Public participatory historical GIS


To cite this article: Don Lafreniere, Luke Weidner, Daniel Trepal, Sarah Fayen Scarlett, John Arnold, Robert Pastel & Ryan Williams (2019): Public participatory historical GIS, Historical Methods: A Journal of Quantitative and Interdisciplinary History

To link to this article: https://doi.org/10.1080/01615440.2019.1567418

Published online: 20 Feb 2019.
Public participatory historical GIS

Don Lafrenierea, Luke Weidnerb, Daniel Trepala, Sarah Fayen Scarletta, John Arnoldab,c, Robert Pastela,e, and Ryan Williamsd

aDepartment of Social Sciences, Michigan Technological University, Houghton, Michigan, USA; bDepartment of Geology and Geological Engineering, Colorado School of Mines, Golden, Colorado, USA; cUpper Peninsula Engineers and Architects, Houghton, Michigan, USA; dGeospatial Research Facility, Michigan Technological University, Houghton, Michigan, USA; eDepartment of Computer Sciences, Michigan Technological University, Houghton, Michigan USA

ABSTRACT

Building historical geographic information system (HGIS) datasets is time consuming and very expensive, especially when built at the scales that permit analysis of the lived experiences of individuals or the morphology of buildings or streets. Further, these datasets are often built exclusively in the academy, with little input from the contemporary communities they represent. In this paper, we review the use of the public in crowdsourcing historical data creation, and using the Keweenaw Time Traveler set in Michigan’s Copper Country as a case study, we call for a new approach to HGIS scholarship that includes a robust public partnership to building HGIS datasets. The creation of a public participatory HGIS approach to HGIS scholarship can increase efficiencies of, public relevance in, and extend the reach of, HGIS projects beyond the academy. We have established a set of best practices that include, incorporating the public in the HGIS interface design, providing immediate public data access, contextualization of spatial data in space-time, comprehensive public history outreach in person and online, and creating affordances for the public to contribute their own historical spatial knowledge through spatial storytelling. Together, these activities can promote the long-term sustainability and success of historical data crowdsourcing projects.

KEYWORDS

Historical GIS; public history; crowdsourcing; public participatory GIS; spatial storytelling

Introduction

Whether you favor historical geographic information system (HGIS), historical geography, spatial history or spatial humanities, as your sub-disciplinary moniker, all share the same interest and need for geographic information, especially GIS-ready historical geospatial data. Historical geospatial data is slow, tedious, and expensive to capture, create and standardize (Gregory and Ell 2007, 41). While the methodology for creating and capturing historical data is different than that of scientific data; the costs for natural, social, physical and health scientists are similar (Sauermann and Franzoni 2015). One response has been to involve the broader public (i.e., “the crowd”) in the research process, especially data collection and creation. With its own set of monikers: citizen science, crowdsourcing, participatory mapping, and public participatory GIS, each differs in their approach to how they involve the public in the research process. This paper reviews the use of the public in historical data creation, and through a comprehensive case study, calls for a new approach to HGIS scholarship that includes a public partnership. The creation of a public participatory HGIS (PPHGIS) approach to historical GIS scholarship can increase efficiencies of, public relevance in, and extend the reach of, HGIS projects beyond the academy to support heritage community building and preservation, environmental history, genealogical research, urban histories, and promote awareness of the past through the lens of geospatial technologies.

Public historians have long understood the value of working with the public to understand, communicate, and discover past environments, peoples, and events (Britton 1997; Cauvin 2016; Clarke 2004). Sociologists and urban geographers have used community-based participatory research for decades, citing benefits such as more quickly identifying appropriate community-relevant research questions, increased support and understanding of academic scholarship, more rapid knowledge translation of findings to community action and benefit (Hacker 2013). Researcher-Public Partnerships have been critical to heritage
organizations that work to balance preservation efforts with community-driven interests in historically-significant properties resulting in grass-roots efforts to raise their communities to a federally-recognized protected status such as that of the Lowell or Keweenaw National Historical Parks (See 2013). Historical demographers have also benefitted from the partnership between the Minnesota Population Center and the genealogical community of the LDS-Mormon church resulting in a significant cost savings to academic researchers in having access to transcribed historical census data (Ruggles 2014). HGIS scholars could also benefit from the same efficiencies as their demography colleagues by working with the public to create GIS datasets, especially acute when National Endowment for the Humanities funding is currently under threat (Farrell 2017). Further, we argue that employing a community-based participatory research approach can improve the public relevance of, and support for, HGIS and spatial humanities scholarship.

**Short history of building historical geospatial datasets**

The use of GIS for historical research was championed by a small cohort of historical geographers in the early 2000s with the first substantial collection of work finding its way into a special issue of *Social Science History*, edited by American historical geographer and early vocal cheerleader of HGIS Anne Kelly Knowles (2000). It was followed quickly by a special edition of *History and Computing*, edited by British historical geographers Paul Ell and Ian Gregory, that outlined the rapid progress and future directions of the emerging discipline (Ell and Gregory 2001; Gregory, Kemp, and Mostern 2001). Early efforts in HGIS research concentrated on the creation of national-scale HGIS projects by digitizing and spatially-referencing large official record sets such as parish or county level censuses (Gregory et al. 2002; Fitch and Ruggles 2003; McMaster and Noble 2005), boundaries (Bol 2008; De Moor and Wiedemann 2001; Kunz and Boehler 2005), and gazetteers (Mostern and Johnson 2008; Southall, Mostern, and Berman 2011). Subsequent projects demonstrated the ability of HGIS to support historical research at very small spatial scales and high spatial resolutions using city directories and tax records (Carrion et al. 2016; Debats 2008, 2009; Dunae et al. 2011, 2013; Gilliland and Olson 2003, 2010; Lafreniere 2014; Lafreniere and Gilliland 2015, 2018; Van Allen and Lafreniere 2016). Several books quickly followed (Gregory 2003; Knowles 2002, 2008) which have captured the attention of non-geographers, specifically historians, and prompted them to explore using GIS for historical inquiries. HGIS has also particularly caught attention of scholars with interests in environmental history and past landscapes (Baeten, Langston, and Lafreniere 2016; 2018; Bonnell and Fortin 2014; Clifford 2017). HGIS scholarship has also inspired humanists to embrace a geospatial approach to their craft, resulting in five recent edited collections (Bodenhamer, Corrigan, and Harris 2010, 2015; Dear et al. 2011; Gregory, Debats, and Lafreniere 2018; Gregory and Geddes 2014; Ridge, Lafreniere, and Nesbit 2013).

HGIS as an academic discipline is now in its adolescence. Researchers have recognized the “revisionist” potential of HGIS-based analysis as a means to challenge established explanations for historical events by examining historical data from previously unavailable perspectives (Knowles 2016; Olson and Thornton 2011). These projects established the value of HGIS as a research approach, but also exposed the biggest challenge to HGIS research – the enormous expense in time and resources required to build the necessary geospatial datasets (Knowles 2016; Southall 2014, 114). This challenge has inspired our conceptualization of a PPHGIS approach which is outlined below after a review of the current state of crowdsourcing and other public participatory scholarly activity in the historical sciences and humanities.

**From historical crowdsourcing to PPHGIS**

Where there has been much less attention by scholars is the integration of HGIS approaches with public-focused academic scholarship. In the few examples where the public has been employed in HGIS scholarship it has been limited primarily to data collection in the form of “volunteered computing,” or as Haklay argues, the most basic of citizen science, “crowdsourcing” (2013, 116). In this section, we outline the contributions and limitations of existing historical-data focused crowdsourcing projects before transitioning to a case study that demonstrates a shift from the more passive public-academic collaborations with crowdsourcing to a more active public participatory historical GIS model.

**Crowdsourcing georeferencing**

The scanning and georeferencing of historical maps are critical steps in the creation of an HGIS; making such digital, spatially-referenced historical maps available to researchers and the public is a key step in maximizing the potential of HGIS (Rumsey and Williams 2002). The earliest crowdsourcing projects were the creation of web applications for
georeferencing historical scanned maps. Among the first is the Map Warper project, created by the New York Public Library’s NYPL Labs team. This successful, free, publicly available, web-based georeferencing tool allows users to upload, georeference, and export virtually any digitally-scanned map to a modern OpenStreetMap basemap (Vershbow 2013). As of late 2018, just over 28,000 maps have been georeferenced by the public and NYPL staff which can be browsed at www.maps.nypl.org/warper.

Another successful web-based georeferencing tool, Georeferencer (found at Georeferencer.com) has been developed by Klokan Technologies as a commercial application (Fleet et al. 2012). It has been utilized by several major map libraries and collections including the Cartography Associates – David Rumsey Map Collection, the National Library of Scotland, the British Library, the Institut Cartographic de Catalunya in Barcelona, and Old Maps Online. Georeferencer also employs the public to assist in georeferencing historic maps to an OpenStreetMap baselayer and makes them searchable in a web-based GIS. As the tool is divided across several institutions, it is unclear how many maps have been georeferenced by public contributors, but it appears to be substantial.

Both projects include detailed how-to documentation and emphasize that accuracy is important, though neither system specifically outlines how they explicitly control for quality. Map Warper presents users with a root mean square error, which may not be well understood by the general public, but they do suggest a target error threshold and encourage users to distribute a minimum of four control points around the extent of the map. Map Warper includes a wiki-style comment section and editing activity tracker that encourages collaborative problem solving of challenging maps to georeference. Maps appear to be infinitely editable, allowing users to edit other users work if they see fit. Georeferencer has fewer quality control tools but it does allow a user to deem a map as finished, presumably removing it from being edited by other users. Both projects present users with a modern basemap for which to georeference against rather than providing the option to georeference against another temporally specific map set, which for some geographies and fine spatial scales, may make georeferencing challenging for public contributors.

Crowdsourcing transcriptions and classifications

A more substantial area where crowdsourcing has been employed is in the transcription of historical documents or records. The leader in this approach is the Citizen Science Alliance, who through their Zooniverse platform, hosts a number of projects including Operation War Diary, Old Weather, and Measuring the ANZACs (Australian and New Zealand Army Corps). Each project includes a customized web-interface with detailed instructions provided to volunteer transcribers. Transcription projects in Zooniverse use the Scribe plugin and ask users to first markup historical documents with shapes by which the transcribed text can be linked to areas on the scans of the historical documents. Transcribing relies on this markup. Once marked up, users can login and only contribute transcriptions of areas marked up by other users – not areas they marked up themselves. Zooniverse is very successful at recruiting volunteer transcribers, for example, the ANZACs project had over one million fields transcribed in less than 2 years (Roberts 2017).

A key limitation to Zooniverse projects however is there is no quality control for the transcriptions. Researchers can require a masked area to be transcribed by up to three unique users, in the hope that errors can be easily identified, however, no text matching or consensus algorithms are available to the public or the researchers to allow them to confirm if there is agreement on the transcription of a particular section of text. Further, users cannot see the transcriptions from other users, nor can they search or query the datasets that they created. Thus, data is collected from public transcribers however the public has no ability to access or use the data created. To date, no Zooniverse projects include a spatial component.

NYPL Labs, the same research team within the New York Public Library that developed Map Warper, has recently launched a transcription and coding project that does include a spatial component. The “Building Inspector” has three crowdsourcing tasks (https://buildinginspector.nypl.org/). The first asks users to note whether a computer-generated digitization of historical building footprints are accurate and if not, invites users to fix them. The second task asks users to transcribe the street addresses noted on the georeferenced historical maps. Users are presented with a random building footprint (generated and confirmed by the first task) and are then asked to place a point on the street numbers on the map, and finally to type in what they see. In the final task users are again presented a random building footprint and they are asked to classify which color the building is which indicates its use. All three tasks encourage users to continue to engage by making the contribution as
quick and simple as possible, creating an addictive gaming effect.

Building Inspector does improve on the Zooniverse model by requiring consensus before information is recorded as accurate. On their website they outline an example using the footprint inspector where at least three different people must agree that a footprint is accurate in order for that data to be classified as accurate. Though it must be noted, that users are not made aware that they have contributed to a classification or transcription reaching consensus. Similar to the transcription projects in Zooniverse, users of the Building Inspector cannot see what others have transcribed or classified, nor can they contribute to improving the data provided by other contributors.

Where Building Inspector also has an opportunity to improve as a publicly-engaged project is in how they explain the purpose and importance of their crowdsourcing activities to their public users. The experience is void of any historical geographical context. Users do not know where they are in space-time, which map there are working with, which neighborhood of New York they are in, or what time period the map represents. Zoom and pan tools restrict users movement to a scale of 1:100, focusing the task but limiting the ability to contextualize the task within the urban geography. The applications also do not allow you to switch to a contemporary basemap which would allow users to place themselves and their crowdsourcing tasks within a familiar modern geography.

The interface is graphically attractive and engaging for users however, similar to Zooniverse transcription projects, public contributors cannot see the data created or interact with it on the webmaps. A data download section is available, however, data is only able to be extracted in GeoJSON or NDJSON, a format that is only usable by expert GIS or computer science trained users.

An exemplary example of a crowdsourcing transcription project is GB1900 (www.gb1900.org) which adds a spatial element to transcription crowdsourcing and makes efforts to engage a wider public beyond the transcription activity itself. The project focuses on the transcription of place names from a large scale UK Ordnance Survey map circa 1900 with complete coverage of Great Britain. The project has drawn 1200 volunteers to contribute, with over 5.5 million transcriptions made (Southall et al. 2017). A consensus model enhances accuracy of the transcribed data, and the use of a score-board type ranking system encourages friendly competition by rewarding more voluminous transcribers with public acknowledgments. The project leaders also took a self-reflective approach to their crowdsourcing, providing other researchers in the field with valuable insights into the strengths and weaknesses of crowdsourcing historical spatial data through interviewing the top contributing transcribers about their motivations and work-habits. GB1900 includes a basic interface that allows users to query the crowdsourced gazetteer or interactively click on a point on the map to see the confirmed place name created through consensus. While innovative, the GB1900 project has disappointedly concluded after only 16 months, with the complete transcription of its source map. The map-based search tool is still accessible on the National Library of Scotland’s webpage (https://geo.nls.uk/maps/gb1900).

These examples cited above are all making significant contributions to the development of the field of publicly-engaged digital or spatial humanities as well as further case studies in the field of crowdsourcing. Collectively, however, they are exclusively unidirectional in their focus and purpose. They employ, as Haklay (2013) outlines, a crowdsourcing or distributed intelligence project design, where citizens serve as basic interpreters of historical sources and as volunteered computers, with the primary benefit being derived by the researcher. Further, the project purpose, questions, and analysis are driven by academic or professional historians rather than using the best practices of shared authority and collaboration that are hallmarks of a public history focus approach to spatial humanities projects (Frisch 1990, 2011; Ridge 2014).

Towards public participatory HGIS

For the benefit of historically-minded scholars interested in incorporating public collaborations into their HGIS project, a set of best practices has already been established by our colleagues in the field of public participatory GIS (PPGIS). PPGIS arose in the mid-1990s in response to a growing critique in the GIS community of the social nature and impact of GIS on society at large. PPGIS also sought to empower less privileged groups in society especially as they related to urban land use decisions and planning (Brown and Kytta 2014; Mukherjee 2015). Over the past 15 years of development, several key topics have emerged as defining characteristics of the field in practice. Dunn (2007), Newman et al. (2010), Poplin (2012), and Sieber et al. (2016) highlight the need for user-centered design of data representation to accommodate...
the unique needs of each user population. Kingston et al. (2000) as well as Haklay and Tobon (2003) collected detailed user feedback on the navigation, data presentation, and overall accessibility of a participatory GIS. Eisner et al. (2012) and Gottwald, Laatikainen, and Kytt (2016) highlighted the need to specifically think about design and usability for one’s target “public,” especially if they have unique obstacles to overcome to fully engage with a GIS, such as the struggles of elderly users of technology.

A core tenant of PPGIS is the ability to combine and contextualize official quantitative data with mapped community perceptions in a consensus-driven dialog with stakeholder groups (Jankowski 2009, Peters-Guarin, McCall, and van Westen 2012; Radil and Jiao 2016). Thompson (2015) argues community-integrated GIS gains value if it can “inform processes and relationships, rather than simply extracting patterns from large volumes of data.” This results in equitable citizen-researcher-government partnerships and increased trust and empowerment by interest groups (Robinson, Block, and Rees 2017). Using PPGIS as a method for fostering close community-researcher partnerships have been demonstrated as successful in projects that focus on developing inner-city neighborhoods (Elwood 2008; Rinner and Bird 2009), tourism development (Stewart, Jacobson, and Draper 2008), conservation planning (Brown 2012; Brown and Weber 2013), invasive species mapping (Hawthorne et al. 2015), flood risk management (Peters-Guarin, McCall, and van Westen 2012) and historic preservation and management (Arnold, Lafreniere, and Scarlett forthcoming; Brisbane 2005).

Our project combines and elevates these best practices of PPGIS by bringing community involvement and feedback into all stages of creating an HGIS. In the next section, we outline our citizen-researcher collaboration in interface design and data collection that put researchers and the community on equal footing in the development of a space-time linked historical GIS, known as the copper country historical spatial data infrastructure (CC-HSDI). We outline how this PPHGIS approach overcomes the limitations of existing projects by having the community design the interface that includes users working directly with historical spatial data, collaborative consensus of classifications and verification of features, contributing their own historical spatial knowledge through spatial storytelling, contextualization of public collected data with “official” historical spatial datasets, and facilitating immediate data access without the need for specialized software or training.

The CC-HSDI and Keweenaw time traveler projects

Michigan’s Copper Country, located on the Keweenaw Peninsula is a singular and nationally recognized example of the lasting sociocultural and environmental effects of the industrialization and deindustrialization experience in North America. Endorsed by the creation of the Keweenaw National Historical Park in 1992, the Copper Country is the oldest and one of the largest copper mining regions in the United States (Figure 1). Much of this industrial landscape is still visible on the cultural landscape and serves to remind residents and visitors alike of the importance of the legacy effects of industrialization and deindustrialization on their lives. This region serves as the community for our project which has two overlapping components (see Figure 2). The CC-HSDI project has created an online deep map that enhances our interpretations of past lives and environments in this region by connecting researchers and the public to recreations of the past in time and space. Our project extends these meaningful linkages to the present day through a public engagement initiative known as the Keweenaw Time Traveler (KeTT) that recruits “Citizen Historians” to help design the interface, build the deep map, and contribute their own spatial stories of the region.

The CC-HSDI is a research infrastructure that allows for the rapid geolocation and relational linking of historical records to facilitate spatio-temporal research. Hundreds of georeferenced and digitized historical maps feature individual buildings and landscape features that are connected in space and through time across multiple big data sets in a spatio-temporal database (Trepal, Lafreniere, and Gilliland forthcoming). These maps and data include immense detail on the region’s built and industrial environments including the location, size, construction material, ownership, civic address, and more, of some 125,000 residential, commercial, civic, and industrial structures for each decade from 1880 to 1950. City directory information are now, and census records, school records, and mining company records will soon be, linked to the built and industrial environments, and across time, to create a comprehensive digital and spatial representation of the social environment of one of the country’s first and largest extractive landscapes. Technical details of how the CC-HSDI was constructed, such as the digitization and 3D modeling process (Arnold and Lafreniere 2018), database structures (Lafreniere and Gilliland 2015), and overall data integration and webGIS interface (Trepal, Lafreniere,
and Gilliland forthcoming) can be found in the aforementioned papers.

The CC-HSDI’s potential as a research tool is enriched by its capacity as an engine for community-engaged public history. Dubbed the Keweenaw Time Traveler (www.mapyourhistory.org) for public audiences, from its inception, our project has been a researcher-community collaborative initiative intended to facilitate the co-production of heritage and data through its PPHGIS applications and public programing (Scarlett, et al. 2018). The project itself was developed in consultation with key history and heritage

Figure 1. Case study location, Michigan’s copper country on the Keweenaw Peninsula. Source: Author.

Figure 2. The complementary projects that when combined create the public participatory historical geographic information system (HGIS) approach.
stakeholders during the National Endowment for the Humanities grant proposal writing process. Heritage professionals from the region’s two National Parks, two major archival repositories, several local museums, and three historical societies were consulted about the scope of the project and the types of datasets that should be captured. Since its launch, the project has engaged a regional community with active in-person programming that promotes the sharing of historical knowledge and a mutual connectedness to place. Our team of undergraduate and graduate student “Time Travelers” attended outdoor festivals and outreach events offering guided use of the apps on touch-screen kiosks and tablets (Figure 3). These events enfranchise marginalized groups including the elderly, children, and those with lower computer literacy or limited access to a broadband internet connection. An active social media agenda and blog cover these public events to connect residents with the Copper Country diaspora, descendants who live elsewhere but maintain a connection with the history and people of this region via family ties and social media networks.

**Community PPHGIS interface design charrettes**

A core tenant of a public participatory approach to historical GIS is the inclusion of community-stakeholders in the design of the HGIS interface. We used a user-centered design approach to design and develop three PPHGIS apps. User-centered design methods seek to develop usable and effective websites by focusing on the users: understanding their needs and goals, and iteratively developing and evaluating the designs by observing the users (Services US Department of Health & Human 2018). Our goal to engage a large portion of the community in contributing and exploring the historical spatial databases implied that a diverse user base ranging from young and tech savvy users to elderly users. We required a design process that could reach our diverse community of users and effectively inform us of their use. Inspired by urban planning charrettes (Lennertz and Lutzenhiser 2006), we developed a design process using focus groups. Employing focus groups during the design and evaluation process is a cheap, efficient and effective technique to reach a large diverse user base (Vredenburg, et al. 2002).

In Winter 2017, we held 10 public design charrettes with a diverse demographic profile including high school students, heritage professionals, senior citizens, tourists, and families. These groups were targeted for the charrettes because they were identified by our heritage partners as the most likely to use the Time Traveler apps and website based on their visitation records, inquiries, and educational missions. The charrette would begin with an overview presentation about the project goals and developments to date and then would breakout into small focus groups. In the focus groups, community members along with project faculty and students would discuss design elements for the application, requirements for search ability, map navigation preferences, speed and efficiency, and how to best represent historical environments. After the initial round of charrettes, the project team worked to implement the public requests into an alpha-version of the application.

Project collaborators and their students with expertise in software design assurance and human-
centered design held a second round of charrettes where community members interacted with the live alpha-version of the community-conceptualized PPHGIS applications on iPads in small focus groups (Figure 4). Additional usability and design feedback was recorded and implemented to create a beta-version of the applications. A third smaller round of charrettes, with a similar diverse group of users as the first round, was conducted to complete final usability testing before the applications were launched live on the web in mid-June of 2017. The charrettes not only provide a mechanism for learning of our user’s needs but they also allowed for a meaningful collaboration between academic researchers and the heritage-minded public, promoted the project beyond the academic circles, and introduced the capabilities of HGIS more broadly to the public. Because they contributed to the design of the online interface, our users became invested in the project which is evidenced by the number of interactions and classifications completed to date (see results section below). A forthcoming publication will outline the charrette procedure, the implementation of the user suggestions, and outcomes in detail.

**The three PPHGIS apps of the Keweenaw time traveler**

The outcome of the interface design charrettes was the creation of three PPHGIS applications (apps) that integrate crowdsourcing-type tasks within an immersive high-resolution set of georeferenced Sanborn fire insurance plans (for more on Sanborn Fire Insurance plans see Wright 1983 and Library of Congress 2011). Each building footprint (~130,000 to date) has been vectorized by our research team and placed as outlines on the plans giving the illusion that the historical map is clickable. All three apps ask users to contribute to the creation of data that is recreating a longitudinal model of the built and industrial environments of the region. Users can choose which time slices they wish to work in, with dates ranging from 1888 to 1949 at approximately 10 year intervals. Each of the three apps outlined below utilize these base layers of temporally specific georeferenced historical maps and building footprints, promoting a historical-geographical contextualized environment that has contributors working directly with historical spatial data, overcoming the limitations of several of the projects noted earlier.

The first app is the *Document Building Material App* which asks users to classify what building material each structure drawn on the maps is made of. Sanborn used a standardized color scheme to denote material, such as yellow for wood, pink for brick, and blue for stone (Library of Congress 2011). Users are first presented with written instructions and a short how-to video on a splash page at initial app load, and then are tasked to choose the corresponding button for the color noted on the map (Figure 5). If a building has multiple materials, such as a brick cladding over wood, users click a multi-button and select the appropriate combination of colors.

The *Document Building Use App* asks community members to classify what type of use a building had, as noted on the fire insurance plans (Figure 6). Categories include dwellings, which include small outbuildings like sheds as well as attached porches, structures for automobiles, stores and other commercial buildings, industrial buildings, and public buildings such as schools, libraries and government offices. These classification categories were established during the charrettes and are an attempt to
balance the researchers’ demands for as specific of classifications as possible and our community groups’ interest in having a small number of options to choose from when completing the classification activity. We learned during the charrettes that having too many categories created choice anxiety in our participants as well as made the app too complex for some participants who had low computer literacy.

The Transcribe the Map App encourages community members to transcribe and interpret the written notations made on the historical maps. Prompts include “What is the building used for?,” which provides greater detail on the building’s use than does...
the Document Building Use App classification categories. For example, in Figure 7, a user is transcribing information about a brewery. In the Document Building Use App this building would be classified as industrial, while in Transcribe the Map App, users can note which part of the complex has the beer cellars. Another prompt asks “Who owns/occupies the building” and a third asks “What other information about the building is present?” We learned in the charrettes that users wanted a limited number of items to transcribe to reduce fatigue and that open ended questions resulted in less reliable data.

Most crowdsourcing projects assign users the specific subjects/objects or class of subjects to work on. The KeTT PPHGIS apps follow this model as it ensures that all objects have the potential of being classified. For our younger users, it creates a gaming effect that was identified in the charrettes as being engaging. We dubbed this Auto Mode and it is the default setting when the Document Building Material and Document Building Use apps load. Users are presented with a random building that needs to be documented or classified. Once they select the appropriate use or color, the app automatically chooses another building that has yet to be complete. For our older users, this rapid map movement was disorienting and the inability to work on the building of their choosing was limiting and disengaging. After charrette #2 we created Manual Mode, which allows users to choose any building they choose to work on. They can pan and zoom the map to any scale, move from town to town or through time, working on buildings of interest to them. This enhances the importance of spatial positioning, spatial relationships, and the placemaking capabilities of the apps. The Transcribe the Map App only operates in Manual Mode as not all buildings have details to be transcribed.

All apps operate with a collaborative consensus model to ensure data quality. For the two documentation and classification apps we require three users to classify a building’s material or use the same way before the building is deemed classified. This follows the same three like-classification model employed in Zooniverse and the Building Inspector project, and is standard practice in scientific experimental design. However, unlike Building Inspector we present users with a message when they make an entry that has reached consensus by changing the selection indicator to “consensus reached.” This notifies users that the data has been confirmed and is now available for use. Once a building reaches consensus it is no longer presented to users in Auto Mode. Users in Manual Mode can click on individual buildings, and if the building has reached consensus, a message appears notifying the user that consensus has been reached and to select another building to classify. A building continues to be available as a candidate for classification until three users have agreed upon the same classification, regardless of how many overall classifications are made. In the Document Building Material app, only
7.5% (4479 of the 59,713 consensus reached buildings to date) have received more than the minimum three required like classifications.

The Transcribe the Map app has additional challenges to reaching consensus as users may have slight spelling, punctuation, or detail differences (i.e., Co. vs. Company in Figure 7). To address these challenges, we use a Levenshtein edit distance algorithm that measures the distance between words as the number of single-character edits needed to change one word into another (Levenshtein 1966). This allows for spelling differences, difference in punctuation, and slight variations in detail to still permit two entries to be deemed similar. Due to these challenges we only require two users to agree for consensus to be reached. Further, previous transcriptions are available to users so they may either correct or improve on other users entries, allowing for more accurate transcriptions and higher rates of consensus. This allows transcriptions to be immediately accessible, linked to spatial data, and available for querying in the Explore App (described below) without the need for post-processing by the research team. This overcomes one of the key limitations of Zooniverse based transcription projects, the ability to automate a quality control mechanism for user-contributed transcriptions.

Keweenaw time traveler explore app

A fourth app, dubbed the Explore App, facilitates immediate access to all of the data created by the community in the three PPHGIS apps outlined above. This is a direct attempt to address what has been noted as limitations in the various crowdsourcing projects noted earlier. Further, researcher-created datasets such as census, city directories, business directories, and school records are automatically linked to community created data. This contextualizes both data types, adding agency and authority to community-created data, while supplementing and enriching researcher-created data. The app is live on the project website but it is still in a development phase (Figure 8). Currently, the app allows users to query and display data in two ways. First is the interactive map display and selection. Users first select the year of interest from a pull down menu. They may pan and zoom to any geographic extent or use spatial bookmarks to zoom to the various towns and villages in the region. Maps can be layered in any order and a combination of transparency slides and a “spyglass” allow users to “peek through time” on the historical maps by comparing two different map years together or a present day satellite image with a historical map.

Any building drawn on the map is clickable. Data associated with that location is shown on the left side of the app. Data created by the community in the three PPHGIS apps about the built environment is immediately available in the Buildings tab. All additional datasets available in the CC-HSDI are also shown in their respective tabs. Currently, the “People” tab returns geocoded city directory data for the specific map year for anyone who lived at the building.
selected. Further development is underway to allow users to follow individuals as they move around the region through time and to more rapidly see who else lived at the same address throughout time. Decennial Census data, school registers, and employee records from the region’s two largest mining companies are soon to be included in the “People” tab. A “Place” tab returns entries from a gazetteer of historical place names created by transcribing places noted in city directories and contemporary place names are searchable thanks to United States Board on Geographic Names database (US Department of the Interior US Geological Survey 2018).

Users can also query the data through four search modes. The “Address” search allows users to search any contemporary or historical address recorded in the CC-HSDI. Historical geocoders are linked to the buildings on the fire insurance plans so the map automatically pans and zooms to the address selected. Any associated data (people, places, stories) available about that address is pre-loaded and a counter above each tab gives users a glance at what additional data is available about the specific address and the specific historical time queried. Users can perform similar queries on the population data, place name gazetteer, and on user-contributed spatial stories.

While the core purpose of the Explore App is to provide immediate access to the data created within the three PPHGIS apps by community members as well as the researcher created data such as city directories and censuses, we have added a beta PPHGIS capability to the Explore App as well. The Explore App includes a qualitative GIS submission tool which allows members of the public to contribute their own historical spatial knowledge through spatial storytelling. Users are prompted to choose a historical map that is closest to the time period that their story took place. After clicking the screen to add a mark on the map, users are walked through a series of question prompts about their story including adding a title, the story narrative, date or date range, and to optionally add authorship to their story. Users can upload any type of image file, as well as .mp3 audio and .mp4 video files. A Facebook plug-in allows users to also post their submissions on their Facebook wall and other users can comment on stories both within the Explore App or via Facebook. Since late 2017 over 625 stories and hundreds of comments have already been submitted to the Explore App.

The focus of stories in the Keweenaw Time Traveler Explore App vary widely from submissions of period photographs of buildings, to newspaper clippings about past events or significant people associated with a location in the region. Others are personal, recalling memories of time spent at hockey games, swimming in waterfalls, and sharing family stories about grandparents’ time spent working in the mines. Online discussions have focused on major events and traditions, such as the annual Winter Carnival, Thanksgiving, and the 4th of July, while others have included sharing your favorite fishing spot on Portage Lake or the passionate debate over what is the appropriate condiment for a local culinary delight, the Cornish pasty.

An important characteristic of our approach to heritage-based storytelling is to ensure that users feel that their memories and stories are valued. Heritage scholars recognize that personal recollections and understandings of the past, even if they are slightly misaligned with scholarly historical knowledge, promote placemaking and community-building and are valuable contributions to our contemporary understandings of history (King, Stark, and Cooke 2016; Cooke 2010). Current research is also demonstrating that heritage projects need to recognize the role that community-engagement instruments can play in perpetuating power structures and silencing hidden narratives (Waterton and Smith 2010). From this perspective, we do not vet or curate any of the user-contributed spatial stories submitted within the Explore App, even if they may be perceived to be “wrong” or exaggerated. Like other online platforms, we expect that if a story is contested, users will comment on the stories using the utilities provided to share their interpretation of events or subjects of the spatial story. A profanity filter is utilized in our Javascript code and a ‘flag this story’ button is available to users who believe stories or images are inappropriate or too egregious. These stories are automatically removed and are manually reviewed by project staff. To date only 21 posts, or less than 3%, of all stories submitted have been flagged by users. Our public programing initiatives aim to encourage traditionally disenfranchized populations to share and engage through this collaborative spatial story-telling platform.

The Keweenaw Time Traveler Explore App is aiming to create a true deep map of an industrial community by providing an immersive, data-rich experience for users. The publically-generated data in the PPHGIS apps are contextualized with the “official” historical spatial datasets such as the fire insurance plans, city directories, gazetteers, and soon the census, school, and employee records. This integration of the
crowdsourced data with researcher-provided data provides real authority to the public in the creation of research and public use data.

**Results**

The three PPHGIS apps were launched in mid-June 2017. As of late fall 2018, the community has completed over 250,000 classifications resulting in just over 78,000 building attributes reaching consensus across the three apps (Table 1). The bulk of the classifications, 188,742, have been completed in the Document Building Materials App resulting in 63,552, or 94% of all buildings classifiable in the HSDI reaching consensus. This app was noted in the design charrettes and the summer festivals as being the most engaging, especially for youth, as the classification of colors was a task that was found to be both fun and easy to do. The classification error rate, or percentage of buildings that reached consensus with more than the minimum number of required unique classifications, was the lowest for the building material classification task at 7.57%.

The Document Building Use App has seen over 65,000 classifications to date with over 18,000 resulting in consensus with a much higher error rate of 16.97%. This higher error rate is expected as we limited the number of use categories in an attempt to reduce choice fatigue in users. Community members disagree on how to classify certain uses that fall between or across our categories, such as blacksmith shops that appear in commercial areas are classified by some users as "stores" and others as "industrial." Similar disagreements occur with businesses that appear to have a commercial store front along with manufacturing activities, such as breweries and wholesale grocers. Further adding to the error rate is the inability for users to select when a building has more than one function such as a commercial building with dwellings in the upper floors of the building. We are working to add the ability to note multiple uses to the app this year.

The Transcribe the Map App, as expected, has seen the lowest number of unique classifications at almost 5000, representing a little over 1100 buildings reaching consensus. It was noted during discussions with community members since the launch last year that the transcription app was the least engaging and the most fatiguing to users for two reasons. First, the task of reading and typing is more challenging, especially for those with slower typing speeds, or for children who do not spell well. Second, not every building on the maps has text that needs to be transcribed, thus removing our ability to create an auto-mode that would randomly issue a prospective building to a user. Thus, this requires users to pan and zoom around the map to locate a building with text that needed to be transcribed. These hurdles limited meaningful interaction with the app and reduced overall completion rates. Despite this limitation, the error rate itself at 16.55% is similar to the Document Building Use App. We contribute this relatively low error rate both to the optimized Levenshtein edit distance algorithm reducing the number of false negatives because of spelling or punctuation, and to the reality that our most dedicated, careful contributors are the ones completing the bulk of the transcriptions.

It must be noted for all three PPHGIS apps that we believe the order they are presented on our website, which is the same order presented in this paper, may also affect, in a small way, the number of classifications completed in each task. We observed at summer festivals and outreach events that users who only spent a couple of minutes in the PPHGIS apps largely did not move beyond the first app. We have recently reordered the apps to study this phenomenon further.

Beyond the distribution of classification by application, we also note distinct patterns in the way community members complete PPHGIS classifications and transcriptions. Observations from the project’s live completion statistics page at (www.keweenawhistory.com/stats.html) indicates that despite being randomly assigned a town and year in which to complete tasks, most users gravitate to working on the earliest year.

**Table 1. Classifications and Consensus Reached in Keweenaw Time Traveler PPHGIS Apps as of November 25, 2018.**

<table>
<thead>
<tr>
<th></th>
<th># Unique classifications</th>
<th>Number of Total buildings possible to be classified</th>
<th>Number of buildings reaching consensus</th>
<th>Number of total possible</th>
<th>Error rate a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Building Material</td>
<td>188,742</td>
<td>63,552(^b)</td>
<td>59,713</td>
<td>94.0%</td>
<td>7.57%</td>
</tr>
<tr>
<td>Building Use</td>
<td>65,539</td>
<td>116,641</td>
<td>18,165</td>
<td>15.6(^c) /o</td>
<td>16.97%</td>
</tr>
<tr>
<td>Transcribe the Map</td>
<td>4961</td>
<td>N/A(^c)</td>
<td>1112</td>
<td></td>
<td>16.55%</td>
</tr>
</tbody>
</table>

\(^a\)Percentage of buildings that reach consensus with more than minimum number of required unique classifications.

\(^b\)Only 63,552 of the total 116,641 in the CC-HSDI are derived from color maps and thus are available to determine the building material.

\(^c\)Not all buildings have map data to be transcribed.
that still needs data created. For the Document Building Use App, the earliest year for most towns in our project are nearly complete. Further, our most dedicated users like to use the Manual Mode to contribute to an entire year/town combination before moving on to make additional contributions. This has resulted in the more heritage focused communities of Calumet and Laurium (home to the Keweenaw National Historical Park) receiving the bulk of the classifications in the project.

However, the most significant driver of classifications and transcriptions appears to be active public history community outreach by the research team and our project heritage partners. Figure 9 illustrates the distribution of unique classifications for all three PPHGIS apps combined. There are three distinct periods of increased activity. In mid-August 2017 we had our most intense period of participation in festivals and outreach events. We held three events at our local national park visitors center, gave walking tours for campus alumni, setup our kiosks at a popular art fair for tourists, and held our end of summer celebration for our cognate research project that provides GIS internships for local high school students. The weekend of the art fair alone saw over 16,000 classifications by community members and tourists alike. The PPHGIS applications were prominent at each of these events.

The second distinct period of activity is from late September until late November 2017. This period also included public history outreach by participating in International Archaeology Day at the national park and a public lecture by the project directors at a local museum. However, the bulk of the activity directly correlates to an increased social media campaign on Facebook where we created hashtags for each app and linked it to a day of the week. #MaterialsMonday gave updates on classifications in the Document the Materials App and encouraged users to complete a given town/year combination. #TranscribeThursday did the same for the Transcribe the Map App, while #TuesdayUseDay supported the Document Building Use App.

Also critical to continued engagement in data collection activities is the ability for the public to access the data collected in a user-friendly way without the need for specialized software or technical skills. Further, we argue that contextualizing public data with other historical spatial datasets further increases engagement. The end of November 2017 saw the launch of the first version of the Keweenaw Time Traveler Explore App. The Explore App initially reduced activity in the PPHGIS apps as community members were eager to turn their attention to the big datasets and query tools that provided access to the data collected through interactive maps. In the first 5 months after the launch of the beta version of the Explore App, over 41,000 queries have been by over 8,000 unique users. Over 625 spatial stories have been submitted along with photographs, newspaper clippings, and other archival material. Further work is underway to improve the Explore App, which will serve as the basis for a forthcoming scholarly paper.

Figure 9. Distribution of unique classifications in public participatory historical geographic information system (PPHGIS) apps, June 15, 2017 to May 15, 2018.
More recently, in spring 2018 a renewed social media effort focused on providing promotional items (mugs and magnets) to significant contributors has again increased interest in the three PPHGIS applications as well as several community events and further local and regional media stories about the project. A clear finding from our work is that community engaged public history, both through in-person programming and online via social media, together with active partnership with community history and heritage groups, is critical to ensure a prolonged engagement of the public in a PPHGIS project.

Conclusion

This paper has called for a new model for historical GIS scholarship that includes a robust collaboration with the public. Borrowing best practices from public participatory GIS and urban planning, we outlined herein a model of public participatory historical GIS that can increase efficiencies of, public relevance in, and extend the reach of, HGIS projects beyond the academy to support heritage community building and preservation, scholarship in environmental or urban history, and promote awareness of the past through the lens of geospatial technologies. PPHGIS is an effective means to collect, share, and create exposure for GIS-ready historical geospatial data. This overcomes the often cited complaint about HGIS projects, that they are slow, tedious, and expensive to capture, create and standardize.

We have outlined some of the challenges and opportunities that still exist when trying to use crowdsourcing methodologies for collecting historical data. We outline the challenges that both georeferencing projects and transcription only projects face in establishing quality control over the datasets created. The ability for users to see their contributions, as well as the contributions of others is shown to increase overall interest in the project. Public contributors want to know that their contributions are meaningful and accurate. Herein we promote the use of a collaborative consensus model to address the verification of classifications and to increase accuracy of manual transcriptions of map data. The contextualization of public collected data with the official historical spatial datasets in the HSDI such as city directories and fire insurance plans improves historical geographic literacy and encourages participation. Further, creating a tool that allows users to contribute their own historical spatial knowledge through spatial storytelling gives the public further agency.

Perhaps most importantly, we demonstrate the importance of collaboration both within the academy and with the public. Our interdisciplinary research team includes over 15 individuals comprised of faculty and students from an array of disciplines including history, geography, GIS, computer science, software engineering, geological sciences, psychology, and the humanities. Scholars with expertise in GIS and computer science work on database schemas, geocode algorithms, and write code for the various apps and websites. Scholars in history and librarians understand the source material, its uses and limitations, and advise on best practices for standardization and coding. Those trained in public history, ethnography, citizen science, and psychology work with the public at community events and online. Project heritage partners such as local archives, libraries, museums, and national parks overcome the divide between local communities and their local universities by connecting researchers to their patrons, promoting project activities, and providing expertise in educational outreach and programming.

Not to be stressed enough is how critical a true citizen-researcher collaboration during the interface design and data collection both improves the quality of the PPHGIS applications but also puts researchers and the community on equal footing, developing loyalty and trust in the research process. The Keweenaw Time Traveler and Copper Country HSDI projects are ongoing. Currently we are preparing additional datasets for including in the Explore App including the full count decennial censuses, detailed employment records from the region’s major mining companies, and a comprehensive school register for all students from 1900 to 1950. With future funding we aim to create space-time record links between all of these datasets to permit examination of micro and macro historical demographic and geographic inquires such as social, occupational, residential mobility, deindustrialization, immigration, urban morphology, and neighborhood or community kinship networks. Our public colleagues are encouraging us to completely rebuild the Explore App to better meet the needs of the community of users. We will include another round of public design charrettes and will evaluate their efficacy using a heuristic evaluation model, which we will be eager to share with other HGIS scholars in a future paper. The integration of public history best practices with innovative HSDI methodologies will continue to provide opportunities for spatial humanities and HGIS scholars, students, and their communities to learn more about the historical
geographical dynamics, populations, and environments that shaped our past, impact the present, and inform the future.

References


Frisch, M. 2011. From a shared authority to the digital kitchen, and back. In In letting go? Sharing historical


Trepal, D., D. Lafreniere, and J. Gilliland. forthcoming. Historical spatial data infrastructures for archaeology: Towards a spatio-temporal big data approach to studying the post-industrial city. *Historical Archaeology* 54(1)


Vredenburg, K., J.-Y. Mao, P. Smith, and T. Carey. 2002. A series *Historical Archaeology* 54(1)