RICHARD TALBERT & SEAN GILLIES

MR. TALBERT: Thank you very much. It's a privilege for us to be here. My only regret is that our colleague Dr. Tom Elliot who has played a leading role in the Pleiades Project is not able to join us. With funding from the National Endowment for the Humanities, the Pleiades Project is building an international community of scholars, teachers, students, and enthusiasts for the study of ancient geography.

At the center of this effort is a web-based collaboration space designed to help community members update, expand, and diversify the spatial and historical reference information assembled by the Classical Atlas Project.

This project, which ran from 1988 to 2000, and produced the Barrington Atlas of the Greek and Roman world with generous support from NEH and many other donors, this is a print product, and I have a specimen here which all may see at their leisure later.

The Pleiades website is built using open source content management and geographic information system software, and is to be available free for reuse by other projects. The aim of the community's work is the perpetual maintenance and improvement of the rich geohistorical collection from the atlas, and its serial publication for use by others in both print and digital formats. A creative commons license will grant anyone wide latitude in the free reuse and redistribution of the information provided, so long as anyone makes their results available for similar use by others, free. Pleiades is on schedule to be ready for public participation early in 2008.

The Barrington Atlas provides approximately 100 maps depicting the ancient landscape with modern contographic graphic conventions. Broadly speaking, the scope is Europe, North Africa, and Western Asia. Both physical and cultural features are highlighted to the extent that modern scholarship can report their details. Wherever possible, physical features like shorelines and watercourses are restored to their ancient aspect.

The atlas maps, some 50,000 historical features for the period of Greek and Roman culture from approximately 1000 BC to 600 AD. In step with the maps directory listings provide both historical and modern names for features as well as time periods and selective scholarly reference. The directory is published in two printed volumes and also in CD-ROM. Pleiades preserves and enhances the data from the atlas. Places and regions are treated as cognitive bundles of information relating to the locations, names and identities of space and place. Features are typed according to the taxonomy employed in the Barrington Atlas. Scholarly references are linked to full bibliographic entries, and where possible to digital texts.

Name information is enhanced beyond that published in the atlas. Most notably original scriptophograpy (phonetic) for each variant is added, as you see here in Greek especially, as our references to salient primary sources.

MR. GILLIES: Viewed from another aspect Pleiades is an open source software initiative. As Richard said, we are developing software that plugs into existing information systems and can be reused in other historical geography enterprises.

The Pleiades project benefits from our experience working on the EpiDoc project, the University of Minnesota's MapServer Project -- which for a number of years was the poster child for the Open Source GIS community, and also on our experience working on the Plone Content Management System. Our philosophy of developing open source software is that of Karl Fogel's "Producing Open Source Software", which we recommend as reading for other history projects that are producing open source software. It's a fine book and overflowing with sound advice based on experience with development of Subversion (the software source control application). Our development infrastructure is based on industry standards for open source software engineering. We use the Subversion source management system (SCM) and Trac, which provides a bug tracker and wiki integrated with the SCM. It's a package which has a lot of promise for other digital humanities projects that use or produce open source software.

Pleiades is networked with the Plone community, 200 of whom are meeting in Napoli next week for their annual conference. I was able to attend last year's conference in Seattle and represent the Pleiades project there in forming a Plone-geography interest group. Pleiades is also a member of the Open Source GIS community, and I presented an overview of the project at the Free and Open Source Software Conference (FOSS4G) in Victoria, Canada last week. I was pleased to see Maurizio Forte mention in one of his slides the GRASS application, which I think has been one of the pillars of Open Source GIS for many years.

At the FOSS4G conference, when I explained the difference between Pleiades and a typical GIS database, I said, Pleiades was the kind of information system that a Roman governor or emperor would have built for himself given our modern technology. That's how I explained it to a GIS analyst or programmer, someone who works in a municipal government GIS shop. Here it's more accurate to say that Pleiades models our understandings of ancient geography. The primary entities of the model are locations, names, and the associations between them that are commonly called "places".

The Pleiades site and web services are based on the Plone Content Management System. Plone is written entirely using the Python language. It has a core of some 20-30 developers from a dozen companies worldwide and very large user base in both academia and government. Our software for Plone is trio of packages. Our entity package is an implementation of our data model. It contains Python classes that represent places, names, locations, et cetera. The geocoder package provides views of entities for the geo-web, using encodings coming from two emerging technologies in the geospatial domain. The first of these is the Keyhole Markup Language (KML), which is the XML format that drives the Google Earth application. Many of us are familiar with that. The second emerging technology is geo-referenced syndication or web feeds. This is a de facto standard coming from the open source GIS community primarily and provides a very simple way of monitoring a site and services for new or modified content in a region of

interest. I will show an example of how to combine feeds from different projects later in the talk. In addition to using Google Earth to look at the places from Pleiades, we are using software called OpenLayers as a map interface. I think that OpenLayers has some promise for other projects in the humanities that might involve annotation of images, whether maps or not.

Pleiades is also pioneering in the field of geospatial web services based on web standards such as HTTP and GIS standards such as GML, the Geography Markup Language.

MR. TALBERT: Pleiades is on schedule to open its website to collaborative update, and improvement early in 2008. Retrospective conversion of a legacy dataset will continue for some time, even as improvements are made to data that is already entered. Participation will be open to all persons with Internet access who can be identified and who agree to the Pleiades contributor agreement and data license.

Under the supervision of the managing editors a group of content editors will rigorously review user's suggestions before they are published for view by others outside the community. To begin with, these editors will be drawn from contributors to the Classical Atlas project and from the staff and affiliates of the Ancient World Mapping Center in Chapel Hill.

But we also expect to promote editors from within the Pleiades user community on the basis of proven work and solid contributions. Interoperability and data interchange must become bedrock strategies for the digitally-enabled scholarly work in the humanities. To achieve these goals many challenges remain, chief among them geography.

From almost every angle, expends, expertise, complexity, and data readiness, geographic visualization and spatial computing present serious hurdles for humanities projects. Particularly those focused on pre-modern cultures.

In fact project-specific impediments to the application of spatial methods are so great that many efforts can do no more than capture idiosyncratic toponymic metadata that then permits only project-specific keywords searching, essentially the practice that has seen little advance, since the late 19th century.

The second phase of Pleiades is designed to address these challenges permanently. The Pleiades community will assist key projects and resources worldwide in collating their geographic information, gazetteers, name lists and so on against the Pleiades data set. For example, our current pilot project aims to normalize the find spot information in the descriptive records for Roman inscriptions from the Province of Moesia Inferior at Epigraphische Datenbank, Heidelberg. At present these records contain bare place names, both ancient and modern drawn from the original publications and other sources and keyed in by hand.

When the pilot is complete each such record will point via a web link to the appropriate record in Pleiades. This arrangement replicated across the various databases and web

services that inform us about primary sources, is an essential prerequisite for accurate comprehensive mapping efforts.

At the same time, Pleiades will cooperate with other projects to develop and improve free and open source and dynamic mapping software that can easily be fitted to the existing websites of many kinds. This software can then draw upon both projects-specific and Pleiades data to provide users with geographic visualization as well as to address spatial queries.

Pleiades has already established liaison with a number of significant projects around the globe. And we continue to identify partners. These two slides list those projects and institutions with whom we have active collaborations already or agreement in principle to participate in Pleiades phase two.

MR. GILLIES: What I'd like to do in the remainder of the talk is demonstrate several applications of the emerging geospatial standards that I mentioned earlier (KML and GeoRSS), and inspire our fellow speakers and attendees to apply these in their own sites and services. I think both KML and GeoRSS provide ready means for collaboration between projects. They are kinds of hypertext and it's really going to be hypertext that binds us all together into an actual web of geographic and heritage projects. First I'd like to talk about searching and viewing Pleiades places using Google Earth and then I'll talk about manipulating Pleiades places using the Yahoo Pipes software, which is a web-based system for combining, manipulating, and mashing up web feeds. Finally, I'll show an example of geoprocessing several feeds using a web service that I am call "Mush."

One of the features of Google Earth and KML that doesn't get as much press as others is that Google (and other search engines in the future) are actually finding and traversing the web of KML documents as they do HTML documents. So, if you have geographic information encoded as KML that links to other KML documents, there is a web there, and the Google bot is traversing these, and indexing these, and building a spatial index for them. There is no direct programmatic API to get to this index but you can get through either the Google Maps application, or through Google Earth. On the upper left here you see an example of typing a query into the box; we are searching for temples, and Pleiades. Google's spatial index is searched for temple type entities of the Pleiades project that are roughly within the bounds of the viewing window. So, there you see the spatial index in action. Here it's popped up six places I believe, shown in a search pane on the left. Clicking on any of those links would then load the place from Pleiades.

(Note: subsequent to the NEH/CNR conference, Microsoft's Live Local site allows local search for GeoRSS and KML.)

Yahoo has built a web application called "Pipes", where you can pick feed processing components, connect them, and build up a processing chain based upon web feeds as input and web feeds as output. Here is a very simple example showing a Pleiades feed where we have Archaic time period places piped through a filter, and then piped back out. This is a just an example of what the future of geo-processing on the web will be.

There will be more and more sites like Pipes where you can mix and match components. Geo-referenced feeds will be one of the most accessible and important means of sharing and remixing data from digital humanities projects. By providing geo-referenced feeds, you create a potential for serendipitous reuse of your information. A group that you don't yet collaborate with, that you may not even know of, is going to find your information and find a clever use for it that could very well enlighten us all.

As another example of what you can do with these geo-referenced web feeds I will show now the spatial intersection between items of the public U.K. plans site (where one must register construction projects) and records from the Oxford Celtic Coin Index. The only thing these two sites have in common is that they provide geo-referenced feeds, but that's all that is needed to begin to compare their contents geographically. The intersection of the two feeds produced by the Mush service shows where digging for utilities or basements might possibly turn up new coin finds. This is what I mean by serendipitous reuse of information from websites and web services. If we all plan for serendipitous reuse, we will find surprising new intersections between our projects.

Thank you very much.

(Applause)

MR. BOBLEY: Okay. Any questions or comments?

MR. BOZZI: I would like only one question for you. You mentioned before that you want to use it to insert in that project also primary sources, I mean, that if you are having a map, a word speaking about a name of a city, of a place, you link this name to some papyri or epigraph in which this place is named. So this I understood correctly.

So, and probably there is also, excuse me, in another second, there is also really the possibility to cooperate with other international project in which that our papyri or documentation manuscripts, which speaks about that the proper names or topographical names.

I don't want to quote specific project, but really there are a lot of information primary sources about the Latin or Greek places.

MR. GILLIES: Yes, we'd like to do that very much what we really need from other projects in order to do this is web URLs, resources that we can point to, and actually together build a web of information that not only human users but also machine users of the future can traverse from node to node. So they can go from places in the Pleiades map to a name on the Pleiades map, and then into a web page for a papyrus text or for an epigraph.

So, what we would like to try to do is promote the use of URLs among the humanities as a way for projects to collaborate, and share information.

MR. TALBERT: Yes, to follow up on that, this is exactly what we have in mind and having with NEH support successfully establish this foundation this methodology, we very much see this as the next logical stage. And those contacts we've made so far, please us very much in that, and that they too want to share that, that many other project have this difficulty.

And Professor Riva also with his Boccaccio's Decameron, for example, it was very noticeable that there one of the clicks was geographic locations in the text. Not so medieval text, but we are doing this for classical antiquity.

And while we don't wish to dominate, nonetheless, we feel that in the area of geography we can help many other projects which have different types of material, and don't necessarily wish to, and shouldn't reinvent the geographic wheel, for themselves, but would like to integrate with a project that specializes in geography, so that a standard can be set.

And there is no intention in setting the standard to set aside other ideas about what the name should be or variance. These can all be recorded, but it does mean that there would be a common pooling of this kind of data and expertise and knowledge and controversy.

But as you yourself Professor Bozzi, have said, much of this depends on whether funding is forthcoming for these expansions and continuations.

MR. BOZZI: In any case congratulations, for the project because -- sorry, because it is really very important I think not only to for the research, but also for teaching classical humanities in the school.

And as my personal offer, I can give you not in the web, but in any case as I said to Bernard Frischer before a compact disc in which there are store all the text of Latin grammarian of the antiquity in which I don't know if there are quoted name of places or towns, but probably yes.

And the information retrieval system I studied and realized probably is very useful for you to find very quickly, the names of the time or the places, and you can find the author who quoted these names. So when we are back in Pisa, I can send you and the rest you can leave me the CD in which you can try to find this questions sometime.

MR. GILLIES: Thank you, I look forward to that.

MR. FRISCHER: I have two questions about -- some basic question about how you are handling time and space. Now, that you are going digital. So space, we know what the scale was of the Barrington Atlas, how do you deal with scale, different scales which can be so much greater with the GIS, I mean, you could be taking archeologist one to one hundred plans and draping them over sites.

Are you interested in doing that or do you have a lower limit or calling an upward limit, that you don't really want to get into, because it's not appropriate to cartography. And then the other thing is about time which is historic GIS, and whether you know, there is ineffective time bar.

I know that you can handle a difference of place names that often is related to time, so Pompey becomes the Colonia, cornelian this sort of thing. But it would be nicer as -- it would be nice, and is it possible to set a kind of a time bar for the ancient world and look at, you know, how things were called in 70 B.C. as opposed to 300 B.C.

MR. GILLIES: Yeah, I am going to leave the time component to Richard, I think. And he can explain how we have settled upon the discrete time intervals that we have.

The spatial domain for Pleiades is restricted to cartographic scales of less than 1:25,000 or so. We are not dealing with a vertical coordinate anywhere. We figured that's the domain of archaeologists, and architects. For any entity where the height above the terrain of the earth becomes important we would like to start linking to archaeological projects, and say, follow this link for more detail. So, for Rome we might have boundaries for different time periods. But we are not going to actually resolve buildings within Rome or streets within Rome, because that's really not our domain, that's your project and we'd like to link to yours and direct users to you for more detail.

MR. TALBERT: Yes, exactly. We don't see a role for -- a large role for Pleiades at very large scales. Yes, that is -- that's simply is beyond its competence and also is being very -- handled very well, by others. In terms of time indeed Professor Frischer is right, that it would not just be nice for us to be able to differentiate different time periods, but it's absolutely essential that we can do this.

And as he will no -- not recall as a good classicist himself the directory for the Barrington Atlas already takes many steps towards this because it records changes overtime and refinements in that regard in more detail, than the Atlas itself can show.

And now we are moving to the digital phase, we really are able to take advantage of that data which almost I could say remains latent, but now can be fully deployed, and I hope further refined overtime in the future, because of course, these shifts through our very long antique period are most important to take into account.

MR. BOBLEY: Okay. Any other questions? In the back, Suzanne Lodato (NEH Director of Preservation and Access)?

SUZANNE: (Off mike)

MR. GILLIES: Well, let me try to -- of course I can't predict the funding future for the project, but I can say that the immediate goal is to have places from the Barrington Atlas digitized and then staged on to the Pleiades sites.

So that will be some 50,000 named places, ethnic areas features on the globe and mountain ranges, et cetera, and roads, I think roads might bring that number up to 60,000 features in the world. And then at that point, a lot of the intensive workload will be done, and we would like the community of scholars at that point to take over, begin to refine the data, start rolling in updates from other related projects.

I am not sure how to answer your question about what would our actual budgetary needs be, but --

SUZANNE: (Off mike) -- to basically maintain and whether you had some views about that was going to be funded?

MR. TALBERT: Well, this is a very good question, and it's a question that all such projects have. This is a project I think which, thanks to very generous NEH support for its first phase has a very clear vision of what could and should be done for a second phase.

We have an outline, a six-year plan with, which involves considerable expansion, but also of course has a fairly sizeable price tag attached to it, and we are sufficiently realist to understand that we may have to take this stage by stage and trim it and adapt it. And I fear as you know, as all of us in the humanities, we are used to living and developing in this way, and thriving.

MR. BOBLEY: I think she was trying to offer you a grant.

(Laughter)

MR. TALBERT: I'd be very glad. I'd be very glad to take a check afterwards.