3 Narrative[1]

3.1 Significance

3.1.1 Long-term benefits to humanities research

An NEH Preservation and Access Research and Development grant (2006-2008), combined with significant institutional investment, sponsored the design and deployment of Pleiades (http://pleiades.stoa.org). Pleiades combines web technologies, social collaboration and open-content licensing in an innovative, online initiative to publish reference information about historically significant places and spaces in the ancient Greek and Roman world. It is poised to give scholars, students and enthusiasts worldwide the ability to use, create and share this information in an environment that encourages contributions from any knowledgeable person in a context of pervasive peer review. Pleiades has no staff offices. Distributed collaboration will decentralize the authoring, review and maintenance of content by embedding these activities in the day-to-day work of scholars in multiple disciplines and institutions, and by sharing it with a wider public whose expectations have been set by the "social software" and "neogeography" movements now sweeping the web.[2]

The present proposal seeks 3 years of funding to help us attract, engage and retain this diverse, essential community of users. The time is right. The definitive reference work in the field, the *Barrington Atlas of the Greek and Roman World*, is now nearly a decade old, and in need of updating and diversification.[3] Digital reference projects for papyrology, epigraphy, prosopography, numismatics and history are seeking a common geographic reference and coordinates for mapping. Individual scholars and students want a reliable, comprehensive and remixable reference resource. With NEH support, we will galvanize their collective action by offering them (for use and for improvement) the compilation materials that underpinned the Atlas. No other project is prepared to put research quality resources for the study of ancient geography on the web for use by the entire web community in so comprehensive and forward-looking a manner. By funding this grant, the National Endowment for the Humanities not only reinvigorates historical geographic study in the Classics, but also sets a new standard for the creation, maintenance and dissemination of humanities reference works.

In so doing, NEH addresses an urgent need, widely acknowledged as crucial to the future of humanities scholarship: new ways to create or retrofit crucial reference works for sustainability. Pleiades' success will transform past investments in print into a digital geographic reference that meets high standards of currency and scholarship on a continuing basis. Moreover, Pleiades will serve as an alternative model for addressing the problem of durability in reference works across the breadth of the humanities. It is time for reference works to come off the shelves and become networked workspaces that not only bring information to users but also enable immediate, incremental correction and improvement by those users.

The *Barrington Atlas* exemplifies the threat. Within years of its appearance, published research findings had begun to undermine its currency. This circumstance certainly should not be seen as a criticism of the worth of the Atlas or of the effort and expertise that went into its development. Comprehensive reference

---

[1] Acronyms are expanded when introduced. A glossary of acronyms, terms and project names is included in the appendices.

[2] Elliott and Gillies treat the neogeographical revolution, with special reference to ancient studies, in their contribution to *Digital Humanities Quarterly*, v.3 n.1 (Winter 2009): "Digital Geography and Classics," http://www.digitalhumanities.org/dhq/vol/3/1/000031.html (a copy has been provided in the Appendices).

[3] R. Talbert (ed.), *Barrington Atlas of the Greek and Roman World,* Princeton, 2000. For background, see Section 3.2.2.1. Extracts and example materials from the Atlas and its companion *Map-by-Map Directory* are included in the appendices.

works like the Atlas have unique value as infrastructural components for entire disciplines, yet they are routinely so expensive to create in print that, for many fields in the humanities, revised editions cannot be produced regularly. Indeed, the publication of the *Barrington Atlas* capped a 12-year, 200-person, $4.5 million effort to redress the then-longstanding lack of a comprehensive classical atlas; the last predecessor had appeared in 1874 (sic).[4] This predecessor constitutes a cautionary tale: because it was expensive to produce and unsupported by any institution or community, there was no way to disseminate revisions in response to discoveries stemming from archaeological field work, scholarly inquiry in widely separated fields, and technological developments in cartography and printing. Several projects did attempt to redress the problem (often for shorter historical periods or smaller regions), but only a Herculean U.S. effort a century later finally succeeded: the American Philological Association's (APA) Classical Atlas Project (CAP), funded by the NEH and many other individuals and foundations.

Unfortunately, many reference works in the humanities struggle with similar constraints. They routinely employ small central staffs of dedicated researchers or, at best, engage a relatively limited group of distributed collaborators, whose expertise must often be compensated. Recent recommendations for sustaining such projects have invoked "business models" and "revenue channels" as potential solutions, but subscription regimes and firewalled resources may subvert the public service mandates of many funders and academic institutions. Moreover, they run the risk of denying access to faculty and students outside Research I universities, and to those in developing countries. Charging models also assume more centralized control over content than is realistic in projects that involve multiple contributing institutions.

An alternative model shifts the focus from revenue generation to cost management, bringing together elements of "open access" licensing and "long-tail" or "social" authoring approaches to drive down core content-related costs. The use of Internet technologies to establish and maintain geographically distributed communities – and to open membership of those communities to a heterogeneous public – can significantly increase the workforce available to develop important content. Enthusiasts and students can be exposed to the approaches of trained scholars, and even work alongside them. Scholars in such communities benefit from the enthusiasm and hard work of non-academically credentialed collaborators. Prompt publication of individual articles circumvents delays endemic to conventional publication projects that must either wait for all content to be completed or adopt a serial publication approach (for example, in separate, alphabetically organized volumes). The ease with which individual components can be updated helps to limit the threat of obsolescence. As leading institutions begin to step up to the challenge of providing a humanities-oriented technical infrastructure, the use of common tools and services by such projects may further help to control costs and promote stability.

Pleiades is ready to forge a persistently accessible and continuously updated ancient geography through open scholarship, neo-geographical methods and unprecedented public involvement. As such, our success will provide a powerful example, not just for classicists or geographers, but also for all humanists who seek to advance scholarship and serve society in an age of networked information.

3.1.2 Nature, size and content of Pleiades

What is it like to use Pleiades? The following user scenario – fictitious in its particulars, but accurate in its depiction of process and technology – was developed in 2002 to frame software design and development. In updated form, it is one of several that still guide our practice today.

> For the past 5 years, Dr. E. (an archaeologist and faculty member at a medium-sized Midwestern college) has been conducting walking surveys in southwestern Turkey. She has collected Global Positioning System (GPS) coordinates for several road stations and small fortification sites and

---

[4] William Smith's *Atlas of Ancient Geography,* London,1874.

made preliminary date estimates on the basis of surface-observable pottery. Dr. E. wishes to compare her data with that previously compiled by Clive Foss for his 1994 compilation of *Barrington Atlas Map 65, "Lycia – Pisidia,"* as well as any subsequent updates and new discoveries. Dr. E. logs in to the Pleiades website with her username and password. As an authenticated user, Dr. E. has several tools to compare her data with Pleiades content: full-text search for places, features, names, and reference citations; customizable collections of search results; and alternate versions of content for analysis in other applications like Google Earth. She can also access Harvard's Digital Atlas of Roman and Medieval Civilization (DARMC) for additional visualization and data options.

Using these tools, Dr. E. reinforces her opinion that seven of the sites constitute new discoveries. While logged in, she uploads her data into a private workspace on the site and adds contextual and descriptive prose. Thanks to epigraphic finds published since Dr. Foss's compilation, Dr. E. can also provide names for two fortifications that had been previously known, but could only be labeled in the Atlas (and in Pleiades) by their modern names. She checks out the corresponding Pleiades resources into a second personal workspace, and adds the names and bibliographic citations to the working copies. When satisfied with the results, she makes the new and modified resources in her workspace viewable to other members of the community.

Dr. E.'s edits garner attention from other members of the Pleiades Community who have been using customized web feeds to stay apprised of changes to Pleiades that intersect with their own interests. Within days, three members have posted comments concurring with Dr. E.'s edits. A particularly helpful response is posted by the director of a joint American-Turkish team excavating a nearby site (Dr. S.). A Norwegian graduate student (Mr. V.) attaches an English translation of Dr. S.'s Turkish-language comments so that non-Turkish-speaking members of the community can understand Dr. S.'s critique of some of Dr. E.'s dating criteria. A French epigrapher, Dr. C., voices doubt about one of the proposed name associations, citing a study of several recently published documents.

After incorporating these suggestions, Dr. E. submits the modified resources for review by the Managing Editors. Pending review, resources are locked against further edits. Upon logging in, an editor immediately sees the pending resources as action items and looks over the content and associated comments from other community members. Dr. E's bibliographic citations are easy to check because they all link directly to online versions (editors routinely ask other community members to assist in checking print citations). The editor contacts the Atlas compiler, Dr. Foss, and asks him to join the Pleiades community to assist in reviewing these new suggestions. Dr. Foss consents to do so and, after familiarizing himself with the on-line environment and community, provides the editors with a review of the suggestions first advanced by Dr. E. This review is shared with Dr. E. and the other substantive participants in the discussion.

Following further on-line discussion, members, reviewers and editors reach agreement on the character of the conclusions. The chief changes are addition of bibliography supplied by Dr. S. and Dr. C., and the recategorization of one of the epigraphically attested placenames as an unlocated place (partly because it was found in a context of reuse). The reviewed resources are sent back to the editing state while these modifications are made. Once finished, Dr. E. resubmits them for review, and the editor promptly commits the changes. This action immediately publishes the new resources to the web where they are available to any anonymous user of the site and are quickly indexed by the major search engines. They are also included in the next dated release of the Pleiades Gazetteer, a standalone Geographic Information System (GIS) dataset that can be downloaded from the Pleiades website and freely reused, just like the site itself, under the terms of an open-content license. In all these contexts, those who made substantive contributions to the identification and refinement of the suggestions are given authorial credit on each entry.

3.1.2.1   Nature and organization of content

Initial Pleiades content derives from the working materials that supported creation of the *Barrington Atlas*. An example set of approximately 2,000 such place resources have been published. During the first two years of the work proposed here, all materials will be processed. Pleiades will then document over 32,000 discrete places, including over 7,000 that the CAP's compilers classified as "unlocated toponyms." In adapting the Atlas content, Pleiades goes beyond print in terms of flexibility and extensibility, and beyond conventional GIS in addressing the special challenges of historical geography and scholarly practice. We have paid particular attention to the challenges presented by incomplete evidence and analytical uncertainty. Materials in Appendix A and B illustrate CAP materials and Pleiades content.

All Pleiades resources include standard fields for title, author, short description, copyright and licensing statement (hereafter "standard metadata"). It is also possible to attach bibliographic citations for secondary literature and extended prose essays to further describe or contextualize objects of interest, or to provide evidence or extended argument. These references can link to either print or digital works by way of an online bibliography that accommodates both print and online publications. Pleiades tracks the editorial history of each resource, providing users the ability to review changes to content over time and revert them as needed.

Our primary resources are classified **Place**, **Feature**, **Location**, or **Name**. Names describe historically-attested variant toponyms and *ethnica*, indicating the accuracy with which a variant is transmitted in known witnesses, the completeness of same (thereby accommodating fragmentary sources) and scholars' level of certainty in associating that variant with a given place. Named time periods may be assigned on the basis of a controlled vocabulary. In addition to the standard secondary bibliography, which here is concerned with linguistic or textual issues relating to the name, references to primary sources are added to indicate witnesses. A **location** represents a region of, or position on, the Earth's surface, its accuracy, and its provenance. It may be represented using any of the standard GIS types: points, lines, polygons, multi-points, multi-lines or multi-polygons and may be assigned to time periods just as names are. A **feature** associates attested locations with attested names. A bridge is a feature; a hilltop, a feature; the nominal center of a settlement extracted from the *Barrington Atlas*, a feature. Interoperability with conventional GIS systems is possible with these entities.

Places essay to model what the ancients knew about their world's geography. A place has no intrinsic location, but collects features and acquires location from them. A trade route is a place; "Athens," a place; "Roman Empire," a place. Other resources in the web of ancient world information will typically link to places for geographic context.

3.1.2.2   Using, creating and sharing geographic information

Unauthenticated visitors to the Pleiades website can find content relevant to their interests by browsing through lists organized by time period or by place type. A textual search capability is also provided. Web pages for individual place resources present the standard metadata and a contextual map, with links to the associated names, features and references. Any interested individual may request membership in the Pleiades community. Once logged in, members may suggest changes to existing content, or add entirely new content reflecting results of their own study. These resources can be built from scratch in the web interface, by uploading geospatial data files or by selecting existing, published content and marking it for revision. Such draft suggestions may be shared (for comment or co-authorship) with other members or the entire community prior to submitting a request for publication. A publication request invites review from all other members and from the Pleiades Editorial College, leading ultimately to a publication decision by the Managing Editors. Contributors retain copyright to their suggestions, but grant Pleiades permission to

reuse, remix and redistribute them. In this way, Pleiades replicates the basic components of scholarly communication and publication in a networked environment.

3.1.2.3   *Audience and community*

Fundamental patterns of information exchange are evolving rapidly. We would be short-sighted to produce a resource designed solely for passive consumption. Rather, we must bring together people, information and web systems in a dynamic way such that our users can not only consume, but also produce new knowledge and share it with others. This dynamic leads us to prefer the word "community" (borrowed from the open-source software and Web 2.0 world) over "audience" or "user base." Because we take an optimistic view of the potential for productive interaction between people from many different walks of life, the nascent Pleiades community cannot be succinctly described in terms of one or two existing constituencies within the conventional hierarchy of the academy. Therefore, it is perhaps clearest to describe that community by moving not up and down the ladder of academic experience and training, but outward from the project's center in the organizations that have invented it through the constellation of external research projects with which we are already collaborating to the broader and more diffuse constituencies on the horizon. In this way, we can enumerate the interests of our community members in an orderly way and highlight their value to the overall endeavor.

3.1.2.3.1 *Core partners: AWMC and ISAW*

Viewed from the center, Pleiades is a collaboration of two organizations that have much to gain from its realization: the Ancient World Mapping Center (AWMC) at the University of North Carolina (UNC-CH), Chapel Hill and the Institute for the Study of the Ancient World (ISAW) at New York University (NYU). Both see Pleiades as an essential element of digital infrastructure supporting their research agendas.

Pleiades will help the AWMC accomplish one of its core missions: the preservation, continuous updating and multi-modal dissemination of the rich research data assembled by the CAP. Pleiades also creates a channel that AWMC's small staff can exploit to collect, evaluate and report new work from around the world, informing a variety of research, mapping and teaching activities at the Center and beyond.

ISAW sees Pleiades as a cornerstone of its commitment to the preservation and long-term vitality of digital scholarly works for ancient studies. This effort, which relies on the NYU Libraries' digital archive infrastructure as well as ISAW's own staff, identifies and rescues at-risk digital documents, datasets and software of potential long-term value. Even after the entire CAP collection is accessioned to Pleiades (a key deliverable of the work proposed here), ISAW expects to use Pleiades as a conduit for the capture, regularization and review of other historical geographic datasets that will then be not only archived for future reference, but also updated and used on a perpetual basis. Beyond the preservation aspect, Pleiades is a central architectural component of ISAW's other digital initiatives. It will provide geographic authority for the description of digital images, primary source editions and secondary materials in the context of multiple archaeological, papyrological and epigraphic publications. In this way we set the conditions for seamless interoperation and reliable citation between all ISAW-sponsored digital publications. With a consistent basis for geographic reference in place, we can reliably and cheaply implement ubiquitous mapping and spatial search services within and among these publications. Users of any ISAW digital publication will benefit from reliable citation of relevant geography and ready access to contextual reference information through direct links to Pleiades.

3.1.2.3.2 *Collaborating Projects*

At Harvard University's Center for Geographic Analysis (CGA), Michael McCormick leads an interdisciplinary team in realizing the DARMC. They are turning paper maps, primary written sources,

and archaeological findings into a spatial database, then using web GIS technology to produce high quality online maps of Roman and Medieval Europe suitable for teaching and research. Pleiades and DARMC are cooperating to improve both resources and link them together to promote discovery, understanding and content development (see further, Section 3.2.1.3).

At King's College, London, born-digital epigraphic publications for Aphrodisias, Roman Cyrenaica and Roman Tripolitania are being crafted to rely upon Pleiades for consistent citation of find spots, modern observed locations and presumed original locations of inscribed objects. Bidirectional links between these corpora and Pleiades are expressed in XML documents (employing the Atom standard) to facilitate cross-resource, geographically aware search services. Development of this approach was funded by a Joint NEH/JISC Transatlantic Digitization Collaboration Grant entitled "Concordia" (see further Section 3.2.2.3). New locations, historical names and other information developed by the epigraphic field teams are expanding and enriching Pleiades.

Duke University leads the Integrating Digital Papyrology (IDP) project, a Mellon-funded international effort to revolutionize the publishing and revision of papyrological editions through digital collaboration methods. A team of scholars from 6 institutions, including ISAW, are upgrading and integrating some of the most venerable and influential digital resources in Classics: the Duke Databank of Documentary Papyri (DDBDP), the *Heidelberger Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens*, and the Advanced Papyrological Information System (APIS; funded by NEH). Via the Concordia interoperability framework, IDP will rely upon Pleiades for geographical authority and reference, and Pleiades will surface IDP content to its users.

A series of meetings (2009-2010) between the ISAW/Pleiades team and the staff of the *Epigraphische Datenbank Heidelberg* will lay the groundwork for similarly tight interoperation (funding from a DFG/NEH Bilateral Digital Humanities Workshop grant).

From early days, key staff of the American Numismatic Society (ANS), the UK Portable Antiquities Scheme (PAS) and the Lexicon of Greek Personal Names (LGPN) have served as technical advisors for Pleiades. All three organizations are developing web services that will facilitate interoperation with Pleiades. We will be able to alert our users to spatially relevant content in their collections, and they will be able to organize content for their users on the basis of Pleiades geography.

3.1.2.3.3 *Unaffiliated projects*

Because Pleiades uses well-known web standards and employs open-access and open-source licensing, external projects with their own technical staff need not confer closely with our team before making use of our content or software. We are unaware as yet of any unaffiliated projects that are making wholesale use of our content (it is, after all, geographically sparse at present and interested humanities projects tend to contact us directly), but components of Pleiades software are in use by non-profit and commercial entities in various fields. Some of these organizations also contributed programming time to the improvement of our software, and we have made substantive, similar contributions to third-party, open-source components that we use in Pleiades. We expect this trend to continue and, as Pleiades content builds out during the grant period proposed here, we anticipate learning of more third-party use.

3.1.2.3.4 *Individual scholars, teachers, students and enthusiasts*

Our artificial journey through the Pleiades user community brings us at last to engagement with individual users on the frontier. Their involvement is essential to Pleiades' success. Vibrant communities of practice form at the intersection of compelling content, useful features and user interest. Pleiades already boasts a small but engaged group of scholarly and technical users and, despite the present limited

amount of content available, continues to attract new interest from individuals around the world. Their potential impact is illustrated by the following unsolicited email:

> Subject: request username
> From: Peter Steadman <peter_steadman@hotmail.com>
> Date: Fri, 12 Jun 2009 09:36:35 +0000
> To: pleiades.admin@nyu.edu
>
> I've never used a wiki before, and I'm not a professional scholar, but I can't help noticing that even a casual user can correct some of your city positions in cases where the ancient walls are visible in Google Earth, or city plans have been published on the web. For example, note that the Pleiades point location for Taucheira/Arsinoe is outside the observable ancient city walls. Even the enthusiast website http://viae-romanae.pbworks.com, though counting many fewer sites, has better accuracy for those sites with visible remains. Speaking as a member of the general public, I would guess that most people will use your site to help them visualize what they see in Google. Placing lines or polygons over ancient features, or at least just city walls, seems to be an effective way to make the ancient world visualizable for the casual viewer.
>
> I'd love to offer to add [Keyhole Markup Language] KML files of outlines of sites to the Pleiades map, but I'm deterred by the technical problem that in Google Earth, outlines can only be drawn continuous, and polygons must be closed. Usually of course not all of a city's walls are visible, or inferable. The resultant inaccuracies are fine for an enthusiast site like viae romanae, but not for a scholarly site. If segments of a line could be drawn in different colors, then, for example, white could stand for archaeologically attested walls, brown for hypothesized. Perhaps there's a program that offers more controls for drawing KMLs than Google Earth? Also, since Thai university libraries are almost devoid of foreign scholarly books, I can't verify archaeologically attested remains that aren't visible on Google.
>
> Please advise.
> Best Regards,
> Peter Steadman
> Bangkok, Thailand

Mr. Steadman's email is delightful not only for its enthusiasm, but also for its sophistication. He has grasped the potential value of his own contribution, enabled by Pleiades. He is clearly engaged by the opportunity to contribute to a scholarly resource, and concerned to make his prospective contribution as precise and serviceable as possible. He has zeroed in on a key use case for Pleiades: exploiting Google Earth to improve horizontal coordinates over the 1-2 km accuracy resulting from the limited scales of the *Barrington* maps (another print constraint). His worries over accuracy in depicting extant remains (and the historicity thereof) highlight an anticipated benefit of Pleiades: he is just the sort of motivated amateur who, by supplying ample draft material and asking probing questions, will prompt lasting engagement from scholars better positioned to provide "ground truth" and informed analysis. Peter Steadman is exactly the kind of "enthusiast" community member we have designed Pleiades to attract.

3.2 History, scope, and duration

Although it traces some strands of its heritage as far back as the early 1980s, Pleiades was set in motion during the decade beginning in 2000. It began with self-funded work at the AWMC. The pace accelerated in 2006 with the award of a two-year Research and Development grant from the NEH. A mix of institutional and grant funding has continued work at AWMC and ISAW since 2008. We now have a solid software platform, collaborative network and interoperability framework that position us to shift to

advancing research through community collaboration. Accordingly, we will focus on content and community development during the period 2010-2013. Subsequent plans are addressed in Section 3.3.10.

3.2.1 Scope and duration of the proposed work

Proposed work will catalyze an animated user community, capable of carrying on collaborative content creation for many years to come. Objectives for the period of performance (May 2010 – April 2013) are laid out in the following sections. An implementation calendar is provided in the work plan section.

3.2.1.1 *Recruit new users*

Please see the "Dissemination, outreach and discovery" Section 3.6.

3.2.1.2 *Reward users by completing rollout of content from the Classical Atlas Project*

A major impediment to turning widespread interest into active engagement remains the limited amount of content currently presented in Pleiades. Pleiades 1 brought point features associated with two *Barrington Atlas* maps into Pleiades (Lycia and Cyrene), and funded preparatory work on two others (Moesia Inferior and southern Britannia). Under the Concordia rubric, features extracted from 5 more maps have been added (north and central Egypt and coastal North Africa west from the Delta to Tripolitania). Faster rollout of legacy data has been impeded to date by the cost of coordinate digitization from raster map images. To overcome this problem, we will fully implement a method we have recently developed for bypassing the costly step of map digitization and feature extraction, laid out in Section 3.3.8.

3.2.1.3 *Engage users by improving and expanding content*

Under this grant, we are budgeting for three activities that will significantly improve Pleiades content beyond the state in which it was inherited from CAP. These activities will demonstrate the range of engaging tasks available to members of the community, as well as their intrinsic value.

DARMC has digitized over 50% of the map locations in the *Barrington Atlas*, and has extended this dataset with coordinates for later features drawn from a number of sources. The project is now well advanced in replacing these with higher-resolution coordinates through visual analysis of imagery publicly available in Google Earth (nominal 20m horizontal accuracy as opposed to ~2km in the original data). According to the latest count, DARMC now holds about 20,000 named points. About 18,000 points are associated with the *Barrington Atlas*, and 2,900 are from the *Tabula Imperii Byzantini* map series. About 3,000 points have been geo-corrected. The DARMC database contains only coordinates and names, not the more detailed information derived from the full CAP compilation materials. With funding from this grant, we will work together to complete the coordinate improvement work, load it into Pleiades and join it to the existing content already there, providing better coordinates and new entries for later sites. Harvard will also incorporate a copy of the newly joined content into their own online mapping environment. These actions will demonstrate not only the value of contributions to Pleiades but also the reusability and remixability of its content. Thereafter, hyperlinks between the two projects will facilitate side-by-side use of their different visualization tools and access to unique additional content in each project. A workshop, to be held at Harvard's Center for Geographic Analysis early in year 1 will enable both teams to efficiently evaluate progress to date, identify challenges and set a detailed agenda for completion of this integration during the first year.

We will continue to pursue productive relationships with other projects and information providers like the ANS, IDP, LGPN and PAS. We will invite them to specify requirements for content and interoperability upgrades early in the project. These tasks will be prioritized alongside upgrade suggestions made by

individual users for implementation during the period of performance. This collaborative group will be expanded to include the Regnum Francorum Online (RFO), a website that geographically presents early medieval Europe 614-840 on maps. Its web application documents persons, cities, institutions and personal names, linking them to primary sources. Its creator, Johan Åhlfeldt, has agreed to attend the Pleiades/DARMC workshop in Boston in year 1 (travel funding requested here). The meeting will give all three projects an opportunity to compare approaches and set an agenda for future collaboration.

~~During the 2009-2010 academic year, Nicola Aravecchia (a Visiting Research Scholar at ISAW) will be laying the foundation for a comprehensive archaeological gazetteer of Christian Egypt. ISAW will contribute the last four months of his support to this project so he can focus on incorporating this information into Pleiades. In the last few decades, scholars interested in Late Antiquity have been engaged in unprecedented, rigorous inter-disciplinary study of Christian sites and monuments in Egypt. This work has substantially increased the amount of evidence at hand for the study of early Christianity in the region, shedding light on key historical, social, economic, and artistic aspects. This information is frequently difficult of access to scholars and graduate students, and practically invisible to the educated public. During May-August 2010, Aravecchia will begin to rectify this deficit by contributing to Pleiades the geographic coordinates for a number of Coptic sites gathered previously through on-site survey using GPS hardware. Many of these sites are not registered in the CAP materials. He will also add geographic names, appropriate bibliography and introductory texts developed during the preceding 8 months. He will further ensure linkage to digital imagery for many of these sites, published to the web via ISAW's Flickr account and archived in the NYU Faculty Digital Archive (FDA). Beyond the intrinsic value of Aravecchia's contributions, his efforts will illustrate Pleiades' utility as a platform for disseminating underrepresented information. Concurrently, Aravecchia and Bagnall will actively recruit other experts in the field of Coptic/Christian Egypt, especially directors of archaeological investigations, to voluntarily contribute to the effort by enhancing resources, writing texts and providing images.~~

During the period of performance, ISAW and AWMC will work together to enhance some of the ancient name information brought forward from the CAP. Over 1,000 name resources will be affected through the addition of minor variants, appropriate Greek orthography and primary source references that attest to the historicity of given variants. This information was assembled during Pleiades 1 and archived pending maturity of the web application. For technical reasons related to the software used to produce some of the *Barrington* maps, Greek names were transliterated using Roman characters. As an economizing measure (given that CAP was conceived to create a print resource), minor variants were omitted from the *Map-by-Map Directory* (e.g., across languages as the Greek "Ephesos" for the more familiar Latinate "Ephesus"). For similar reasons, primary sources containing variants were not cited when an authoritative secondary work likely to provide them could be cited instead (e.g., a monograph or comprehensive article in a scholarly dictionary or journal). Filling these print-necessary holes will make digital Pleiades content more flexible. Users will be more likely to find matches for searches and emerging web services will be exploitable for direct links to primary and secondary sources alike. Pleiades will also be more useful as an authoritative source of names, suitable for use in geoparsing. Moreover, the editors have imposed a requirement to cite at least one witness for each new name variant suggested by members for publication. Therefore an effort to provide these for those brought forward from CAP is essential.

3.2.1.4 *Improve the user experience*

Community members must feel confident in the value and visibility of their contributions if they are to continue their engagement with Pleiades. The ease with which they can alter or add information, and use contents for their own purposes, will play a major role in building that confidence. We have already identified several high priority software improvements that will receive attention if the present proposal is funded. They will surely be joined by more as the community works with the production system. We expect this process to continue throughout the period of performance, and will take an agile approach to

implementing community-generated requirements. This entails working closely with the community member(s) who have identified the need to design, develop and deploy an upgrade or solution with minimal turnaround time. A continuously updated list of identified tasks and priorities is maintained in the Pleiades online issue tracker, where it can be reviewed by any interested party.[5] The identification and elaboration of specific requirements will be managed using the tracker. Priorities will be set by the PI in consultation with the Chief Engineer, editors and community members at large.

Documentation and training that helps community members learn how to use the system is critically important. We will conduct webinars and webcasts, and add textual content on our development wiki, to reach out to our users (example webcast at http://www.atlantides.org/video/search-and-browse.mov). Priorities will be set by the community and tickets for the completion of same created in the issue tracker. Wherever possible, we will engage the community in helping us design and execute these materials. This activity will begin in earnest following our first quarterly Community Review.

3.2.1.5 Migrate bibliography to Zotero

The Pleiades bibliography (http://www.atlantides.org/pleiades) is published using custom-designed software that extracts citations from the CAP compilation materials, verifies them against a bibliographic database developed by the AWMC and publishes them to the web (see further, Section 3.3.3). This solution does not scale to the contribution of new bibliography by the community. Fortunately, the open-source Zotero citation management system, as of its June 2009 release, provides for stable identifiers, collaborative authoring and open access. We will modify our software to produce Zotero-ready data files from the CAP bibliography. These will be used to populate a Pleiades group on the Zotero server (http://www.zotero.org/groups/pleiades). When people join the community, they will also be given access to the Pleiades group where they can post full citations for anything they cite. Pleiades software will be modified to facilitate linkage to content in Zotero.

3.2.1.6 Facilitate reuse and remixing of content

Although Pleiades boasts a rich array of mechanisms for dynamic, web-based sharing of content, many parties have requested a downloadable version suitable for use in conventional GIS. We will address this need by implementing export to ESRI Shapefiles. Although the Shapefile format is proprietary, it is used around the world in a variety of commercial and open-source GIS systems and can be readily decoded by third-party and open-source software. Pleiades Shapefiles will include name and other information drawn from Pleiades content, as well as the URI that identifies the corresponding place resource in Pleiades.

3.2.1.7 Lead users to excellent related content on other sites

We will modify the Pleiades user interface so that, as external resources link to those of Pleiades via the Concordia interoperability framework, they will be discovered and surfaced to authenticated users. Users may create reciprocal links to excellent, relevant resources, and thereby contribute to the emergence of linked ancient world data. In this way, we hope to do our part to catalyze a federation of ancient world web sites (see further Section 3.2.2.3 regarding Concordia).

3.2.1.8 Protect users' investment through digital archiving

Pleiades community members will contribute content because they hope it will be useful to other members of the community and to scholarship in general. This expectation naturally extends forward in time. Although we continue vigorous efforts to ensure the operational longevity of Pleiades on the web,

---

[5] http://atlantides.org/trac/pleiades/roadmap; a printed summary of tracker contents as of July 2009 is included in Appendix F.

prudence demands steps to secure the long-term digital archiving of Pleiades content. During the period of performance, we will integrate Pleiades workflow with ISAW's existing digital archive pipeline. On an annual basis (or more frequently as judged appropriate by the Managing Editors) all published Pleiades content will be exported to a combination of eXtensible Hypertext Markup Language (XHTML) and KML files and accessioned, along with appropriate, complete metadata, to the FDA. A copy of each deposit package will also be uploaded to the Internet Archive for additional security. Software modifications to realize these packages and associated metadata will be implemented during the period of performance.

3.2.2 History of the entire project

3.2.2.1 *Background and early developments (1980 – 2006)*

Pleiades traces its origins to 1980. That year marked the publication of the final report of a committee chartered by the APA to make recommendations for the improvement of basic research tools for the study of Classics.[6] The report, edited by Roger Bagnall, singled out historical cartography – and in particular the lack of a comprehensive, up-to-date reference atlas – as a matter of deep concern requiring immediate action. The committee's recommendation led to the formation of the CAP, which was directed by Richard Talbert. Bagnall served as a regional coordinator for this effort, supervising the work of scholarly compilers who researched and prepared information for the maps. During the period 1988-2000 and with support from a professional cartographic firm and scholars worldwide, the CAP took in hand the process of researching, compiling and producing the *Barrington Atlas,* assembling for the first time in over a century a comprehensive synthesis of scholarly findings in the area of Greek and Roman geography.

The AWMC was founded in 2000 by the College of Arts and Sciences at UNC-CH. A key motivating factor was the realization that the pace of technological change was opening up new horizons in historical geographic study. The Center was given the mission of updating the information assembled by CAP in the light of new scholarship, discoveries and technologies. Tom Elliott was appointed as the AWMC's founding director, and continued in this position through early 2006, when he stepped down to assume leadership of Pleiades itself. During his tenure as Director, Elliott drew on his experience and training as a software developer, a Roman historian and a veteran of the CAP to plan for the transition of the CAP's records to an all-digital information system to support the Center's mission. Preparatory steps included the development of a bibliographic database embracing all works cited in the Atlas and *Map-by-Map Directories*, as well as the creation of prototype software to parse and clean up the full lists of name variants, periods of occupation, location descriptions and bibliographic citations previously assembled.

By 2003 it was clear that creating a digital gazetteer from the CAP compilation materials was an urgent need. Technological change, already noted during the final years of the atlas project, continued to quicken. The passage of time was threatening the currency of the information itself. These concerns, coupled with the realization that a small Center at a single state-funded institution could not hope to keep pace with all relevant developments across the entire spatial, linguistic and methodological footprint of the *Barrington Atlas*, motivated interest in a collaborative solution. The example of the *Suda Online Project* (SOL) was particularly helpful. SOL combines web-based collaboration with active editorial control and open review to translate a massive Byzantine encyclopedia of significant historical value. It has achieved remarkable success. In consultation with the late Ross Scaife (then SOL editor), Talbert and Elliott devised a plan that would bring SOL's model together with the editorial practices employed by CAP around a structured, digital gazetteer. Interested parties around the world would be invited to participate in the improvement of the gazetteer as an increasingly essential tool. Since reviewers of an

---

[6] Roger Bagnall (ed.), *Research Tools for the Classics: The Report of the American Philological Association's Ad Hoc Committee on Basic Research Tools*, Scholar's Press, 1980, ISBN: 0891304525.

initial NEH proposal recommended we pre-select the basic software framework for our web application before receiving funding, we investigated and came to the conclusion that the open-source Plone content management system provided the best base platform. A second proposal for research and development funding was granted, and work began on Pleiades 1 in February 2006.

3.2.2.2   Pleiades 1 (February 2006 – June 2008)

For Pleiades 1 we promised to produce a "spatially-enabled community support system enabling interested parties worldwide to participate in the maintenance, diversification and beneficial reuse" of the CAP project data. The work plan called for two years of iterative design, development and testing (February 2006 – January 2008). A no-cost extension stretched the period of performance through June 2008. The final report is available online at http://www.atlantides.org/reports/pleiades1/.

During 2008, ISAW recruited Elliott to be Associate Director for Digital Programs. Consequently, the locus of technical development has shifted there. ISAW is better equipped to ensure Pleiades' long-term success. It is a center for scholarly research and graduate education that cultivates comparative and connective investigations of the ancient world. ISAW's digital programs are fundamental to its mission. ISAW's ambitious slate of digital initiatives extend well beyond the institute. They emphasize the delivery of primary and secondary texts and images, as well as geographic and archaeological reference information. Pleiades is key to ISAW's goal of serving the field of ancient studies by promoting the durability of digital publications and the sustainability of the projects that produce them.

3.2.2.3   Concordia (April 2008 – March 2010)

Before applying for a second major grant to transform Pleiades from a prototype to a production-level resource, it seemed useful not only to have the basic system up and available, but to test its utility in a practical way by linking it to other types of ancient-world digital publications. The Concordia Project (http://concordia.atlantides.org) is elaborating a standards-based interoperability framework for ancient studies. It will connect papyrological documents, epigraphic texts, images and historical geographic information. A collaboration between the Pleiades team and the Centre for Computing in the Humanities at King's College, London, Concordia is supported through a Joint Transatlantic Digitization Collaboration grant by the NEH and the Higher Education funding Council for England of the United Kingdom acting through the Joint Information Systems Committee.

The Concordia framework employs Atom web feeds that describe the resources published by participating web applications and the links between them. These links are annotated in a standard way to express relationships of interest to humanists (e.g., "findspot" or "isEditionOf"). Once indexed and fitted with a web-facing query interface, these typed links enable a variety of discovery services, for example: proximity search, browse-by-findspot, or dynamic mapping. Terms used for these link types are documented at http://atlantides.org/trac/concordia/wiki/ConcordiaThesaurus . Example feeds showing links to Pleiades place resources from the standard digital publication of the inscriptions of Aphrodisias in Turkey may be examined at http://www.atlantides.org/trac/concordia/wiki/ConcordiaAtomFeeds.

3.3 Methodology and standards

3.3.1 Software engineering

Pleiades follows the best practices of modern software engineering. Code is managed in a Subversion (SVN) repository, the revision control system used by the Apache Software Foundation. We design our open-source software for reuse by broader web and GIS communities, encouraging in-kind contributions

and improvements. It is rigorously tested, and deployed using an automated and reproducible process. Application server load balancing and reverse proxy caching provide a high level of service.

3.3.2 Organization of content and data structure

Pleiades follows the architecture of the Web itself and so inherits its strengths: ubiquity, scalability, and capacity to evolve with time. Places, features, names, and locations are identified by permanent Uniform Resource Identifiers (URIs) and have multiple representations. Data-centric representation variants, including JavaScript Object Notation (JSON) and Keyhole Markup Language (KML), serve the needs of other web applications, including Google Earth. Text-centric variants like Hypertext Markup Language (HTML) are for consultation by individuals. Atom syndication feeds allow human users, as well as digital agents such as search engines, to discover the resources of Pleiades.

3.3.3 Scholarly bibliographic citation

Pleiades editorial policy mandates appropriate citation of primary and secondary sources for all content. New research may be published (with appropriate argument) in Pleiades, but reliance on prior work must be signaled by citing supporting secondary literature in the relevant resource. New name resources may not be published without verifiable citation of a reliable witness (primary source). Within the Pleiades web application, both types of citation are supported with simple data objects that contain reference details like page numbers together with URIs corresponding to bibliographic records for the works cited. For Pleiades 1, we created bibliographic records for individual works in accordance with the Metadata Object Description Schema (MODS) developed by the Library of Congress. These, and XHTML versions thereof, are provided for users via the Pleiades bibliography. The XHTML pages include embedded bibliographic data in the OpenURL Context Objects in Spans (COinS) specification to facilitate harvesting with the free Zotero citation tool. If funded, Pleiades 2 will migrate all bibliographic records to a shared group on the Zotero server (see Section 3.2.1.5).

3.3.4 Pleiades thesaurus (controlled vocabularies):

Several fields in Pleiades resources take discrete values. Eight vocabularies establish standard codes for languages and writing systems; type, accuracy, and completeness of historical names; types of places and features; named historical periods and levels of confidence assigned to them; and levels of certainty with regard to the association of attested names with locations. Files are maintained in the Pleiades SVN repository in a standard format for ease of use with the vocabulary manager, but are also published directly on the Web for use by other digital humanities projects (http://pleiades.stoa.org/thesaurus/). We have also adapted the Pleiades editorial workflow to the thesaurus. This permits community members to recommend new terms that can then be reviewed just as other suggestions are.

3.3.5 Tagging schemas and character encoding

Pleiades encodes information with standard schemas and languages. All web pages on the site conform to the World-Wide Web Consortium's (W3C) XHTML 1.0 Transitional schema. If funded, work proposed here will enrich these documents by annotating them with semantic markup using the W3C's Resource Description Framework in Attributes (RDFa) recommendation. Web feeds for resources and collections employ the widely used Atom Syndication Format, which is recognized by both the W3C and the Internet Engineering Task Force (IETF). Geographic coordinates are included in Atom feeds with the GeoRSS application profile, which itself uses the Geography Markup Language (GML) promulgated by the Open Geospatial Consortium (OGC). Pleiades provides ready interoperation with Google Earth via documents encoded in KML, another OGC standard. The Pleiades Thesaurus is encoded in the Vocabulary Definition Exchange (VDEX) format. In preparing legacy CAP content for inclusion in Pleiades, the

project employs a hybrid XML schema designed to ensure unambiguous, lossless transition of content. This schema draws upon several well-known namespaces for appropriate content elements, including the Dublin Core Metadata Element Set, the Text Encoding Initiative Guidelines for Electronic Text Encoding and Interchange and the Alexandria Digital Library Gazetteer Content Standard. Regardless of the language or file format, all character text is encoded in the 8-bit Unicode Transformation Format (UTF-8) established by the Unicode Consortium. Metadata concerning languages and writing systems (including the associated controlled vocabulary) uses the standard web syntax and identifier system.[7]

3.3.6 Updating, participation, editorial revision and workflow

Pleiades employs a broadly collaborative, transparent editorial process that cultivates contributions and critiques from the widest possible range of contributors. The bar to initial participation is purposely set low to encourage participation by scholars, students and enthusiasts alike, regardless of their degree status, institutional affiliation or skill level. All that is required is a verifiable email address and acceptance of a contributor agreement governing issues of professionalism, mutual respect, intellectual property, editorial policy and assertion of identity. See Appendix C.

Pleiades workflow is described in Section 3.3.6 and illustrated at the beginning of Section 3.1.2.

Review and improvement of content is a community concern. Every authenticated user is empowered to endorse or critique content developed by others. The Managing Editors, in consultation with the Senior Editors, set the policies that govern community review, and they control final publication decisions. They also, from time to time, name community members to the Pleiades Editorial College in recognition of their subject-matter expertise, history of substantive contributions and demonstrated collaborative spirit. Although the members of the Editorial College acquire no additional power over the mechanics of the review and publication process, the Managing Editors may call upon them to provide expert opinions in difficult cases or to work with other members in bringing their proposals up to publication readiness.

Roger Bagnall and Richard Talbert serve as Senior Editors. Tom Elliott and the AWMC Director serve as Managing Editors. Senior Editors may name Managing Editors as workload increases and suitable candidates are identified among the members of the Editorial College. Veterans of the Pleiades 1 Steering Committee and the CAP are automatically named to the Editorial College.

3.3.7 Hardware and software

Pleiades is entirely open source. We use the standard combination of Linux and Apache for our operating system and web server. The rest of our "software stack" is written in the Python programming language. Plone is a leading open source content management system that we use to provide the framework and base functions for our web application. It incorporates the Zope Object Database (ZODB), which provides object-oriented database storage and retrieval for all content. OpenLayers, the leading open source web map toolkit, provides contextual maps in the user interface. We have developed two general-purpose Python GIS software components, Shapely and Rtree, that are used throughout the system. Reuse of Pleiades software by other projects is documented at http://trac.gispython.org/lab/wiki/UserTestimonials.

The Pleiades web application and development environment run on a server leased from tummy.com ltd. The server is a dedicated piece of hardware and sits in tummy.com's data center in Fort Collins, Colorado. It is equipped with an Intel Core Duo processor, 1 gigabyte (GB) of Random Access Memory

---

[7] IETF *Best Current Practice*, no. 46 = A. Phillips and M. Davis (eds.), *Tags for Identifying Languages*, RFC 4646, September 2006, http://rfc.net/rfc4646.html ; *Internet Assigned Numbers Authority (IANA) Language Subtag Registry*, http://www.iana.org/assignments/language-subtag-registry . Other formats are defined and referenced in the Glossary.

(RAM) and dual 500GB hard drives, configured as a Redundant Array of Independent Disks (RAID). Tummy.com provides 24-hour, 7-day-week system and network monitoring, UPS and generator-protected power, operating system configuration and upgrade and regular backups. The Pleiades development team has secure, root-level configuration and management access to the machine.

3.3.8 Implementation of data processing

Three data processing tasks will be executed during the period of performance. All legacy CAP content will be migrated to Pleiades using a streamlined technique that is quicker and more affordable than the one employed for the sample materials treated in Pleiades 1. Once this task is complete, coordinates from the DARMC database will be loaded and either joined to existing content or used to create new entries. Separately, Pleiades bibliographic records will be migrated to Zotero-compatible formats.

Entries will be parsed from the CAP *Map-by-Map Directory* documents (MS-Word format) and the data used to initialize Pleiades name, location, feature, and place resources. Coordinates derived from modern locations using the open-content GeoNames database and from CAP grid squares will be used in preliminary locations. This method permits us to bypass manual digitization of map coordinates and entry of names from CAP's map sheets. Given that the nominal accuracy of CAP coordinates is on the order of 1-2 km (a function of original map scale) and that DARMC has already made significant strides in sourcing coordinates with much greater accuracy, we feel strongly that retrospective CAP feature extraction should be done only on a selective basis. We are happy to defer this task – which should only be necessary for those places that have not been treated by DARMC and are not visible in Google imagery – until after place resources have been created for all CAP content. Software to parse the *Map-by-Map Directory* files has already been written and tested. A modest amount of development time will be required to add a module capable of implementing coordinates based on grid squares and GeoNames.

An existing module for upload of the files produced by our parsing software will be used to bring legacy content into Pleiades. More significant programming time will be needed to add support to Pleiades for named relationships between places. This modification will permit modeling of roads and other network features in the absence of accurate coordinates for all points: we will create a named relationship for road connections between places. With this modification in place, we will be able to capture the bibliography and other information that CAP compilers associated with groups of road segments by way of multi-branched itineraries. The Pleiades community will then be able to assign more accurate coordinates to the nodes in these networks, and to the routes of the roads themselves.

Joining of DARMC coordinates to Pleiades content will be done by matching CAP labels across the two collections after limiting the list of possible matches on the basis of geographic proximity. Modification to the Pleiades web application will be required to support posting of GIS data that creates new features linked to existing places and to resolve conflicts between incoming and pre-existing features. This will entail improvements to the user interface for individual workspaces so that conflicting data is collected there for review and manual intervention. These enhancements will be useful to all Pleiades users.

Pleiades bibliographic records are already in one of the formats supported by Zotero; however, Pleiades encodes information about "parent" works (e.g., journal details in an article citation) using a different set of options than those recognized by Zotero. We will write software to modify Pleiades encoding to ensure records are transferred to Zotero with no loss of information.

3.3.9 Preservation

In collaboration with the NYU Digital Libraries Technology Services group, ISAW has established an archival workflow for its scholarly digital assets. This workflow takes advantage of NYU's FDA, an

institutional repository system built on the same infrastructure as the Library's own digital preservation repository. Backed by an institutional mandate to ensure the long-term availability of all deposited content, the FDA provides robust storage both on-and-offsite, persistent Uniform Resource Locators (URLs) via the Handle System for discovery and citation, and templates for standard Dublin Core metadata. To enhance the longer-term utility of ISAW assets, we have implemented the full preservation metadata standard used internally by the libraries for archiving projects such as the Afghanistan Digital Library. ISAW's archival workflow is already in place to support its digital imagery project. If this proposal is funded, it will be expanded to support Pleiades content as outlined in Section 3.2.1.8.

Adoption of widely used standard formats, Web 2.0 publication models, software engineering best practices and institutional embedding establishes a firm foundation for the future. Should it prove necessary to re-host Pleiades, moving from tummy.com should pose no difficulties. We have a solid system for reproducible deployments and can recreate our production environment on any Linux/Unix system in less than 30 minutes. The ZODB is portable. We would copy it to the new production environment and relaunch. We believe that both Plone and our dissemination variants will give us good mileage; however, change must be expected in any web application. Formats will continue to evolve and we will adjust our interfaces appropriately, as we have done several times already. It is possible that Pleiades could outgrow the ZODB, but we would be surprised if this happens within the next decade. We can easily move the database file to a dedicated, higher performance server when necessary. If the geographic component of our data grows too large, we could partition it to a separate Postgres/PostGIS database. The Plone community has experience with such hybrid storage operations, and eagerness to share it. Any migration would use XML as a format. We are already using a set of XML namespaces for data processing and input that will serve us through any migration. XML is a common element of our shared infrastructure, facilitating archiving and interoperation (via Concordia) and even underpinning an increasing number of the primary source publications with which Pleiades will be interacting.

3.3.10 Sustainability: maintenance and support beyond period of grant

Pleiades is built on the idea that distributed, open collaboration can beat centralized content-creation approaches. When this grant period ends, day-to-day operation of Pleiades will be jointly embedded in the on-going activities of the AWMC, ISAW and our users. Senior and managing editorship are already shared jointly by the two institutions. The project is designed to require, when mature, no central academic staff beyond that already associated with the two institutions. A robust community of collaborators around the world will share the costs and benefits of on-going content development.

It is also our view that reliable institutional technology hosts are required for stability of online reference services. ISAW is committed to providing this function for Pleiades. ISAW has already expended its own funds to facilitate Elliott's work on the project prior to the start of the Concordia grant, and has continued to do so for both Elliott and Gillies (alongside other duties) since their support ran out on Concordia in May 2009. ISAW has shouldered, from January 2009, the on-going costs of hosting Pleiades development and production environments and is contributing same to the total project costs on this grant.

A growing number of projects and resources originating both at ISAW and elsewhere are to be built around the interoperability framework and comprehensive geographic authority services Pleiades will provide. ISAW's digital image collection, now in pre-production, as well as the IDP project and the Epigraphic Databank in Heidelberg are maturing plans of this sort. These are the first fruits of ISAW's commitment to the preservation and long-term vitality of born-digital scholarly works for ancient studies. Drawing on permanent staff and on partnerships with other institutions, ISAW will foster the creation and conversion of uniquely valuable digital resources in open, standards-based formats recommended by the archival community. Wherever possible, copies of these resources will be placed into collaborative software environments and opened to broadly collaborative update and maintenance in order to reap the

benefits of public, persistent creation and curation. A combination of projected internal, endowment and spend-down funds will support these allied services.

During the decade following completion of the work described in this proposal, we anticipate that Pleiades (together with allied projects) will have periodic recourse to additional public and foundation funding for specific tasks beyond the scope of day-to-day operations. Key milestones will include 1-2 major software upgrades to keep pace with rapidly changing technology. We will also, from time to time, seek smaller grants to finance the creation, completion or conversion of unique and valuable collections that can be enhanced or preserved by addition to Pleiades or integration with the Concordia framework.

3.3.11   Intellectual property

Intellectual property considerations have been fully addressed. Compilers' contributions to the CAP were compensated and governed by contracts with the APA. The APA has granted to the AWMC and its partners permission to use and distribute these materials for the benefit of the field with the proviso that any profits deriving from said use are to be shared with the APA. Free distribution of content under Creative Commons license will produce no profit. Should a third party approach Pleiades to negotiate use of content under other terms and for a fee, appropriate negotiations would involve the APA. Contributions of content or software by individual Pleiades users are governed by the Pleiades Contributor agreement, explicitly leaving any intellectual property inherent in the contribution in the possession of the contributor. Pleiades and its supporting institutions are granted non-revocable rights under appropriate license to reuse, remix and redistribute the content (see Appendix C).

3.3.12   Geospatial data uniqueness and standards conformance

Pleiades content is unique and will only grow more so over time. In accordance with NEH proposal guidance, we conducted a thorough search of the Geospatial One-Stop (GOS) Portal. We have verified that there are no geospatial datasets registered therein that remotely approach the coverage, completeness and scholarly value of the materials that will be incorporated into Pleiades under this funding.

Applicable standards promulgated by the FGDC will be followed. As Pleiades deals with locations outside the United States and as much of its content is historical, the majority of FGDC guidance does not apply; however, provisions of the base CSDGM have been carefully examined for relevance. The architecture, goals and content of Pleiades mark it as much more than the standalone "spatial datasets" that concerned the original FGDC drafters. It is conceived as a live, constantly improving resource, designed to work in concert with other systems of interest to humanists: digital libraries, linked data services and semantic web components. Where Pleiades content is expressed as a dataset – namely, for archiving – appropriate CSDGM documentation will be created by the Managing Editors and included in the package (see further Section 3.2.1.8). When live in the web application, the Pleiades data model itself provides documentation at the level of individual resources. Pleiades already records Dublin Core standard metadata for identification, attribution, rights information, and content history. Each position resource links to an accuracy assessment, which provides spatial metadata for the geometry: horizontal accuracy, a prose description of origin and, where possible, a copy of the original spatial data file that was uploaded when the position was created.

3.4 Work plan

The following schedule sets benchmarks for completion of objectives outlined in the scope of work. Programmatic milestones (e.g., reporting) are also included. Staff, volunteers and service providers with major roles in each task area are listed. As PI, Elliott will provide oversight and design input on all tasks; his name has been included only in those areas where he will take a substantial hands-on or technical role.

PROJECT WORK PLAN

	2010	2011	2012	2013
	M J J A S O N D	J F M A M J J A S O N D	J F M A M J J A S O N D	J F M A

GRANT YEAR 1 | **GRANT YEAR 2** | **GRANT YEAR 3**

MEETINGS
- DARMC workshop — Elliott, Gillies, CGA, Ålfeldt
- qrtrly community review (Gillies, Elliott)

OUTREACH (Elliott/Gillies)
- survey collab groups
- blogging/posting
- community status report
- APA/AIA
- DH2011

PROGRAMMATICS
- NEH interim reports (Elliott, Gillies)
- NEH final report (Elliott, Gillies)
- Accounting setup and review (Elliott)

USER-DRIVEN UPGRADES
- agile response to community — Gillies, Elliott, AWMC

CONTENT
- editorial oversight — Elliott, AWMC with Bagnall, Talbert
- roll out CAP content — Gillies
- DARMC data/integration — Gilles, Elliott, CGA
- Name improvements — Elliott, AWMC

USER EXPERIENCE
- enhance workspaces/kml — Gillies
- improve search — Gillies
- upgrade map interface — Gillies
- add webcasts — Gillies, Elliott
- support for relationships — Gillies

INTEROP/DISCOVERY
- federation/Concordia interface — Gillies
- RDFa/search engine optimization — Gillies
- Shapefile export — Gillies

INFRASTRUCTURE
- archive deposit (minimal metadata) — Elliott, AWMC / Elliott, AWMC
- develop full metadata package — Gillies, Elliott
- archive deposit (full metadata) — Elliott, AWMC
- refine full metadata package — Gillies, Elliott
- Zotero integration — Gillies, Elliott

3.5 Staff

ISAW will directly employ 3 individuals working on this project as indicated in the budget:

- **Dr. Tom Elliott: Associate Director for Digital Programs / Senior Research Scholar (ISAW)**
  Elliott is the **Principal Investigator** for the Pleiades Project. As such, he will also exercise the role of co-managing editor (alongside the Director of the AWMC). He will direct all aspects of Pleiades operation, enhancement and community building to include supervision of the Chief Engineer (Gillies). Elliott will devote 20% of his work time weekly to the project.

- **Sean Gillies: Chief Software Engineer (ISAW)**
  Gillies will continue to lead and coordinate all aspects of Pleiades software development and data conversion as project **Chief Engineer**. He will work closely with Elliott to prioritize user requirements and translate them into working code. He will continue to work with users and developers both within and outside the Pleiades community to produce the best possible user experience and ensure all necessary improvements and releases of open-source software. 100% of Gillies' time will be devoted to this project.

- **Nicola Aravecchia: Visiting Research Scholar (ISAW)**
  See Section 3.2.1.3 for details of Aravecchia's work. 100% of his time from May-August 2010 will be devoted to this project.

As explained above, Pleiades relies upon the contributed work of a growing community of collaborators worldwide. Their day-to-day contributions cannot be quantified directly in advance, and so are not treated specifically in the budget; they undertake Pleiades work as part of their day-to-day scholarly activities. We do, however, provide information here about the following key individuals from this class:

- **Prof. Roger Bagnall: Director / Professor of Ancient History (ISAW)**
  Bagnall serves as Co-Senior Editor, providing policy guidance and editorial assistance to the Managing Editors and the community at large.

- **Prof. Michael McCormick: Francis Goelet Professor of Medieval History (Harvard)**
  McCormick serves as a member of the Editorial College, and advises the Managing Editors on publication decisions related to information contributed from the DARMC project.

- **Prof. Richard Talbert: Kenan Professor of History and Classics (UNC-CH)**
  Talbert serves as Co-Senior Editor with Bagnall. He also provides liaison to the Advisory Board of the AWMC, whose endorsement helped initiate the Pleiades project.

- **Director, Ancient World Mapping Center (UNC-CH)**
  The incumbent AWMC Director (currently Brian Turner) serves as Co-Managing Editor, sharing policy and publication decisions with Elliott.

3.6 Dissemination, outreach and discovery

Pleiades provides free, online access to all its digital resources. On-site search, ubiquitous web feeds and customizable virtual collections facilitate content discovery and change alerting for existing users. The first challenge for our community-building phase is enabling those not already familiar with the project to discover it. We have a three-pronged strategy for this: aggressive outreach, expansive interoperation with other online resources, and search engine optimization.

Because our potential community base spans multiple constituencies, we cannot confine outreach to a single venue. We propose to continue an established tradition (from CAP days) of communicating the project's progress to professional classicists and archaeologists through presentations and ad hoc focus groups at the annual joint meetings of the APA and Archaeological Institute of America (~~January 201~~ and 2012). Already fruitful connections with humanists engaged in online publication and research will be reinforced and expanded through a similarly paired presentation/focus group at the international Digital Humanities conference in 2011. Community members and others already aware of the project will be kept informed about changes in content, feature upgrades and other relevant matters through quarterly Community Reports. Like a changelog for a software release, these reports will concisely summarize all major changes since the last report, linking to additional information as necessary. These reports will be announced via the community email list, the PI's blog (http://horothesia.blogspot.com) and twitter. Staff will hold online meetings in the Pleiades Internet Relay Chat (IRC) channel to gather reactions and feedback from interested parties.

To reach a more dispersed and heterogeneous audience of both international scholars and enthusiasts, we will expand our reliance on digital publication. The Pleiades project staff will continue their practice of regular blogging and posting to a wide variety of discussion lists and online fora (aggregated here: http://planet.atlantides.org/pleiades/). We will also continue to maintain and elaborate the publicly accessible documentation and design wiki, issue tracker and SVN installation established and used during Pleiades 1. Under this grant we will expand our online offerings to include short "webcasts": narrated videos showing users how to carry out common tasks with our web application. We will also continue our practice of announcing software release packages via the Python Package Index, the Plone Product Release List and other appropriate venues.

Our interoperation strategy is built on the Concordia framework outlined in Sections 3.2.1.7 and 3.2.2.3.

Prospective Pleiades users are already finding their way to our content via the major search engines. The web crawling software (bots) used by Google, Microsoft and Yahoo! all parse and properly index the content variants surfaced by our web applications. This means that a search for "Ninoe" on bing.com, search.yahoo.com or google.com will produce a link to the corresponding place resource in Pleiades (http://pleiades.stoa.org/places/638753). Location-aware search services, like the "Geographic Web" layer in Google Earth, also yield links to Pleiades from placemarks in matching locations on the globe. During this grant, we will enrich our web pages with semantic tagging (see Section 3.3.5), a move that will generally improve search relevance and increase page rank in the search engines. Reliable, relevant search results will bring individuals to Pleiades who have no prior knowledge of it. Finding there information of interest, and an opportunity to improve it through their own effort, some will stay and help grow the community.