CREATING HISTORICAL KNOWLEDGE SOCIALLY:
NEW APPROACHES, OPPORTUNITIES, AND
EPISTEMOLOGICAL IMPLICATIONS OF UNDERTAKING
RESEARCH WITH CITIZEN SCHOLARS

Conference at the German Historical Institute Washington (GHI), October 26-28, 2017. Organized by the GHI, the Roy Rosenzweig Center for History and New Media (RRCHNM), and Maryland Institute for Technology in the Humanities (MITH). Made possible by support from the Deutsche Forschungsgemeinschaft (DFG). Conveners: Matthew Hiebert (GHI), Simone Lässig (GHI), and Trevor Muñoz (MITH). Participants: Samantha Blickhan (Adler Planetarium), Denise Burgher (University of Delaware), Jim Casey (Princeton University), Constance Crompton (University of Ottawa), Laura Coyle (National Museum of African American History and Culture), Fabian Cremer (Max Weber Foundation), Meghan Ferriter (Smithsonian Institute), Aleisa Fishman (United States Holocaust Memorial Museum), Carl-Henry Geschwind (Independent Scholar/Historian), Silke Glitsch (Göttingen State and University Library), Katharina Hering (Georgetown University), Franklin Hildy (University of Maryland), Elizabeth Hopwood, (Loyola University Chicago), Rebecca Kahn (Humboldt Institute for Internet and Society), Joseph Koivisto (University of Maryland), Ursula Lehmkühl (University of Trier), Patrick Manning (University of Pittsburgh), Atiba Pertilla (GHI), Mia Ridge (British Library), Andrew Taylor (University of Ottawa), Jennifer Serventi (National Endowment for the Humanities), Raymond Siemens (University of Victoria), Heather Wolfe (Folger Shakespeare Library), Vladimir I. Zadorozhny (University of Pittsburgh).

This conference brought together historians, digital humanists, archivists, and librarians to discuss the growing trend of knowledge co-creation involving citizen scholars, and the implications for our conceptions of history, methodology, and how the discipline produces knowledge. The event sought to bring into critical international dialog research involving methods, systems, and standards designed and implemented to ensure quality of results and inclusivity of perspective; best approaches and consideration in bringing scholars and the public into collaboration; research into how historical knowledge is formed within groups; and the effects of citizen scholarship on societies, communities, contributors, and historical research fields.

The first panel, “Developing Communities of Practice” chaired by Carl-Henry Geschwind, explored different models for bringing
together scholars and members of the public into research communities and their effects. Mia Ridge in her paper, “How Can Crowdsourcing and Citizen History help teach Historical Thinking?”, showed ways in which research involving the public can develop core historical competencies amongst participants. Fostering skills such as source familiarity, paleography, and name disambiguation extends to the development of analytic skills and individual research interests through web forums and other community tools allowing exchange with experts. Ridge perceives communities of practice within crowdsourcing projects as social, situated learning systems. In her paper “Pelagios Commons: Decentralizing Knowledge Creation in the Web of Historical Data,” Rebecca Kahn presented Pelagios Commons as infrastructure for linked open geodata in the humanities and a decentralized community connected through shared sources and tools. The project aggregates data of external resources into an environment with search, social annotation, exploratory, and community affordances, enabling new modes of collaboration and registers for critical thinking. Ray Siemens in his paper titled “Thoughts towards an Engaged and Open Social Scholarship” outlined the trend towards “methodological commons,” advancing problem-based research through the representation, analysis, and scholarly communication of data. Collaborative networks and shared infrastructures that organize data for such purposes are fostering mutual engagement and shared skills generative of communities of practice. Together, these trends have led to greater public participation in academic data creation and analysis, and underpin a concept of open social scholarship motivating the nascent partnership-based Canadian Social Knowledge Institute (C-SKI), with its active initiatives in research, tool development, and DH education.

Chaired by Jennifer Serventi, the second panel, “Knowledge Production through Partnership and Engagement” approached ways to undertake collaborative knowledge production through partnerships and engagement with organizations outside of academia. Denise Burgher, Jim Casey, and Elizabeth Hopwood in their paper “Collective Histories and Communities of Knowledge: Transcribe Minutes and the Colored Conventions,” discussed an imperative ethics of radical inclusion, care, and partnership in approaching the collaborative transcription of minutes and other archival materials for the over 200 national, regional, and local Colored Conventions of free and once captive Black people held between 1830 and 1900. The presenters conveyed that addressing their challenges of gathering and archiving
new knowledge depends upon teamwork, a culture of consensus and acknowledgement that deemphasizes traditional academic hierarchies, and learning and respecting how stakeholder communities themselves are organized. Franklin Hildy in his presentation, “The Theatre-Finder Project: Sustainability, Efficiency, and the Challenge of Crowd Sourcing with Citizen Scholars,” discussed the challenges of collaboratively building and sustaining a comprehensive online guide to all theaters, from the classical period up until the twentieth century. Hildy stressed the importance of peer review, technical sustainability, and data interoperability. To prevent content becoming “silod” and avoid duplication of labor, Hildy proposes that the resources of his site and other indexes in the field partner in a Wikipedia-based solution.

In his keynote address that evening, entitled “Frontiers of Digital History: Citizen Scholars, New Knowledge, and Public Debate,” Patrick Manning foregrounded his ongoing project involving citizen scholars, Collaborative for Historical Information and Analysis (CHIA), by recounting his work in digital history over the past two decades as “achievements in failure.” His “Migration in World History” (2000), provided curated historical documents, but was caught between technological obsolescence of the CD-ROM format and the infeasibility of putting an extensive resource on the web. The “World History Network” (2004) had difficulties assembling a comprehensive source base, while committed institutional support for digital projects was lacking. CHIA, which aims to collect global data at scale for research on environmental degradation, human inequality, and human-natural interaction, must cope with the reluctance of historians to contribute their data for public editing and analysis. Turning to citizen science, the project has enlisted undergraduate students and others to submit and edit historical documents. Beyond significant contributions to large-scale research projects in this way, Manning perceives citizen scholarship as helping produce an informed citizenry able to speak knowledgeably on crucial issues of policy, helping bridge gaps between specialists and the public.

The third panel, focusing on “Social Research Methods and Publics,” was chaired by Ursula Lehmkuhl, and looked to explore societal-level and societal-inclusive research initiatives. Samantha Blickhan, in her paper “Crowdsourcing Public History: Text Transcription, Historical Documents, and the Zooniverse,” discussed the history of Zooniverse, a platform with well over a million registered volunteers, and the
challenges in adapting for humanities research a system established to crowdsource natural and physical science data. Challenges include aggregating varying genres of text, and forthcoming affordances for transcribing audio archives. Blickhan highlighted the effectiveness of the site’s “talk” discussion feature in stressing that a great user interface should allow volunteers not only to participate, but to learn and acquire new skills. In his paper, “Crowdsourcing Data in the Mass Observation Project (1937–1965),” Matthew Hiebert presented archival research on the methodology of the early Mass Observation project (1937–1965). Founded as a study of British society in response to the perceived false news of mass media, the project innovated early participatory social research methods derived from ornithology, enlisting 1300 volunteers to contribute diaries, questionnaires, observational reports, and other materials in conducting “anthropology at home.” Hiebert discussed methodological challenges of the project, its adoption of computation, and impacts on British society and policy. Vladimir I. Zadorozhny in his presentation “Towards Computational Social Sustainability,” discussed the challenges of integrating different sorts of historical data—which may be distributed, heterogeneous, sparse, aggregated differently, inconsistent, or unreliable—and the necessity of doing so for analyzing and predicting societal change. The Collaborative Data Fusion (Col*Fusion) infrastructure project, being undertaken in cooperation with CHIA, provides a crowdsourcing-based model for large-scale data integration, curation, and reliability assessment.

The fourth panel, “Standards, Authority, and Participation,” chaired by Fabian Cremer, explored approaches to quality control and knowledge authority in citizen science projects. Laura Coyle in her paper “No Time Like the Present: Launching the National Museum of African American History and Culture’s Crowdsourced Research Projects,” discussed the crowdsourcing initiatives of the National Museum of African American History (NMAAH) using the online Smithsonian Transcription Center. Coyle attests that volunteers do high-quality work, also contributing alongside staff in error checking processes. Volunteer transcribers of the Freedman Bureau Records and other important indexed archival projects help fulfill the missions of the NMAAH and the Smithsonian, “literally helping us rewrite history.” Meghan Ferriter in “Ready — Set — Know: What You Learn When You Invite the Public to Pilot an Open Source Framework,” presented the crowdsourcing workflow based on Scribe and iterative development of the Library of Congress pilot project “Beyond Words,”
which invites the public to identify cartoons and photographs in historic newspaper collections. The effectiveness of its consensus model for assuring quality was discussed, and its tradeoffs in prioritizing accuracy over negotiation and potentially richer participation. Important considerations regarding the role of volunteers as “laborers or learners” were raised by Ferriter, questioning how knowledge, interpretation, and design decisions may also be granted to citizen scholars. Joseph Koivisto in his presentation, “Crowdsourcing as a Means of Authority Assessment and Enhancement for Cultural Heritage Description,” examined the effects of crowdsourcing on cultural heritage authority records. In the context of developing a linked open data iconographic thesaurus within the Project Andvari initiative, and using the open-source PYBOSSA application, limitations of inherited classification schemas were discovered in the inability of citizen scholars to attribute suitable terms from extant vocabularies when describing iconographic content, suggest additional terms. Koivisto observed that crowdsourcing can be an effective approach towards authority creation and refinement if clear instruction and expert review are provided.

The fifth panel, “Social Formations of Knowledge,” chaired by Atiba Pertilla (GHI), presented projects using crowdsourcing and other digital methods to ascertain how knowledge is formed within groups. The presentation given by Constance Crompton, “People’s Patterns: Social Prosopography and the Gay Liberation Movement,” discussed the Lesbian and Gay Liberation in Canada project which explores the potential of digital research environments to recover Canadian contributions to the gay liberation movement. Using Text Encoding Initiative (TEI) standards, the team has encoded Donald W. McLeod’s annotated chronology. “Lesbian and Gay Liberation in Canada” (1964-1975), creating graph network visualizations of people, prosopographical data, places, and other information in researching the structure of the movement. Crompton advocated looking to communities who have successfully used social knowledge creation within community-led activism as a research paradigm. In her paper, “History Unfolded: U.S. Newspapers and the Holocaust,” Aleisa Fishman discussed the activities of citizen historians in the History Unfolded project of the United States Holocaust Memorial Museum. More than 2000 participants from across the U.S. have submitted more than 14,000 articles from local newspapers of the 1930s and 1940s with news or opinion about 34 different Holocaust-era events in an effort to determine what Americans could have known about
the Nazi threat during the period. In the process, Fishman observed these volunteers have come to see history as a process of discovery, in learning about Holocaust history, using primary sources in historical research, and in challenging assumptions about American knowledge of and responses to the Holocaust. Silke Glitsch in her paper titled “Exploring the Visual History of Göttingen University: A Citizen Science Approach” first recounted recent citizen science developments in Germany, which include the Citizens Create Knowledge (GEWISS) program and platform (2014); the Citizen Science Strategy 2020 for Germany Green Paper (2016); the Directive on the Promotion of Citizen Science in Germany (2016); and the commitment of the Federal Ministry of Education and Research to “continue strengthening Citizen Science in the future” (2017). Her project, facilitated at the Göttingen State and University Library, explores the role of visual aspects in the self-representation and public perception of Göttingen University by providing a citizen science framework and 4600 photos of university events from the 1930s to the 1950s for shared analysis. Bringing together multiple institutions and players with complementary knowledge domains for the project, Glitsch underscored the effectiveness of libraries as facilitators and co-creators in citizen science research projects, possessing source material, expertise, technical infrastructure, and their function as an interface between institutions, academic departments, and the public.

The final panel, “Implications of Knowledge Co-creation,” chaired by Trevor Muñoz, sought out implications of participatory research for both volunteers and disciplinary fields. Katharina Hering in her presentation, “Punch Cards for Social Change: Participatory Research in History,” discussed two participatory research projects of the 1970s and early 1980s that fostered community organizing and social change. Activist Barry Greeter used punch card databanks in working with the Highlander Center on the Appalachian Land Ownership Study, where teams of local people were trained in how to search land records to investigate local property ownership and power structures. The Baltimore Neighborhood Heritage Project, co-coordinated by oral historian Linda Shopes, conducted more than 200 interviews with longtime residents of four Baltimore City neighborhoods, designing the project to promote community involvement. Hering stressed that such research projects aiming towards progressive social change by design benefitted rather than exploited participants. Heather Wolfe in her presentation “Shakespeare’s England: What the Folger Has Learned from Developing our Crowd,” discussed the large-scale Early
Modern Manuscript Online (EMMO) project at the Folger which provides, with the Text Creation Partnership, web access to XML/SGML encoded transcriptions, images, and metadata for English manuscripts from the sixteenth and seventeenth centuries. Their project Shakespeare’s World uses Dromio, EMMO’s transcription, mark-up, collation, and vetting tool, to allow members of the public to transcribe a variety of manuscripts from Shakespeare’s time written in secretary hand, with results then entered into the EMMO database. Measurable impact on early modern scholarship and expansion of interpretative range is already discernible. Through numerous public “Transcribathons” held on campuses around the U.S., a community of thousands of readers of early modern English handwriting has been created, able to teach, transcribe, and conduct research with manuscript sources.

The concluding discussion took up salient threads that had emerged in the productive panel discussions that followed each set of presentations. Discussion often coalesced around issues of power in knowledge production, with an emergent theme that de-colonizing the archive is not a matter extracting new stories, but about furthering communities’ self-determination through ensuring public access to and engagement with their own histories. Questions of how crowdsourcing may or may not empower participants in advancing research is related to how knowledge co-creation processes are talked about, and the conference throughout attempted to interrogate the nuances of our own critical terminology. Discussion also gravitated towards how to better facilitate the involvement of citizens scholars in the very design of co-created research projects. Issues of access were also of concern, for while Open Access policies are key to knowledge co-creation, research generated from citizen science may ironically end up behind academic journal paywalls. Citizen science research may require new forms of rigor in respect to peer review, with increased general interest in open review as a deserved corollary of open access.

The importance of project sustainability was a recurrent theme. Funding was discussed openly, with a number of projects presented supported by the National Endowment for the Humanities, the Social Science and Humanities Research Council of Canada, and the event itself by the Deutsche Forschungsgemeinschaft. Continued and expanded funding internationally was perceived crucial for new and sustained co-creation initiatives. Differences between the developments of citizen science in North America and Europe also became
more apparent through the event, with Germany now having clear and strong national-level research policy in support, yet relatively fewer projects in current development. North American projects were also more likely to involve students or incorporate a pedagogical dimension. Participants were in agreement that the conference had shown the benefit of assembling a multi-country, multidisciplinary group, to engage in a rich discussion that focused not only on outcomes, data, and technical systems, but on the processes of co-created knowledge production, including honest treatment of failure and new challenges.

Matthew Hiebert (GHI)