

Interim Performance Report

Grant Number: HK-230973-15

Project Title: Pleiades 3

Project Director: Tom Elliott

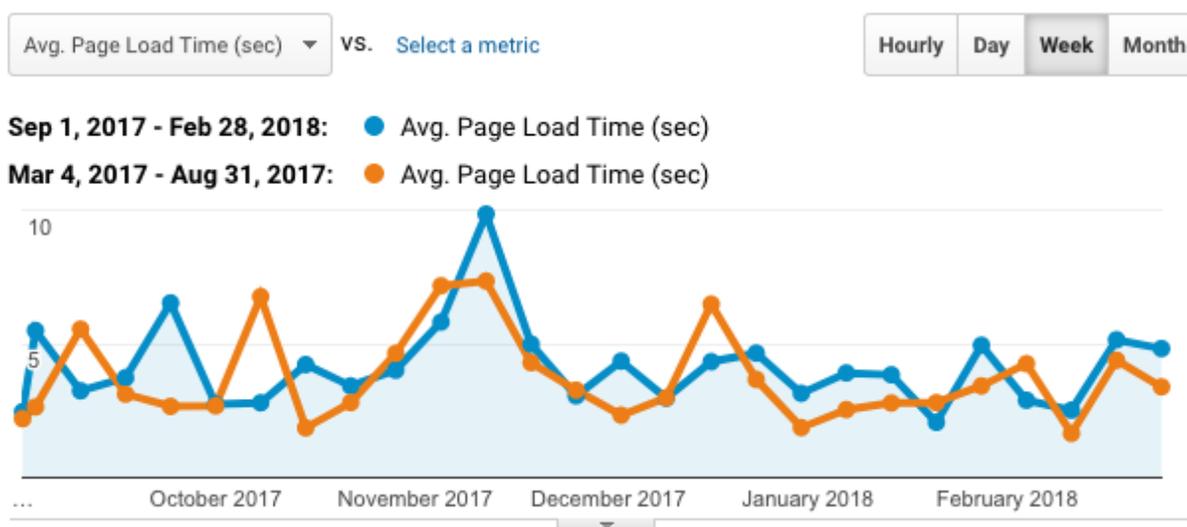
Grantee Institution: New York University

Date Submitted: 30 March 2018

In this report, we summarize project activities during the first half of its final, scheduled year, as well as plans for the successful completion of the grant. As anticipated in the original proposal, we have much less to report in the way of grant-funded activity, since the majority of expenditure was deliberately front-loaded in order to permit uptake and reaction from the user community ahead of any final arrangements. Accordingly, we report on both grant-funded and concurrent external/volunteer activity, clearly distinguishing each.

#### System Performance and Upgrades

Site speed and availability, as reported by *Google Analytics* during the current reporting period, remained statistically consistent with data reported in the previous two biannual performance reports, demonstrating the lasting effectiveness of the system upgrades enabled by this grant. Despite heavier usage of advanced search and other compute-intensive functions of the web application, as well as a 10% increase in non-bot page views over the previous six-month period (an increase from 163,155 to 167,844), average users continued to experience average page load times on the order of 3-5 seconds. The following graph illustrates the weekly average page load time (in blue). The weekly average for the preceding 6-month period is overlaid in orange.



We aim to secure additional performance improvements and more headroom for future expansion with a second round of hardware replacement for our production server as discussed in the previous report. Work on this change is in progress, with completion now expected by the end of April 2018. The new computer, which is already in place in the commercial data center in Colorado that has been home to *Pleiades* since 2006, boasts several advantages over the current host, including more central processor cores. The biggest change is a switch from conventional hard drives to solid-state drives. This change will speed database-intensive tasks like complex search queries and scripted content updates. Our software services contractor (Jazkarta, Inc.) and managed hosting provider (Tummy.com) will be under contract to provide support as needed during cutover.

Outreach, Training, and External Evaluation

As anticipated in our original proposal and work plan, the PI and the all-volunteer *Pleiades* Editorial College have interacted intensively with each other and with the broader *Pleiades* contributor and user community during the reporting period. Much of this interaction has been focused on mentoring contributors and soliciting feedback about the user interface, content interactions, editorial workflow, and documentation. For example, the PI has organized a series of on-line meetings to discuss and implement improvements to *Pleiades* resources documenting the settlement of Aphrodisias in Turkey.<sup>1</sup> As Aphrodisias is a UNESCO World Heritage Site with significant archaeological remains and has an on-going archaeological mission, it provides an excellent test bed for recent improvements to *Pleiades* and its documentation, as well as a rich training ground for new and infrequent contributors. This and similar efforts, conducted both in person and on-line, have yielded new bug reports and feature requests, which are documented on the public “*Pleiades* Gazetteer Issue Tracker” at GitHub.<sup>2</sup> The PI has been able to address a number of these concerns already, but some will require contractor assistance. Accordingly, we anticipate during the next 6-month period the execution of a new purchase order for Jazkarta, Inc., drawing on remaining grant funds in accordance with project scope.

A *Pleiades* workshop was conducted by members of the *Pleiades* Editorial College at the Joint Annual Meetings of the Archaeological Institute of American and the Society for Classical Studies. Unfortunately, a major winter storm caused widespread travel disruptions, preventing many participants (including the PI) from reaching Boston for the event. The PI and other remote presenters and participants used a Google Hangout to join those who had made it to Boston, but only 10-15 participants other than presenters were able to attend and the planned “hands on” portion of the workshop had to be abandoned. In addition, a planned collaboration meeting with staff of the Aga Khan Documentation Center at the Massachusetts Institute of Technology had to be canceled. Because travel for several participants did not occur, budgeted grant funds for this travel were not expended. The PI has subsequently committed a portion of these funds for travel to meet with several high-volume users and contributors at university and conference locations in France, Germany, Italy, and the U.S. during March and May 2018.<sup>3</sup> A summer date for rescheduling the MIT meeting is also under discussion. The PI is also exploring the possibility of conducting an informal *Pleiades* workshop at the 2019 joint meetings (San Diego in January 2019); if this idea proves viable or if the MIT meeting needs to occur later than July, the PI in April 2018 will seek a no-cost extension to the current grant so that the in-scope travel and meeting activities may be completed.

Interactions such as those described above provide avenues for informal, but actionable evaluation of every aspect of *Pleiades*, including software, data, documentation, and contributor/editor process. Our open, on-line approach to software development and data

---

<sup>1</sup> <https://pleiades.stoa.org/places/638753>

<sup>2</sup> <https://github.com/isawnyu/pleiades-gazetteer/issues/>

<sup>3</sup> Major content expansion initiatives in line with those predicted in the original proposal are anticipated as a result of these meetings. Announcements will be made via the *Pleiades* News Blog once coordinated with new partners and these will also be addressed in a future performance report.

dissemination makes our code, issue tracker, website, and export datasets freely accessible online as a matter of daily practice. Consequently, third parties can easily explore and enquire about any issue of interest. For example, recently the PI received an email from a team of Semantic Web researchers in Europe previously unknown to the project. They had downloaded *Pleiades* data in the Resource Description Format<sup>4</sup> and tested it in software they were developing for a collaborative research project with scholars of early Christianity at another institution. They provided suggestions for improving our data representation to take advantage of recent enhancements to Semantic Web standards for RDF. These suggestions are now in line for implementation.

#### Data Management, Dissemination, and Archiving

Regular deposit of a comprehensive, open data export from *Pleiades* into third-party archives and into the hands of other digital projects is a major deliverable outlined in the project proposal. In February 2018, the PI completed work on a “*Pleiades Datasets*” package for this purpose, building on capabilities established earlier in the period of performance.<sup>5</sup> This package of data files and documentation is publicly available via GitHub, where its contents are updated every few days based on the latest publications in *Pleiades*. A first formal release of this data was made in late February 2018 and deposited with Zenodo.org where it automatically acquired a Digital Object Identifier (DOI).<sup>6</sup> The PI plans to conduct a second release in April 2018, with concurrent deposit not only at GitHub and thence into Zenodo, but also directly into the NYU Faculty Digital Archive. As anticipated in our Data Management Plan, “dark” (i.e., non-public) copies of this release will be also distributed for safe-keeping to the Perseus Digital Library, the German Archaeological Institute, and OpenContext, as well as to individual members of the Editorial College.

#### Other Technical Deliverables

---

<sup>4</sup> <https://pleiades.stoa.org/downloads#rdf>

<sup>5</sup> <https://github.com/isawnyu/pleiades-datasets>

<sup>6</sup> <https://doi.org/10.5281/zenodo.1193921>

### 1.1. Connections between places

During this reporting period, the PI made changes to the user interface to improve the presentation and clarity of “connections” between place resource. The direction and type -- e.g., “part of (administrative)” -- of each connection is now shown on in the listing of connections on each place page as illustrated by the screen capture at right. Users can now click through not only to the connected place page, but also to the connection page itself, making it easier for all users to examine details of a connection and for authenticated users to make suggested modifications to same.

**Aphrodisias makes connections with:**

- Aphrodisias → connection → Caria (unspecified date range)
- Aphrodisias → connection → Caria (unspecified date range)

**Aphrodisias receives connections from:**

- Hadrianic Baths → at → Aphrodisias (unspecified date range)
- Walls → at → Aphrodisias (unspecified date range)
- Martyrion → at → Aphrodisias (unspecified date range)
- Sebasteion → at → Aphrodisias (unspecified date range)
- Temple/Church → at → Aphrodisias (unspecified date range)
- South Agora → at → Aphrodisias (unspecified date range)
- Theater → at → Aphrodisias (unspecified date range)
- Tetracylon → at → Aphrodisias (unspecified date range)
- Basilica → at → Aphrodisias (unspecified date range)
- North Agora → at → Aphrodisias (unspecified date range)
- Odeon → at → Aphrodisias (unspecified date range)
- Stadium → at → Aphrodisias (unspecified date range)

**Place type:**  
settlement

The addition of more attribute fields to the connections data model has been mooted and will be refined and prioritized in coming months, with possible implementation during the period of performance.

### 1.2. Modeling places below the settlement level

Changes to editorial policy, documentation, and the data model necessary to support the treatment of buildings, structural complexes, monuments, and archaeological zones as full-fledged *Pleiades* “place resources” – as well as their relationships (connections) to other, larger entities as discussed above – have been in place for a number of reporting cycles. Contributors have been making increasingly constructive use of them. The Editors continue to work through a backlog of such features that were previously modeled only as “Location resources,” upgrading them to full “place resources.” Automation for facilitating repetitive, formulaic tasks associated with such upgrades was developed during an earlier period by Jazkarta and has been used successfully by the PI to prepare items for completion by other editors and contributors. Some enhancements to this code are anticipated in April 2018, and other editors will be trained in its use.

### 1.3. Bibliographic Integration with Zotero

Bibliographic references continue to be an area of focused interest, now that mechanisms for loading bibliographic references from a Zotero library have been added.<sup>7</sup> The PI has developed the “Pleiades RefBot,” a collection of Python scripts, to clean up existing references.<sup>8</sup> Related work is on-going as a collaborative effort between the PI and the Editorial College, including: improvements and expansion of the Zotero library, regularization of short titles for editions of primary sources, and updates to the “*Pleiades* Citation Guide.”<sup>9</sup> More advanced Zotero-enabled features have been requested and these will be implemented during the remainder of the performance period if sufficient resources are available once all pending bug reports and feature requests have been prioritized.

#### Contribution and Editing

As of March 1, 2018, *Pleiades* has published 35,787 place resources for which 32,133 historical and modern names and 39,503 locations are also published. The Editors continue to work through a significant review backlog, most of which dates back to the period before 2016 when system performance inhibited efficient review and the desire for feature and policy changes encouraged deferred publication of content exhibiting unexpected characteristics. The majority of these resources were submitted by members of the Editorial College, rather than contributors at large. Eight hundred forty-four place resources remain on the backlog along with 978 names and 2,029 locations. Most of the names and nearly half of the locations are subordinate to the pending places, having been submitted as packages of information. The rest of the locations are largely drawn from *OpenStreetMap* and provide more accuracy for already-published place resources than data originally digitized from the *Barrington Atlas*.

The Editorial College uses a variety of complementary methods to address current contributions and work through the backlog. These include: individual weekly or biweekly checks of the review list; weekly “review parties” during which available members of the college work together online to address pending items; creation and assignment to individual reviewers and editors of regional queries that show pending content only in a certain area; and establishment of a private issue tracker and wiki on GitHub where collaborative notes and conversations can be stored when working on a complex or problematic review. In addition, the managing editors continue to mentor reviewers for editorial roles in future and to identify and recruit contributors and external parties as reviewers. In other words, the Editorial College is now keeping pace with new contributions and making progress on clearing the backlog thanks to the technological improvements enabled by this grant and the process improvements they have enabled.

---

<sup>7</sup> The “*Pleiades* Zotero Library” is online at <https://www.zotero.org/groups/2533/pleiades>. The steps required to add a new reference using the library are in the *Pleiades* help documentation at <https://pleiades.stoa.org/help/add-new-reference>.

<sup>8</sup> <https://github.com/isawnyu/pleiades.refbot>

<sup>9</sup> <https://pleiades.stoa.org/help/citation-guide>