



NATIONAL ENDOWMENT FOR THE HUMANITIES

OFFICE OF DIGITAL HUMANITIES

Narrative Section of a Successful Application

The attached document contains the grant narrative and selected portions of a previously funded grant application. It is not intended to serve as a model, but to give you a sense of how a successful application may be crafted. Every successful application is different, and each applicant is urged to prepare a proposal that reflects its unique project and aspirations. Program guidelines also change and the samples may not match exactly what is now required. Please use the current set of application instructions to prepare your application.

Prospective applicants should consult the current Office of Digital Humanities program application guidelines at <https://www.neh.gov/grants/odh/digitalhumanities-advancement-grants> for instructions.

Applicants are also strongly encouraged to consult with the NEH Office of Digital Humanities staff well before a grant deadline.

Note: The attachment only contains the grant narrative and selected portions, not the entire funded application. In addition, certain portions may have been redacted to protect the privacy interests of an individual and/or to protect confidential commercial and financial information and/or to protect copyrighted materials.

Project Title: World Historical Gazetteer: Toward a Digital Epistemology of Place

Institution: University of Pittsburgh

Project Director: Ruth Mostern

Grant Program: Digital Humanities Advancement Grants, Level III

List of Personnel

Project Team

Ruth Mostern, Principal Investigator & Project Director
Associate Professor of History
Director, World History Center
University of Pittsburgh

Karl Grossner, Technical Director & Lead Developer
Research Affiliate, World History Center
University of Pittsburgh

Alexandra Straub, Project Manager & Managing Editor
Research Associate, World History Center
University of Pittsburgh

Nathan Michalewicz
Digital World History Postdoctoral Fellow, World History Center
University of Pittsburgh

Catherine Fratto
Engagement Coordinator, Asian Studies Center
University of Pittsburgh

Narrative

Overview

This is a proposal for funding to develop infrastructure, content, and community for Version 3 of the [World Historical Gazetteer](#) (WHG), a software platform for linking knowledge about the past via place. In 2021, the World Historical Gazetteer won a 2,910-ballot vote for “Best DH Tool or Suite of Tools” in the annual [Digital Humanities Awards](#) which recognize talent and expertise in the digital humanities community.

Spatial humanists and others have long recognized the enormous integrative potential of using places as common points of reference for heterogeneous information. To reach this goal, collections of named places must be abundant, diverse, collectively assembled, and historically deep. Given that the names and attributes of places vary over time and between communities, often in the context of struggles over power and authority, ambitious development of linked, diverse, and multilingual gazetteers is absolutely vital. Maps depend upon geometry, but dots and shapes on a map tell only a small part of the story of a place. The rest of the information about it – for instance, what it has been called, by whom, and why; who has been there and what has occurred there; who has contended for authority over it, and what texts have referred to it – has historically been the province of gazetteers.

The WHG is already a powerful tool for scholarly collaboration and a part of backend architecture for named entity recognition, digital mapping, and library search. It is on a trend of expansion in content and community and is now positioned to enter a new stage of development. An NEH DHAG grant will permit the WHG to transition beyond being a successful and relatively visible single-institution experiment. It allows the WHG to become the piece of core digital humanities infrastructure that we have always envisioned: a platform, supported by a large user base and by numerous and diverse stakeholders, that is sufficiently stable, well-documented, and rich with features and content to persist and grow into the future through a subscription or consortium model.

At present, the WHG is a rich-featured environment that includes tools for Searching and Exploring open access historical place data; Uploading data and then Extending it by reconciling it to Wikidata, Linking to data in the WHG index and Publishing it, Teaching with WHG-focused lesson plans, Building and Sharing collections of thematically linked place information, and Integrating WHG data into other projects by means of an application programming interface. The WHG platform allows data creators to establish links between their individual place records and those from other sources already in the WHG index, a capability that distinguishes it from the metadata catalogs of repositories and clearinghouses. WHG is a platform for assembling rich, multivocal, connected, and distributed descriptions of place using Linked Open Data methodologies. The WHG is a digital humanities platform that supports storytelling about power and resistance by connecting place names across time and space.

This proposal fulfills the aims of the NEH call for a “More Perfect Union” by surfacing suppressed place names and difficult histories and by supporting peoples’ discoveries about past places. It allows genealogists and others to discover common historical connections to

places, even if ancestors had different experiences at them and may have called them by different names. A visitor to the WHG who searches for Ayers Rock finds information about Uluru. A search for Tenochtitlan returns Mexico City and Ciudad de México (and vice versa), and a search for Batavia links to Jakarta as well. This proposal also meets the NEH goal to “Advance, Support, and Empower Libraries and Museums” by permitting search and discovery with reference to a thesaurus of multiple names for places.

WHG community and content are on a growth trajectory. The WHG currently has 250 registered users, a number that has almost doubled within the last year. We have had 8,490 distinct visitors since launching in July 2020, a number has tripled since last year. 2,755 visitors come from the US. The remainder come from the UK, China, Germany, the Netherlands, and 134 other countries. As of June 2022, the project has 783 Twitter followers. In addition to its use in the WHG, our [Linked Places data interconnection format](#) is seeing increasing uptake in other projects including [PeriodO](#) and the British Library’s [Locating a National Collection](#).

Since launching, the World Historical Gazetteer has published approximately 2 million place records, of which approximately sixty thousand are temporally scoped. The remainder includes 1.8 million records from the Getty Thesaurus of Geographical Names and from sources of physical geography. The number of uploaded datasets has tripled since this time a year ago. Moreover, we have a queue of over 70 potential contributions at various stages of preparation or consultation totaling approximately 200,000 to 300,000 records. The datasets include six about Africa, eighteen about Central and Eastern Europe and Central Eurasian, eight about East Asia, five global, five Latin American, eleven Middle East and North African, and eleven Western European. **Appendix A lists these existing and prospective contributions.**

Places make meaning in connection with one another, so we are also creating collection building tools that make the WHG a true spatial history tool (Appendix B, Figs. 3, 4). Our Dataset Collection feature enables individuals or collaborative teams to link multiple datasets into focused domains for given regions, periods, or themes. Our Place Collection feature allows authors to link place records drawn from multiple WHG datasets. In Fall 2022 we will introduce Region Collections, to link all indexed places that lie within the asserted bounds of a historical region or to depict historical regions with unknown or unbounded extent. In Fall 2022 we will also add the ability to share collections within user-defined groups, a key request of the school teachers whose input has shaped our collection design efforts.

Our Teaching page, which launched in Spring 2022, currently includes eight pilot lesson plans. Our API to publish place records has been extended based on developer requests, enabling integration with other platforms. We have held numerous advisory panel and community feedback meetings, teacher focus groups, and public workshops. PI Ruth Mostern and Technical Director Karl Grossner have delivered dozens of conference presentations and invited talks and have also consulted privately with many project teams. **Appendix C lists some key events, meetings, and publications.**

Developing WHG Version 3 is urgent now because there is so much momentum in the humanistic linked data community. Participants in spatial humanities communities increasingly recognize that creating maps and working with GIS are not necessarily the most appropriate

end goal for many spatial history projects and that spatial history does not equate to historical GIS. Scholars leading specialist projects are eager to connect to one another via the WHG by using names as a common point of reference. The WHG has engendered an eager community of interest, evinced by our recognition with a Digital Humanities Award. It is imperative that we maintain the trust and attention of our community.

Enhancing the Humanities

The WHG team is uniquely positioned to create tools, platforms, content, and community for the history of place at the global scale. Indeed, we are the only group that currently occupies this domain. Through PI Mostern and her leadership of the University of Pittsburgh World History Center, the WHG team is closely affiliated with the intellectual mission and scholarly community of world and global history. As world history teaches, and as the WHG instantiates, peoples have always come into contact with each other. In doing so, they create new named places, new spatial networks, and new spatial relations. The WHG is unique as a digital humanities project that is global in scope and that is methodologically informed by the large, diverse, and dynamic teaching and research field of world history. The global content of the WHG offers a platform for conducting world historical reasoning and teaching.

The WHG is conceptually impactful for the spatial humanities field. Spatial humanists are fascinated by maps, and understandably so: they can be compelling and beautiful. However, much humanistic spatial information appears in texts and takes forms that are difficult to map, since it refers to named places that hold significance for individuals and groups of people – on their own, in relationship to one another, and with heterogeneous monikers and associated features. Gazetteers, not maps, are the key platform and methodology for engaging with humanistic information about meaning-making places. Spatial history will more effectively attain its objective to promote “rich” and “deep” humanistic approaches through sophisticated attention to the concept of place and linked places.

Our platform also extends best practices in humanistic digital collaboration. Dataset contributors and collection authors assemble heterogeneous and even contradictory attestations from diverse sources without overwriting or merging any of it and without privileging any one history of a given place (Appendix B, Fig. 1). These tools offer capabilities for people to gather asynchronously around shared interest in places and to platform suppressed or contentious histories. Version 3 will build out tools for accessioning and peer review, ensuring that these practices are embedded in a scholarly ecosystem.

The WHG is becoming an invaluable tool for enhancing other spatial history projects. The work of resolving place references found in historical sources is increasingly central to many digital humanities research projects, but it can be difficult. At present, projects often duplicate one another’s efforts. The WHG saves projects costly repetition in finding coordinates, alternative names, and other attributes of the places they identify. The WHG also makes it considerably easier for developers of historical place datasets to publish them so they can be readily re-used, returning results in standardized formats as opposed to repositories of files in ad hoc, project-specific formats. This makes it much easier to access and reuse granular and specialized

gazetteer data, either from WHG directly via its graphical web interface or APIs, or from annotation and visualization tools like [Recogito](#) that can make use of the data we index. Larger projects also benefit from contributing to WHG by exposing their work via links back to their typically rich records and associated thematic content. Of course, there is benefit to the knowledge commons generally whenever researchers share data, and there is benefit for diversity when many contributors from a range of institutions and backgrounds link place name records together.

Environmental Scan

Nearly all existing gazetteers, map navigation systems, and geolocation services are non-historical, and among those existing systems that are explicitly historical, none aspire to global coverage, and none provide a comparable range of features and services. Although a number of web-based digital resources catalog place information, none are directly comparable to the WHG, which is distinctive in having seven characteristics in combination: 1) global geographic coverage; 2) structured, granular temporal information for both place records as a whole and important individual attributes; 3) records that are attestations of place drawn from historical sources by researchers; 4) multiple attestations for a given place that are linked using tools embedded in the system, thereby linking research products across disciplines and projects (Appendix B, Fig. 1); 5) faceted search across datasets at the level of the record, not metadata-level; 6) serving as a Linked Data publishing platform for smaller, specialist historical gazetteers; and 7) providing machine access via an application programming interface (API).

Appendix D is a bibliography of related projects.

History of the Project

The WHG is based at the University of Pittsburgh [World History Center](#) (WHC), Pitt's home for the study of the past at the global scale, with over a decade of focus on digital methods for the study of world history. WHC director Ruth Mostern is the PI for this grant. As the administrative home for the WHG, the WHC offers in-kind staff support, intellectual guidance, publicity, institutional sustainability, and access to a large constituency of users and contributors.

An NEH Digital Humanities Level I Start-Up grant from 2014 funded an organizing meeting of experts in the field of gazetteer development, social scientists, information scientists, and linked data specialists. The two-day meeting laid the groundwork for the development of a subsequent NEH Preservation and Access Grant proposal, "World-Historical Gazetteer," awarded in 2017 for work over a three-year period (PW-253719-17) and resulting in the Version 1 release of the WHG. From its inception, the WHG project team has cultivated a global community of interest invested in the long-term success and sustainability of this important resource. **Appendix A**, which lists indexed datasets and datasets in development, and **Appendix E**, which lists participants in open meetings in 2021-2022, reflects the breadth and diversity of our global community. The WHG team also collaborates with contributors of place data on every continent, including, for instance, [Enslaved: Peoples of the Historical Slave Trade](#) on the history of African and African-descended populations on both sides of the Atlantic, which includes an index of over 6,300 places, the [China Historical GIS](#), and the [Historical Middle East Data Alliance](#). Since our Preservation and Access Grant concluded, we have released Version 2 in August 2021,

supported with funding from the Royal Netherlands Academy of Arts and Sciences' Humanities Cluster ([KNAW-HuC](#)), and 2.1-beta in April 2022, supported by the Pitt Asian Studies Center.

The WHG has numerous and close links with Pitt's [University Center for International Studies](#) (UCIS), which offers leadership for Pitt's global programs and partnerships. We collaborate with multiple area studies centers, two of the three of which are Title VI funded National Resource Centers ([Asian Studies](#), [Latin American Studies](#), and [Russian, East European and Eurasian Studies](#)). We are jointly planning workshops with representatives from the [Title I](#) public schools and [HBCU](#) campuses with which the UCIS centers have developed relationships, building on a Spring 2021 pilot focus group for teachers held in collaboration with the Asian Studies Center. PI Mostern will present at the UCIS [Summer Institute for Global Educators](#) in July 2022.

In 2021 we developed a Memorandum of Understanding with KNAW-HuC, the institutional home of a vast content collection including the complete records of the Dutch East India Company. [The International Institute for Social History](#) (IISH), part of KNAW-HuC, is one of the most prestigious social history research institutes in the world. The IISH hosts the Indian Ocean and Maritime Asia Slave Trade Database and actively engages in international debates and outreach related to slavery and coerced labor history. KNAW-HuC also fosters digital humanities research and develops digital infrastructure for it. It is a founding partner in [CLARIAH](#), the Dutch Common Lab Research Infrastructure for the Arts and Humanities. KNAW-HuC funding supported WHG Version 2 software enhancements and assistance in accessioning datasets from the KNAW-HuC research community. A new KNAW-HuC contract is supporting integration between the WHG and a range of open-source text and map annotation software libraries. KNAW-HuC is currently evaluating the potential of integrating the WHG into the Dutch national digital infrastructure. The WHG is also a founding and active partner in the [Pelagios Network](#), a global association devoted to creating and sharing historical geodata. The WHG has a Memorandum of Understanding with the Pelagios Network Association to ensure that content and code are highly visible across multiple projects.

Activities and Project Team

The proposed work falls into **three interlocking categories of activity** to improve governance, expand content and community including establishing an editorial workflow, and complete needed software improvements. Each proposed task within these categories addresses our vision to transform the WHG from a small-team project into a robust piece of global infrastructure with an effective and diverse management structure, integrated with large-scale library and institutional platforms and workflows, and widely used in classrooms. The **Work Plan** and **Sustainability** documents describe these activities in more detail.

Governance

By the end of the grant period, the WHG will have a community of core stakeholders from multiple institutions serving in key roles and individuals transitioning in and out of project team roles without disruption. The Technical Director will transition out of software development and into directing the work of contract developers. The Project Director/Principal Investigator will become the Chair of a Board, diverse in race, gender, and national origin, which solicits content,

raises funds and visibility, guides the WHG vision, hires staff and contractors, and conducts outreach. Bylaws will describe an organizational structure, decision making protocols, and leadership rotation.

We will also consult with digital humanities infrastructure projects that have successfully transitioned out of grant funding and PI leadership and into long term financial and governance success and growth momentum. **Appendix F lists projects that have achieved sustainability by several means.** Following broad-based consultation, we will identify suitable subscription and governance models for the WHG. We are already exploring these topics with our colleagues at KNAW-HuC and with the Dean of the School of Arts and Sciences at Pitt.

Accomplishing these goals requires enhancement of software features and system-level development tasks including code containerization, updates and maintenance of core software dependencies, improved codebase organization and documentation to facilitate onboarding of new developers, and implementation of a regular automated depositing of WHG data to the Pitt Libraries repository.

Content and Community

The WHG currently has a contribution bottleneck. We will make the steps for contributing simpler and clearer, and we will reorganize staff time to ensure that we are assisting contributors at each of the points of friction that we have identified. As the WHG and its community grows, it is also now urgent to implement peer review for datasets, collections, and lesson plans. **Appendix G** describes in more detail how we will ensure that contributions move through uploading, accessioning, indexing, and peer review with support from documentation, software, and various human roles. Success in this category of activity means that we can actively solicit content in full confidence that it will move smoothly through a workflow and through review procedures that accord with scholarly best practices.

An appealing, vast, and diverse index of place name data is the core of the WHG mission. We will devote substantial upfront investment in developing datasets and collections from diverse and global populations by collaborating with contributors, fostering communities organized around particular regions or themes, and accessioning publicly available data. We will conduct two workshops annually and develop a template for dataset contributor workshops. We have about 120,000 records pending in near-term contributions. The owners of an additional 200-300,000 records have expressed interest in contributing. We are confident that we can add those, and additional hundreds of thousands of records beyond those, during the grant period. We will also add significant content in the form of Dataset Collections, Place Collections, and Region Collections.

This category of activity also includes our plan for teacher workshops and educational content development, especially focusing on teachers in Title I schools and HBCU college campuses, who will receive honoraria to collaborate in developing lesson plans, collections, and syllabi. The workshops will explain how to use the WHG to connect Americans to their ancestries and their diverse personal and family histories. We will also develop a template for WHG teacher

workshops. We will conduct two teacher workshops per year, each with one to two dozen participants.

These tasks require new software features for managing, finding, displaying, and sharing content as well as refactoring the site for internationalization and initiating crowdsourced translation of site text language. Several other software development tasks associated with these activities will make the contribution pipeline more efficient and user-friendly, enhance and extend existing features to make the project web site more useful and usable, and extend integration possibilities. Software development and documentation tasks that will enhance contributor experience and make the WHG into a true data publication platform include more robust data validation functions, progress tracking tools, and administrative screens for use by the WHG editorial staff.

Additional Software Development Tasks

Some needed software development tasks, many related to the website frontend experience, fall outside the two categories of activity detailed above. The WHG platform has six maps, all of which need some combination of enhancements, including enabling temporal controls, better interaction with accompanying data tables, and consistent symbolization. The public presentation of collections and references to them in Place Portal pages need several usability improvements, and collections need to be better incorporated into the site's search, browse, and API features. We will also complete work recently begun that will support a full-featured integration with the open-source [RecogitoJS](#) and [AnnotoriousJS](#) annotation software modules (for text and map images respectively), permitting remote read and write capability with WHG data stores. Users of any web software implementing those modules will be able to georeference documents directly against WHG data, and to create and add to WHG datasets as they work.

Project Team

The core team consists of PI *Professor Ruth Mostern*, who directs aspects of the project that concern scholarly content, educational applications, assessment and evaluation, scholarly ecosystems, relationships with other responsibility centers at Pitt, and governance and financial management, and who is the Director of the World History Center. || *Dr. Karl Grossner*, the Technical Director of the WHG, is its software architect and lead developer, and is a global leader in information architecture in the GeoHumanities. Grossner sets WHG objectives and transforms them into design and development achievements and also works closely with content developers. || *Dr. Alexandra Straub*, the Research Associate of the Pitt World History Center, is a permanently funded staff person who tracks milestones, organizes meetings, circulates minutes, documents accomplishments, organizes funding submissions, and serves as Managing Editor and Project Manager for our publication functions. || *Dr. Nathan Michaelwicz*, the Digital History Postdoctoral Fellow at the World History Center, works with dataset contributors on data transformation and contribution workflow and assists with web map design. || *Ms. Catherine Fratto* is a UCIS staff person and former schoolteacher who arranges workshops with teachers and serves as a liaison between the WHG team and the educational community.

Grant funds will permit us to hire three additional individuals as well. A *Lead Developer* will initially work under the direction of the Technical Director and will assume many of the duties now performed by him, including feature design, backend and frontend coding, and server maintenance. || A *student employee with information science expertise* will assist with advanced data management tasks, principally transforming incoming structured historical data into formats required for upload to WHG. || *Student employees with domain expertise* will assist in accessioning data in various foreign languages. || A *student employee with assessment expertise* will assist in designing, guiding, and analyzing results for teacher workshops and will assist with other assessment and user testing activities including accessibility and universal design as well as other needed administrative and research tasks.

In order to hear a range and diversity of advisory voices, we have decided not to empanel formal boards for this grant period. Via Zoom, we will hold public meetings at least twice a year that we will announce to our email list and that we will post on the website and social media. We have already begun doing this, and **Appendix E** lists the attendees who have attended our public meetings this year. The WHG community also overlaps with the 60 partners in the [Pelagios Network](#). We will develop strategies for ensuring that we are maximizing opportunities to involve our community in decision making and feedback. We will monitor participation for diversity of gender, race, national origin, and career stage, and conduct additional outreach as needed. We will also continuously monitor how effectively the project is benefitting everyone involved: through financial compensation for developers, student employees, and teachers who participate in workshops; through professional development opportunities for students and early career scholars; and, for our community members, through opportunities to advance the scholarly enterprise, to publish articles and data, and to extend heterogeneous and diverse scholarly networks.

Final Products and Dissemination

The **Work Plan** section of this proposal includes a detailed account of tasks and final products of the above activities and their dissemination: hosting meetings, developing software and cultivating a development community, accessioning content and establishing an editorial workflow, and connecting with students and disseminating insights. The overarching goal of all these activities is to make the WHG into a self-sustaining venture that does not require grant-by grant funding after the conclusion of the NEH-DHAG grant period. We will accomplish this goal by empaneling and cultivating boards with internal plans for governance such as membership terms, bylaws, and a leadership structure; by making it easy for new software developers to get involved in coding initiatives, and by ensuring institutional commitments to sustainability distributed between Pitt, KNAW-HuC, and other institutions. We also plan a robust program of publication in a mix of genres including scholarly articles, tutorials, and blog posts. PI Mostern is writing a book about place and history, tentatively entitled *Where We've Been*. The **Budget** for this grant permits travel to conferences and workshops to conduct tutorials and deliver presentations.

Work Plan

This work plan is organized around the three categories of effort discussed in the Activities and Project Team section of the Project Narrative: Governance, Content and Community, and Software. It delineates specific tasks, responsible staff, and timeframes. Tasks entirely or partially covered by potential matching funds are annotated with a **[m]** symbol.

Unforeseen events such as potential travel restrictions and staffing delays could impact the schedule. PI Mostern and Technical Director Grossner, both highly experienced project managers, will monitor progress and adjust timing and priorities as required to ensure that we meet the goals for each activity category. Some reprioritization is normal in projects of this size and complexity, and the WHG team has an excellent track record of delivering promised work products for initiatives funded by grants and contracts.

Responsible staff and participants are indicated by the following initials. **RM**: Ruth Mostern; **KG**: Karl Grossner; **LD**: Lead Developer (tbh); **AS**: Alexandra Straub; **NM**: Nathan Michalewicz; **CF**: Catherine Fratto; **RA**: student research assistants (tbh); **GB**: governing board; **WHG**: project team; **TW**: teaching workshop participants.

Governance

The specific tasks for this activity category can be grouped as follows: a) designing and implementing a sustainable governing structure, b) designing and implementing a more flexible and sustainable staffing structure, and c) bringing what has been a single-institution pilot platform up to international standards of organization and stability .

Establish a governing board (GB) with bylaws and sustainability strategy		Y1	Y2	Y3
Identify and consult with other projects in similar situations at various stages	RM, KG, RA			
Draft white paper delineating issues and potential solutions	RA, RM			
Specialist meeting of invited project leaders/managers to discuss the general issue(s) and advise on WHG strategy	WHG, 15 invitees			
Recruit board members (emails, online meetings)	RM, KG			
Board meeting to draft bylaws	WHG, GB			
Document, disseminate	AS, NM			

Reorganize staffing		Y1	Y2	Y3
Hire and onboard new lead developer	KG/LD, RM			
Hire and onboard student employees	RM, NM, AS			

Governance, cont'd

Systems-level development tasks		Y1	Y2	Y3
Updates and maintenance of core software dependencies	KG/LD			
Improve codebase organization and documentation; containerize for portability	KG			
Script scheduled deposits to Pitt Libraries D-Scholarship digital repository	KG			
(potential) Permissions system for subscription and/or freemium access	LD/KG			

Content and Community

Many of the activities and specific tasks for increasing and diversifying content and for growing user communities are closely interrelated. More content and improved features for using it will lead to more consumers of it. That requires an easier and more efficient contribution and publication workflow, which are made possible by software tools and better staff organization. We will also attend and host an increased number of open meetings, conferences, and workshops since these also lead to more numerous and more diverse users.

Acquire and publish content (datasets, collections, lesson plans)		Y1	Y2	Y3
Identify and process candidate datasets from public sources	RA, WHG			
Perform ingest, reconciliation, and accessioning tasks	RA, AS, KG/LD			
Ongoing consultations with current, near-term, and prospective dataset contributors and collection creators	KG/LD, AS, RM, NM			
Assist contributors in data conversions	KG/LD, NM, AS, RA			
Review datasets and collections for publication	KG/LD, NM, AS, RA			
Create and publish lesson plans [m]	CF, AS, TW			

Mitigate pipeline/workflow issues		Y1	Y2	Y3
Design project management software tooling for tracking contribution progress through the system	KG, AS, RM			
Implement and iteratively improve that PM tooling	KG/LD			

Expand communities of contributors, research users, and teachers		Y1	Y2	Y3
Online public meetings (2 per year)	WHG			
Virtual and in-person conferences and spatial humanities community meetings	RM, KG/LD, NM			
Host WHG contributor workshops (2 per year); develop contributor workshop script/template	WHG			
Host Teaching workshops (2 per year); develop teaching workshop script/template [m]	CF, NM, AS			
Publications (articles, conference papers, blog posts)	RM, KG, NM			

Content and Community, cont'd

Related software features and usability upgrades		Y1	Y2	Y3
Improve data upload and validation functions, error trapping and reporting	KG/LD			
Refactor codebase for language internationalization (i18n)	LD/KG			
Additional API endpoints for integration with other systems and resources	KG/LD			

Additional Software Development

Several software enhancement tasks not directly tied to the first two categories will make existing features more useful and usable. We expect this will over time help to expand the WHG user base in size and diversity.

		Y1	Y2	Y3
Enhancements to six site maps, including heat maps to handle very large datasets and collections, interactive controls for temporal attributes, and more consistent symbolization	KG/LD, NM			
Design and implement improved representation and search for and across all types of collections: Place, Dataset, and Region	LD/KG			
Integration with open-source annotation software, RecogitoJS and AnnotoriousJS, for use with texts and maps respectively	KG/LD			

Dissemination

The WHG team will continue presenting papers and hosting workshops at relevant conferences and provide regular updates to our constituent communities in the WHG blog and on social media. The near-term conferences most suited to informing and growing the WHG community include the American Historical Association (2023, Philadelphia; 2024, San Francisco), ADHO/Digital Humanities (2023, Graz), Linked Pasts Symposia (annually in December), and the NEH directors meeting (2023, Washington D.C.).

A conference of special note is the World History Association, which meets annually, and which will come to Pittsburgh in June 2023. PI Mostern is playing a key role in organizing this event, which typically draws hundreds of world historians from both schools and universities. This will be a milestone event for workshops, focus groups, and mapathons, and an opportunity to center the WHG at a major event for some of its core consumers.

Ruth Mostern will publish academic articles about place and history and the WHG throughout the grant period and will also work on a book manuscript about place and history. In collaboration with other team members, she will also write guides to gazetteer design for publications such as *The Programming History*. The WHG team will share the task of completing other needed pieces of writing such as improved documentation, website content, and blog posts.

Data Management Plan

Roles and Responsibilities

Principal Investigator and Project Director Ruth Mostern has overall responsibility for the World Historical Gazetteer (WHG) project including its data products, and will conduct periodic reviews of data management practices and the status of contributions.

Technical Director Karl Grossner oversees the acquisition, storage, and backup of contributed datasets and related material in uploaded files, a relational database, and an index. He also leads software design, is responsible for developing and documenting the project's software code, and maintains the project cloud server.

Managing Editor Alexandra Straub oversees the editorial process: reviewing incoming contributed Place datasets and monitoring their progress through the accessioning process. She also solicits and guides collection development and dataset contributions to the World History Center's Dataverse instance.

The University of Pittsburgh Library will store a periodically updated snapshot version of the WHG index in a library-managed repository that includes rich metadata and a permanent DOI. The data will be visible in the library catalogue and will appear in search results as a citable version of the WHG index.

Expected Data

Place data. WHG aggregates and publishes data about historical places contributed by researchers studying the past within and across numerous disciplines. Registered users can upload data files in one of our contribution formats (details below). These are ingested into a relational database, and self-managed within each contributor's private workspace.

- Data within the database can be deleted or updated with subsequent uploads by its owner, but at this time not edited directly. The original uploaded files are kept for reference for as long as the dataset remains in the system.
- Uploaded records are each assigned a permanent *place_id* identifier that remains stable for as long as the dataset remains in the system. The database and uploaded files are stored on a dedicated commercial cloud server, where they are backed up regularly. Snapshots are taken of the entire server monthly (OS, software, databases, index); these can be restored efficiently if necessary.
- WHG offers a reconciliation service for augmenting uploaded data with geometry and concordance identifiers discovered by matching contributors' place records with those in Wikidata. To support this activity, we have indexed a large subset of place data from Wikidata (3.6m records). We plan to automate the regular updating of this resource.
- When a dataset is flagged as public, its records are made accessible for search, for download, and for queries via an API, effectively publishing them as Linked Open Data under a CC-BY-4.0 licence. The further step of accessioning into the WHG *union index* links multiple attestations of closely matched places. Graphical search and a separate API make those union records publicly accessible as well.

- The WHG union index resides on the same server, is backed up regularly, and is also included in the server snapshot mentioned above.

Collections are sets of public WHG records and/or datasets. A collection definition therefore consists only of its metadata and rows of collection_id and place_id or dataset_id pairs. These reside in the relational database, subject to the same management policies as the place data itself. Accompanying PDF essays and images are stored in the server file system and maintained for as long as the collection exists.

Teaching resources such as lesson plans and syllabi are stored in the server file system and maintained in the same manner as data files and collection materials.

Software. The WHG platform software code is maintained in [a public GitHub repository](#) under an open source GPLv3 licence. WHG is principally a Django framework project (v2.2.20) and makes use of PostgreSQL databases and Elasticsearch indexes.

Period of Data Retention

Both data stores—database and index—are considered permanent (*see Preservation of Access below*). Uploaded data and related files will be maintained in the system unless and until its creator deletes the associated object(s), and is backed up and replicated as described above.

Data Formats and Dissemination

Formats. To permit data from multiple sources to be searched and used together, the WHG project has led development of a standardized interconnection format for representing places and their dynamic attributes, the [Linked Places \(LP\) format](#). LP format is valid JSON-LD (a syntax of RDF), and valid GeoJSON, extended to enable temporal scoping of an entire feature and individual attributes such as names, extents, types, and relations to other places. An alternative delimited file format, [LP-TSV](#), can be used for relatively simple data. Because uploads to WHG must take one of those two forms, a transform from the contributor's working format is required.

Metadata. WHG serves as a data publisher for smaller specialized gazetteers. Our existing metadata record created by data owners for contributed datasets will be expanded to reflect this, adopting a subset of the W3C/DCAT vocabulary. The choice of attributes will reflect requirements for scholarly evaluation and attribution.

Dissemination. WHG makes all data holdings available via an application programming interface (API) and several types of downloads: individual records, search results, entire datasets, and collections, in the Linked Places format and CSV. Endpoints for the API are added continually to facilitate integration with other systems.

Data Storage and Preservation of Access

All data in both the database and index are held indefinitely on the WHG server and will be replicated and periodically updated in the [D-Scholarship digital repository](#) at the University of Pittsburgh, along with accompanying metadata.

Access to WHG data is provided principally by the graphical interface of its web platform. The Narrative and Sustainability Plan documents describe in detail our intentions and strategy for providing ongoing support for the platform's continued enhancement and maintenance.

Sustainability Plan

A sustainable World Historical Gazetteer is a permanent, free, and globally accessible resource that grows continually via contributed content. In addition to sustaining data and software, realizing this vision requires multiple communities of interest to invest in its long-term success and to offer a modest level of ongoing financial support.

Sustaining Data and Software

Data gathered into the WHG database and index will be replicated as data objects with accompanying rich metadata on a regular schedule to Pitt Libraries' [D-Scholarship digital repository](#), ensuring that it is stored permanently and sustainably according to the highest institutional standards for digital scholarship. The index will be assigned a DOI, will be visible in the library catalog, and will appear in search results, supporting use of WHG data in scholarly communication ecosystems by creating a permanently accessible and citable version that can be used elsewhere at any future time under a CC-BY-4.0 license. This aligns closely with the best practices outlined in the NEH-funded *Socio-Technical Sustainability Roadmap* (STSR) developed by our colleague at Pitt, Alison Langmead. Apart from that, the database and index are backed up regularly such that they can be quickly recovered in the event of a server malfunction.

We will also evaluate the library's open access journal publishing (OJS) platform as a platform for managing our editorial workflow and peer review process. This would offer the possibility of even tighter integration with the library's systems.

The open-source software code for the WHG platform is maintained in [a public repository on GitHub](#) under a GPLv3 license. The codebase will be made more amenable to prospective development collaboration by means of tasks detailed in the Narrative, the Data Management Plan, and the Work Plan.

Sustaining Governance and Finances

A sustainable WHG requires a commitment to funding and governance both at the University of Pittsburgh and across our global community. We are targeting an annual permanent budget of around \$120,000. Now is an opportune time to pursue this goal. As of our Version 2.1 release, the WHG has become a fully featured platform and a viable and valuable candidate for hard money support.

We have already taken many of the steps recommended in the Ithaka S+R Briefing Paper referenced in the NEH DHAG call. The administrative home for the WHG, the [World History Center](#) (WHC), offers ongoing staff support and access to server space and connections with services and synergistic projects throughout Pitt and in national and global communities of world historians. The WHC has a compensated Director and Associate Director, Research Coordinator, postdoctoral fellow, and graduate student researcher. The WHC also devotes several thousand dollars of funds annually to WHG activities. The Pitt Center for Research Computing administers virtual machine server space for the WHC. In recognition of the importance of the WHG to the research mission of the WHC and the stature of the WHC at Pitt, the Dietrich School of Arts and Sciences is offering a \$15,000 match in support of this grant application.

We have been invited to approach the Dietrich School Dean about the potential of permanent WHG funding. She has also recommended that we reach out to the Associate Vice Chancellor for Research for the Humanities and Social Sciences, the Vice Provost for Global Affairs, and the University Librarian. Other responsibility centers at Pitt are also committed to the long-term success of the WHG. The University Library will support regular uploads from the WHG index to a library repository. The University Center for International Studies (UCIS), directed by the Vice Provost for Global Affairs, leads Pitt's global programs and partnerships, advances the university's global vision, and supports multidisciplinary research with global impacts. Several UCIS area studies centers are presently writing support for the WHG into their Department of Education Title VI grant applications.

At the same time that we seek sustainable permanent funding at Pitt, we are also taking steps toward sustainability as a global project. Our partnership with the [KNAW Humanities Cluster](#) in the Netherlands includes a possible commitment to long term maintenance of the WHG codebase and future contributions to development. However, in order to justify a potentially significant budget for this, they are requesting that we transition to governance by a globally distributed Board.

To that end, we are drawing on our significant network of partnerships. The WHG is a founding and active partner in the [Pelagios Network](#), a global association for creating and sharing historical geodata. This international partnership ensures that best practices are visible and shared throughout a broad scholarly ecosystem. The Pelagios Network is itself part of a broader Linked Pasts international community, which has since 2015 staged an annual symposium that draws dozens of participants each year. WHG Technical Director Grossner has been an active organizer and participant in that series, a key venue for identifying contributors, board members, enthusiastic supporters, and potential code collaborators.

Appendices A and E suggest the further breadth of expertise and institutional affiliation among our community of supporters, content contributors, and attendees at open meetings. These are key individuals for contributing to our momentum and success, for ensuring growth and uptake of the project, for joining governance boards, and potentially for contributing financially to the future of the WHG. Most of these individuals are affiliated with large institutions, and we will identify possibilities for them to contribute to sustainability by funding discrete tasks or joining boards. We will create clear paths for identifying and onboarding collaborators into these roles.

During the grant period, as reflected in our Narrative and Work Plan, we will be converting these significant strengths into a tangible governance and financing structure. We will survey project sustainability exemplars, consult with domain leaders, draft a white paper, and hold a meeting. At that point, we will finalize and implement a plan. We will identify how our growing and potential individual and institutional user base of historians, libraries and museums, schoolteachers, and genealogy hobbyists may be able to engage with a subscription model.

Appendix F lists some of the sustainability exemplars that we are currently researching in the areas of software platforms, data publication, and crowdsourced spatial and historical content.