Dealing with post-excavation data: the Omeka S TiMMA web-database

Bastien Rueff

Volume 4 (2024), article e5

https://doi.org/10.24072/pcjournal.366

Abstract

This paper reports on the creation and use of a web database designed as part of the TiMMA project with the Content Management System Omeka S. Rather than resulting in a technical manual, its goal is to analyze the relevance of using Omeka S in the frame of a post-excavation project, gathering a wide team of researchers from several countries, working in their own language and having their own specialty. Designed to assist organizations and institutions in creating and managing digital collections, Omeka S offers a number of commodities from the perspective of both the administrator(s) and the users, the most significant being its efficiency in capitalizing on linked data standards for items description. Additionally, one of the successful achievements of this platform is to benefit the open-science tools more broadly, permitting for instance, as was done in this project, to import images with the iiiF API and bibliographical references with the Zotero module, thus avoiding the creation of new digital files and metadata. Designed with simplicity in mind, the graphic interface of Omeka S makes it a particularly appropriate tool for collective projects, permitting each user to perform a specific role and record data in their chosen language. This user-friendliness extends up to the data recording and publishing, as well as to the website designing which requires no specific skills in code. The TiMMA project also faces the limitations of Omeka S, such as the lack of efficiency of the search engine, which makes the use of web-semantic vocabularies difficult, a few bugs, and the absence of convenient functionalities in archaeological projects (e.g. the automatic creation of primary keys, a statistic dashboard, cells' automatic filling). Despite these limitations, the TiMMA project contributes to showing that, without being a complete archaeological system recording, Omeka S is adapted to post-excavation projects.
The TiMMA project and the need for a web-database

The Aegean world in the 3rd and 2nd millennia B.C. is notably characterized by the emergence of two intertwined civilizations: the Minoan and the Mycenaean. Since the first archaeological research that led to their discovery in the second half of the 19th century, architectural studies have proposed numerous reconstructions of buildings, with the most intensively investigated being the monumental structures known as palaces. Within this context, stone, and to a lesser extent, mudbrick, are the most extensively studied architectural materials. Conversely, due to the challenging preservation conditions of wood, timber has been largely overlooked. One of the primary reasons for this oversight was the apparent lack of evidence for its use in architecture. However, recent studies, building on earlier research, have revealed that timber was extensively used—from the smallest houses to massive palaces—and could play a crucial structural role, in addition to its anti-seismic properties (notably Devolder 2019; Pomadère and Hilbert 2019; Rougier-Blanc 2019; Tsakanika-Theohari 2006; Tsakanika-Theohari 2009).

Launched in 2022 for a duration of four years, the TiMMA project aims to investigate the role of timber in Minoan and Mycenaean architecture (3200-1100 BC). Through an interdisciplinary and integrated approach, the project’s goal is to reconstruct the Minoan and Mycenaean’s exploitation of wood, from tree felling to the use of timber in architecture. One of its specificities is to bring together over twenty specialists, from five different countries, including architects, archaeologists, historians of techniques, civil engineers, and archaeobotanists. Therefore, it also deals with a compilation of archaeological, architectural, archival and archaeobotanical sources. Data is collected from sites already or currently excavated (fig. 1), namely the Bronze Age settlements of Malia, Ayia Triada, Phaistos and Zakros (Crete), Pylos, Mycenae (Greek mainland) and Akrotiri (Thera).

The evidence collected is predominantly architectural, encompassing imprints, holes, chases, traces of door leaves on thresholds, cavities on walls, stone bases, mortises, curved bedding on ashlar walls, and socket-pivot holes (fig. 2). Charcoals associated with architectural remains may also be present: preserved mineralized wood and imprints in the ash can be found in Akrotiri, where the exceptionally good preservation of remains is owed to the volcano eruption. Indeed, this site is often referred to as the "Pompeii of the Aegean." In addition to the archaeological finds, the project also involves a vast collection of metadata resulting from fieldwork and research in archives. This includes thousands of photographs, plans, and other illustrations, as well as an extensive bibliography related to the excavations and, finally, timber elements and architectural reconstructions sometimes identified in excavation notebooks from the first excavators of the studied sites.

To address this variety and volume of data and metadata, there was a necessity to build a web-based database that is both user-friendly and multilingual, facilitating the work of multiple specialists involved in the project. This database needs to serve as an analytical tool for processing queries, requiring the creation of complex links between different tabs. Additionally, it should function as a heuristic tool, aiding in the conceptualization and understanding of data. Finally, it must align with the Findable, Accessible, Interoperable, Reusable (F.A.I.R) principles to promote open science (Wilkinson et al. 2016).

The choice for Omeka S: related works and advantages

AIR and Heurist are two open source web database systems that also meet our expectations and are specifically designed for archaeological purposes (e.g., Derudas, Nurra, and Svensson, 2023). In the early stages, when considering which solution should be retained, we explored the use of these tools as well; these also notably allow input from multiple users with controlled access in different languages. These systems are also linked to a web-publishing platform. However, our final decision was to work with Omeka
S (https://omeka.org/s/), a Content Management System (CMS) designed to assist organizations and institutions in creating and managing digital collections. Funded by The Andrew W. Mellon Foundation (https://www.mellon.org/) and the Institute of Museum and Library Services (https://imls.gov/), Omeka S shares the same goals and principles as Omeka Classic (https://omeka.org/classic/). Unlike AIR and Heurist, it was not specifically designed for archaeological fieldwork projects but rather for heritage purposes and online exhibitions. Projects using Omeka S in History, Anthropology and Literacy are numerous in various fields, ranging from prosopography repositories (https://exhibits.lib.utah.edu/s/century-of-black-mormons/page/welcome) to musical instrument collections (https://omeka-s.grinnell.edu/s/MusicalInstruments/page/welcome).

A distinctive feature is the emphasis on the web-editorialized platform, designed with simplicity in mind, featuring entry by caption. This is why Omeka S is said to be more suitable for managing heritage, textual, and museum collections, rather than archaeological finds (Bérard et al., 2020; Lombardo et al., 2020). However, behind the editorialized platform, Omeka S relies on PHP and uses MySQL as a database storage system. It can thus be used as a proper database with interconnected tabs and has the ability to process requests through an advanced request engine module. Yet, it has rarely been used for archaeological projects. One of the few exceptions, however, applied to the site of Tobiotsuka Kofun in Japan, demonstrates well its efficiency to link together archaeological finds and stratigraphic units (https://bearchaeo.unito.it/omeka-s/s/bearchaeo-resources-site/page/welcome). One of the reasons we chose this solution is its distinctive feature: Omeka S operates on downloadable modules, allowing for adaptation and personalization for any project. Moreover, as TiMMA is not an excavation project but deals with post-excavation data, the diverse functionalities offered by AIR and Heurist would not be fully used in

Figure 1 – Map of the Aegean world and of the settlements analyzed in the project.
this context. Additionally, Omeka S offers a highly useful mobile version, facilitating convenient access to data from mobile phones. Last but not least, the multilingual version of Heurist is still under preparation. Given the variety of languages spoken by the different specialists involved in the project, it has not been able to provide a suitable solution.

The architecture of the database

The TiMMA database currently comprises five main tabs, two of which have multiple sub-tabs (fig. 3). A tab, or resource template, refers to a thematic grouping of items or entries. These tabs aim to describe locations, such as sites and buildings, as well as spaces where evidence of timber and structural units have been found. There are also tabs dedicated to describing the structural units themselves and the evidence of timber, in addition to the charcoals and preserved wood identified through archaeobotanical analysis.

The creation of these tabs implied a significant amount of work in a preliminary phase. This was accomplished by gathering several specialists, members of the project, who had already worked on case studies at Ayia Triada, Malia, and Knossos (Pomadère and Hilbert 2019; Tsakanika-Theohari 2006).
Regarding the "Structural unit" and "Timber evidence," this step proved to be essential since, collectively, we were able to anticipate all possible situations one may face in the field.

Let us consider the example of two different structural units expected in the Aegean architecture: colonnades with timber columns and door frames. These structural units share 21 fields, mainly dedicated to the description of the structure, dimensions, as well as connections with other tabs (Space tab, Location tab). However, the former has an additional 18 fields specifically intended for its precise description (table 1). Following this logic, we personalized every sub-tab according to the architectural characteristics to be described (table 2).

<table>
<thead>
<tr>
<th>Colonnades with timber columns and door frames: fields in common</th>
<th>Fields only found in the sub-tab “Door frame”</th>
</tr>
</thead>
</table>
| • Code name (Unique Key)  
• Type of structural unit  
• List of related timber evidence  
• Location code  
• Space code  
• Chronological phase  
• Position in the space  
• General dimensions of the structural units  
• General dimensions of the saved part of the door frame  
• Preserved diameter  
• Preserved height  
• Preserved length  
• Preserved width  
• Description of the connection of the structural unit with other structural units  
• State of preservation when excavated  
• Current state of preservation  
• Type of destruction  
• Method of documentation  
• Description  
• Media legend and metadata  
• Bibliography | • Type of door (structural)  
• Number of door leafs  
• Type of door jamb bases  
• Shape of the door jamb bases  
• Presence of threshold  
• Type of threshold (architectural)  
• Shape of the threshold  
• Type of door (architectural)  
• Distance between bases and the horizontal timbers of the lintel  
• Presence of transversal horizontal timber on the stone base  
• Presence of transversal horizontal timber at other position of the timber frame of the door (intermediate)  
• Number of transversal horizontal timber at other position of the timber frame of the door (intermediate, lintel levels)  
• Position of vertical timbers at each jamb pier of the timber frame of the door  
• Number of vertical timbers at each jamb pier of the timber frame of the door  
• Number of horizontal timbers of the lintel  
• Type of lintel  
• Type of infill between vertical timbers  
• Plaster on the door |
<table>
<thead>
<tr>
<th>3a. Floor</th>
</tr>
</thead>
<tbody>
<tr>
<td>3b. Flat roof</td>
</tr>
<tr>
<td>3c. Inclined roof</td>
</tr>
<tr>
<td>3d. Staircases</td>
</tr>
<tr>
<td>3e. Free-standing column</td>
</tr>
<tr>
<td>3f. Free-standing pier</td>
</tr>
<tr>
<td>3g. Pier inside or at the free end of a wall</td>
</tr>
<tr>
<td>3h. Colonnade with timber columns</td>
</tr>
<tr>
<td>3i. Colonnade with timber piers</td>
</tr>
<tr>
<td>3j. Colonnade with timber columns and piers</td>
</tr>
<tr>
<td>3k. Vertical timber reinforcement system of masonry</td>
</tr>
<tr>
<td>3l. Horizontal timber reinforcement system of masonry</td>
</tr>
<tr>
<td>3m. Polythyron, pier-and-door partition / cupboard partition</td>
</tr>
<tr>
<td>3n. Door frame</td>
</tr>
<tr>
<td>3o. Pier and window partition</td>
</tr>
<tr>
<td>3p. Window frame</td>
</tr>
<tr>
<td>3q. Partition wall with timber frame</td>
</tr>
<tr>
<td>3r. Parapet</td>
</tr>
<tr>
<td>3s. Unknown structural unit</td>
</tr>
<tr>
<td>4a. Imprint on mudbrick / mortar / plaster / ash</td>
</tr>
<tr>
<td>4b. Hole</td>
</tr>
<tr>
<td>4c. Chase / cavity on a wall</td>
</tr>
<tr>
<td>4d. Stone base without mortises or bedding</td>
</tr>
<tr>
<td>4e. Stone base with mortises and/or curved bedding</td>
</tr>
<tr>
<td>4f. Ashlar course with mortises and/or curved bedding</td>
</tr>
<tr>
<td>4g. Ashlar wall with mortises and/or curved bedding</td>
</tr>
<tr>
<td>4h. Use-wear traces (door leaf / window leaf on floor / threshold / door sill / other)</td>
</tr>
<tr>
<td>4i. Socket-pivot hole</td>
</tr>
<tr>
<td>4j. Other timber element</td>
</tr>
<tr>
<td>4l. Reconstructed timber element</td>
</tr>
</tbody>
</table>

Semantic web vocabularies and interoperability: a Linked Open Data-based project

The significant accomplishment of Omeka S is its use of a set of design principles for sharing machine-readable interlinked data on the web, known as linked data standards for item description. To ensure the longevity of data, each recorded item in the TiMMA database is described in Dublin Core format (https://www.dublincore.org/specifications/dublin-core/dcmi-terms/), including the title and textual description, within a set of concepts and categories referred to as an ontology and provided by default in Omeka S. The database is also linked to the Getty Museum's Thesaurus Art & Architecture vocabulary (https://www.getty.edu/research/tools/vocabularies/aat). Specifically, this vocabulary is used to describe the wood species identified through anthracological analysis. This approach avoids the creation of redundant value lists and ensures the accuracy of the terms used in English and their scientific names in...
Latin. The "Value Suggest" module (https://omeka.org/s/docs/user-manual/modules/valuesuggest/) is helpful in this context, allowing users to quickly find a term by typing its first letters in the extensive list of existing species (fig. 4).

![Figure 4](https://omeka.org/s/styles/mediumsized/s/VALUE-SUGGEST.png)

**Figure 4** – The "Value Suggest" module, linked to external vocabularies, avoids creating new lists of values and provides multiple choices by simply typing the first letters of the value to be inserted. In this example, it is used in connection with the Getty Museum’s Thesaurus Art & Architecture. Typing the "Quercus" species provides a wide range of taxa.

Although semantic web vocabularies like Getty's are extremely helpful, they can be too extensive to be used with efficiency by a team of over 20 individuals. This may require refining to achieve homogenization, which itself can lead to confusion when conducting searches in the absence of a closed list of values. As a solution, the TiMMA Project Ontology has been designed specifically for the TiMMA project using the "Custom Ontology" module (https://omeka.org/s/modules/CustomOntology/). This allows for the creation of ontologies in the Dublin Core format and makes them available for any similar project. Currently, this new ontology comprises 252 properties that are specifically dedicated to describing wood resources and timber use in the architecture of archaeological settlements.

Additionally, Omeka S provides a convenient platform for users to upload various file formats, including PDF, JPG, PNG, TIFF, and videos, which can be directly attached to the corresponding items in the database. Users can add multiple media files to each item within the storage capacity limit. Notably, Omeka S also supports images in the iiiF format (https://iiif.io/), a way to standardize audio and visual files from servers to different environments on the Web where they can then be viewed and interacted with in many ways. Therefore, it is possible to import images by copy-pasting their Uniform Resource Identifier (URI). This feature is particularly useful for importing high-definition images directly from Archimage, a web database designed and maintained by the French School at Athens (https://archimage.efa.gr/?kroute=accueil), and demonstrates the value of Omeka S for open science initiatives by eliminating the need to create new digital files and metadata. This turned out to be extremely convenient for the Malia site, excavated under the auspices of the French School at Athens, as some of the archives have already been digitized. For instance, we could copy and paste the digitized images of the North Portico of the Palace at Malia from the TiMMA database using their iiiF link (fig. 5).

Another key feature of Omeka S is the ability to import bibliographical references from Zotero using the "Zotero Import" module (https://omeka.org/s/docs/user-manual/modules/zoteroimport/). Zotero is a free, open-source reference management software that allows users to manage bibliographic data and related research materials such as PDF files. It also enables the organization and accurate citation of references, as well as document sharing with groups.
The TiMMA team uses Zotero to import each reference and its metadata only once, instead of manually typing it every time it is needed. A Zotero group library has been created for this purpose, allowing everyone to add references from their standalone interface. This functionality saves a significant amount of time and ensures a standardized bibliography, even with a large team contributing to the database. Thanks to this tool, references are imported and described using the terms available in the "Bibliographic ontology." Their metadata can be consulted just like any other item (fig. 6). In the example below, the space item 360 (Palace at Pylos) is linked to the publication of this site through the Zotero module.

**Data treatment: the perspective of the user**

Omeka S is particularly well suited for collective projects, as it allows each user to be assigned a specific role. Access is granted by the general supervisor(s), who assign each user an editor or administrator role, along with a personal password to ensure the security of the data. The general supervisor(s) also set the language preference for each user. This is perhaps what distinguishes it the most from other web based database systems such as Heurist. Omeka S thus offers the ability to display fields in a wide range of languages, including English, Greek, French, and Italian, which is particularly important for the TiMMA project. It facilitates the development of a shared list of concepts and their cultural meanings across these languages. These aspects are crucial for an international project, as they allow users from around the world to contribute to the database with personalized access. Additionally, the user interface is designed with simplicity in mind, making it easy for inexperienced users to manage. In the case of the TiMMA project, a manual was created to enable users to independently populate the database, and it is accessible on the website linked to the database (https://timma.efa.gr/s/timma/page/how-to-fill-the-database). This is
complemented by a personal meeting with each individual responsible for recording data. Based on our experience, this is sufficient for individuals to work autonomously.

**Figure 6** – Linking a space item to a reference imported through the Zotero module: the case of Spa360.

Data recording in Omeka S can be achieved in two ways. First, it is possible to import a Comma Separated Value(s) (.csv) file using the "CSV Import" module ([https://omeka.org/s/docs/user-manual/modules/csvimport/](https://omeka.org/s/docs/user-manual/modules/csvimport/)). Filling can also be done manually by typing information in the corresponding cell. The TiMMA database uses two types of cells: free text cells and cells with a list of values. Free text cells can be used for description and dimension fields, while most fields use cells with a list of values. For a collaborative and international project, it is essential to agree on the terms to be used and to define a shared thesaurus. The "Custom Vocab" module ([https://omeka.org/s/docs/user-manual/modules/customvocab/](https://omeka.org/s/docs/user-manual/modules/customvocab/)) makes this possible by allowing users to create and edit lists of values.

One significant advantage of Omeka S is that users can easily create fields with multiple values by clicking the "Add value" button without modifying the resource templates themselves. Users can also create links between tabs using the "Omeka S resources" button. Multiple links can be created for each item, which is particularly useful for archaeological projects where finds may come from multiple spaces, structures, or deposits. For example, Structure 115 (Str115) in figure 7 is composed of two timber evidence (Tim257, 258) respectively identified as chase on the floor and imprint of threshold and separated into two different spaces (Spa87 and Spa89) as a door frame.

One major constraint of Omeka S, however, is that Primary Keys, intended to attribute a single identifier to each item, cannot be automatically incremented, which makes data management more complex and time consuming. In the case of TiMMA, a shared sheet has been created in which a unique key number is
recorded for each created item. However, this solution is not optimal and may lead to mistakes in recording.

**Figure 7** – Example of structure (door frame) composed of several timber elements and separating two distinct spaces.

Additionally, the process of publishing and unpublishing items is designed with simplicity in mind. The MySQL database is linked to a web-publishing platform, which means that every created item is visible online by default. However, both administrators and users can choose to unpublish items they do not wish to make available. Not only do we use the TiMMA website to browse collections, but also to centralize information about the project and the team members, to process queries and to publish project-related news that might be of interest to the community of scholars (https://timma.efa.gr/s/timma/page/timma-project).

The website's graphic interface does not require any specific skills for design. Omeka S provides the ability to download themes and modify them directly in the Cascading Style Sheets (CSS) file or using the "CSS editor" module (https://github.com/omeka-s-modules/CSSEditor). This is an extremely interesting feature that saves time, emphasizing data recording over graphic design. For example, in the case of TiMMA, we initially applied a theme developed by the Omeka team (fig. 8, A). However, we encountered a limitation due to our lack of skills in graphic design—specifically, only the first picture could be displayed for each item. As a solution, we installed a theme created by Federico Nurra and the Service Numérique de la Recherche of the French History of Art National Institute. This theme, inspired by the Roy Rosenzweig Centerrow and under Creative Commons license 4.0 (fig. 8, B) (https://github.com/INHAParis/omeka-theme-DigitalMuret), not only possesses aesthetic qualities but also allows for the display of images in the form of a carousel, a very convenient tool for data consultation.

Finally, the "Advanced Search" module (https://gitlab.com/Daniel-KM/Omeka-S-module-AdvancedSearch) allows users to execute both simple and complex queries by analyzing data from multiple tabs. Simple queries can be executed using operators such as "contains," "starts with," "ends with," "has any value," and "has no value." Users can also combine multiple search criteria to execute complex queries using operators such as "AND," "OR," and "NOT." However, the standardized format of the search engine
may seem restrictive, as it relies on a vocabulary provided by Omeka S, described in Dublin Core, and is solely based on labels such as Author, Creator, Subject, Description, and Class. Any other vocabulary is omitted in the research engine. This means that it is possible to conduct proper complex research.

A) Theme originally adopted

Theme developed by the Omeka team. Version 1.6.1.

B) Theme finally retained

Theme Digital Muret, INHA, under a Creative Commons license 4.0. Version 2.0.0 Federico Nurra (inspired by Roy Ronenzweig Centerrow) pour l’INHA.

Figure 8 – A) Theme originally adopted. In this theme, only the first picture of each item could be displayed. B) This theme offers the ability to display pictures in a carousel; it also features an aesthetically pleasing design that enhances the appreciation of data during consultation.

To overcome the limitations of the search engine, users may prefer to export data and process queries through more specialized software such as Excel or LibreOffice Calc. Fortunately, Omeka S offers the ability to download data in XML format through its "Export" module (https://github.com/lauraleif/Export). However, metadata linked to the media is not currently downloadable, which is a significant issue given the amount of media in the database. Further investigation is required to address this issue.
In the case of TiMMA, to prevent each user from performing analysis from their own .csv file, we decided to export the content of the database and import it into a Heurist platform. Indeed, Heurist offers a thorough and powerful research engine that allows for processing complex queries.

Conclusion

To conclude, this case study demonstrates that Omeka S is indeed a suitable tool for managing post-excavation data. The TiMMA project encounters various challenges, including the participation of numerous individuals, the use of multiple languages, and, most significantly, the processing of diverse and voluminous data. Just like other web databases such as Heurist or AIR, Omeka S enables each user to work from their own laptop, in their preferred language, providing secure access and a user-friendly graphical interface designed for simplicity.

What sets Omeka S apart as a distinctive tool and particularly suitable for the TiMMA project is its modularity and adaptability to project needs. The community of volunteer contributors and developers dedicated to enhancing modules, themes, and Omeka S in general facilitates the tool's use in projects dealing with post-excavation datasets.

Building on TiMMA's experience, we would like to emphasize two aspects mentioned earlier. Firstly, the ease with which this tool allows for the use of Linked Open Data and open vocabularies described in standardized formats, such as the Art & Architecture Getty Museum, Cidoc-CRM, and others, along with open data tools like Zotero or the iiIF API. Secondly, the graphic interface can be changed by simply downloading pre-defined sets of themes, reducing the effort required for web graphic development. In the TiMMA project, it took some time to agree on the final theme, but changing the overall appearance was always straightforward, without encountering bugs.

The TiMMA project also encountered limitations in the software, suggesting the need for collaborations with other tools like Heurist and AIR, seen here as complementary tools. This was exemplified by the challenges faced in processing requests and the use of Heurist as a solution. We believe this serves as a meaningful lesson, demonstrating that rather than being in competition, various open web-based databases benefit from synergies, working together and fostering collaborations. In this sense, the TiMMA project is indebted to all the communities that made these tools available, bringing a thorough renewal to the field of research.

Acknowledgements

I am grateful to Sylvie Rougier-Blanc, head of the TiMMA project, Louis Mulot, Catherine Bouras and Federico Nurra for their technical assistance and their careful reading. I would also like to warmly thank Claire Dupin-Debeysat and Miled Rousset for their help, Bruno Morandière and Medi Jones-Williams for their thorough perusal, as well as the Omeka S community, the OpenTheso and the Getty Museum organizations for providing tools, resources and assistance. Finally, I would like to express my gratitude to the anonymous reviewers of this article who significantly helped improve the first version submitted. Preprint version 3 of this article has been peer-reviewed and recommended by Peer Community In Archaeology (https://doi.org/10.24072/pci.archaeo.100341, Hanna 2024).
Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Funding

This work was supported by a grant from the French National Research Agency (ANR) as part of the “TiMMA” Programme (ANR21-CE27-0029, 2022-2025).

References


Poursat, J.-Cl. And Knappett, C. 2005 Fouilles exécutées à Malia : le quartier Mu. IV, La poterie du minoen moyen II : production et utilisation, Athènes, Études crétoises


Tsakanika-Theohari, E. 2006. Ο δομικός ρόλος του ξύλου στην τοιχοποίηση των ανακτορικού τύπου κτιρίων της Μινωικής Κρήτης. Εθνικό Μετσόβιο Πολυτεχνείο (ΕΜΠ), Σχολή Αρχιτεκτόνων Μηχανικών. https://doi.org/10.12681/eadd/16329