California's Geospatial Records: 
Archival Appraisal, Accessioning, and Preservation

1. What is the purpose of the project and what are its goals?
The project has two interrelated goals necessary to move forward both the Electronic Records program in the California State Archives specifically, and research in electronic records in general. During this two-year project, we plan to adopt the necessary hardware and software infrastructure at the California State Archives (CSA) to preserve California's geospatial records stored by the California Spatial Information Library (CaSIL) as part of the California Resources Agency's California Environmental Resources Evaluation Systems (CERES) office. We are interested in testing the emerging preservation infrastructure developed at the San Diego Supercomputer Center (SDSC) through the NHPRC's funded Persistent Archives Testbed and applying the results within the broader technological structure of the state of California. This preservation infrastructure is based on “data grid” and “persistent archives” technology. Initially, we will deploy technology at three sites (CaSIL, CSA, and SDSC). Because of this project, the California State Archives will have the appropriate infrastructure for the management of electronic records transferred from a state agency to ensure their archival preservation.

Simultaneously, the State Archives will appraise and accession the archival geospatial records held by CaSIL into CSA. This accessioning will constitute the first structured receipt of electronic records by CSA. To accomplish this goal, we will incorporate the existing methods of electronic records appraisal developed by the National Archives and other states to appraise California's geospatial records. In addition, because of the many formats for the geospatial records stored by CaSIL, we plan to experiment with data conversion and migration to ensure that we understand how to preserve the integrity of these records. Based on these experiments and on the results of studies from other archives and institutes, we will adopt a set of guidelines for geospatial records to be preserved by CSA and develop a set of transmission requirements for geospatial records. With an appraisal completed and these standards in place, we will transmit selected archival record series from CaSIL to CSA's implemented technological infrastructure. The large size of, complexity of, and different uses for these geospatial records will permit us to test the effectiveness of this infrastructure.

As a result, over the two-year grant period, the project will both implement a preservation strategy that utilizes persistent archives for electronic records and test that strategy by appraising and transferring archival GIS records to the State Archives and the SDSC. In the distributed environment, additional tests will establish the integrity of the accessioned data, supplement metadata to record changes to the data, and ensure that our infrastructure design serves the needs of the California State Archives and the state's citizens. We will build on the lessons learned from other projects such as: Columbia University's CIESIN's study of how to preserve geospatial data, the GeoArchives in Maine, VanMAP in InterPARES 2, the National Archives guidelines in spatial data accessioning and the Persistent Archives Testbed at SDSC. The project will create a unique project by adapting the best practices from these studies to the state of California and implementing a data grid framework. This test of an infrastructure model by accessioning geospatial records will result in a set of model archival practices for California and others who need to work with complex sets of archival electronic records.
This project develops from the identified need of the California State Archives to address our handling of electronic records. Likewise, our partners at California Resources Agency's California Environmental Resources Evaluation Systems (CERES) office recognize the need for a long-term preservation strategy for the CaSIL data that they supply to governments and citizens (see www.gis.ca.gov and www.gio.ca.gov). We will collaborate with San Diego Supercomputer Center's Sustainable Archives and Library Technologies (SALT) Laboratory so that we can build on their extensive research into infrastructure decisions and help us establish methods that will help with the appraisal and accessioning of these geospatial records. In addition, we have assembled a group of experts, including experienced archivists in electronic records issues and leading GIS professionals for California's agencies, to serve as an Advisory Board. This board will monitor our research, experimentation, and implementation, as well as advise us where our decisions may not fit the needs of institutions in other structures (see Project Staff section for details on their qualifications; for a list of participants, see Appendix 1). The SALT staff and many of the Advisory Board members are currently involved in research that addresses aspects of geospatial records preservation.

Over the course of this grant, the California State Archives, working with the CaSIL, and the San Diego Supercomputer Center, will:

1. Establish the necessary hardware and software infrastructure at the California State Archives to preserve California's geospatial records stored by the California Spatial Information Library (CaSIL), a division of the California Resources Agency.
2. Use results from the San Diego Supercomputer Center's NHPRC-funded Persistent Archives Testbed (PAT) project to determine appropriate hardware and software resources necessary to set up a “geospatial archivist grid” distributed across appropriate state agencies (initially CSA, CASIL, and SDSC).
   a. Install the appropriate hardware and software at the State Archives and tests its use in a distributed environment.
3. Appraise and accession archival geospatial records held by CaSIL into CSA.
   a. Build on California's records appraisal program as well as existing methods for such appraisal developed by other institutions to appraise CaSIL's geospatial records.
   b. Apply this research to define an item and a series of geospatial records.

---

1 As part of the California Historical Records Statewide Planning Project, the California Historical Records Advisory Board (CHRAB) commissioned an assessment of California State Archives programs; this resulting 2001 study recommended numerous changes in the Archives' appraisal and description process in order to respond to electronic records (see History Associates Incorporated, State Government Records Assessment Report (January 2001), available at: http://www.ss.ca.gov/archives/pdf/stategovtrecrept.pdf). In addition, the CHRAB final summary of that assessment process listed "Electronic Records" as crucial to the State Archives' long-term planning and declared "No challenge is greater today than the need to develop programs, create standards, and train staff for the management of electronic records" (CHRAB, Historical Records in the Golden State: A Strategic Plan for Preserving California's Documentary Heritage (February 2002), 8 at http://www.ss.ca.gov/archives/pdf/strategicplan.pdf).
c. Experiment with conversion of data in CaSIL to Geographic Markup Language and Spatial Data Transmission Standard to test whether the integrity of the records can be preserved during data migrations.\(^2\)

d. Based on these experiments, adopt a reliable format standard for geospatial records to be preserved by CSA and develop a set of transmission requirements for geospatial records based on models coming from CIESIN, the Maine GeoArchives project and NARA.

e. Using these requirements, transmit selected archival record series from CaSIL to CSA's storage systems and create multiple access points for project staff. We will use the Storage Resource Broker developed at the SDSC to monitor these copies, track changes, and preserve authenticity during the accessioning process.

4. Develop next steps for long-term access to geospatial records stored at State Archives, and extrapolate, where appropriate, our results to other electronic records issues facing CSA and related institutions.

a. Establish the necessary requirements for the description of and authentic access to these geospatial electronic records in ways that meet the needs of historians, resource managers, and policy makers.

b. Identify how the technological infrastructure, programmatic, and Archives' employee skills developed for this project can provide models for the CSA's appraisal, accessioning, and description of other electronic records.

These project steps and goals respond to a combination of challenges facing archivists at the California State Archives in particular, but at many other institutions as well. First, this project seeks to develop and implement the specifications for a robust technological infrastructure for storing electronic archival records and their associated metadata. Researchers working on digital preservation issues in a variety of frameworks have concluded that redundancy and version control are among the most central requirements for a secure digital preservation system.\(^3\) This part of the project will build on previous research by the San Diego Supercomputer Center on distributed networks in connection with the Persistent Archives Testbed project and in their research for the National Archives. We will initially set up a persistent archives network consisting of three storage sites (SDSC calls these “grid bricks”), one at CSA, one at CASIL, and one at SDSC. Our investigations should help determine the feasibility of this type of infrastructure for archival facilities of a variety of sizes.

\(^2\) The National Archives has declared a preference for geospatial records transferred to their repository to come in SDTS or GML (see http://www.archives.gov/records_management/initiatives/digital_geospatial_data_records.html). The CaSIL staff tends to use GML as their most common format.

In addition, we are convinced of the value of focusing on GIS records for preservation. The California State Archives has an extensive map collection already and these records will be important ancillaries to their predecessors. The CaSIL office is widely known for the quality of the geospatial data that it makes available to the public. Among their holdings are wide varieties of datasets that document administrative divisions, cultural geography, facilities, and physical geography with an emphasis on the related natural resources. CaSIL's staff has already done an initial assessment of datasets that they feel document important aspects of California (included in Appendix 2). The project archivists will use that list as a starting point as they develop their appraisal procedures and then select an initial set of 5-10 varied series to accession.

Though many of the basic principles of archival preservation, arrangement, and access are replicated in the digital environment, CSA’s planning for electronic records suggests that we must adjust our methods of appraisal to deal with this format of records. In California, our appraisal has been mainly done at the level of the review of retention schedules that are prepared by departments' records managers, submitted to the California Department of General Services' Records Management Program, and then forwarded on for review by the State Records Appraisal Program at the State Archives. In certain cases, we have worked directly with program staff and records managers to identify other records of value. But neither method is suited to electronic records. Indeed, outreach efforts to encourage records managers to schedule electronic records have yet to result in significant change because California's culture of records management tends to emphasize records that staff can measure physically. A telling example is that in the State Archives' records appraisal database, 7,025 records series are flagged for archival review at the end of their retention period, but only 157 series are in any electronic format. Since only 60% of departments have current record schedules, the Appraisal Program cannot even guess at the number of significant electronic records that ought to be reviewed. The current records retention schedule for CaSIL identifies two cubic feet of electronic records that CERES staff assumes are their backup tapes. Such a vague indication on the schedule obviously cannot guide our appraisal of the geospatial data CaSIL holds. The State Archives' Electronic Records Program expects that we can use this project to leverage our work with individual records managers and the State Records Management program in scheduling, appraising, and then accessioning electronic records.

In addition, the size of the CaSIL collection suggests why our focus on infrastructure is also essential. CaSIL has over 35 terabytes of data distributed across various servers. In the coming year, they plan to add more data including a complete flyover of the state at the 1 meter level; this new project should add another 3 terabytes alone. The State Archives knows that much of that data are not records of CaSIL, but merely other agencies' data sets that they help distribute to the public. Nevertheless, even one terabyte of data is more data than the State Archives has ever actively managed. Paul Veisz, a GIS expert for California Department of Fish and Game and a member of the Project Advisory Board, remarked that the exponential growth and replication of geospatial data means that it will be difficult and unwise for a single location to store even the fraction of the California geospatial records created by state agencies deemed to have archival value. This remark encourages us that it is necessary to experiment further with a distributed architecture that still allows for archival standards and methods. We will use two terabyte grid

---

bricks to store the selected records at distributed locations and test how our infrastructure works in data-intensive experiments. This gives us the unique opportunity to go beyond the experiments done during the PAT project to show that archivists can continue to control the archivally-valuable data that requires accessioning and description.5

Project Origins
This collaborative project grows out of a series of meetings and studies that have involved staff from the State Archives, the Resources Agency, and the San Diego Supercomputer Center. Beginning in fall 2002, the California State Library and the State Archives began collaborating to address the challenges posed by digital publications, websites, and electronic records.6 Because of this collaboration, we grew more conscious of CaSIL's mission to develop and distribute geospatial data. Since then, the State Archives, State Library, and CaSIL have become involved with the San Diego Supercomputer Center's various initiatives to research and develop solutions for persistent archival preservation of electronic records. In February 2005, David Harris of the California Environmental Resources Evaluation System (CERES) office that oversees CaSIL's data and infrastructure, Janet Coles of the State Library, and Lucy Barber of the State Archives attended the Persistent Archives Testbed (PAT) project meeting in San Diego. This meeting educated us about the nature of collaborations in other states and about possible approaches to implementing an archival electronic records program in California. At a subsequent planning meeting, we resolved to continue our collaboration. The proposed study is just part of our broader effort to increase awareness among state leaders and state agencies about the need for archival preservation of California's electronic records in order to sustain the state's history of providing its citizens with access to its records.

The SDSC and the CSA have worked together regularly over the past three years, and are currently among the partners in the NHPRC-funded effort to develop a Persistent Archives Testbed. This is a demonstration of the potential of grid technology for the preservation of electronic records in a distributed, collaborative framework. In addition, in our role as partners in the NHPRC's e-legislature grant to the Minnesota Historical Society that began in April 2005, staff members from the California State Archives and the State Library are engaging in further discussions of these technologies and their use for archival storage and access. Because of these previous and on-going collaborations, we are confident this project will achieve its ambitions.

Value of the Records
The geospatial records that CaSIL stores and distributes document the political, social, and environmental history of California through geographically-oriented data sets. On one hand, the CERES office uses the CaSIL system so that the state, federal, and local agencies, as well as interested private organizations and business with common interests in resource development or protection could share their data sets. At the same time, CERES inherited responsibility for some of the earliest and most used geospatial layers created by the state. They continue to create new original datasets for special projects of high priority to the state and which often meet needs that cross traditional departmental responsibilities.7 It is these records they create -- rather than

5 This is an NHPRC funded project, for details, see www.sdsc.edu/PAT.
7 Before 2001, there was a GIS division within California's Teale Data Center that created and supplied this information. After that office was disbanded, CASIL became the distributor of California's physical and cultural
all the data they distribute -- that the State Archives is concerned with preserving. These records document California's political, cultural, and environmental history that will be essential to geographers, social scientists, and historians in the future. Many have archival value to the California State Archives whose appraisal guidelines emphasize interest in records with clear "evidential and informational values" and that reflect major trends in California, political, policy, and social history. Because CaSIL both creates records and distributes data created by other providers, a careful records appraisal will determine which geospatial datasets are actual records of the Resources Agency and, of those, which have archival value.

Project Results
The results of the project will extend to other tasks facing the State Archives and the state of California in general. In terms of the preservation of geospatial records, the CaSIL office is just one of many creators of such records for the state of California. We will develop the infrastructure and methods for handling these records with careful attention to the need to accession, describe, and provide access to additional archival geospatial records. At the same time, we are convinced that the solutions we develop in this project will be scalable to the larger issues of electronic records preservation facing the California State Archives. Though geospatial records pose some unique questions, many are similar to those associated with the appraisal and accessioning of other relational database sets or appraisal of centralized electronic document management systems. As a result, we think the combination of a skilled set of partners, the value of the records themselves, and the transferability of our decisions to other electronic records projects in California and elsewhere makes this a project worthy of the National Historical Publications and Records Commission’s support.

2. What is the significance of the project in relation to the NHPRC’s programs and goals?
This project overlaps with a number of the NHPRC’s goals. In the NHPRC’s 2004 strategic plan, the Commission placed a priority on supporting institutions that acquire, preserve, and promote ready use of electronic records, especially those that promise to be sustainable and that are built upon collaboration and open systems and support research projects in electronic records. Of the 10 questions specifically articulated in the Electronic Records Research Agenda of 1991, the primary ones that apply, and which this project will address, are questions 2, 3, 4, and 6:

2. What are the technological, conceptual, and economic implications of capturing and retaining data, descriptive information, and contextual information in electronic form from a variety of applications?

We will address this question through our development of a robust infrastructure for the State Archives' electronic records program and then testing our skills by accessioning geospatial records that come in a variety of formats and require considerable documentation to preserve.

geospatial information. Because that office no longer exists, CaSIL is now both the creator and distributor of these datasets. In addition, for special projects, they have also created or added significant value to other geospatial datasets.

3. How can software-dependent data objects be retained for future use?
This question is addressed through our close examination of existing commercial-methods of converting geospatial records from one format to another.

4. How can data dictionaries, information resource directory systems, and other metadata systems be used to support electronic records management and archival requirements?
CaSIL already uses the Federal Geographic Data Center's standard for collecting metadata about the collections it hosts. We will determine what additional metadata elements are necessary to ensure the preservation and migration of this data.

6. What policies best address archival concerns for the identification, retention, preservation, and research use of electronic records?
The project addresses this crucial issue by planning to both experiment with but also implement solutions to the preservation of electronic records. Much of the concern, in our view, results from a lack of experience in handling these records. In addition, by training staff at both the Archives and geospatial departments about our methods, we hope to educate data providers about best practices.

In addition, the NHPRC has a long record of supporting projects that preserve map collections in general as well as more specific projects that have focused on geospatial records in particular.  

3. What is the plan of work for the grant period
If the grant is awarded, the project will begin in January 2006 and be completed in December 2007. The plan includes four phases. Tasks related to establishing the necessary hardware and software infrastructure as well as developing State Archives' skills in appraisal and accessioning geospatial records dominate the first two phases. Tasks related to accessioning appraised records from CaSIL, experimenting with their replication on a data grid, and validating their integrity dominate the third phase. These tasks continue during the fourth phase, but project staff also will concentrate on disseminating results, receiving feedback on work products, and assessing how to use these products for accessions of other types of electronic records.

---


10Map preservation is part of the stated goals of the NHPRC and numerous funded projects have dealt with preservation and access to non-digital maps (for example: Hawaii (80-077), Alabama (90-034)). In 2003, the Commission awarded the Columbia University's Center for International Earth Science Information Network funds "to identify and disseminate practical policies, techniques, standards, and procedures to manage, preserve, and provide access to electronic records that have significant geospatial components, especially those generated by a Geographic Information System." ([http://www.archives.gov/grants/funded_endorsed_projects/states_and_territories/ny.html](http://www.archives.gov/grants/funded_endorsed_projects/states_and_territories/ny.html)) (2003-038) More recently, in 2004, the NHPRC's Electronic Records Program funded the Maine State Archives' "Creating the GeoArchives, a collaboration among the state archives and other state agencies that are creating the Maine Library of Geographic Information (GeoLibrary) to develop the GeoArchives. The project plans to use the expertise and infrastructure of the GeoLibrary to create a repository for archival geo-spatial data created by state and local governments. (2004-084)"

Project staff will base their approaches on the lessons learned and best practices developed from such projects as: the GeoArchives being created by the Maine State Archives (funded by NHPRC), Columbia University's CIESEN's dissemination of preservation policies for archival geospatial data (funded by NHPRC), the VanMAP project for InterPARES 2, the National Archives’ guidelines on the preservation of spatial data and on-going experimentation with the accessioning of spatial data, and the Persistent Archives Testbed at SDSC (funded by NHPRC with NARA as a partner). CIESIN, VanMAP and GeoArchives will have completed most of their work by the start of this project. We will collaborate with their researchers to validate their results and integrate their geospatial techniques with the latest data grid technologies to create a robust infrastructure for California State Archives' electronic records.

1. Project staff establish project web site.
2. Project staff develop hardware and software specifications for California State Archives to manage transferred geospatial records.
3. Project Advisory Board meeting at SDSC in connection with the ESRI User Group meetings. The ESRI User Group meeting attracts all major geospatial creators in the United States; the project staff and Advisory Board will learn from the latest trends in geospatial systems about ongoing developments in the field that will shape our solutions.
4. CSA staff and SDSC staff research definition of geospatial data in archival terms based on the results of the studies from previous archives-based geospatial preservation projects.
5. CSA staff appraises geospatial holdings of CaSIL to identify those records with archival value.

July - December 2006: Infrastructure Installation and Geospatial Records Appraisal
1. CSA staff procure and install specified hardware and software.
2. Project staff test converting geospatial records into different open-source formats, with especial focus on the SAFE company's FME File Conversion Suite, which is one of the most widely-used programs. Based on similar tests of content interoperability, these tests will allow the State Archives to determine the steps necessary to preserve the authenticity of the records we preserve. A key focus will be on the migration of records in proprietary formats into open-source formats.
3. Project staff determine necessary level of documentation of such conversions to ensure records' authenticity and preservation and experiment with validators to help CSA authenticate transferred records.
4. CSA staff document their geospatial appraisal guidelines and data transfer standards.
5. CaSIL staff and project advisory board evaluate the appraisal guidelines, data transfer standards, and conversion test results and staff revise as suggested.

January - June 2007: Infrastructure Testing and Geospatial Records Accessioning
1. After archival records are identified, CaSIL staff and CSA staff transfer selected series to Archives.
2. CSA staff accession records into CSA's collection management system and add additional metadata determined to be required for geospatial data.
3. CSA staff document the records stored at CSA and replicate them at an additional location within the state's infrastructure and at SDSC.
4. CSA and SDSC continue to test transfer procedures and reliability of data transferred to the archive with validators developed by CSA and SDSC. These validators will help automate the process of testing for lossless migration and format interoperability.

July - December 2007: Dissemination, Refinement, and Next Steps
1. Project staff present initial findings at appropriate forums including state records managers meetings, regional geospatial and archival meetings, and national archival and records managers conferences and receive feedback from their archival and GIS colleagues.
2. CSA staff revise as necessary the transfer protocol and accession additional series.
3. CSA and SDSC staff examine how lessons learned translate to other electronic records expected to be transferred to CSA.
4. Final project advisory board meeting to evaluate project achievement, solicit advice on ensuring sustainability of the project, especially in terms of outreach to policy makers, GIS creators, and records managers in California.

4. What products or publications will be produced during the grant period?
This project will produce the following documentation and specifications that will be used in documenting our efforts and providing guidelines to archivists, records managers, and records creators.

- Software and Hardware Specifications and Documentation for California's Geospatial Electronic Records, including instructions for on-going management.
- Definition of geospatial archival record, series and items.
- Appraisal guidelines for geospatial records.
- Format specifications and transfer protocols for geospatial archival records.
- Metadata requirements for geospatial records transferred to California State Archives.

Research Studies, suitable for presentation at professional conferences or publication in GIS or Archival publications

- Reliability of GIS Data Conversion Process for Preservation Purposes.
- How Should Archivists Preserve Geospatial Records?
- GIS Record Series Preservation Strategies as Model for Preservation Strategies for other Electronic Archival Records.

In addition, the project will:
- Maintain a project web page, with all written materials (e.g., project reports, documentation, research studies) available online during and after the grant period.
- Present reports on the project at professional meetings, for example, the 2006 ESRI User Groups Conference and the 2007 Society of American Archivists' meeting, as well as regional archival and GIS meetings;
- Develop recommendations for enhancement to the technology infrastructure that serves the state of California in order to facilitate distributed storage networks for archival data.
Throughout the course of the project, the project staff will:

• Make every effort to educate other archivists about the questions raised by the project's focus on infrastructure design.
• Meet with geospatial records creators and users to educate them about the need to preserve some geospatial records with archival standards.
• Make presentations to policy makers and other resource allocator.
• Submit reports and articles to professional journals.

5. What are the qualifications of the personnel?
(Resumes for all named staff are in appendix 3; letters of commitments for staff outside the State Archives who are contributing to the cost-share of the project are in appendix 4).  

STAFF:

**Dr. Lucy Barber**, Archivist II, at the California State Archives will dedicate one-third of her time to this project. She will serve as the project director, liaison with project partners, facilitate the installation of the infrastructure by IT staff that serve the California State Archives, monitor the work of our project staff people, and facilitate interchanges with the project advisory board. Lucy Barber has worked at the State Archives since 2001. She has developed the CSA's electronic records program. In that capacity, she has received training in digital document management and electronic records management. She was the State Archives' representative on a project led by the Government Publication Section of the California State Library to examine the state of government publishing in the digital age; as part of that project she evaluated RFP for consultants, offered comments on work products, and facilitated a working group on technology issues surrounding publication digital preservation. This was one of the first statewide collaborative projects on digital preservation issues and resulted in the State Archives becoming a regular participant in subsequent conversations on the importance of preserving California's born-digital records. She presented on aspects of electronic records management at the Government Technology Conference of 2004, resulting in more contact with records creators. She organized an extremely well-attended seminar for records managers and IT professionals to start a conversation on electronic records preservation in the state. She regularly advises records managers and program managers on strategies for improving the long-term viability of their electronic records systems.

Two other professionals from the State Archives will also contribute to the project. **Laren Metzer**, deputy state archivist, has over 20 years of experience as an archivist, has served as the deputy director of California's State Historical Records Advisory Board, and now handles the daily administrative matters of the State Archives. For 5% of his time, he will provide administrative oversight for the project to ensure that other divisions of the Office of the Secretary of State provide effective support where necessary. In addition, he will offer his expertise on how to publicize this program to state decision-makers to ensure its long-term sustainability. **Linda Johnson**, Archivist II, has over 10 years of experience as an archivist working in special collections, local historical societies, and the State Archives. Since 1999, she has co-directed CSA's State Records Appraisal Program. With 10% of her time, she will contribute her expertise in appraising maps and other collections to the geospatial records and assist Barber in the preparation of appraisal guidelines and transfer documents. In addition, she will help facilitate communication with the State Records Management Program staff and agency
records managers so that we will be able to learn from them about how our project can fit their needs. She will also take part in meetings of the Project Advisory Board.

The State Archives will also hire a half-time graduate student intern who is planning on a career in archives to assist on this project (a job advertisement is attached in appendix 5). The graduate student will provide research and administrative support to the Project Director. In addition, the graduate student will be trained in the appraisal of geospatial records, in the management of the distributed storage network system infrastructure, and in the accessioning of the records into the State Archives. In this way, the project will have the added benefit of providing intensive experience in electronic records issues to a beginning archivist.

Our principal staff person at CERES is David Harris, their internet systems specialist and chief security officer for the Resources Agency. The CERES office is responsible for the acquisition, deployment, and access to the geospatial data held by CaSIL. Harris has extensive experience in managing innovative projects that provide California citizens with more access to information. He has been an active participant in both the State Archives and State Library's effort to lay the foundation for a reliable infrastructure that provides people with access to the public records and publications they need. He will contribute 5% of his time. Part of this time will be spent in the coordination of a graduate student intern who will work with this project exclusively. CERES regularly uses geography or IT graduate students to assist their work and this will ensure that a starting professional in that field will gain extensive experience in geospatial records preservation issues. This intern will work at 50% time (a job advertisement is attached in appendix 6).

From the SDSC, Richard Marciano will serve as the technical advisor and liaison for the project. He will devote 20% of his time to the effort over the course of the project. If this grant is approved, the CSA and the SDSC will finalize a contractual agreement for Marciano’s time and the SDSC’s services. Over the past decade, the SDSC, with Marciano playing a principal role, has participated in most of the notable research efforts addressing the management of large volumes of digital content. Staff have principally supported the development of grid technology, the California Digital Library, InterPARES, the National Archives’s ERA and numerous other projects. These have led to the development of a specialist lab led by Marciano: Sustainable Archives and Library Technologies. With the development of the Storage Resource Broker (SRB), the SDSC has provided a proven and practical tool for the long-term management and preservation of digital information. Allied with its unparalleled computing capacities, these achievements all make the SDSC a uniquely valuable partner in this project.

In addition, Reagan Moore, Distinguished Scientist at San Diego Supercomputer Center, will provide management support for interactions between the SRB software development group and this project; specifically helping with the installation of the SRB servers at each site, and the generation of reports and presentations. Moore is the principal investigator on projects for the development of prototype persistent archives for NARA, a collaborator on projects with NSF for the creation of persistent archives for the National Science Digital Library, the principal investigator on projects with NSF for the development of the Storage Resource Broker data grid, and the principal investigator at SDSC on projects for the application of the technology for the California Digital Library. Moore has published numerous papers and book chapters on the use
of data grid technology to implement persistent archives and digital libraries, and has given numerous presentations and tutorials on the SRB. Moore is the organizer of the Persistent Archive Research Group within the Global Grid Forum, which is attempting to define the set of features needed within data grids to support persistent archives. Moore and Marciano are currently conducting a case-study for InterPARES2 on preservation of the GIS records of the city of Vancouver (VanMAP). Moore will commit 5% of his time to the project.

ADVISORY BOARD:
In the course of developing this proposal, we have attracted the support and advice of a group of experts on the issues that we seek to solve. This group of seven will serve as an advisory board for the project. They will help evaluate the pertinence and applicability of the infrastructure we select for the State Archives, and our methods of appraising and accessioning the geospatial records of CaSIL. This oversight committee will ensure that the methods and tools being developed for this project will have applicability, both inside and outside of state government for other archivists. This advisory board will review our infrastructure decision will help us preserve the integrity of the archival electronic records and that our methods of appraising and accessioning geospatial records respect their original context, record the necessary metadata, and will be sustainable beyond the grant period.

GIS experts who work within California state government make up one half of our board. All of them have volunteered their time and expressed their enthusiasm for this project. John Ellison, Agency Technology Officer for the California Resources Agency and director of CERES, is also coordinating the state's strategic GIS planning. He is a nationally-recognized leader in the development and management of geospatial data and he brings to the project a keen understanding of the broader policy issues surrounding technology in California. He will contribute 5% of his time to the project. Michael Byrne, research specialist for the California Office of Statewide Health Planning and Development, has spent his entire career in GIS, and now is the lead researcher on an innovative study to improve the use of geospatial data in public health planning. He will also contribute 5% of his time. Paul Veisz, a GIS research manager for California Department of Fish and Game, also has extensive practical experience in GIS development and maintenance. He has also written and presented regularly on issues in GIS in California. Initial conversations with all three have shown that they have historical knowledge of geospatial data development in California; active participation in current projects; awareness of the technical challenges and the vital need for a preservation strategy to be developed. Because of their positions within the state, these three individuals will help ensure that our solutions can translate into California's technology policy.

For the other members of the Board, we have brought in archivists who work in other institutional settings to ensure that that project builds on the latest archival practices in electronic records, especially for geospatial information. Robert Horton, state archivist for the State of Minnesota, has been involved with many of NHPRC's electronic records programs. He brings his knowledge of sound archival practices to the digital environment with a creativity that results in clear successes. In addition, he is active in a GIS project in Minnesota that is adding geographical information to historical data. Mary Taylor has a MLIS from the University of Texas at Austin, and is the University of Nevada Reno's Metadata Services Coordinator, where she works with issues that include management of geospatial data and long-term preservation of
digital collections. Before coming to UNR, she managed the "Sacramento History Online" project (www.sacramentohistory.org) and did consulting for La Galeria Posada's ARTchives collection (www.galeriaposada.org/artchives.htm). She recently presented at Nevada GIS meetings on issues of digital preservation. She is eager to contribute her knowledge of the issues facing the preservation of the data stored at the W. M. Keck Earth Sciences and Mining Research Information Center at UNR. The university has agreed to let her spend 5% of her time on the project. **David Keller** brings the perspective of a records manager and archivist for the Metropolitan Water District of Los Angeles. He is already actively involved in managing electronic records in electronic document management systems and is working with the GIS experts in his office on their preservation needs. Finally, **Dr. Robert Downs** of Columbia University's Center for International Earth Science Information Network, is their senior digital archivist and co-project director on the Center's project to develop guidelines for archivists in the northeastern region on Managing and Preserving Geospatial Electronic Records. He will share with us what he learned through interviews with a wide range of geospatial users, creators, and archivists about this type of data and the preservation requirements. These experts will help us test our assumptions about appropriate archival practice and the applicability of California State Archives' technological solutions to other repositories.