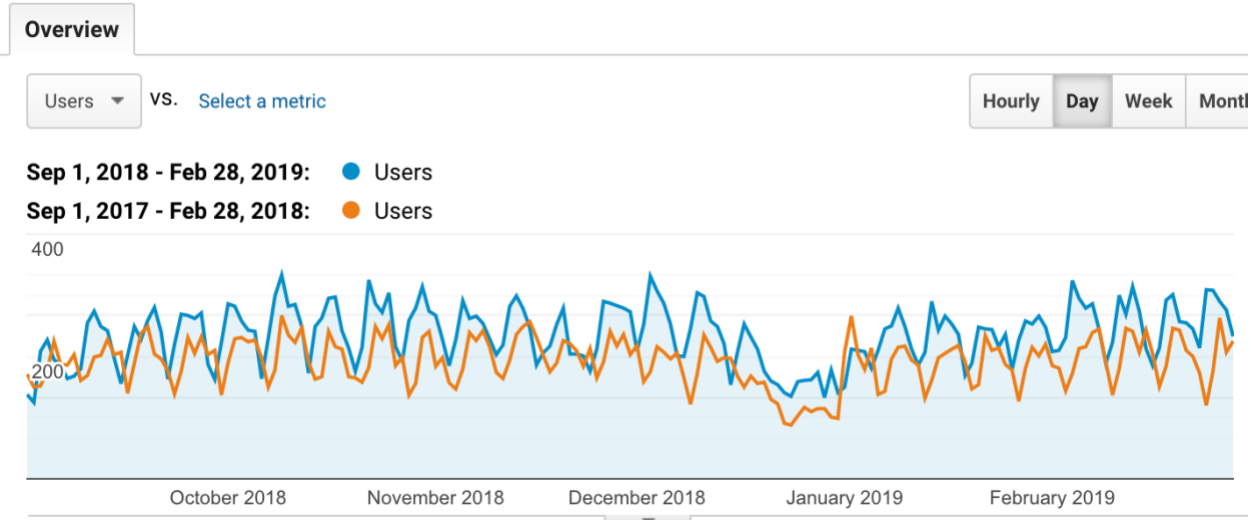
Interim Performance Report
Grant Number: HK-230973-15
Project Title: Pleiades 3
Project Director: Tom Elliott
Grantee Institution: New York University
Date Submitted: 1 April 2019

In this report, we summarize project activities during the second half of the fourth and final year of *Pleiades 3*, as extended under a no-cost extension approved 6/6/2018. As anticipated in the original proposal, we have much less to report this period 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.

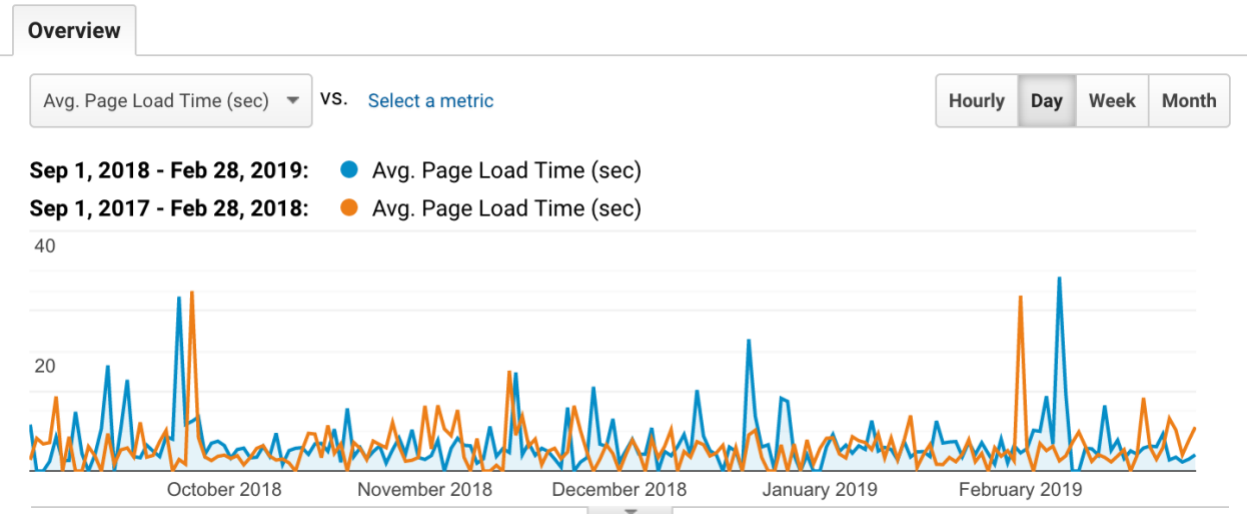
System Performance

No grant-funded work on the system was performed during the last 6 months; however, we offer the following assessment of site performance as continuing evidence that earlier grant-funded work on site performance continues to yield good results.

According to data collected and classified through Google Analytics, Pleiades user load during the period 1 September 2018 - 28 February 2019 increased by nearly 23% over the same 6-month period the preceding year: from 27,218 discrete users to 33,380. Despite this user increase, page views dropped by nearly 10%, average session duration by almost 27%, and bounce rate by 6%. We interpret these statistics to indicate that, although more users were visiting the site, they were finding the information they sought more quickly.



By most metrics, the site performance measures put in place earlier in the grant period seem to be holding up well. Server connection times and page load times remained statistically unchanged period-to-period. Among the component metrics of overall page load time, initial server response time added, on average, a few tenths of a second, but page download speeds increased by a compensatory amount. We believe these fluctuations to be largely the result of variations in network speed, user locations, and particular pages visited, rather than indicators in changes in software behavior on the site.



One component metric of the user experience does stand out: average page load time (i.e., the total amount of time measured from the start of a user attempt to view a page to the moment the browser completes loading of all page components). This number remains higher than we would like: approximately 4 seconds; however, closer analysis mitigates much of our concern.

First of all, the data sample is small. Only 1,234 of the total 152,091 page views logged during the reporting period resulted in transmission from the client browser to Google Analytics of the transaction metadata required for "page speed" analysis: a sample size of less than 1%. This small sample size not only calls into question the statistical significance of the page speed performance metrics, but also cautions us to beware of the possibility that a single user visit involving a slow connection beyond our control could contribute significantly to the high average. Moreover, the "broad and flat" profile of audience interest in Pleiades content also increases the likelihood of a single, idiosyncratically slow visit surfacing in the page timing statistics. In our previous report, we discussed some incidents of this sort, and we detect them again this period.

For example, the slowest page listed in the Google Analytics "Page Timing" analysis is the Place Resource for "Phraaspa?/Phraata?/Vera?" (https://pleiades.stoa.org/places/884193) for which a single visit in early February 2019 recorded a 112-second page load time. Google was unable to determine the originating location of that visit, so we cannot make the assumption (as we can with some other pages) that a poor connection in a developing country is the cause of a long page-load time reported by Google; however, Google detected 18 other visits to this particular Pleiades page during this reporting period, each from a discrete user and representing at least 9 different originating locations. Unfortunately, page-load timing for these visits were not gathered by Google, but subsequent testing we conducted from various locations in the United States shows no untoward delays loading this page. Moveover, Google's own "Speed Suggestions" tool, which analyzes performance on demand, places this particular page in the 77th percentile web-wide for performance. This same tool did make suggestions for improving load performance on this page by as much as a full second and, as these suggestions

are applicable to all Pleiades pages, we will investigate addressing same in future as resources permit.

In our previous report, we raised the possibility of serving *Pleiades* through the Cloudflare content delivery network in order better to serve our users outside the continental United States. We have decided not to pursue such a course at this time. Overall, our numbers are good. Moreover, Google's "Page Timing" analysis (discussed above) identifies existing Cloudflare dependencies (some JavaScript libraries we use) as the slowest "blocking resources" on our place pages. We believe that the combination of pending hardware upgrades (see below) and adjustments to such blocking resources, will provide additional performance improvements.

Hardware and Software Upgrades

In our last report, we indicated our intent to rehost *Pleiades* on a new, faster leased server on a revised timeline in December 2018 or January 2019. Administrative and scheduling challenges slowed issuance of a purchase order to our preferred subcontractor for this work, but we are now confident this work will be complete by the end of April 2019.

Additional software work is also slated for the upcoming, final 6 months of the project. Problems and feature requests identified by the Pleiades user community (including its reviewers and editors) have been documented in the *Pleiades Gazetteer Issue Tracker* on GitHub.[1] Candidates for work during the remainder of the grant have been grouped into a "Milestone" in that tracking system.[2] The issue tracker's "labels" functionality will be used to indicate priorities and its "projects" functionality will be used to plan and manage week-long development iterations, beginning the final week of March 2019.[3] Use of these systems permits all community members to follow and participate, if desired, in the process.

Data Management, Dissemination, and Archiving

Arrangements for dissemination of *Pleiades* data remain as described in the previous report. Pleiades export data continues to be generated on a nightly basis for download from the site, and JSON datasets in GitHub have continued to be updated weekly, on average.[4] Due to resource and scheduling constraints, the PI was unable to make a release of the GitHub data to Zenodo and the NYU Faculty Digital Archive in November, 2018 as planned. We now expect this to occur in April 2019 and quarterly thereafter.

---

[1] https://github.com/isawnyu/pleiades-gazetteer/issues.

[2] https://github.com/isawnyu/pleiades-gazetteer/milestone/15.

[3] https://github.com/isawnyu/pleiades-gazetteer/projects.

[4] https://pleiades.stoa.org/downloads and https://github.com/isawnyu/pleiades-datasets.

Contribution, Reviewing, and Editing

As of 7 March 2019, the all-volunteer *Pleiades* editorial college has published 36,540 place resources.