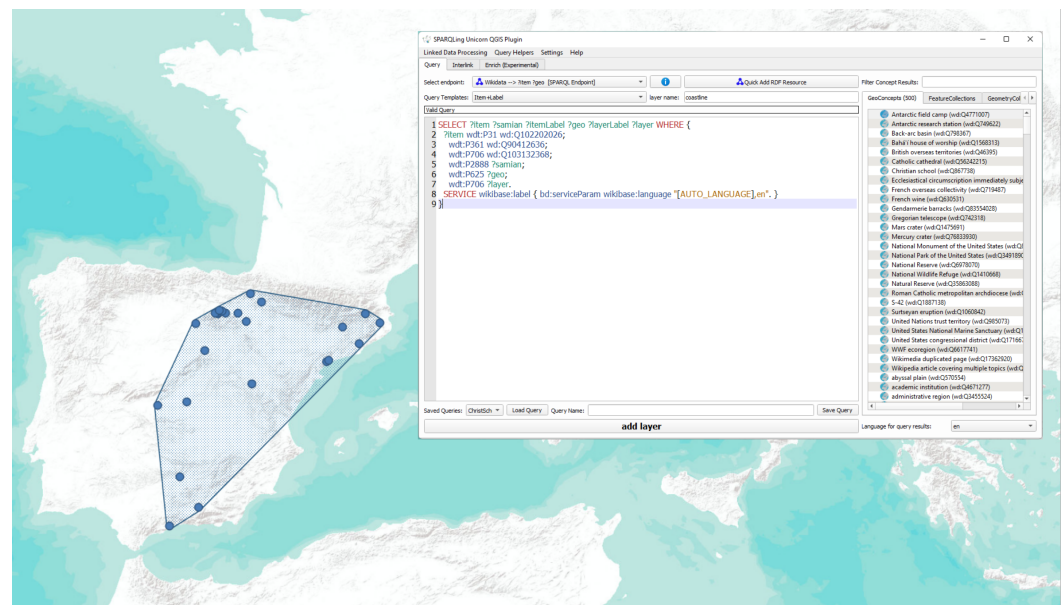
Within Wikidata, custom classes were prepared beforehand, such as the geospatial classes ‘Samian Ware Discovery Sites’ (Q102202066, currently 3886 items included, c.f. Figure 9), production centres (kiln sites) as ‘Samian Ware Production Centres’ (Q102202026, currently 103 items included), and kiln regions as ‘Samian Ware Kilnregion’ (Q102201947, currently 11 items included). Each of the geospatial resources is also categorised as an ‘archaeological site’ (Q839954).

For inclusion of individual records, forms have been prepared with the Cradle tool [118] to ensure a consistent set of statements for the respective Wikidata items. Further properties used to describe individual resources are exemplified in Table 2.

**Table 2.** Samian Ware in Wikidata; examples of discovery sites, production centres and kiln regions.

|                            | Discovery Site  | Production Centre  | Kiln Region  |
|----------------------------|---|--|--|
| <b>QID</b>                 | Rheinzabern (Q103191516)  | La Graufesenque (Q102763431)   | South Gaulish Samian Ware Kilnregion (Q102764958)  |
| <b>instance of (P31)</b>   | Samian Ware Discovery Site (Q102202066) & archaeological site (Q839954) | Samian Ware Production centre (Q102202026) & archaeological site (Q839954) & manufactory (Q380342) | Samian Ware Kilnregion (Q102201947) & archaeological site (Q839954) & Economic region of production (Q5333539) |
| <b>part of (P361)</b>      | Samian Research (Q90412636)   | Samian Research (Q90412636)  | Samian Research (Q90412636)  |
| <b>located in (P706)</b>   | n/a   | South Gaulish (Samian Ware Kilnregion) (Q102764958)  | n/a  |
| <b>coordinate (P625)</b>   | 49°7'5.016" N, 8°16'41.016" E   | 44°6'0.000" N, 3°4'59.999" E   | 44°4'44.666" N, 2°36'8.561" E  |
| <b>Geoshape (P3896)</b>    | n/a   | n/a  | SamianKilnregionSouthGaulish   |
| <b>exact match (P2888)</b> | lado:samian/loc_ds_1004152  | lado:samian/loc_pc_2000001   | lado:samian/loc_kr_131462  |
| <b>Pleiades ID (P1587)</b> | 109362  | n/a  | n/a  |
| <b>Cradle form</b>         | samian ware discovery site  | samian ware productioncentre   | samian ware kilnregion   |
| <b>SPARQL query</b>        | <a href="https://w.wiki/52Tu">https://w.wiki/52Tu</a>                   | <a href="https://w.wiki/4pKz">https://w.wiki/4pKz</a>  | <a href="https://w.wiki/4pL4">https://w.wiki/4pL4</a>  |

For further analyses easy-to-use Open Source tools have to be developed, an endeavour which, e.g., the Research Squirrel Engineers Network pursues [119,120]. One such tool is the ‘SPARQLing Unicorn QGIS Plugin’ [121] ([122] pp. 90–100) which helps to create SPARQL queries to populate QGIS vector layers. The Samian Ware data in Wikidata can be queried in QGIS using this plugin, e.g., Figure 10 shows Spanish production centres located in the ‘Spanish Samian Ware Kilnregion’ (Q103132368) loaded into QGIS.



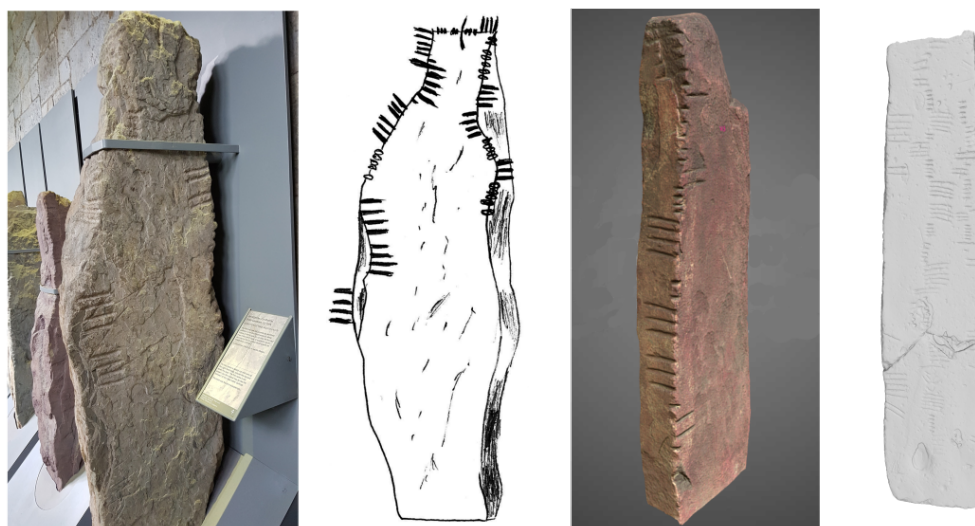
**Figure 10.** Spanish production centres located in the Spanish Samian Ware Kilnregion in QGIS, background ESRI Terrain. The data was imported with the ‘SPARQLing Unicorn QGIS Plugin’ using the queries <https://w.wiki/4pVW> and <https://w.wiki/4pVY>, Florian Thiery, CC BY 4.0.

This project shows a reproducible workflow of transforming expert databases into LOD resources for further dissemination and research opportunities.

#### 4.1.2. Linked Open Ogham: Connecting Catalogues Using Wikidata

The [Linked Open Ogham Data project](#) [123] was started in 2019 by the [Research Squirrel Engineers Network](#). It aims at providing and integrating legacy data from printed and digital catalogues about Ogham Stones in Wikidata, which is used as a community hub [65,122,123]. Furthermore, tools for exploring and enriching Ogham data are being created [124]. Here, we will showcase how the different concepts and descriptions pertaining to the same stone can be modelled and linked under a common denominator in Wikidata.

Ogham stones (e.g., Figure 11) are Early Medieval stones inscribed with the Ogham script created between the 6th and 9th centuries A.D. Linked Open Ogham Data primarily focuses on the digitisation and publication of the analogue Ogham stone catalogues ‘Corpus inscriptionum Insularum Celticarum’ (CIIC) by Macalister [125], ‘A Guide to Ogam’ by D. MacManus [126], and ‘An Archaeological Survey of South Kerry’ by O’Sullivan and Sheehan [127]. Information from existing online resources such as the Celtic Inscribed Stones Project (CISP) and the [Ogham in 3D Project](#) are being incorporated and interlinked as well.

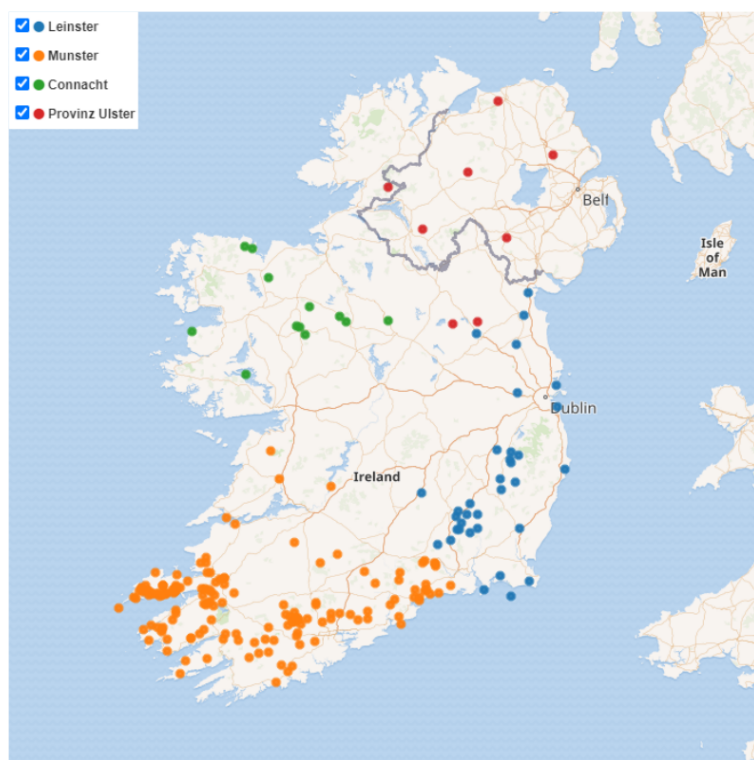


**Figure 11.** Left: Ogham Stone 4 (CIIC 81), Stone Corridor University College Cork (UCC); left-middle: drawing of CIIC 81, CC BY 4.0, Florian Thiery, via [Wikimedia Commons](#); right-middle: [Ogham Stone CIIC 215. Whitefield I, Co. Kerry](#) as 3D view created by the Ogham in 3D Project, CC BY-NC-SA 3.0 Ireland; right: [Ogham Stone CIIC 241. Kilbonane, Co. Kerry](#) as 3D model created by the Ogham in 3D Project, rendered using MeshLab, screenshot by Florian Thiery, CC BY-NC-SA 3.0 Ireland.

The respective project page on Wikidata is the WikiProject [Irish Ogham Stones](#). Dedicated Wikidata items created for this project include ‘Ogham Stone Concept’ (Q106602575), ‘Ogham Site’ (Q72617071), ‘Ogham Person Concept’ (Q110897921) and ‘Ogham Stones Cluster Region’ (Q110897622). As of June 2022, Wikidata contains 234 sites, 538 person concepts, 26 cluster regions, and 1238 stone concepts related to Ogham stones. Figure 12 depicts Ogham sites in Ireland.

Ogham Stones are modelled as ‘Ogham Stone Concepts’ (Q110897921) to distinguish the physical archeological artefacts from their descriptions in the analogue or digital sources. Currently, four stone concept types from four sources have a Wikidata item: ‘Ogham Stone Concept (RAS Macalister)’ (Q106602599), ‘Ogham Stone Concept (CISP Database)’ (Q106602627), ‘Ogham Stone Concept (Ogham in 3D Project)’ (Q106602633), and ‘Ogham Stone Concept (O’Sullivan & Sheehan)’ (Q111442203). The ‘Ogham Stone Concept (Research Squirrel Ogham Project)’ (Q106602643)—also known as ‘Squirrel Stone’—serves as an umbrella for Ogham stone concepts referenced in several sources. Another umbrella for Ogham stone concepts in regard to findspots is the geographical reference, the ‘Ogham

Site'. A dedicated tool to view Ogham stones grouped by stone concepts and sites was developed: the [Ogham Search Lookup Tool](#). How the different stone concepts from these sources are interlinked, can be best shown with an example like [OSS 908](#), which is an Ogham Stone Concept by O'Sullivan & Sheehan [127]. In Wikidata, OSS 908 ([Q111442223](#)) is linked to the Ogham Stone Concept by the Research Squirrel Ogham Project 218 via the property 'partially coincident with' ([P1382](#)). The latter then links to related stone concepts from CIIC, CISP, and the Ogham in 3D project.



**Figure 12.** Ogham sites in Ireland, colored by the four Irish Provinces via <https://w.wiki/4pjf>, CC0 (Public Domain).

In this project, Wikidata has been used to link different resources about Ogham stones, facilitating future research on this topic. By differentiating between Ogham stone concepts as digitally described representations of the stones and the actual stones themselves, it is possible to create umbrella IDs in Wikidata that bring together differing descriptions.

#### 4.1.3. Bibliography of Bronze Age Aegean Seals

In Section 3 we described that – although done in an analogue way – linking of information and sources in archaeology has always been part of the research process and output. An elementary tool in this regard are printed bibliographies with thematically and geographically indexed references. One problem inherent to printed bibliographies is that they cannot be easily updated in a timely manner.

The project '[A Linked and Open Bibliography for Aegean Glyptic in the Bronze Age](#)' aims at closing this gap in research about Minoan and Mycenaean Bronze Age seals by including the reference information recorded in '[A Bibliography for Aegean Glyptic in the Bronze Age](#)' [128] with all related keywords into Wikidata. The bibliography is part of the '[Corpus der Minoischen und Mykenischen Siegel](#)' (CMS) established in 1958. Younger's publication serves as a basis for an interactive and online bibliography that can be extended and corrected by anyone. Additional information will be included, such as links to the full text or to the digital representation of the mentioned seals in Arachne [31], as well as more



recent references that are dispersed across the [webpage of the CMS](#), the [Open Library for Aegean Archaeology](#), and [Nestor](#).

A [dedicated web application](#) is being developed to present all references in a user-friendly format and allow for export. Visualisations of bibliographic information from Wikidata are also provided with the tool [Scholia](#), which include co-occurrence of topics, author networks, or citation networks for a given topic, such as ‘[Aegean glyptic](#)’ ([Q58681669](#)).

The dataset being imported consists of more than 1200 references indexed with about 500 keywords, places and periods. Automated parsing (done with [AnyStyle](#)) and extensive manual checking are required because the references are not available in a structured bibliographic format, such as bib or RIS. The dataset import is documented on a [dedicated Dataset Imports page](#) on Wikidata. Since every piece of information, such as keywords, authors, publishers or even journals, has to be represented with its own item on Wikidata, the import has to be done in several rounds. [OpenRefine](#) is used for the preparation of the data, which is then batch-processed into Wikidata with [QuickStatements](#).

Future work in this endeavour will include adding information about the archaeological objects referred to and including cross citation information to allow for the investigation of citation networks. The former was already tested with the seal CMS XII 087 ([Q61293075](#)) which is now included as the main subject ([P921](#)) for the article ‘[Seals and Script II](#)’ ([Q61292379](#)). The latter will require significantly more work. A comparable undertaking has been done with the topic ‘[archaeological excavations in Ephesos](#)’ ([Q93429379](#)), which already provides visualisations for citation networks on [Scholia](#).

This project exemplifies how Wikidata can be used as an open and collaborative bibliometric research tool.

#### 4.1.4. Source Attribution of a Finds Database in Wikidata

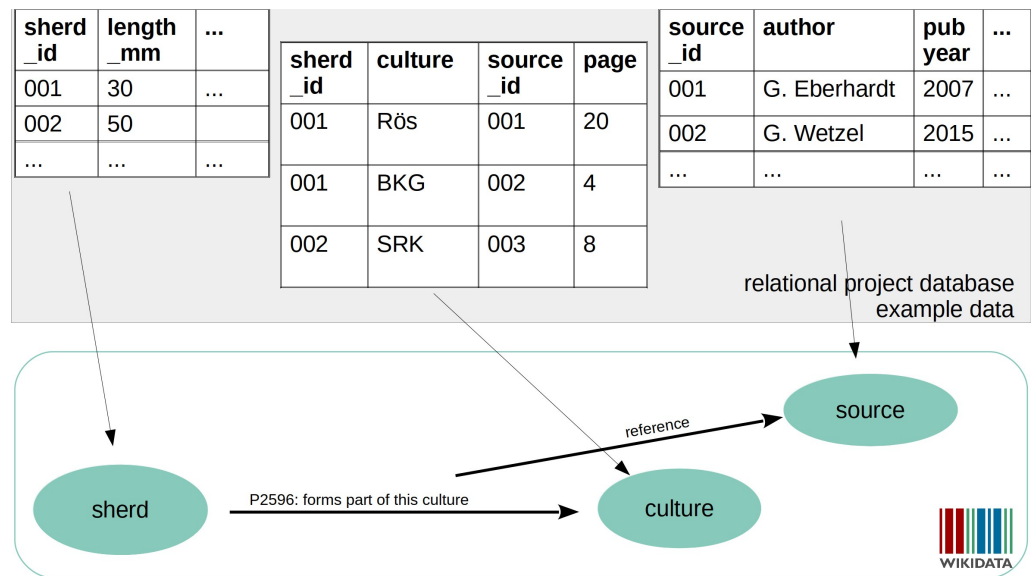
References included in Wikidata, e.g., those described in Section 4.1.3, can be (re-)used within Wikidata as sources for statements. The project presented here represents a practical application of this mechanism.

The framework of this undertaking is a project that aims to reconstruct the development of the post-linear band cultures (post-LBK) of the 5th millennium BC in Brandenburg (North-East Germany) by analysing ceramic finds. The finds have been (in some cases varyingly) attributed to the Stroke-Band Pottery and Rössen Culture, Guhrauer Group, and Brześć-Kujawski Culture. It is intended to import the database records into Wikidata at the end of the project and to link to external vocabularies thereby adding the data to the LOD Cloud.

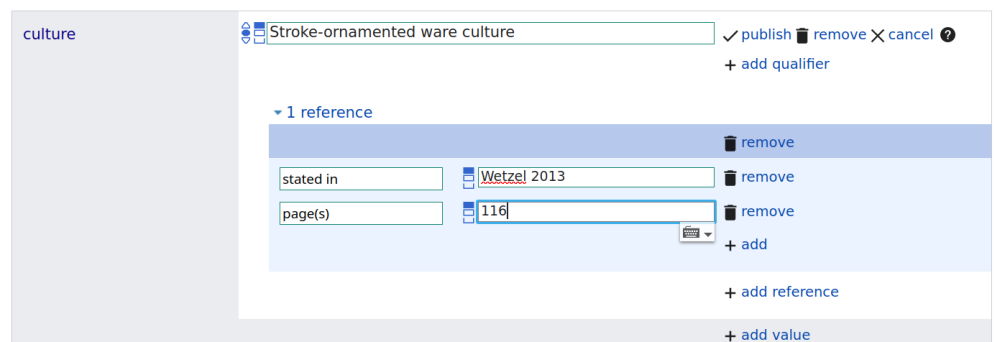
Sources for this project include survey results from professional and volunteer archaeologists, recent excavation documentation, archival information, and legacy information. The latter originates from finds made before 1926, where the original material was partially lost due to World War II. In order to adequately record the changes in cultural attributions of finds resulting from scholarly discourse, the database schema has to be able to properly reflect which scholar was responsible for which cultural attribution. The approach used in the database is similar to the use of references for Wikidata statements as described in Section 2.1 (Figure 13).

In the practical implementation, a relational database is developed with two schemas: One for input, which follows standards used for archaeological databases. The second schema, though, consists only of cross tables with three columns: subject, predicate, and object. These will be more easily transferred to a LOD format at the end of the project for publication in Wikidata. Upon import to Wikidata, the sources of the respective scholarly attributions will be included as references in the matching statements (Figure 14). For this to happen, the source (e.g., a journal article) first needs to be a Wikidata item, as described above in Section 4.1.3. External links are created by including [Open Streetmap IDs](#) and [GeoNames](#) for the nearest settlement of a findspot, and [PeriodO \[87\]](#) information for the dates given.





**Figure 13.** Example visualisation on how to transfer a relational database to the reference attribution in Wikidata.



**Figure 14.** Reference attribution in Wikidata: The statement is given a source, which needs to be an item in Wikidata. It is possible to add a page number. Screenshot 14th April 2022.

By creating this collection in Wikidata, it can act as an information source not just for researchers, but also for local historians and citizen scientists.

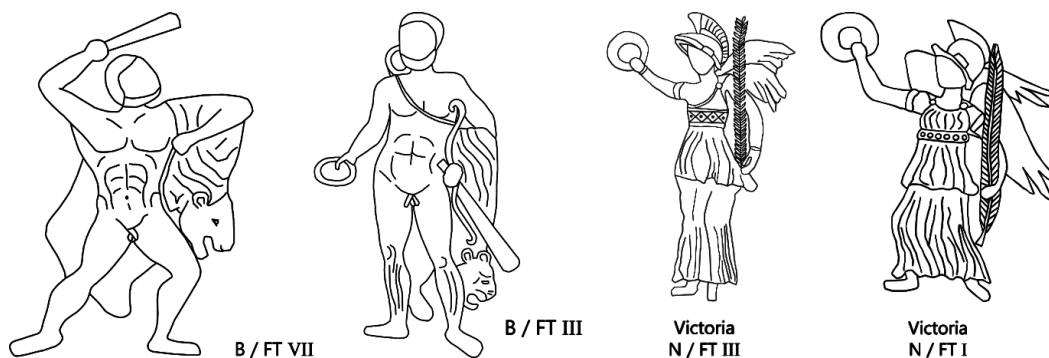
#### 4.1.5. Iconography in Wikidata: Linked Open African Red Slip Ware

African Red Slip Ware (ARS) is a category of terra sigillata produced in the province of Africa Proconsularis, mostly in the region of modern Tunisia. ARS dates from the 3rd to the 7th century AD and is characterised by the use of relief decorations, called appliqués. These appliqués depict a large variety of topics, e.g., mythological and biblical scenes, or depictions of circus games [129]. The African Red Slip Ware digital project (ARS3D) digitised and linked information in a cooperation between the RGZM and the i3mainz [130]. A total of 336 objects in different preservation conditions were processed within the framework of ARS3D, including various vessel types as well as models and stamps for manufacturing the objects [131].

LOD provided as RDF was used to realise the FAIR principles for this dataset [112]. Wikidata serves as a secondary publication and community exchange platform for ARS and iconographic items (WikiProject African Red Slip Ware Digital). This approach enables the research community to work collaboratively on the referencing and interlinking of several iconographic typologies in combination with openly licensed images, hosted on Wikimedia Commons. Furthermore, the project profits from already present items on Wikidata, such as the iconographic items ‘Hercules’ (Q240679) and ‘Victoria’ (Q308902),

which are described with several statements, including external links to, e.g., the [Iconclass classification system](#) [132].

Iconographic items, such as those of Hercules and Victoria described by Sophie zu Löwenstein ([133] pp. 457–489, 632–634) and depicted on the two bowls [O.39446](#) and [O.40718](#) are represented in Wikidata with the properties listed in Table 3. As the classification and description of a specific iconographic item depend on the scholar doing this work, multiple names for the same item might exist. This is, e.g., the case for Hercules ‘Löwenstein B / FT VII’ ([Q110892406](#)), which was also described by Meg. A. Armstrong [134] and is named ‘Armstrong 8.108’ ([Q110892542](#)). Although these two examples are included as individual items in Wikidata, they are related to each other with the property ‘said to be the same as’ ([P460](#)). Object O.40718 depicts ‘Löwenstein B / FT VII’ and ‘Löwenstein N / FT I (Victoria)’ ([133] pp. 460, 632–633), Object O.39446 ‘Löwenstein B / FT III’ and ‘Löwenstein N / FT III (Victoria)’ ([133] pp. 459, 634) shown in Figure 15.

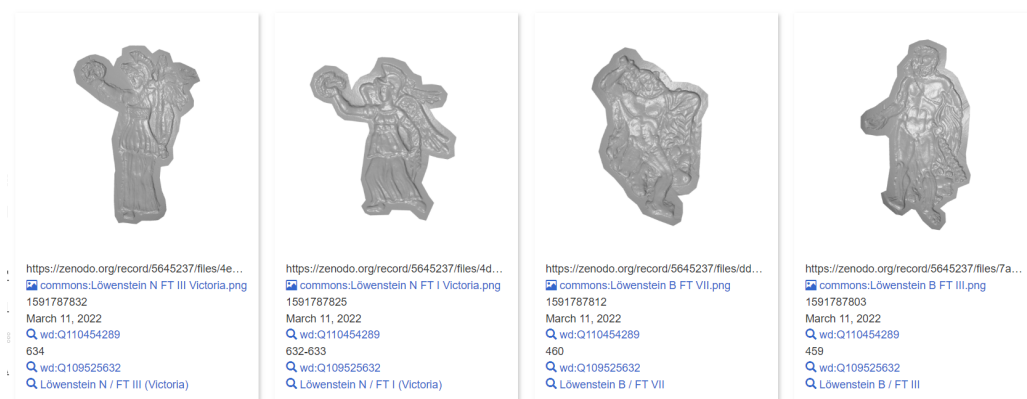


**Figure 15.** Examples for representations of Hercules (B) ([133] p. 702) and Victoria (N) ([133] p. 719) on Roman terra sigillata, reproduced with permission from Sophie zu Löwenstein, published in [133].

**Table 3.** African Red Slip Ware in Wikidata. Examples according to Sophie zu Löwenstein [133].

|                           | B / FT VII   | B / FT III   | N / FT III (Victoria)  | N / FT I (Victoria)  |
|---------------------------|--|--|--|--|
| <b>QID</b>                | <a href="#">Q110892406</a>   | <a href="#">Q110892402</a>   | <a href="#">Q110892417</a>   | <a href="#">Q110892415</a>   |
| <b>instance of (P31)</b>  | iconographic item ( <a href="#">Q109525730</a> ) & work of art ( <a href="#">Q838948</a> )   | iconographic item ( <a href="#">Q109525730</a> ) & work of art ( <a href="#">Q838948</a> )   | iconographic item ( <a href="#">Q109525730</a> ) & work of art ( <a href="#">Q838948</a> )   | iconographic item ( <a href="#">Q109525730</a> ) & work of art ( <a href="#">Q838948</a> )   |
| <b>part of (P361)</b>     | Löwenstein (2015) ( <a href="#">Q109525632</a> ) & ARS3D project ( <a href="#">Q105268778</a> )  | Löwenstein (2015) ( <a href="#">Q109525632</a> ) & ARS3D project ( <a href="#">Q105268778</a> )  | Löwenstein (2015) ( <a href="#">Q109525632</a> ) & ARS3D project ( <a href="#">Q105268778</a> )  | Löwenstein (2015) ( <a href="#">Q109525632</a> ) & ARS3D project ( <a href="#">Q105268778</a> )  |
| <b>has creator (P170)</b> | Sophie z. Löwenstein ( <a href="#">Q110454289</a> )  | Sophie z. Löwenstein ( <a href="#">Q110454289</a> )  | Sophie z. Löwenstein ( <a href="#">Q110454289</a> )  | Sophie z. Löwenstein ( <a href="#">Q110454289</a> )  |
| <b>collection (P195)</b>  | Mythologische Darstellungen auf Gebrauchsgegenständen der Spätantike. Die appliken- und reliefverzierte Sigillata C3/C4 ( <a href="#">Q109525632</a> ) [133] | Mythologische Darstellungen auf Gebrauchsgegenständen der Spätantike. Die appliken- und reliefverzierte Sigillata C3/C4 ( <a href="#">Q109525632</a> ) [133] | Mythologische Darstellungen auf Gebrauchsgegenständen der Spätantike. Die appliken- und reliefverzierte Sigillata C3/C4 ( <a href="#">Q109525632</a> ) [133] | Mythologische Darstellungen auf Gebrauchsgegenständen der Spätantike. Die appliken- und reliefverzierte Sigillata C3/C4 ( <a href="#">Q109525632</a> ) [133] |
| <b>depicts (P180)</b>     | Hercules ( <a href="#">Q240679</a> )   | Hercules ( <a href="#">Q240679</a> )   | Victoria ( <a href="#">Q308902</a> )   | Victoria ( <a href="#">Q308902</a> )   |
| <b>same as (P460)</b>     | Armstrong 8.108 [134] ( <a href="#">Q110892542</a> )   | Armstrong 8.109 [134] ( <a href="#">Q110892540</a> ), Atlante 135 [135] ( <a href="#">Q110892520</a> )   | Armstrong 8.100 [134] ( <a href="#">Q110892537</a> )   | Armstrong 8.101-103 [134] ( <a href="#">Q110892533</a> ; <a href="#">Q110892534</a> ; <a href="#">Q110892535</a> )   |
| <b>image (P18)</b>        | Löwenstein B FT VII.png  | Löwenstein B FT III.png  | Löwenstein N FT III Victoria.png   | Löwenstein N FT I Victoria.png   |

The ARS3D project also added iconographic elements to Wikidata which are not explicitly mentioned in the literature [136]. This additional iconography catalogue is modelled in Wikidata as the collection ‘African Red Slip Ware - Additional Iconography Catalogue’ (Q111370392) and, e.g., includes ships/boats as physical items ‘G / FT I’ (Q111370381), ‘G / FT II’ (Q111372289), ‘G / FT III’ (Q111372294) and hybrids like the human-horse hybrid, ‘M / FT I’ (Q111372221). The data, especially the iconographic items, can be visualised with the Wikidata Query Service. The Service provides various display options, including an image grid as shown in Figure 16.



**Figure 16.** Visualisation of iconographic items (Q109525730) from the ARS3D project (Q105268778) as an image grid in the Wikidata Query Service on 19/04/2022 queried by <https://w.wiki/54zr>. Wikidata Community, CC0 (Public Domain).

The data collected and published by this project can now be easily used by scholars to find and refine information on the modelled iconographic items, thereby facilitating further research.


#### 4.1.6. Wikidata as an Open Reference Collection for Citizen Science

The projects described above, focus on Wikidata as a resource for scholars. By its open and accessible nature, the platform also lends itself well to citizen science projects. Together with a group of volunteer field walkers, a small project was designed that aimed at creating an online reference collection for the volunteers in Wikidata. The project *Zerschlagenes Geschirr—Archäologische Quellen in Wikidata* (‘Smashed dishes—archaeological sources in Wikidata’) is still ongoing. It employs the Cradle tool [118] to create a [form to collect information about prehistoric ceramic](#) which enables an intuitive and standardised entry of data in Wikidata (Figure 17).

As this is a citizen science project and to keep the entry barrier low, the use of scientific jargon is avoided. Instead, e.g., the description of the decoration can be made with simple terms like ‘incision’, ‘point’, or ‘wave’ (see Table 4). Multiple terms can be added to the record of a sherd. Spatial information is given by providing the nearest settlement as a findspot, which will be interesting for field walkers in the region. Images of the finds are hosted by Wikimedia Commons. The information entered via the cradle tool can later be searched for by using the Wikidata Query Service [137].

A manual on how to contribute, as well as dedicated queries, will be provided for the users on [Wikidata:WikiProject Prähistorische Keramik](#) and distributed to the community. Queries could, e.g., yield all items where a certain decorative element such as ‘incision’ was used, or list all items from a specific region. This way inexperienced volunteer archaeologists can compare the sherd they found to a growing number of items in Wikidata and Wikimedia Commons. An example query is <https://w.wiki/3BWg>.

## prähistorische keramik

|   |  |     |
|---|--|-----|
| Labels *                                | de SBK-Scherbe Hohenbrück  | +   |
| Aliasse                                 | Sprache Text   | +   |
| Beschreibungen                          | Sprache Scherbe, evtl eines Tiergefäßes aus Hohenbrück   | +   |
| ist ein(e)                              | Wandscherbe <input type="text" value="Wandscherbe"/> <input type="text" value="Q106990172"/>                 | +   |
| Material *                              | Keramik  | +   |
| Bild                                    | SBK-Scherbe Hohenbrück.jpg  | +   |
| hat dekoratives Muster                  | Stempel <input type="text" value="Stempel"/> <input type="text" value="Q96093273"/>                          | -   |
|   | <input type="text" value=""/> <input type="text" value=""/>  | + - |
| Kultur                                  | <input type="text" value=""/> <input type="text" value=""/>  | +   |
| Entdecker oder Erfinder                 | <input type="text" value=""/> <input type="text" value=""/>  | +   |
| Fundort                                 | <input type="text" value=""/> <input type="text" value=""/>  | +   |
| Zeitpunkt der Entdeckung oder Erfindung | time <input type="text" value=""/>   | +   |

- Einschnitt
- Stempel
- Kerbe
- Dreieck
- Linie
- Punkt
- Viereck
- Kreis
- Streifen
- Wellenform
- Fingerkniff
- Wulst
- Knubbe

**Figure 17.** Cradle tool form for input of information regarding found sherds in Wikidata in German (project language). Screenshot 14th April 2022. For English translation of the fields see Table 4).

### 4.2. Tools for Wikidata

Wikidata has a vibrant community, which creates and maintains a diverse **ecosystem of tools** that aid in all possible aspects of working with Wikidata: input, editing, output, querying, and visualisation. At this point, we would like to give more information on the tools that were used in the projects described above in Section 4.1. For a selection of tools that aid in the LOD workflow in Wikidata, refer to [138].

**Cradle** [118,139] allows the creation of forms for data entry, which facilitate data input for beginners and help in standardising specific item types or classes. The tool aids in formalising a specific data model for individual item classes, which otherwise is not possible in Wikidata. With Cradle, mandatory and optional statements can be designed, as well as pre-defined value lists provided. The Cradle tool has already been used for archaeology related data, e.g., for **ancient ceramicists and vase painters**, **coin types**, or **historic buildings in Germany**. Dedicated Cradle forms were created for the projects presented in Sections 4.1.6 and 4.1.1.

**Table 4.** List of properties included in the form [prehistoric ceramics](#) prepared with the Cradle tool.

| Property                                    | Data Entry Options   | Example: <a href="#">Q111600125</a>   |
|---|--|---|
| <b>instance of (P31)</b>                    | <b>mandatory:</b> archaeological find ( <a href="#">Q10855061</a> ) and pottery ware ( <a href="#">Q17379525</a> )   | archaeological find ( <a href="#">Q10855061</a> ) and pottery ware ( <a href="#">Q17379525</a> )  |
| <b>made from material (P186)</b>            | <b>mandatory:</b> ceramic ( <a href="#">Q45621</a> )   | ceramic ( <a href="#">Q45621</a> )  |
| <b>instance of (P31)</b>                    | <b>drop-down-menu:</b> rim sherd ( <a href="#">Q106990428</a> ), wall sherd ( <a href="#">Q106990472</a> ), base sherd ( <a href="#">Q106990489</a> ), ...   | wall sherd ( <a href="#">Q106990472</a> )   |
| <b>image (P18)</b>                          | <b>free field</b> to link to Wikimedia file  | <a href="#">SBK-Scherbe_Hohenbrück.jpg</a>  |
| <b>has pattern (P5422)</b>                  | <b>drop-down-menu:</b> incising ( <a href="#">Q6014696</a> ), stamp ( <a href="#">Q96093273</a> ), triangle ( <a href="#">Q19821</a> ), line ( <a href="#">Q1228250</a> ), quadrilateral ( <a href="#">Q36810</a> ), ...                                 | stamp ( <a href="#">Q96093273</a> ) and quadrilateral ( <a href="#">Q36810</a> )  |
| <b>culture associated with item (P2596)</b> | <b>drop-down-menu:</b> Linear Pottery Culture ( <a href="#">Q806348</a> ), Stroke-ornamented Ware Culture ( <a href="#">Q1932196</a> ), Rössen Culture ( <a href="#">Q1886212</a> ), ...   | Stroke-ornamented Ware Culture ( <a href="#">Q1932196</a> )   |
| <b>discoverer or inventor (P61)</b>         | <b>free field</b> to link to Wikidata item   |   |
| <b>location of item at discovery (P189)</b> | <b>free field</b> to link to Wikidata item   | Hohenbrück ( <a href="#">Q1623655</a> )   |
| <b>time of discovery (P575)</b>             | <b>free field</b> to enter date  |   |
| <b>part of WikiProject (P5008)</b>          | <b>mandatory:</b> Prehistoric Ceramics ( <a href="#">Q107588426</a> )  | Prehistoric Ceramics ( <a href="#">Q107588426</a> )   |
| <b>stated in (P248)</b>                     | <b>drop-down-menu:</b> volunteer ( <a href="#">Q24716636</a> ), Heritage Management of Baden-Württemberg ( <a href="#">Q1541782</a> ), Heritage Management of Bavaria ( <a href="#">Q812412</a> ), ... all other German Heritage Management institutions | Heritage Management of Brandenburg ( <a href="#">Q897952</a> )  |
| <b>described by source (P1343)</b>          | <b>free field</b> to link to Wikidata item   |   |
| <b>described by URL (P973)</b>              | <b>free field</b> to enter URL   | <a href="https://brandenburgikon.net/index.php/de/sachlexikon/stichbandkeramik">https://brandenburgikon.net/index.php/de/sachlexikon/stichbandkeramik</a> |

The **Wikidata Query Service** [137,140,141] provides a graphical user interface for human users to query data in Wikidata from a SPARQL endpoint. Because of several built-in features, including extensive examples and the Wikidata Query Builder [142], special knowledge of the query language SPARQL is not required. The service provides a plethora of visualisation formats for the queries, which are also used in such tools as Scholia (described below). All projects detailed in Section 4.1 use the service for querying and analysing data included in Wikidata.

The **SPARQLing Unicorn QGIS Plugin** [121], ([122] pp. 90–100) ([Q74005133](#)) is a plugin for the free and Open Source Geographic Information System QGIS. It offers users an easy way to query and import geospatial as well as associated data from Wikidata and other triplestores into QGIS, which can be used for further data processing and analysis beyond the capabilities of Wikidata or the Wikidata Query Service. The plugin is developed and maintained by the Research Squirrel Engineers Network [143]. The projects described in Sections 4.1.1 and 4.1.2 make use of this plugin.

**QuickStatements** [114,144] is a tool for batch creating or editing Wikidata items, thus allowing for the import of large datasets to Wikidata without the time-consuming manual steps. It works by providing a simple set of text commands, which can also be created from within OpenRefine (see below). The projects described in Sections 4.1.1 and 4.1.3 make use of QuickStatements.

**OpenRefine** [145,146] is a versatile programme for working with structured data, such as tabular data. Its functionality is especially aimed at cleaning messy data, transforming it into other formats, and enriching data from external sources. Local data can be matched and enriched with Wikidata items using the Wikidata reconciliation service. Data in OpenRefine can easily be transformed into Wikidata statements by creating a dedicated export schema that maps each column to respective items and properties. Data prepared in this way can directly be processed with QuickStatements, thus allowing for editing and creating of



Wikidata items directly from within OpenRefine. The tool was extensively used for data preparation in the bibliographic project described in Section 4.1.3.

**Scholia** [147] (Q45340488) is a web service built on top of Wikidata to handle and visualise scientific bibliographic information contained within Wikidata for scientometrics without requiring any technical knowledge from the user. The data to be displayed is collected via pre-built SPARQL queries from the Wikidata Query Service and visualised with the output formats provided by the service. Visualisations include lists of publications, different charts to represent publications per year, graphs for co-author or citation networks, and images to augment a topic graph. Within Scholia, different kinds of information, called ‘aspects’ ([147] p. 5) can be examined, e.g., individual authors, works, organisations, or topics. It is also possible to, e.g., combine multiple topics for viewing, such as ‘archaeology’ (Q23498), ‘linked data’ (Q515701), and ‘linked open data’ (Q18692990), which renders a page that includes a [list of works on any combination of these topics](#). Scholia is a central tool used in the bibliographic project described in Section 4.1.3.

#### 4.3. Summarising the Role of Wikidata

The presented projects deal with turning (legacy) data into LOD by using Wikidata as an already established knowledge hub that is multilingual and interdisciplinary by design and aims at connecting different sources. While the projects seem completely different, there are commonalities in their workflows: developing a suitable schema or ontology for data representation both in dedicated databases and in Wikidata; cleaning up and enriching legacy data; reconciling data with Wikidata to avoid duplicates; preserving provenance information such as sources and references for cultural attributions or classifications; enriching the data with external links and further information. The added value lies in the (re-)useability of the enriched and interlinked datasets that are open for researchers and citizen scientists alike.

Though we agree in general with Franca et al., who state that “*individual researchers and research groups should may [sic!] not be thought of as a primary focus of Linked Data initiatives. Managers of digital archives for the research community and institutional repositories are much more relevant target groups*” ([61] p. 16), we show that by using Wikidata individual researchers are enabled to contribute to the LOD Cloud in archaeology.

Although using Wikidata to provide LOD is much easier than developing your own framework, it still poses a few challenges for less technically savvy users. Many tools, of which a selection is presented in Section 4.2, have been implemented for Wikidata to aid in dealing with various tasks.

In practice, two key features of Wikidata could pose obstacles. One is that Wikidata is by definition open to everyone, both for reading and editing. Secondly, all facts are available under the free licence CC0. The former might lead to (unintended) editing of items, meaning someone may enter false data, tamper with formerly correct information or delete a statement they consider irrelevant. This is not the experience of any of the authors—on the contrary: many Wikidata items have profited from enrichment with further information and external links. In some cases, sensitive information not suitable for open publication, e.g., exact in situ location, is either only entered with vague statements or not at all. In addition, not all kinds of information can be included in Wikidata due to the use of the CC0 licence. These especially include longer texts, such as abstracts, that might be under copyright.

## 5. Conclusions

Due to their heterogeneity, their scattered nature, and, in some cases, their incompleteness and lack of documentation, legacy data is usually more difficult to process and analyse than data collected by oneself. Nonetheless, these challenges have to be continuously addressed because of the core practice of archaeological research that builds on original documentation and former publications. We believe that if we do not use contemporary digital publishing media and structures for data nowadays, we continue to

create discrete legacy data that will just increase these challenges for future generations of researchers. Digitisation of legacy data is being undertaken in several institutions and projects, e.g., [148–150]. Digitisation poses its own challenges; most notably, loss of information due to financial or curatorial issues or deliberate decisions, as well as faulty metadata that can only be identified as such and corrected when access to the original source is possible. While online publication of data and results increases its accessibility, as, e.g., travelling to archives and warehouses in person is not a must anymore, we argue that this is not sufficient to meet current research methodologies and needs.

As described in Section 3, LOD represents the consequent further development of scientific citation and referencing practice in the digital decentralised space. LOD is what enables linking of data with synthesis and argument [151], as well as ultimately connecting different sources, e.g., [56] and disciplines, e.g., [58]. When LOD is properly implemented, computer-aided processing and inference become possible ([8] pp. 177–194) [63]. Nonetheless, we recognise that an adequate implementation of LOD principles is not without challenges (cf. Section 2.2), with additional costs of human labour for creation and the need of a sustainable infrastructure being the most pressing ones.

We understand that providing and using LOD via dedicated APIs requires technical skills not every archaeologist has or will have. This problem has to be addressed from two sides. On one side, human-readable (and -friendly) interfaces help lower the barriers to using LOD from, e.g., Wikidata in daily work. Tools like those presented in Section 4.2, which might even be included in familiar working environments, such as the SPARQLing Unicorn QGIS Plugin, have the same effect. On the other side, another important step towards lowering barriers related to technical know-how is described by Kansa and Kansa:

*“If one aim of preserving data is for others to be able to use it, we need to increase practitioners’ skills in accessing and using data. Broadening data literacy skills over the next decade will help us realize the full potential of archaeological data.”* ([152] p. 82)

Increasing data literacy skills requires awareness of LOD and its potential as well as sustainable infrastructures. An important building block in the direction of these requirements is an active user community that includes researchers as well as citizen scientists. Only with an active user group, such frameworks as Wikidata, Pelagios, and other LOD providing systems can become sustainable. This also applies to research software [153]. Regarding the infrastructure side, a commitment by national and international funding bodies is necessary. Although the problem has been widely recognised and is being addressed, most recently, e.g., the funding of a German national research data infrastructure (NFDI) by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) with the NFDI4Objects consortium dedicated to material remains of human history, some core infrastructure components such as data repositories still require action [7].

Ultimately, just as LOD links data from different sources, it should also bring together researchers from a wide range of disciplines to work together on interdisciplinary research questions and collaborate towards sustainable (LOD) infrastructures.

**Author Contributions:** Concept: S.C.S., M.T. and F.T.; Understanding Linked Open Data (LOD): M.T., S.C.S. and F.T.; LOD in Archeology: S.C.S., M.T. and F.T.; Wikidata: M.T., S.C.S. and F.T.; Early Neolithic Ceramic Sherds: S.C.S.; Roman Ceramics: F.T.; Medieval Ogham Script: F.T. and S.C.S.; Bibliography of Bronze Age Aegean Seals: M.T.; Tools for Wikidata: S.C.S., M.T. and F.T.; Conclusions: M.T., S.C.S. and F.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** Parts of the projects were funded by the Wikimedia Foundation Germany within the Open Science Fellows Program in the periods 2018/2019 (Section 4.1.3) and 2020/2021 (Sections 4.1.6 and 4.1.2). The Evangelische Studienwerk Villigst e.V. funds the project described in Section 4.1.4. ARS3D was funded by the Federal Ministry of Education and Research Germany (BMBF), Förderkennzeichen: BMBF-01UG1888AX, BMBF-01UG1888BX. Linked Open Samian Ware and the CeraTyOnt Ontology is supported and maintained by the department of scientific IT at the RGZM.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data described in this paper is available on [Wikidata](#). The respective projects and their datasets are included in the overview at [WikiProject Archaeology](#). Further data is available on Zenodo and GitHub for the Linked Open Samian Ware project (DOI: [10.5281/zenodo.4305708](#) and <https://github.com/RGZM/samian-lod>), the ARS3D project (DOI: [10.5281/zenodo.5722941](#) and <https://github.com/RGZM/ars-lod>), CeraTyOnt (DOI: [10.5281/zenodo.5767082](#) and <https://github.com/RGZM/ceramiptypologies-lod>), and the Ogham project (DOI: [10.5281/zenodo.4765603](#) and <https://github.com/ogi-ogham/ogham-datav1>).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

|           |  |
|-----------|--|
| ADS       | Archaeology Data Service   |
| ARS       | African Red Slip Ware  |
| ARS3D     | African Red Slip Ware digital project  |
| CAA       | Computer Applications and Quantitative Methods in Archaeology  |
| CC0       | Creative Commons Licence 0 (Public Domain)   |
| CC BY     | Creative Commons Licence Author Attribution  |
| CeraTyOnt | Ceramic Typologies Ontology  |
| CIDOC     | Comité International pour la Documentation<br>(Eng. International Committee for Documentation)             |
| CIDOC CRM | CIDOC Conceptual Reference Model   |
| CIIC      | Corpus Inscriptionum Insularum Celticarum  |
| CISP      | Celtic Inscribed Stones Project  |
| CMS       | Corpus der Minoischen und Mykenischen Siegel<br>(Eng. Corpus of Minoan and Mycenaean Seals)                |
| DANTE     | DAtendrehzscheibe für Normdaten und TERminologien<br>(Eng. data hub for authority files and terminologies) |
| DFG       | Deutsche Forschungsgemeinschaft (Eng. German Research Foundation)  |
| FAIR      | Findable, Accessible, Interoperable, Re-usable   |
| Getty AAT | Getty Art & Architecture Thesaurus   |
| HTTP URI  | Uniform Resource Identifier for an object on the World Wide Web  |
| LD        | Linked Data  |
| LIDO      | Lightweight Information Describing Objects   |
| LOD       | Linked Open Data   |
| LOUD      | Linked Open Usable Data  |
| NFDI      | Nationale Forschungsdateninfrastruktur<br>(Eng. national research data infrastructure)                     |
| QID       | Wikidata Q-Identifier  |
| RDF       | Reference Description Framework  |
| RGZM      | Römisch-Germanisches Zentralmuseum in Mainz  |
| SPARQL    | SPARQL Protocol And RDF Query Language   |
| URI       | Uniform Resource Identifier  |
| URL       | Uniform Resource Locator   |
| W3C       | World Wide Web Consortium  |

## References

- Richards, J.D.; Jakobsson, U.; Novák, D.; Štular, B.; Wright, H. Digital Archiving in Archaeology: The State of the Art. Introduction. *Internet Archaeol.* **2021**, *58*. [[CrossRef](#)]
- Bauer, B.; Ferus, A.; Gorraiz, J.; Gründhammer, V.; Gumpfenberger, C.; Maly, N.; Mühlegger, J.M.; Preza, J.L.; Solís, B.S.; Schmidt, N.; et al. Forschende und Ihre Daten. Ergebnisse Einer Österreichweiten Befragung (PDF Full Report DE): Report 2015. Available online: <https://phaidra.univie.ac.at/o:407513> (accessed on 1 June 2022).
- Heinrich, M.; Sieverling, A.; Schäfer, F.; Jahn, S.; Altertumswissenschaften, I.F.F.A. *Stakeholderanalyse zu Forschungsdaten in den Altertumswissenschaften*; IANUS—FDZ Archäologie & Altertumswissenschaften: 2015. Available online: <https://www.ianus-fdz.de/projects/ap3-community/wiki/Stakeholderanalyse> (accessed on 1 June 2022). [[CrossRef](#)]

4. Schmidt, S.C.; Backhaus, H.; Keller, C.; Rokohl, L.; Thierry, F. Preliminary Report on the NFDI4Objects Survey. 2020. Available online: <https://osf.io/zcexm/> (accessed on 27 January 2022).
5. Geser, G.; Richards, J.D.; Massara, F.; Wright, H. Data Management Policies and Practices of Digital Archaeological Repositories. *Internet Archaeol.* **2022**, *59*, 1–52. [CrossRef]
6. Marwick, B.; Birch, S.E.P. A Standard for the Scholarly Citation of Archaeological Data as an Incentive to Data Sharing. *Adv. Archaeol. Pract.* **2018**, *6*, 125–143. [CrossRef]
7. Jakobsson, U.; Novák, D.; Richards, J.D.; Štular, B.; Wright, H. Digital Archiving in Archaeology: The State of the Art. *Internet Archaeol.* **2021**, *58*, 1–5.
8. Berners-Lee, T.; Fischetti, M. Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor. 1999. Available online: [http://archive.org/details/isbn\\_9780062515872](http://archive.org/details/isbn_9780062515872) (accessed on 14 April 2022).
9. Berners-Lee, T. Linked Data—Design Issues. 2006. Available online: <https://www.w3.org/DesignIssues/LinkedData.html> (accessed on 7 January 2022).
10. Vrandečić, D.; Krötzsch, M. Wikidata: A free collaborative knowledgebase. *Commun. ACM* **2014**, *57*, 78–85. [CrossRef]
11. Isaksen, L. Archaeology and the Semantic Web. 2011. Available online: <https://eprints.soton.ac.uk/206421/> (accessed on 12 April 2022).
12. Geser, G. ARIADNE WP15 Study: Towards a Web of Archaeological Linked Open Data. 2016. Available online: [http://legacy.ariadne-infrastructure.eu/wp-content/uploads/2019/01/ARIADNE\\_archaeological\\_LOD\\_study\\_10-2016-1.pdf](http://legacy.ariadne-infrastructure.eu/wp-content/uploads/2019/01/ARIADNE_archaeological_LOD_study_10-2016-1.pdf) (accessed on 21 February 2022).
13. Blaney, J. Introduction to the Principles of Linked Open Data. *Program. Hist.* **2017**, *6*. [CrossRef]
14. What Are Linked Data and Linked Open Data? 2022. Available online: <https://www.ontotext.com/knowledgehub/fundamentals/linked-data-linked-open-data/> (accessed on 14 April 2022).
15. Opendata.Swiss. Handbook.Opendata.Swiss: Linked Open Data. Available online: <https://handbook.opendata.swiss/de/content/glossar/bibliothek/linked-open-data.html> (accessed on 24 May 2022).
16. Bizer, C.; Heath, T.; Berners-Lee, T. Linked Data—The Story So Far. *Int. J. Semant. Web Inf. Syst.* **2009**, *5*, 1–22. [CrossRef]
17. Berners-Lee, T. Semantic Web Roadmap. 1998. Available online: <https://www.w3.org/DesignIssues/Semantic.html> (accessed on 14 April 2022).
18. Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web. *Sci. Am.* **2001**, *284*, 34–43. [CrossRef]
19. Gartner, R. Breaking the Silos. In *Metadata*; Springer International Publishing: Cham, Switzerland, 2016; pp. 87–96. [CrossRef]
20. Wikipedia Contributors. Machine-Readable Data—Wikipedia, The Free Encyclopedia. 2021. Available online: [https://en.wikipedia.org/w/index.php?title=Machine-readable\\_data&oldid=1057115179](https://en.wikipedia.org/w/index.php?title=Machine-readable_data&oldid=1057115179) (accessed on 14 April 2022).
21. W3C. Semantic Web. 2015. Available online: <https://www.w3.org/standards/semanticweb/> (accessed on 7 January 2022).
22. Manola, F.; Miller, E.; McBride, B. W3C: RDF 1.1 Primer. 2014. Available online: <http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140624/> (accessed on 17 May 2022).
23. Aranda, C.B.; Olivier Corby, S.D.; Feigenbaum, L.; Gearon, P.; Glimm, B.; Harris, S.; Hawke, S.; Herman, I.; Humfrey, N.; Michaelis, N.; et al. W3C: SPARQL 1.1 Overview. 2013. Available online: <https://www.w3.org/TR/sparql11-overview/> (accessed on 14 April 2022).
24. Wikipedia. Phaistos Disc. Page Version ID: 1081215059.2022. Available online: [https://en.wikipedia.org/w/index.php?title=Phaistos\\_Disc&oldid=1081215059](https://en.wikipedia.org/w/index.php?title=Phaistos_Disc&oldid=1081215059) (accessed on 14 April 2022).
25. Berners-Lee, T. What a Semantic Can Represent. 1998. Available online: <https://www.w3.org/DesignIssues/RDFnot.html> (accessed on 14 April 2022).
26. Wood, D.; Zaidman, M.; Ruth, L.; Hausenblas, M. *Linked Data: Structured Data on the Web*; Manning: Shelter Island, NY, USA, 2014; OCLC: ocn828182162.
27. The Linked Open Data Cloud. 2020. Available online: <https://lod-cloud.net/> (accessed on 27 May 2022).
28. Bray, T.; Hollander, D.; Layman, A.; Tobin, R.; Thompson, H.S. W3C: Namespaces in XML 1.0 (Third Edition). 2009. Available online: <http://www.w3.org/TR/2009/REC-xml-names-20091208/> (accessed on 14 April 2022).
29. W3C. Ontologies. 2015. Available online: <https://www.w3.org/standards/semanticweb/ontology> (accessed on 14 April 2022).
30. Harpring, P. Introduction to Controlled Vocabularies. 2010. Available online: [https://www.getty.edu/research/publications/electronic\\_publications/intro\\_controlled\\_vocab/index.html](https://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/index.html) (accessed on 20 April 2022).
31. iDAI. Objects/Arachne. 2022. Available online: <https://arachne.dainst.org/> (accessed on 19 April 2022).
32. Institute for the Study of the Ancient World (NYU); Ancient World Mapping Center (UNC-CH). Pleiades. 2009. Available online: <https://pleiades.stoa.org> (accessed on 14 April 2022).
33. Simon, R.; Isaksen, L.; Barker, E.; de Soto Cañamares, P. The Pleiades Gazetteer and the Pelagios Project. In *Placing Names: Enriching and Integrating Gazetteers*; OCLC: ocn933437838; Berman, M.L., Mostern, R., Southall, H., Eds.; The Spatial Humanities, Indiana University Press: Bloomington, IN, USA, 2016; pp. 97–109.
34. Open Knowledge Foundation. Open Definition 2.1. 2015. Available online: <http://opendefinition.org/od/2.1/en/> (accessed on 14 April 2022).
35. Shafranovich, Y. RFC 4180. Common Format and MIME Type for Comma-Separated Values (CSV) Files. 2005. Available online: <https://datatracker.ietf.org/doc/html/rfc4180> (accessed on 17 May 2022).



36. Creative Commons. About CC Licenses. 2019. Available online: <https://creativecommons.org/about/cclicenses/> (accessed on 14 April 2022).
37. Siebes, R.; Coen, G.; Gregory, K.; Scharnhorst, A. Linked Open Data. 2019. Available online: <https://librarycarpentry.org/Top-10-FAIR//2019/09/05/linked-open-data/> (accessed on 24 May 2022).
38. Frosterus, M.; Hansson, D.; Dadvar, M.; Kyriazis, I.; Zapounidou, S.; Grant, G. Best Practices for Library Linked Open Data (LOD) Publication. Available online: <https://libereurope.eu/wp-content/uploads/2021/02/LOD-Guidelines-FINAL-Feb-2021.pdf> (accessed on 24 May 2022).
39. Hyland, B.; Ateazing, G.; Villazón-Terrazas, B. Best Practices for Publishing Linked Data. 2014. Available online: <http://www.w3.org/TR/2014/NOTE-ld-bp-20140109/> (accessed on 24 May 2022).
40. Suominen, O.; Hyvönen, N. From MARC silos to Linked Data silos? *O-Bib. Das Offene Bibl.* **2017**, *4*, 1–13. [CrossRef]
41. Rossenova, L. Examining Wikidata and Wikibase in the Context of Research Data Management Applications. 2022. Available online: <https://blogs.tib.eu/wp/tib/2022/03/16/examining-wikidata-and-wikibase-in-the-context-of-research-data-management-applications/> (accessed on 12 April 2022).
42. Internet Engineering Task Force (IETF). *RFC 8259. The JavaScript Object Notation (JSON) Data Interchange Format*; Bray, T., Ed.; Internet Engineering Task Force (IETF): Fremont, CA, USA, 2017. Available online: <https://datatracker.ietf.org/doc/html/rfc8259> (accessed on 17 May 2022)
43. DuCharme, B. The Wikidata Data Model and Your SPARQL Queries. 2017. Available online: <http://www.snee.com/bobdc.blog/2017/04/the-wikidata-data-model-and-yo.html> (accessed on 24 January 2022).
44. MediaWiki. Wikibase/DataModel/Primer. 2022. Available online: <https://www.mediawiki.org/wiki/Wikibase/DataModel/Primer> (accessed on 19 April 2022).
45. WikidataCommunity. Wikidata:Identifiers. 2021. Available online: <https://www.wikidata.org/wiki/Wikidata:Identifiers> (accessed on 24 January 2022).
46. WikidataCommunity. List of Properties in Wikidata by Data Type: External Identifier. 2022. Available online: <https://www.wikidata.org/w/index.php?title=Special:ListProperties/external-id&limit=50&offset=0> (accessed on 24 January 2022).
47. Haak, L.L.; Fenner, M.; Paglione, L.; Pentz, E.; Ratner, H. ORCID: a system to uniquely identify researchers. *Learn. Publ.* **2012**, *25*, 259–264. [CrossRef]
48. Paskin, N. Digital Object Identifier (DOI®) System. 2015. Available online: [https://www.doi.org/overview/DOI\\_article\\_ELIS3.pdf](https://www.doi.org/overview/DOI_article_ELIS3.pdf) (accessed on 25 January 2022).
49. GeoNames. GeoNames. 2022. Available online: <http://geonames.org/> (accessed on 25 January 2022).
50. Neubert, J. Wikidata as a Linking Hub for Knowledge Organization Systems? Integrating an Authority Mapping into Wikidata and Learning Lessons for KOS Mappings. In Proceedings of the 17th European Networked Knowledge Organization Systems Workshop, Thessaloniki, Greece, 21 September 2017; pp. 14–25.
51. Erxleben, F.; Günther, M.; Krötzsch, M.; Mendez, J.; Vrandečić, D. Introducing Wikidata to the Linked Data Web. In *The Semantic Web—ISWC 2014*; Mika, P., Tudorache, T., Bernstein, A., Welty, C., Knoblock, C., Vrandečić, D., Groth, P., Noy, N., Janowicz, K., Goble, C., Eds.; Springer International Publishing: Cham, Switzerland, 2014; Volume 8796, pp. 50–65. [CrossRef]
52. The Linked Open Data Cloud: Wikidata. Available online: <https://lod-cloud.net/dataset/wikidata> (accessed on 27 May 2022).
53. Meta. LinkedOpenData/Strategy2021—Meta, Discussion about Wikimedia Projects. 2022. Available online: <https://meta.wikimedia.org/w/index.php?title=LinkedOpenData/Strategy2021&oldid=23106873> (accessed on 27 May 2022).
54. Wikidata:Data Access. 2022. Available online: [https://www.wikidata.org/w/index.php?title=Wikidata:Data\\_access&oldid=1637691714](https://www.wikidata.org/w/index.php?title=Wikidata:Data_access&oldid=1637691714) (accessed on 28 May 2022).
55. Binding, C.; May, K.; Souza, R.; Tudhope, D.; Vlachidis, A. Semantic Technologies for Archaeology Resources: Results from the STAR Project: Computer Applications and Quantitative Methods in Archaeology (CAA2010), Granada, April 2010. In Proceedings of the 38th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Granada, Spain, 6–9 April 2013; BAR Publishing: Oxford, UK, 2013; pp. 555–561.
56. Gerth, P.; Sieverling, A.; Trognitz, M. Data Curation: How and Why. A Showcase with Re-use Scenarios. *Stud. Digit. Herit.* **2017**, *1*, 182–193. [CrossRef]
57. Evans, T. Linking It All Together. Available online: <http://www.archaide.eu/blog/-/blogs/38396326> (accessed on 24 May 2022).
58. LeFebvre, M.J.; Brenskelle, L.; Wiczorek, J.; Kansa, S.W.; Kansa, E.C.; Wallis, N.J.; King, J.N.; Emery, K.F.; Guralnick, R. ZooArchNet: Connecting zooarchaeological specimens to the biodiversity and archaeology data networks. *PLoS ONE* **2019**, *14*, e0215369. [CrossRef] [PubMed]
59. Network (CHIN). Linked Open Data—Benefits and Challenges. Available online: <https://chin-rcip.github.io/collections-model/en/resources/current/lod-benefits-challenges> (accessed on 24 May 2022).
60. Wenige, L.; Ruhland, J. Retrieval by recommendation: Using LOD technologies to improve digital library search. *Int. J. Digit. Libr.* **2018**, *19*, 253–269. [CrossRef]
61. Debole, F.; Meghini, C.; Geser, G.; Tudhope, D. D15.2: Report on the ARIADNE Linked Data Cloud. 2017. Available online: [http://legacy.ariadne-infrastructure.eu/wp-content/uploads/2019/01/D15.2\\_Report-on-the-ARIADNE\\_Linked\\_Data\\_Cloud\\_Final.pdf](http://legacy.ariadne-infrastructure.eu/wp-content/uploads/2019/01/D15.2_Report-on-the-ARIADNE_Linked_Data_Cloud_Final.pdf) (accessed on 1 June 2022).



62. Hofmann, K.; Grunwald, S.; Lang, F.; Peter, U.; Rösler, K.; Rokohl, L.; Schreiber, S.; Tolle, K.; Wigg-Wolf, D. Ding-Editionen. Vom archäologischen (Be-)Fund übers Corpus ins Netz. 2019. Available online: <https://publications.dainst.org/journals/efb/article/download/2236/6674/> (accessed on 30 May 2022).
63. Abhayaratna, J.; van den Brink, L.; Car, N.; Atkinson, R.; Homburg, T.; Knibbe, F.; McGlinn, K.; Wagner, A.; Bonduel, M.; Rasmussen, M.H.; et al. OGC Benefits of Representing Spatial Data Using Semantic and Graph Technologies. Available online: <http://www.opengis.net/doc/wp/using-semantic-graph> (accessed on 24 May 2022).
64. Nayak, A.; Božić, B.; Longo, L. (Linked) Data Quality Assessment: An Ontological Approach. 2021. Available online: <http://ceur-ws.org/Vol-2956/paper17.pdf> (accessed on 30 May 2022).
65. Thiery, F.; Homburg, T.; Schmidt, S.C.; Voß, J.; Trognitz, M. SPARQLing Geodesy for Cultural Heritage—New Opportunities for Publishing and Analysing Volunteered Linked (geo-)data. In Proceedings of the FIG e-Working Week 2021—Smart Surveyors for Land and Water Management—Challenges in a New Reality Virtually in the Netherlands, Virtual, 21–25 June 2021; FIG: Copenhagen, Denmark, 2021. [CrossRef]
66. Nomisma. Nomisma.org. 2022. Available online: <http://nomisma.org> (accessed on 19 April 2022).
67. Huvila, I. Being FAIR When Archaeological Information Is MEAN: Miscellaneous, Exceptional, Arbitrary, Nonconformist. 2017. Available online: <https://www.istohuvila.fi/files/IstoHuvilaCDH2017-handout.pdf> (accessed on 31 May 2022).
68. Rula, A.; Maurino, A.; Batini, C. Data Quality Issues in Linked Open Data. In *Data and Information Quality*; Springer International Publishing: Cham, Switzerland, 2016; pp. 87–112. [CrossRef]
69. Gruber, E. Numishare: On Stable URIs at the British Museum. Available online: <http://numishare.blogspot.com/2018/02/on-stable-uris-at-british-museum.html> (accessed on 31 May 2022).
70. Marcus Smith. Is the British Museum’s Endpoint Working? (@bm\_lod\_status)/Twitter. Available online: [https://web.archive.org/web/20220211180616/https://twitter.com/bm\\_lod\\_status/status/1492197022413336584](https://web.archive.org/web/20220211180616/https://twitter.com/bm_lod_status/status/1492197022413336584) (accessed on 31 May 2022).
71. Thiery, F.; Mees, A.; Arera-Rütenik, T. TRAIL4.2: Implementing Mapping Processes for Vocabularies Related to Site and Object Protection; Zenodo: Geneva, Switzerland, 2021. [CrossRef]
72. SPARQL Query Service/WDQS Backend Update/Blazegraph Failure Playbook. 2022. Available online: [https://www.wikidata.org/w/index.php?title=Wikidata:SPARQL\\_query\\_service/WDQS\\_backend\\_update/Blazegraph\\_failure\\_playbook&oldid=1618872307](https://www.wikidata.org/w/index.php?title=Wikidata:SPARQL_query_service/WDQS_backend_update/Blazegraph_failure_playbook&oldid=1618872307) (accessed on 1 June 2022).
73. Homburg, T.; Klammt, A.; Mara, H.; Schmid, C.; Schmidt, S.C.; Thiery, F.; Trognitz, M. Recommendations for the review of archaeological research software. *Archäologische Inform.* **2020**, *43*, 357–370. [CrossRef]
74. iDAI World. 2022. Available online: <https://idai.world/> (accessed on 19 April 2022).
75. Archaeology Data Service. 2022. Available online: <https://archaeologydataservice.ac.uk/> (accessed on 19 April 2022).
76. re3data.org. Archaeology Data Service. 2012. Available online: <https://www.re3data.org/repository/r3d100000006> (accessed on 19 April 2022). [CrossRef]
77. Fernández González, J.M.; Polo Marques, A.; Cerrillo Cuenca, E. Bases for the Creation of Ontology in the Context of Archaeology. In *The World Is in Your Eyes. CAA2005. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 33rd Conference, Tomar, March 2005*; CAA Portugal: Tomar, Portugal, 2007; pp. 285–290. Available online: [https://proceedings.caaconference.org/paper/42\\_fernandez\\_et\\_al\\_caa\\_2005/](https://proceedings.caaconference.org/paper/42_fernandez_et_al_caa_2005/) (accessed on 19 April 2022).
78. Isaksen, L.; Martinez, K.; Gibbins, N.; Earl, G.; Keay, S. Linking Archaeological Data. In *Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference Williamsburg, Virginia, United States of America, March 22–26 2009*; Archaeopress: Oxford, UK, 2010; pp. 130–136. Available online: [https://proceedings.caaconference.org/paper/18\\_isaksen\\_et\\_al\\_caa2009/](https://proceedings.caaconference.org/paper/18_isaksen_et_al_caa2009/) (accessed on 19 April 2022).
79. May, K.; Binding, C.; Tudhope, D.; Jeffrey, S. Semantic Technologies Enhancing Links and Linked Data for Archaeological Resources. In *Revive the Past. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 39th International Conference, Beijing, April 12–16 April 2011*; Pallas Publications: Amsterdam, The Netherlands, 2012; pp. 261–272. Available online: [https://proceedings.caaconference.org/paper/29\\_may\\_et\\_al\\_caa2011/](https://proceedings.caaconference.org/paper/29_may_et_al_caa2011/) (accessed on 19 April 2022).
80. Dentamaro, F.; De Luca, P.G.; Genco, L.; Perrino, G.; Cannito, C.; Stufano, M.A.; Sibilano, M.G. A CIDOC CRM-Based Ontology System. In *Digital Discovery. Exploring New Frontiers in Human Heritage. CAA2006. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 34th Conference, Fargo, United States, April 2006*; Archaeolingua: Budapest, Hungary, 2007; pp. 437–444. Available online: [https://proceedings.caaconference.org/paper/cd45\\_dentamaro\\_et\\_al\\_caa2006/](https://proceedings.caaconference.org/paper/cd45_dentamaro_et_al_caa2006/) (accessed on 19 April 2022).
81. Cripps, P.; May, K. To OO or not to OO? Revelations from Ontological Modelling of an Archaeological Information System. In *Beyond the Artifact. Digital Interpretation of the Past. Proceedings of CAA2004, Prato 13–17 April 2004*; Archaeolingua: Budapest, Hungary, 2010; pp. 59–63. Available online: [https://proceedings.caaconference.org/paper/08\\_cripps\\_may\\_caa\\_2004/](https://proceedings.caaconference.org/paper/08_cripps_may_caa_2004/) (accessed on 19 April 2022).
82. Gonzalez-Perez, C.; Parcerero-Oubiña, C. A Conceptual Model for Cultural Heritage Definition and Motivation. In *CAA2011—Revive the Past, Proceedings of the 39th Conference in Computer Applications and Quantitative Methods in Archaeology, Beijing, China, 12–16 April 2011*; Mingquan, Z., Romanowska, I., Wu, Z., Xu, P., Verhagen, P., Eds.; Pallas Publications: Leiden, The Netherlands, 2012; pp. 234–244.
83. LIDO Overview. 2022. Available online: <https://cidoc.mini.icom.museum/working-groups/lido/lido-overview/> (accessed on 20 April 2022).

84. CIDOC. CIDOC CRM. 2022. Available online: <https://www.cidoc-crm.org> (accessed on 14 April 2022).
85. The Getty Research Institute. Art & Architecture Thesaurus. 2017. Available online: <https://www.getty.edu/research/tools/vocabularies/aat/> (accessed on 14 April 2022).
86. FISH. Terminology (FISH—Forum on Information Standards in Heritage). 2022. Available online: <http://www.heritage-standards.org.uk/terminology/> (accessed on 14 April 2022).
87. PeriodO. PeriodO—Periods, Organized. 2022. Available online: <https://perio.do/en/> (accessed on 19 April 2022).
88. Golden, P.; Shaw, R. Nanopublication beyond the sciences: The PeriodO period gazetteer. *PeerJ Comput. Sci.* **2016**, *2*, e44. [CrossRef]
89. Gruber, E.; Pett, D.; Tolle, K.; Heath, S.; Wigg-Wolf, D.; Meadows, A. Semantic Web Technologies Applied to Numismatic Collections. In *Archaeology in the Digital Era, Proceedings of the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, UK, 26–29 March 2012*; Earl, G., Sly, T., Wheatley, D., Romanowska, I., Papadopoulos, C., Murrieta-Flores, P., Chrysanthi, A., Eds.; Amsterdam University Press: Amsterdam, The Netherlands, 2012; Volume II, pp. 264–274. [CrossRef]
90. Pelagios—The Digital Classicist Wiki. 2022. Available online: <https://wiki.digitalclassicist.org/Pelagios> (accessed on 19 April 2022).
91. Tolle, K.; Wigg-Wolf, D.; Gruber, E. An Ontology for a Numismatic Island with Bridges to Others. 2018. In *Oceans of Data. Proceedings of the 44th Annual Conference on Computer Applications and Quantitative Methods in Archaeology*; Archaeopress: Oxford, UK, 2018; pp. 103–108. Available online: <https://www.archaeopress.com/Archaeopress/download/9781784917302> (accessed on 19 April 2022).
92. Tolle, K.; Wigg-Wolf, D. Improving Data Quality by Rules: A Numismatic Example. In *CAA: Digital Archaeologies, Material Worlds (Past and Present). Proceedings of the 2017 CAA Annual Meeting, New York, NY, USA, 14–16 March 2017*; Universität Tübingen: Tübingen, Germany, 2020; pp. 193–201. [CrossRef]
93. Thiery, F. Linking Potter, Pots And Places: A LOD Approach To Samian Ware, In *Proceedings of the CAA 2014, Paris, France, 22–25 April 2014*. [CrossRef]
94. Gruber, E.; Gondek, R.; Smith, T.J. Kerameikos.org: Linked Open Data for Greek Pottery. In *Proceedings of the CAA 2019: Poster Session, Krakow, Poland, 23–27 April 2019*; Zenodo: Geneva, Switzerland, 2019. [CrossRef]
95. Cuy, S.; Schmidle, W.; Thiery, F.; Kallas, N. Linking periods: Modeling and utilizing spatio-temporal concepts in the ChronOntology project. In *Proceedings of the 44th Conference on Computer Applications and Quantitative Methods in Archaeology, Oslo, Norway, 30 March–3 April 2016*; Zenodo: Oslo, Norway, 2016. (forthcoming).
96. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 160018. [CrossRef]
97. Sanderson, R. LOUD: Linked Open Usable Data. 2019. Available online: <https://linked.art/loud/> (accessed on 7 January 2022).
98. Discover inspiring European cultural heritage | Europeana. 2022. Available online: <https://www.europeana.eu/en> (accessed on 19 April 2022).
99. Welcome—Ariadne Portal. 2022. Available online: <https://portal.ariadne-infrastructure.eu/> (accessed on 19 April 2022).
100. Digital Index of North American Archaeology (DINAA) | Heritage Bytes. Available online: <http://ux.opencontext.org/archaeology-site-data/> (accessed on 19 April 2022).
101. Kansa, E.C.; Kansa, S.W.; Wells, J.J.; Yerka, S.J.; Myers, K.N.; DeMuth, R.C.; Bissett, T.G.; Anderson, D.G. The Digital Index of North American Archaeology: Networking government data to navigate an uncertain future for the past. *Antiquity* **2018**, *92*, 490–506. [CrossRef]
102. Jeffrey, S.; Xia, L.; Richards, J.D.; Bateman, J.; Kintigh, K.; Pierce-McManamon, F.; Brin, A. The Transatlantic Archaeology Gateway: Bridging the Digital Ocean. In *CAA2011—Revive the Past, Proceedings of the 39th Conference in Computer Applications and Quantitative Methods in Archaeology, Beijing, China, 12–16 April 2011*; Mingquan, Z., Romanowska, I., Wu, Z., Xu, P., Verhagen, P., Eds.; Pallas Publications: Leiden, The Netherlands, 2012; pp. 198–208.
103. May, K.; Binding, C.; Tudhope, D. Barriers and opportunities for Linked Open Data use in archaeology and cultural heritage. *Archäologische Inform.* **2015**, *38*, 173–184. [CrossRef]
104. Bibby, D.; Bruhn, K.C.; Busch, A.; Dührkohp, F.; Himmelmann, U.; Höke, B.; Keller, C.; Lang, M.; Mees, A.; Metz, S.E.; et al. Digitales Forschungsdatenmanagement in der Archäologie und die Initiative NFDI4Objects. *BLiCKpunkt Archäologie* **2021**, *2021*, 150–164. [CrossRef]
105. Thiery, F.; Mees, A.; Wienand, J.; Börner, S. TRAIL2.5: A workflow for enhancing iconography authority data in the Wikimedia Universe. In *NFDI4Objects TRAILS*; Zenodo: Geneva, Switzerland, 2021. [CrossRef]
106. Mees, A.W.; Thiery, F.; Weisser, B. Digitale Vernetzung von Sammlungsdaten. *Squirrel Papers* **2021**, *3*, 1. [CrossRef]
107. Samian Research Community. Samian Research. 2022. Available online: <https://www.rgzm.de/samian> (accessed on 19 April 2022).
108. Hartley, B.R.; Dickinson, B.M. *Names on Terra Sigillata: An Index of Makers' Stamps & Signatures on Gallo-Roman Terra Sigillata (Samian Ware)*; Number 102 in Bulletin of the Institute of Classical Studies; Institute of Classical Studies, University of London: London, UK, 2008.

109. Thiery, F.; Mees, A.; Gottwald, D. Linked Open Samian Ware—Unveiling the hidden Data Dragons and uncovering temporal vagueness with the help of Little Minions. In Proceedings of the Joint Chapter Meeting 2020 of Computer Applications and Quantitative Methods in Archaeology by the CAA chapters CAA-NL-FL and CAA-DE (JCM2020), Deventer, The Netherlands, 3–4 December 2020; Zenodo: Geneva, Switzerland, 2020. [CrossRef]
110. Thiery, F. Semantic Web und Linked Data: Generierung von Interoperabilität in Archäologischen Fachdaten am Beispiel Römischer Töpferstempel. Master's Thesis, Fachhochschule Mainz, Mainz, Germany, 2013. [CrossRef]
111. Thiery, F.; Mees, A.W.; Gottwald, D. Linked Open Samian Ware: External Linking. 2021. Available online: <https://rgzm.github.io/samian-lod/linking/> (accessed on 7 April 2022).
112. Thiery, F.; Rokohl, L. Linked Open African Red Slip Ware. *Squirrel Papers* **2021**, *3*, 1. [CrossRef]
113. Thiery, F.; Mees, A.; Gottwald, D. *Linked Open Samian Ware*; Zenodo: Geneva, Switzerland, 2020. [CrossRef]
114. Quick Statements. 2022. Available online: <https://quickstatements.toolforge.org/#/> (accessed on 19 April 2022).
115. Online, V.P. DFD Using Yourdon and DeMarco Notation. 2022. Available online: <https://online.visual-paradigm.com/knowledge/software-design/dfd-using-yourdon-and-demarco> (accessed on 1 June 2022).
116. Thiery, F.; Homburg, T. *Linked Pipes @ Linked Pasts 7: Introduction*; Zenodo: Geneva, Switzerland, 2021. zenodo.5781274. [CrossRef]
117. Thiery, F.; Homburg, T.; Trognitz, M. *Linked Pipe: Linked Open Samian Ware*; Zenodo: Geneva, Switzerland, 2021. [CrossRef]
118. Cradle. 2022. Available online: <https://cradle.toolforge.org/> (accessed on 19 April 2022).
119. Trognitz, M.; Thiery, F. *Wikidata—A SPARQL(ing) Unicorn?* Zenodo: Geneva, Switzerland, 2019. [CrossRef]
120. Thiery, F.; Schmidt, S.C.; Homburg, T.; Trognitz, M. *The SPARQL Unicorn: An introduction*; Zenodo: Geneva, Switzerland, 2020. [CrossRef]
121. Homburg, T.; Thiery, F. *SPARQLing Unicorn QGIS Plugin*; Zenodo: Geneva, Switzerland, 2021. [CrossRef]
122. Bogdani, J.; Montalbano, R.; Rosati, P. *ArcheoFOSS XIV 2020: Open Software, Hardware, Processes, Data and Formats in Archaeological Research*; Archaeopress Publishing Ltd: Oxford, UK, 2021; OCLC: 1274200091.
123. Schmidt, S.C.; Thiery, F. SPARQLing Ogham Stones: New Options for Analyzing Analog Editions by Digitization in Wikidata. *CEUR Workshop Proc.* **2022**, *3110*, 211–244. [CrossRef]
124. Thiery, F. *My Little Linked Open Data Ogham Minion: Visualising Graph Data Connections Using SPARQL Endpoints, Talk at CAA 2021, Limassol, Cyprus*; Zenodo: Geneva, Switzerland, 2021. [CrossRef]
125. Macalister, R.A.S. *Corpus inscriptionum insularum Celticarum*; Stationery Office: Dublin, Ireland, 1945; Volume I, QID: Q70256237.
126. MacManus, D. *A Guide to Ogam*; Number 4 in Maynooth monographs; An Sagart: Maynooth, Ireland, 1997; OCLC: 248434889; QID: Q70310399.
127. O'Sullivan, A.; Sheehan, J. *The Iveragh Peninsula. An Archaeological Survey of South Kerry*; Cork University Press: Cork, Ireland, 1996.
128. Younger, J.G. *A Bibliography for Aegean glyptic in the Bronze Age; Vol. Beiheft 4, Corpus der Minoischen und Mykenischen Siegel*; Gebr. Mann: Berlin, Germany, 1991. [CrossRef]
129. Hayes, J.W. *Late Roman Pottery*; British School at Rome: London, UK, 1972; QID: Q50262763.
130. Thiery, F.; Raddatz, L.; Boochs, F. Close to the Original-Erfassung archäologischer Objekte und ihre webbasierte semantisch modellierte Bereitstellung zur fachwissenschaftlichen Analyse. In *Photogrammetrie—Laserscanning—Optische 3D-Messtechnik: Beiträge der Oldenburger 3D-Tage 2022 [In Print]*; Wichmann: Berlin, Germany, 2022; OCLC: 1102430437.
131. Thiery, F.; Rokohl, L. African Red Slip Ware digital (ARS3D)—The Portal. *Squirrel Papers* **2021**, *3*, 1. [CrossRef]
132. Kabashi, A. ICONCLASS—Classification System for Art and Iconography. 2019. Available online: <https://urn.nsk.hr/urn:nbn:hr:131:711297> (accessed on 19 April 2022).
133. zu Löwenstein, S. *Mythologische Darstellungen auf Gebrauchsgegenständen der Spätantike: Die Appliken-und Reliefverzierte Sigillata C3/C4*; Number 48 in Kölner Jahrbuch; Gebr. Mann: Berlin, Germany, 2015; pp. 397–823.
134. Armstrong, M. *A Thesaurus of Applied Motives on African Red Slip Ware*. Ph.D. Thesis, New York University, New York, NY, USA, 1993; OCLC: 60852088; QID: Q109525251.
135. Anselmino, L.; Carandini, A.; Pavolini, C.; Sagui, L.; Tortorella, S.; Tortorici, E. *Atlante Delle Forme Ceramiche: Ceramica Fina Romana nel Bacino Mediterraneo; Vol. 1, Enciclopedia Dell'arte Antica Classica e Orientale*; Istituto della Enciclopedia Italiana: Roma, Italy, 1981; OCLC: 1124005431; QID: Q109525400.
136. Thiery, F. African Red Slip Ware - Additional Iconography Catalogue. *Squirrel Papers* **2022**, *4*, 1. [CrossRef]
137. Wikidata Query Service. 2022. Available online: <https://query.wikidata.org/> (accessed on 19 April 2022).
138. Wikidata: Linked Open Data Workflow. 2021. Available online: [https://www.wikidata.org/w/index.php?title=Wikidata:Linked\\_open\\_data\\_workflow&oldid=1493989706](https://www.wikidata.org/w/index.php?title=Wikidata:Linked_open_data_workflow&oldid=1493989706) (accessed on 28 May 2022).
139. Manske, M. Cradle. 2022. Available online: <https://github.com/magnusmanske/cradle> (accessed on 19 April 2022).
140. Wikidata Query Service/User Manual—MediaWiki. 2022. Available online: [https://www.mediawiki.org/wiki/Wikidata\\_Query\\_Service/User\\_Manual](https://www.mediawiki.org/wiki/Wikidata_Query_Service/User_Manual) (accessed on 19 April 2022).
141. Wikidata:SPARQL Query Service/Wikidata Query Help—Wikidata. 2022. Available online: [https://www.wikidata.org/wiki/Wikidata:SPARQL\\_query\\_service/Wikidata\\_Query\\_Help](https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/Wikidata_Query_Help) (accessed on 19 April 2022).
142. Wikidata Query Builder. 2022. Available online: <https://query.wikidata.org/querybuilder/> (accessed on 19 April 2022).

143. Thiery, F.; Homburg, T. SPARQLing Unicorn QGIS Plugin. 2022. Available online: <https://github.com/sparqlunicorn/sparqlunicornGoesGIS> (accessed on 19 April 2022).
144. Help: QuickStatements—Wikidata. 2022. Available online: <https://www.wikidata.org/wiki/Help:QuickStatements> (accessed on 19 April 2022).
145. Open Refine. 2022. Available online: <https://openrefine.org/> (accessed on 19 April 2022).
146. Wikidata: Tools/OpenRefine—Wikidata. 2022. Available online: <https://www.wikidata.org/wiki/Wikidata:Tools/OpenRefine> (accessed on 19 April 2022).
147. Nielsen, F.Å.; Mietchen, D.; Willighagen, E. Scholia and scientometrics with Wikidata. In Proceedings of the 1st International Workshop on Scientometrics and 1st International Workshop on Enabling Decentralised Scholarly Communication (SciEDSC), Portorož, Slovenia, 28 May–1 June 2017; Springer: Cham, Switzerland Aachen; 3; urn:nbn:de:0074-1878-8.
148. Robinson, M. Illuminating Hopewell Legacy Data: A Case Study of Mound 23 at Hopewell Mound Group. 2016. Available online: <https://digitalcommons.unl.edu/anthrotheses/46> (accessed on 14 April 2022).
149. Laneri, N.; Brancato, R.; Figuera, M.; Cristofaro, S.; Spampinato, D.; Asmundo, N.; Santamari, D.F. Towards an ontology of the Museum of Archaeology of the University of Catania: From the digitization of the legacy data to the Semantic Web. ArcheoFOSS XIV 2020: Open Software, Hardware, Processes, Data and Formats in Archaeological Research. In Proceedings of the 14th International Conference, Online, 15–17 October 2020; Archaeopress: Oxford, UK, 2021; pp. 128–137.
150. Jones-Cervantes, S.A.; Blinman, E.; Tauxe, L.; Cox, J.R.; Lengyel, S.; Sternberg, R.; Eighmy, J.; Wolfman, D.; DuBois, R. MagIC as a FAIR repository for America’s directional archaeomagnetic legacy data. *Earth Space Sci. Open Arch. ESSOAr* **2021**, *25*. <http://dx.doi.org/10.1002/essoar.10507730.1>.
151. Opitz, R. Publishing Archaeological Excavations at the Digital Turn. *J. Field Archaeol.* **2018**, *43*, S68–S82. [[CrossRef](#)]
152. Kansa, E.; Kansa, S.W. Digital Data and Data Literacy in Archaeology Now and in the New Decade. *Adv. Archaeol. Pract.* **2020**, *9*, 81–85. [[CrossRef](#)]
153. Anzt, H.; Bach, F.; Druskat, S.; Löffler, F.; Loewe, A.; Renard, B.Y.; Seemann, G.; Struck, A.; Achhammer, E.; Aggarwal, P.; et al. An environment for sustainable research software in Germany and beyond: Current state, open challenges, and call for action. *F1000Research* **2021**, *9*, 295. [[CrossRef](#)]