Title
The Digital Index of North American Archaeology: networking government data to navigate an uncertain future for the past

Journal
Antiquity, 92(362)

ISSN
0003-598X 1745-1744

Authors
Kansa, Eric C
Kansa, Sarah W
Wells, Josh J
et al.

Publication Date
2018-04-24

DOI
10.15184/aqy.2018.32

Data Availability
The data associated with this publication are available at:
https://doi.org/10.6078/M7N877Q0

Author's Pre-print
June 2017: Accepted for publication in *Antiquity.*
Please cite the version of record.

The Digital Index of North American Archaeology (DINAA): Networking Government Data to Navigate an Uncertain Future for the Past

Kansa, E. C.*, Kansa, S. W., Wells, J. J., Yerka, S. J., Myers, K. N., DeMuth, R. C., Bissett, T. G., & Anderson, D. G.

*Corresponding author contact: kansaeric@gmail.com

Introduction

Government policies and bureaucracies shape the practice of archaeology in much of the world, including the United States. Laws, regulations, and government agencies oversee heritage management influencing where archaeologists work, how intensely they record sites, and the manner in which archaeological sites and their documentation see long-term curation. Despite this importance, the impact and outcomes of government policies in archaeology remain largely opaque to research communities and other public stakeholders. Furthermore, the nation's investment in archaeology and historic preservation has produced a literature that, while vast, is often inaccessible. The Digital Index of North American Archaeology (DINAA) builds critically-needed information infrastructure to make information about archaeology across North America accessible and useful to scholars and the public alike (Wells et al. 2014). DINAA (ux.opencontext.org/archaeology-site-data/) provides the most comprehensive and detailed database documenting human settlement in North America currently available by aggregating archaeological and historical data from state and tribal governmental authorities that manage U.S. cultural resources. To date, the public can download nearly 500,000 site file records (with precise location and other sensitive data redacted) free of charge, and free of intellectual property restrictions. These records, available via Open Context (opencontext.org), an open access data publishing service (Figure 1), cross-reference an ever-growing number of reports, museum collections, bibliographic references, and other online datasets. This provides researchers, museums, libraries, government offices, tribal officials, and members of the public with a powerful gazetteer of all documented historical and archaeological sites in the United States, and this integrated content can facilitate cross-disciplinary research on a continental scale.

Making the Most of Public Investment in Archaeology

Altschul & Patterson (2010) conservatively estimate that US taxpayers invest over $500 million per year to comply with federally-required historical and archaeological protection measures. This level of public investment nearly matches the total combined budgets of the Institute of Museum and Library Services (ca. $240 million in 2015), National Endowment for the Humanities (ca. $140 million in 2015), and National Endowment for the Arts (ca. $150 million in 2015), and is thirty times the annual National Science Foundation budget (ca. $16.5 million in 2013) for archaeology and archaeometry (Rocks-Macqueen 2014). These surprising numbers demonstrate archaeology's relative importance in public cultural heritage investments. Unfortunately, much of this work goes unnoticed: decades of investment in managing and protecting America's archaeological heritage have led to few publicly-accessible impacts. Cultural resource management (CRM) largely takes place within relatively opaque bureaucratic processes that regulate construction and development. CRM work has resulted in an estimated 350,000 reports nationwide as of 2004 (NADB 2016), but because of poor access and cataloging, irreplaceable cultural heritage documentation in these "grey literature" reports is under-appreciated or ignored. Thus, CRM findings typically see little external reuse in research or publication venues recognized by professionals. Similarly, this literature has very little accessibility to other communities, including Native American descendant communities.

"Open Government" reform efforts attempt to promote greater public access to information resulting from public investments like CRM archaeology. These initiatives complement so-called "Open Science" and other "Open Data" efforts with broad goals of promoting greater "democratic accountability" to research and public policy through broader public access to information (Lake 2012; Kansa 2012). Appropriate investments in data management can help this information reach its potential by making it accessible and usable (Kansa & Bissell 2010; Kansa & Kansa 2011, 2013; Kansa et al. 2014; Kintigh et al. 2014a, 2014b; Wells et al. 2014; Raviele 2015; Anderson 2017) Indeed, as we discuss later in this paper, recent political upheavals in the US further highlight the importance of public policy in shaping access to scientific data.

DINAA has open science and open government goals and motivations. DINAA leverages this tremendous public investment in cultural heritage for wider public impacts by building a comprehensive, free and open access inventory of U.S. archaeological and historical resources. DINAA compiles, cleans, and publishes site file data aggregated from state and other agencies that enforce U.S. historical protection laws. Site files include information about periods of use, associated artifact collections and documents, preservation condition, National Register of Historic Places eligibility status, and a host of other variables useful for research and management purposes. In 1993, the last time primary site file data was compiled nationally through the National Park Service's National Archaeological Database (NADB) effort, just under one million archaeological sites had been recorded (NADB Maps 1993). This total has grown over the past two decades (see also Anderson & Horak 1995, Anderson & Sassaman 2012:32) and we estimate the number now to be nearly 2.5 million sites.

DINAA has integrated archaeological site data from nearly 500,000 sites from more than one dozen states, encompassing the rich chronological, cultural, and anthropological metadata used by government compliance officials and the research community (Figure 2). Over the next two years, with support from the Institute of Museum and Library Services and the National Science Foundation, DINAA will continue this work, adding an estimated one to two million archaeological sites to encompass the remainder of the United States.

As discussed above, contemporary Native American communities largely lack access to CRM outcomes documenting their past. Prior to requesting any data from state offices, we contacted every federally-recognized tribe with ancestral homeland in the portions of the Eastern US (DINAA's current area of coverage). In the planning and the execution of the new DINAA-related grant projects, we expanded outreach and consultation efforts to Native American nations in the remainder of the US. We staffed a booth at the recent meeting of the National Association of Tribal Historical Protection Officers (NATHPO, a professional organization for tribal government heritage officials) and met with the Native American Advisory Council of the Phoebe A. Hearst Museum in Berkeley. Finally, two of our team members (Myers, Yerka) have current or past heritage management employment with tribal governments. Based on such consultation and work on behalf of tribal governments, DINAA can play an important role in supporting the work of tribal government heritage officials (THPOs). Because of colonial displacements and forced migrations, ancestral homelands can span multiple state boundaries. By making certain (low-sensitivity, as discussed below) data available, DINAA can provide tribal officials with information needed to better manage ancestral territories.

DINAA records associate every site with its Smithsonian trinomial identifier, an administrative code assigned by the state, with mapping, chronology, and other metadata. To eliminate the risk of accidental or malicious disclosure of sensitive data, DINAA only stores and releases spatial coordinates at a reduced level of geographic precision. To prevent looting and vandalism, a US federal law, the Archaeological Resource Protection Act (ARPA), requires protection of site location data. In discussion with state historic preservation offices (SHPOs) and agency personnel, we arrived at a generally-acceptable consensus on levels of spatial precision for public disclosure. In most cases, DINAA allocates sites to a roughly 15-20 km grid cell (Figure 3). This consensus grid applies to all states currently in DINAA, with the exception of Ohio, which required DINAA to spatially reference sites to county centroids, not the general DINAA grid. Our negotiations about spatial data highlight how the understanding of risk among government officials plays a key role in shaping public records, and we hope DINAA will provide more experience in ways to balance information security needs with information usability needs. As we describe in the next section, DINAA's current approach to spatial data still permits important research programs and applications.

Research Applications of DINAA

In the face of rapid and far-reaching global changes, the concept of the "Anthropocene", though contested, has captured the imaginations of scientists, cultural heritage professionals, policy makers, and public communities. While variably defined, the Anthropocene refers to the period when human agency began to measurably alter climate and biota at regional, continental, and global scales (Lane 2015; Lightfoot & Cuthrell 2015; Zalasiewicz et al 2015). Whether or not the term has much explanatory value, public awareness about the Anthropocene helps motivate mobilization of data. Part of the justification for requesting mass data dumps from state officials comes from research questions around large-scale human impacts. For instance, documenting the threat of sea-level rise to tens of thousands of sites known through DINAA, demonstrates how unprecedented data access can reveal the vast scope of cultural preservation challenges resulting from accelerating "Anthropocene" climate change (Figure 4, Table 1).

From an information-management perspective, DINAA's greatest value centers on Linked Open Data applications; that is, the ability to participate in the growing body of related data shared openly via the Web (Figure 5). The next section provides examples of how DINAA data already is being combined and visualized with other Web-based data sources. In short, because DINAA publishes sites through Open Context, it benefits from Linked Open Data technologies that Open Context employs and continually develops. Open Context emphasizes the use of stable Web Uniform Resource Identifiers (URIs, i.e. stable URLs that serve as universally unique "primary key" identifiers) to identify concepts and other entities so they can be easily and precisely referenced and related across different data collections on the Web. DINAA uses Open Context and the EZID service to mint persistent URIs for each site files record. In archaeology and historical geography, the "site" is a key organizational entity. Minting stable Web URIs and offering rich temporal, geographic, and cultural metadata about sites will create significant Linked Open Data resources essential for broadly integrating museum, library, and scientific datasets.

Archaeological Linked Data and DINAA

The DINAA project is not alone in developing information systems to integrate large-scale cultural heritage data. Mega-Jordan (http://www.megajordan.org/), an early and ongoing effort, served as the basis for continued software development with Arches, an open source heritage data management project (http://www.getty.edu/conservation/our_projects/field_projects/arches/). Arches mainly serves the needs of government administrators, and uses the CIDOC-CRM (International Council of Museums, International Committee for Documentation, Conceptual Reference Model; http://www.cidoc-crm.org/) as a standard ontology to organize data. As such it is an excellent tool for creating standards-compliant cultural heritage data.

In contrast, DINAA publishes legacy data created in each state by government personnel (typically site file managers working under SHPO/state archaeologist/state professional archaeological council direction), and without reference to the CIDOC-CRM. Loading each dataset into DINAA involves what data managers describe as an "ETL" (extract, transform, load) process. ETL processes migrate data from one database to another, often involving transformations in formats and data organization. We have yet to encounter a state site-file database with any sort of API (application program interface) that we could use to automate requests for data. For DINAA, the ETL process involves obtaining tabular data "dumps" manually generated by the database managers of each state database. DINAA then redacts sensitive data and cleans inconsistencies (misspellings in controlled vocabularies, non-numeric text in numeric fields, etc.) and converts geographic coordinates (reduced to a low-level of spatial resolution as discussed above) to the WGS-84 standard as is common to Web mapping services. The DINAA team also creates additional metadata (especially date ranges for different periods defined by a dataset). Finally prior to publication online, the state data need modeling according to Open Context's general and highly abstracted database schema. While schema mappings can be reused, most of the ETL process needs to be repeated to add new records to DINAA as state sitefile databases expand over time.

Open Context uses a very general and abstracted database schema in order to preserve the wide variety of attributes, relationships, and vocabularies used in source datasets. Use of a more specific data organization standard like the CIDOC-CRM would require additional investments of time, effort and expertise in the ETL workflow. The DINAA project explored mapping data to the CIDOC-CRM, but we decided against it for practical and conceptual reasons (see discussion below). Our ETL workflow mints URIs for each entity, controlled vocabulary concept, and descriptive attribute. For DINAA, this means Open Context mints a stable URI for each site file record, and those URIs can be used for Linked Data applications that cross-reference archaeological site information reported in disparate sources, such as the Federal Register, JSTOR publications, the National Register of Historic Places, the Canadian Archaeological Radiocarbon Database (CARD), and other databases. The following examples illustrate how DINAA's approach to linking and integrating archaeological datasets can facilitate research.

Example 1: Mapping government heritage administration

The current DINAA dataset is cross-linked with the Federal Register, a U.S. government outlet that provides notifications of decisions and other news relating to the administration of laws and regulations (see http://ux.opencontext.org/2016/12/02/dinaa-and-the-federal-register/). Regulatory processes greatly impact and shape the practice of archaeology in the US, and the Federal Register offers a key information resource for understanding governance of the archaeological past. DINAA allows archaeologically-meaningful context to be added to Federal Register notifications via Linked Open Data methods. Archaeological sites in Federal Register notifications are typically listed with Smithsonian Trinomials. By themselves, these trinomials are just strings of letters and numbers with very little meaning. However, because DINAA curates Smithsonian Trinomial identifiers along with geospatial, chronological, and other metadata, documents listing Smithsonian Trinomials can be matched with DINAA, thus adding rich spatial, chronological, and other metadata to government documents. This added context helps make the Federal Register a more meaningful window onto how we regulate the archaeological past.

Example 2: Mapping publications

Since the 1960s, many researchers have published scholarly papers and books identifying historical and archaeological sites with Smithsonian Trinomial identifiers. The DINAA team recently text-mined the literature in JSTOR to find trinomials and associate them with DINAA records. Open Context describes links between DINAA sites and JSTOR articles that reference those sites with concepts from Dublin Core Terms, a widely-used digital library standard. This linking powers map-based search and browse interfaces to discover scholarly literature related to sites. DINAA can display heat-maps of this content to show where academic scholarship has focused, helping illustrate the history of research. This highlights how text-mining and entity identification can enhance discovery and analysis of archaeology's publication record (see Jeffrey et al. 2009; Kintigh 2015). This exercise will guide future entity identification efforts beyond JSTOR, to encompass the HathiTrust (digitized books), reports in tDAR (a digital repository for North American archaeology), and other document archives that contain poorly-catalogued grey-literature reports.

Example 3: Cross-referencing DINAA with external data sources

By matching Smithsonian Trinomials, we have established links between DINAA site file records and metadata records in other datasets and repositories. These include the Paleoindian Database of the Americas, or PIDBA (Anderson et al. 2010, 2015; Anderson and Miller 2017), the Eastern Woodlands Household Archaeological Data Project (White 2014), the Federal Register, and tDAR (McManamon & Kintigh 2010). Figure 6 shows all sites in DINAA that are referenced by records or web resources held in all four of these external data repositories. This helps people discover potentially useful research content in any online collection with related and linked material. Essentially, DINAA acts as a vast series of pegs on which to hang external content (Sheehan 2015). As more content from across the Web is hung on each peg over time, online resources will enhance their metadata and broaden the impact and reach of their collections.

Example 4: Contributions to the PeriodO temporal gazetteer

As noted above, the CARD team has agreed to cross-reference with DINAA. This will help better integrate North America's settlement history with our understanding of absolute chronologies. Nevertheless, relative chronologies remain important in archaeology, and because of their methodological significance, relative chronologies need appropriate data modeling. The PeriodO project (http://perio.do/), an initiative to develop a gazetteer of scholarly definitions of temporal periods, represents another aspect of DINAA's contributions to Linked Open Data. PeriodO provides Web URIs and a common schema to model temporally and geographically scoped period concepts (Rabinowitz 2014; Shaw et al. 2016). Ariadne (http://ariadne-infrastructure.eu/), a major European archaeological data integration effort uses PeriodO as a framework work for temporal interoperability. DINAA contributed over 700 period entities used to describe the chronologies of individual states to the initial PeriodO database. Thus, through PeriodO, DINAA's North American collections have some common modeling and cross-referencing with key European collections. Eventually, as PeriodO develops aggregation features, Open Context will use PeriodO to enhance temporal indexing of DINAA and other datasets.

Importantly, PeriodO does not demand agreement where agreement does not exist. Because PeriodO uses a common schema to model periods, it can serve as a basis for interoperability between different chronological schemes defined by different institutions and scholars. Representatives of different communities can use PeriodO to author and model new and alternative periodization schemes. In a colonial context such as North America, archaeological chronologies used by government agencies often do not reflect the needs or perspectives of Native American communities. Rich traditions from oral histories and other indigenous ways of knowing can more broadly contextualize a public dataset like DINAA to make it more meaningful and valued among members of contemporary descendant communities (Cochrane et al. 2008, Nicholas 2008). PeriodO can help open the door for indigenous community perspectives to reorganize the histories represented in DINAA site files. Participatory research in archaeology has allowed indigenous knowledge to be applied not only to tribal land and sites, but also to the existing archaeological record across the landscape of the U.S.. Known sites are increasingly visited and utilized by contemporary descendant groups, and as memoranda of agreement and understanding are developed for site maintenance and historic resource management, the intersection of traditional knowledge and scientific archaeological data deepens. A unique opportunity exists for referential terms familiar to traditional communities, including those describing origin stories and or historic events curated through oral history, to be linked to temporal or cultural terms already in use by state and federal agencies. Bringing these disparate conventions of naming can open new pathways for archaeological and indigenous knowledge systems to works in concert. Indigenous communities can reference and reclaim sites through the use of their own vocabularies.

Discussion: Interpretation and Integration

The examples above highlight how DINAA's approach to Linked Open Data can facilitate resource discovery as well as some research applications. However, we recognize that using the CIDOC-CRM for more comprehensive data harmonization would enable more sophisticated querying and analysis of aggregated data (Binding et al. 2008; May et al. 2015). But, greater semantic harmonization involves difficult challenges and levels of effort beyond the DINAA project's current scope. The legacy data sources aggregated by DINAA use informally defined concepts and schemas sometimes inconsistently (see also May et al. 2015 for United Kingdom and European archaeological data). Greater comprehensive semantic harmonization would require much more effort to investigate and resolve the ambiguities and inconsistencies in the legacy data. Without such care, mapping to the CIDOC-CRM would only make data seem more harmonized and comparable than they really are. While we regard CIDOC-CRM mapping to be an important long-term goal, for the immediate future, DINAA forgoes full semantic harmonization. As argued by Isaksen et al. (2014), the Pelagios project demonstrates how Linked Open Data programs can provide useful levels of data integration even without reference to a complex ontology like the CIDOC-CRM (see also Kansa 2015). Leveraging the "charm of weak semantics" (Baker & Sutton 2015) can still support efficient publication and indexing of structured data. This approach has enabled Open Context to leverage bioinformatics ontologies and datasets (chiefly the Uberon ontology and the Encyclopedia of Life) to support research outcomes based on the integration of zoorchaeological datasets (Arbuckle et al. 2014; Kansa et al. 2014).

Networking Data for Sustainability in Turbulent Times

As we discussed above, public policy, including recent "Open Government" and "Open Science" reform movements, largely shapes the creation, accessibility, and content of a dataset like DINAA. Similarly, public policy will play an important role in shaping the future of these data. The sustainability of digital data management involves a host of practical, political, financial and institutional challenges (e.g., Kintigh & Altschul 2010). Charging for access to data created by public agencies, funded by taxpayer dollars, poses ethical problems and would undermine the public benefit of this project. Since much government-created content carries no copyright restrictions, DINAA makes all data open access under a Creative Commons Zero (CC0) public domain dedication. We employ various complementary approaches to promote the sustainability of DINAA beyond the lifespan of this proposed project, including income from Open Context's research data management services, consulting and training services, and philanthropic donations. Beyond common digital preservation methods (open file formats, widely-understood documentation metadata, and support of digital repositories) we recognize that sustainability needs extend beyond preservation of data created by this project. As a comprehensive map of U.S. archaeological sites, and in its key role in linking diverse museum, archive, and library collections, DINAA will play a central role in the stewardship of cultural heritage across North America. Once completed, DINAA will require ongoing maintenance, curation, and updates as SHPOs register new sites, as well as organizational support and a governing body.

In recent years, federal agencies responsible for managing public lands have come under increasing pressure. The future for archaeology in the U.S. is even more uncertain with the new Trump administration. Archaeology, generally mandated by regulatory requirements, will suffer if enforcement of those requirements ceases. Granting bodies such as the National Science Foundation and the National Endowment for the Humanities as well as memory institutions like museums, libraries and digital archives, also face a more hostile political and financial climate. Government agencies maintain key information resources of archaeological significance, and these datastores can disappear with loss of funding. Beyond financing, the treatment of climate science and scientists by the current administration and its allies provide clear indications that it will seek to undermine freedom of inquiry and scholarly independence.

The political turmoil now experienced in the U.S. challenges existing assumptions about digital preservation and archiving. The Archaeology Data Service, a national archive for the United Kingdom, has provided a powerful and successful model to similar such efforts, including Digital Antiquity in the U.S. However, the model of a centralized "national repository" assumes that the nation will not fail or become hostile to research and the preservation of research. The rise of far-right and neo-fascist movements highlights under-recognized risks in conventional (i.e., centralized as opposed to distributed) approaches to digital preservation. The rapid changes to the U.S. political system since the election of Donald Trump have already motivated another key digital repository, the Internet Archive, to start building a backup in Canada. The Internet Archive preserves broadcast news, Websites, and government information, including climate science data. Much of this content has become or may become highly politicized. Duplicating this information in Canada highlights how distributing research data across international borders can help guard against national political risks.

The Internet Archive's recent efforts show how decentralized and globally distributed strategies for digital preservation may better weather the rise of hostile national governments (see also Findlay 2015). With respect to DINAA, and Open Context more generally, we are taking steps to archive content in multiple institutions globally. In addition to archiving with the California Digital Library, we now have a fully mirrored instance of Open Context and DINAA hosted by the German Archaeological Institute. We are also archiving data with the Internet Archive so that it can be globally replicated along with other climate science and US government information. However, these initial steps toward distributed preservation can be taken much further. Blockchains (see Miller et al. 2014; Findlay 2015), implemented by decentralized systems such as the Interplanetary Database (https://ipdb.foundation/), have the potential to help support archival and data management services across globally distributed networks.

This content is relatively easy to preserve in distributed networks because the data are small (in a technical sense, though large in a thematic sense), encoded in open formats, and have no security or intellectual property restrictions. These data can and should be used to experiment with new decentralized institutional and technical means for digital preservation. As discussed, Linked Open Data provides pathways for different datasets, curated by different communities, to enrich and more broadly contextualize each other. We should explore ways to build better data preservation approaches into Linked Open Data's globally distributed collaboration. Globally networked professional societies, library and museum "memory institutions", and even enthusiast communities can use Linked Open Data as a basis to discover and document relevant content to backup and secure. In this way, efforts like DINAA can not only survive national political turbulence, but by networking related information together, DINAA may also facilitate the preservation of digital cultural information more broadly.

Conclusion

DINAA offers a case study in migrating key archaeological information from centralized government controlled repositories to more distributed civil society networks. In making these data easily referenced on the Web and accessible to different public stakeholders, DINAA provides a basis for integrating a host of other information in various government, museum, and library contexts. These Linked Data efforts help to explicitly define relationships between disparate parts of the published and digitally-curated archaeological record. At the same time, in building bridges with tribal heritage officials and in aggregating and internationally disseminating data created by administrative offices, DINAA helps make these data more integral to wider segments of civil-society and more resilient to local and national political upheavals. In doing so, DINAA highlights how "bottom-up" and globally-networked groups can help play an important role in safeguarding digital aspects of archaeology through turbulent times.

ACKNOWLEDGMENTS

This research is funded by the National Science Foundation (1216810, 1217240, 1623621, 1623644) and the Institute of Museum and Library Services (LG-70-16-0056-16). We are very grateful for the enthusiastic participation of the growing number of DINAA project partners.

REFERENCES

Altschul, J. H., & Patterson, T. C. 2010. Trends and employment in American archaeology, in W. Ashmore, D. Lippert, & B. J. Mills (ed.) Voices in American archaeology: 291-316: Washington, D.C.: SAA Press.

Anderson, D. G., 2017. Using CRM data for "big picture" research, in F. P. McManamon (ed.) *Perspectives on Cultural Resource Management*. Oxon, U.K.: Taylor and Francis. In Press.

Anderson, D. G., & V. Horak (ed.) 1995. *Archaeological site file management: a southeastern perspective*. Atlanta, Georgia: Interagency Archeological Services Division, National Park Service, Southeast Regional Office.

Anderson, D. G., & D. S. Miller. 2017. PIDBA (Paleoindian Database of the Americas): Call for Data. *PaleoAmerica* 3(1):1-5.

Anderson, D. G., D. S. Miller, S. J. Yerka, J. C. Gillam, E. N. Johanson, D. T. Anderson, A. C. Goodyear & A. M. Smallwood. 2010. PIDBA (Paleoindian Database of the Americas) 2010: current status and findings. *Archaeology of Eastern North America* 38: 63-90.

Anderson, D. G., A. M. Smallwood & D. S. Miller. 2015 Pleistocene human settlement in the southeastern United States: current evidence and future directions. *Paleoamerica* 1(1): 7–51.

Arbuckle, B. S., S. W. Kansa, E. Kansa, D. Orton, C. Çakirlar, L. Gourichon, L. Atici, A. Galik, A. Marciniak, J. Mulville & H. Buitenhuis. 2014. Data sharing reveals complexity in the westward spread of domestic animals across Neolithic Turkey. *PLOS ONE*, 9(6): p.e99845.

Baker, T. & S. A. Sutton. 2015. Linked data and the charm of weak semantics: introduction: the strengths of weak semantics. *Bulletin of the American Society for Information Science and Technology*, 41(4): 10-12.

Binding, C., K. May & D. Tudhope. 2008. Semantic interoperability in archaeological datasets: data mapping and extraction via the CIDOC CRM, in *International Conference on Theory and Practice of Digital Libraries*: 280-290. Berlin: Springer.

Cochran, P. A. L., C. A. Marshall, C. Garcia-Downing, E. Kendall, D. Cook, L. McCubbin & R. M. S. Gover. 2008. Indigenous ways of knowing: implications for participatory research and community. *American Journal of Public Health* 98(1). January:22–27.

Findlay, C. 2015. Decentralised and inviolate: the blockchain and its uses for digital archives. *Recordkeeping Roundtable*. Archived at: https://web.archive.org/web/20170105012522/https://rkroundtable.org/2015/01/23/decentralised-and-inviolate-the-blockchain-and-its-uses-for-digital-archives/ [accessed 5 January 2017).

Isaksen, L., R. Simon, E. T. Barker & P. de Soto Cañamares. 2014. Pelagios and the emerging graph of ancient world data. *Proceedings of the 2014 ACM conference on Web science*: 197-201.

Jeffrey, S., J. Richards, F. Ciravegna, S. Waller, S. Chapman, & Z. Zhang. 2009. The Archaeotools project: faceted classification and natural language processing in an archaeological context. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 367(1897): 2507-2519.

Kansa, E. C. 2012. Openness and archaeology's information ecosystem. *World Archaeology* 44(4): 498–520.

Kansa, E.C. 2015. Contextualizing digital data as scholarship in eastern Mediterranean archaeology. *CHS Research Bulletin* 3 (2). Available at: http://nrs.harvard.edu/urn-3:hlnc.essay:KansaE.Contextualizing_Digital_Data_as_Scholarship.2015 (accessed 23 January 2017).

Kansa, E. C. & A. Bissell. 2010. Web syndication approaches for sharing primary data in “small science” domains. *Data Science Journal* 9: 42-53.

Kansa, E. C. & S. W. Kansa. 2011. Towards a do-it-yourself cyberinfrastructure: open data, incentives, and reducing costs and complexities of data sharing, in E.C. Kansa, S.W. Kansa & E. Watrall (ed.) *Archaeology 2.0: new approaches to communication and collaboration*: 57-91. Los Angeles, CA: Cotsen Institute of Archaeology Press.

Kansa, E.C. & S.W. Kansa. 2013. We all know that a 14 is a sheep: data publication and professionalism in archaeological communication. *Journal of Eastern Mediterranean Archaeology and Heritage Studies* 1(1): 88–97.

Kansa E. C., S. W. Kansa & B. Arbuckle. 2014. Publishing and pushing: mixing models for communicating research data in archaeology. *International Journal of Digital Curation* 9(1): 57– 70. doi:10.2218/ijdc.v9i1.301

Kintigh, K. W. & J. H. Altschul. 2010. Sustaining the digital archaeological record. *Heritage Management* 3(2): 264-274.

Kintigh, K. W., J. H. Altschul, M. C. Beaudry, R. D. Drennan, A. P. Kinzig, T. A. Kohler, W. F. Limp, H. D. G. Maschner, W. K. Michener, T. R. Pauketat, P. Peregrine, J. A. Sabloff, T. J. Wilkinson, H. T. Wright & M. A. Zeder. 2014a. Grand challenges for archaeology. *American Antiquity* 79(1): 5- 24.

Kintigh, K. W., J. H. Altschul, M. C. Beaudry, R. D. Drennan, A. P. Kinzig, T. A. Kohler, W. F. Limp, H. D. G. Maschner, W. K. Michener, T. R. Pauketat, P. Peregrine, J. A. Sabloff, T. J. Wilkinson, H. T. Wright & M. A. Zeder. 2014b. Grand challenges for archaeology. *Proceedings of the National Academy of Sciences* 111(3): 879-880.

Kintigh, K. W. 2015. Extracting information from archaeological texts. *Open Archaeology* 1(1). doi: 10.1515/opar-2015-0004.

Lake, M. 2012. Open Archaeology. World Archaeology 44(4): 471–478.

Lane, P. J. 2015. Archaeology in the age of the Anthropocene: A critical assessment of its scope and societal contributions. *Journal of Field Archaeology* 40: 485–498.

Lightfoot, K. G., & R. Q. Cuthrell. 2015. Anthropogenic burning and the Anthropocene in late-Holocene California. *The Holocene* 25: 1581–1587.

May, K., C. Binding & D. Tudhope. 2015. Barriers and opportunities for Linked Open Data use in archaeology and cultural heritage. *Archäologische Informationen*, 38(1): 173-184.

McManamon, F. P. & K. Kintigh. 2010. Digital Antiquity: transforming archaeological data into knowledge. *The SAA Archaeological Record* 10(2): 37-40.

Miller, A., A. Juels, E. Shi, B. Parno & J. Katz. 2014. Permacoin: repurposing bitcoin work for data preservation. *2014 IEEE Symposium on Security and Privacy*: 475-490.

NADB 2016. *National Archeological Database, Reports Module*. https://www.nps.gov/archeology/tools/nadb.htm (Accessed 26 January 2016).

NADB Maps. 1993. *NADB-Maps Archeological Site Counts (State Historic Preservation Officers)*. Archeology Program, National Park Service, Washington, D.C. Original URL http://cast.uark.edu/other/nps/maplib/USsittot.1993.html archived by University of Arkansas Libraries with Archive-It.org and available online at http://cast.uark.edu/cast-archived-page.php (Accessed 5 August 2016).

Nicholas, G. 2008. Native peoples and archaeology (indigenous archaeology), in D. Pearsall (ed.) *The Encyclopedia of Archaeology* Vol. 3:1660–1669. Oxford: Elsevier.

Rabinowitz, A. 2014. It's about time: historical periodization and Linked Ancient World Data. *ISAW Papers* 7(22). Available at: http://dlib.nyu.edu/awdl/isaw/isaw-papers/7/rabinowitz/ (accessed 23 January 2017).

Raviele, M. E. 2014. Archaeology and an interdisciplinary Digital Age. *Institute of Museum and Library Services*. Available at: https://www.imls.gov/news-events/upnext-blog/2014/05/archaeology-and-interdisciplinary-digital-age (accessed 16 December 2015).

Rocks-Macqueen, D. 2014. The actual amount of National Science Foundation funding for archaeology. *Doug's Archaeology Investigating the Profession and Research.* Available at: https://dougsarchaeology.wordpress.com/2014/03/06/the-actual-amount-of-national-science-foundation-funding-for-archaeology/ (accessed 22 December 2016).

Shaw, R., A. Rabinowitz, P. Golden & E. Kansa. 2016. A sharing-oriented design strategy for networked knowledge organization systems. *International Journal on Digital Libraries* 17(1): 49-61. doi:10.1007/s00799-015-0164-0

Sheehan, B. 2015. Comparing digital archaeological repositories: tDAR versus Open Context. *Behavioral & Social Sciences Librarian* 34(4):173-213. doi:10.1080/01639269.2015.1096155

Wells, J. J., E. C. Kansa, S. W. Kansa, S. J. Yerka, D. G. Anderson, T. G. Bissett, K. N. Myers & R. C. DeMuth. 2014. Web-based discovery and integration of archaeological historic properties inventory data: The Digital Index of North American Archaeology (DINAA). *Literary and Linguistic Computing* 29(3): 349–360. doi:10.1093/llc/fqu028

White, A. A. 2014. Inter-linking with DINAA; database search; dataset clean up. *Eastern Woodlands household archaeology data project*. Available at: http://www.householdarchaeology.org/whats-new/inter-linking-with-dinaa-database-search-dataset-clean-up (accessed 21 December 2016).

Zalasiewicz J., C. N. Waters, M. Williams, A. D. Barnosky, A. Cearreta, P. Crutzen, E. Ellis, M. A. Ellis, I. J. Fairchild, J. Grinevald, R. Leinfelder, J. McNeill, C. Poirier, D. Richter, W. Steffen, D. Vidas, M. Wagreich, A.P. Wolfe, and A. Zhisheng. 2015. When did the Anthropocene begin? A mid-twentieth century boundary level is stratigraphically optimal. *Quaternary International* 383:204-207.

FIGURES

Figure 1: DINAA partnerships as of December 2016 with dot density plot showing distribution of cultural resources at low resolution within states whose data has been received thus far. Dots do not refer to exact site locations, but to groups of five sites whose position has been randomly distributed within 20x20 km grid cells.

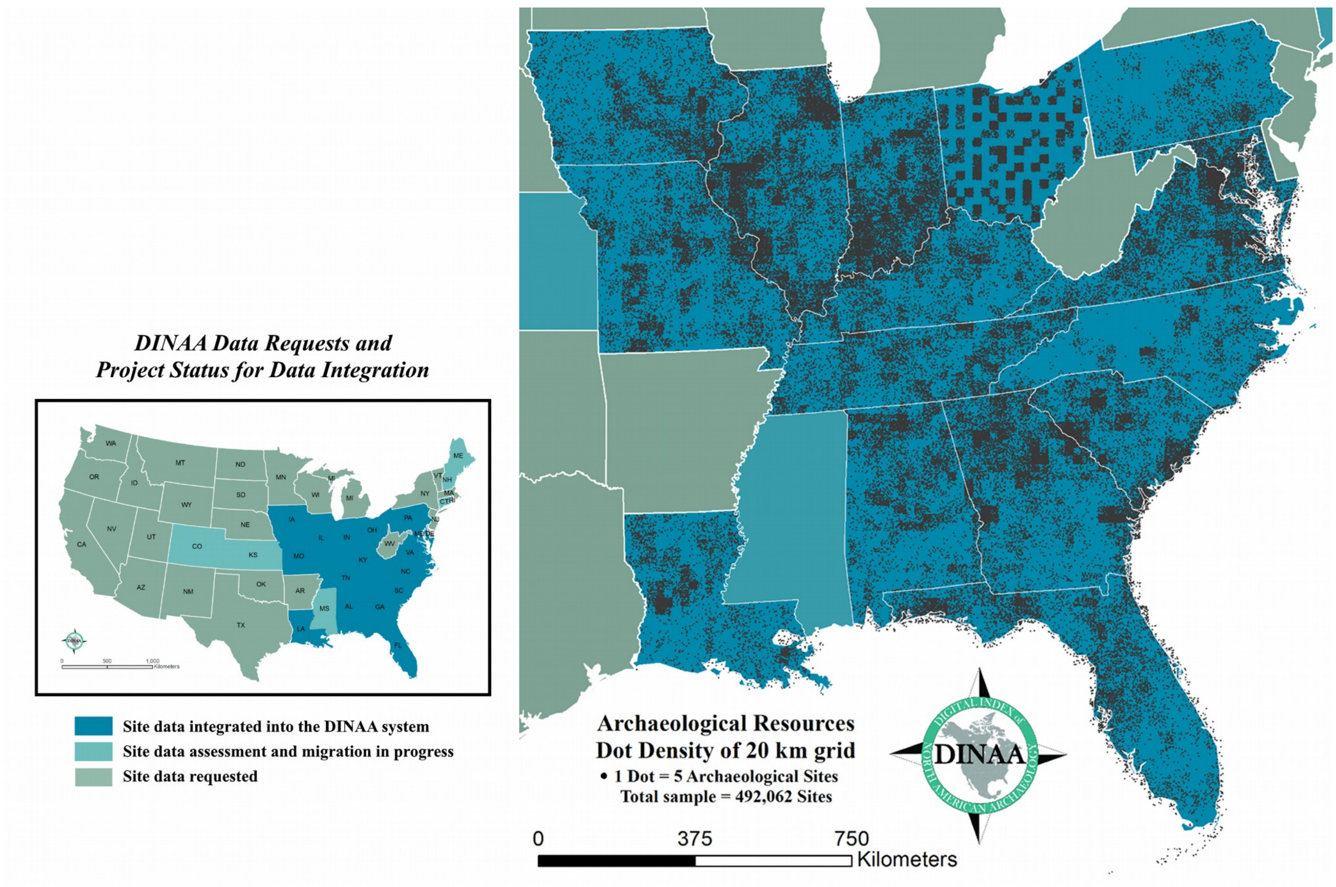


Figure 2. DINAA area of effect as of December 2016 with dot density plot showing distribution of cultural resources at low resolution within states whose data has been received thus far, or is being delivered (n=21). Dots do not refer to exact site locations, but to groups of five sites whose position has been randomly distributed within 20x20 km grid cells. Ohio data is at county-level resolution.

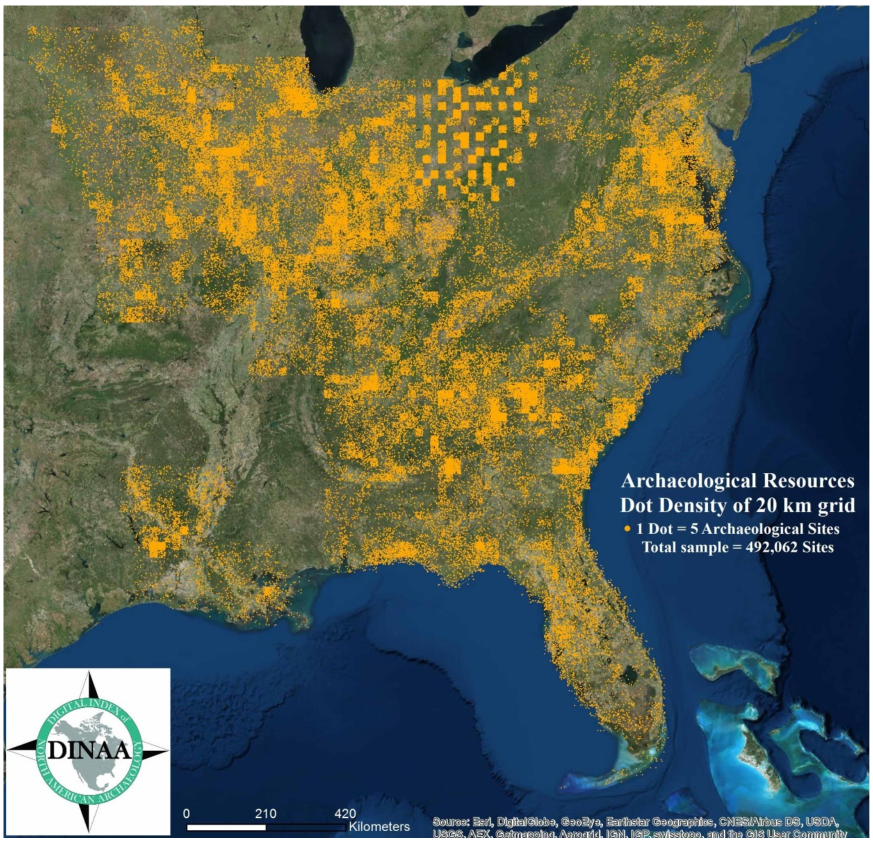


Figure 3. Open Context map showing how DINAA site file records are displayed online. The display employs shaded 20x20 km grid cells to protect site locational information.

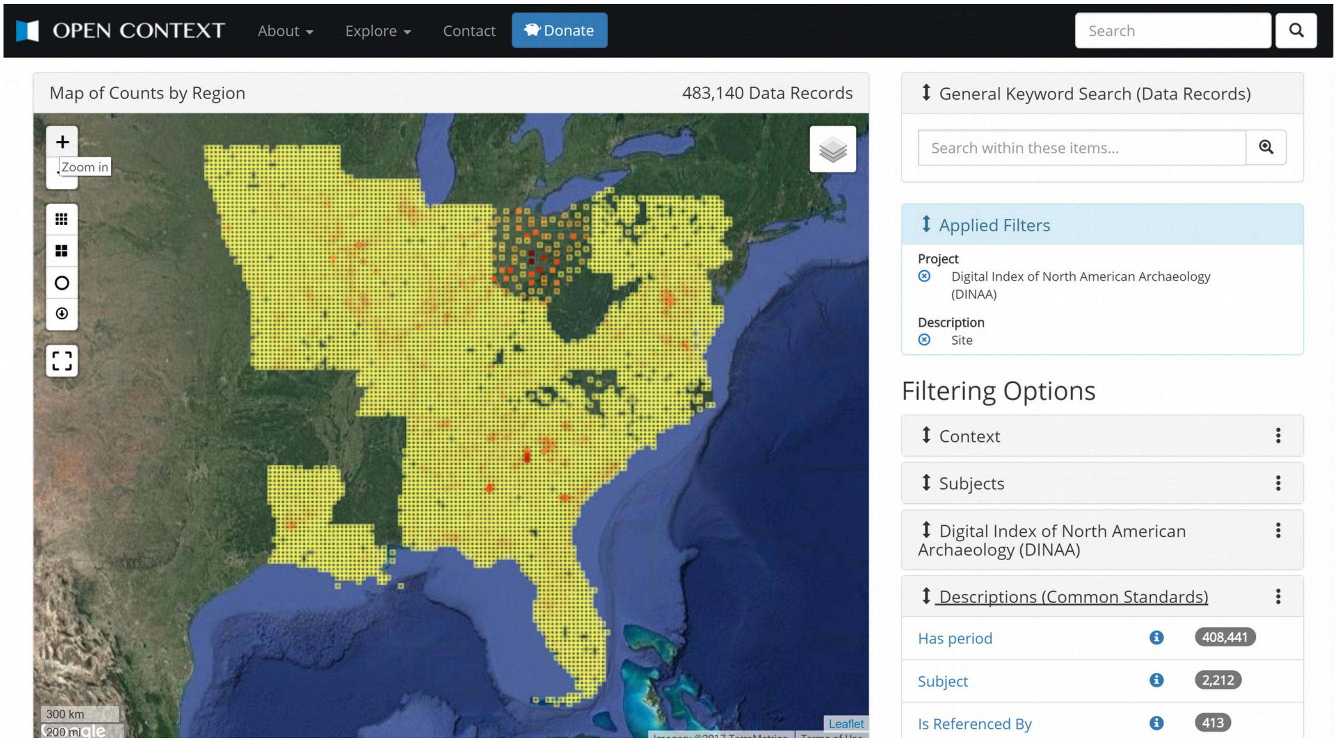


Figure 4. Map showing site density as it relates to potential loss from sea-level rise, and grouped by elevation in meters above present mean sea level, illustrating all sites within a buffer of 200 km from the present coastline in grey.

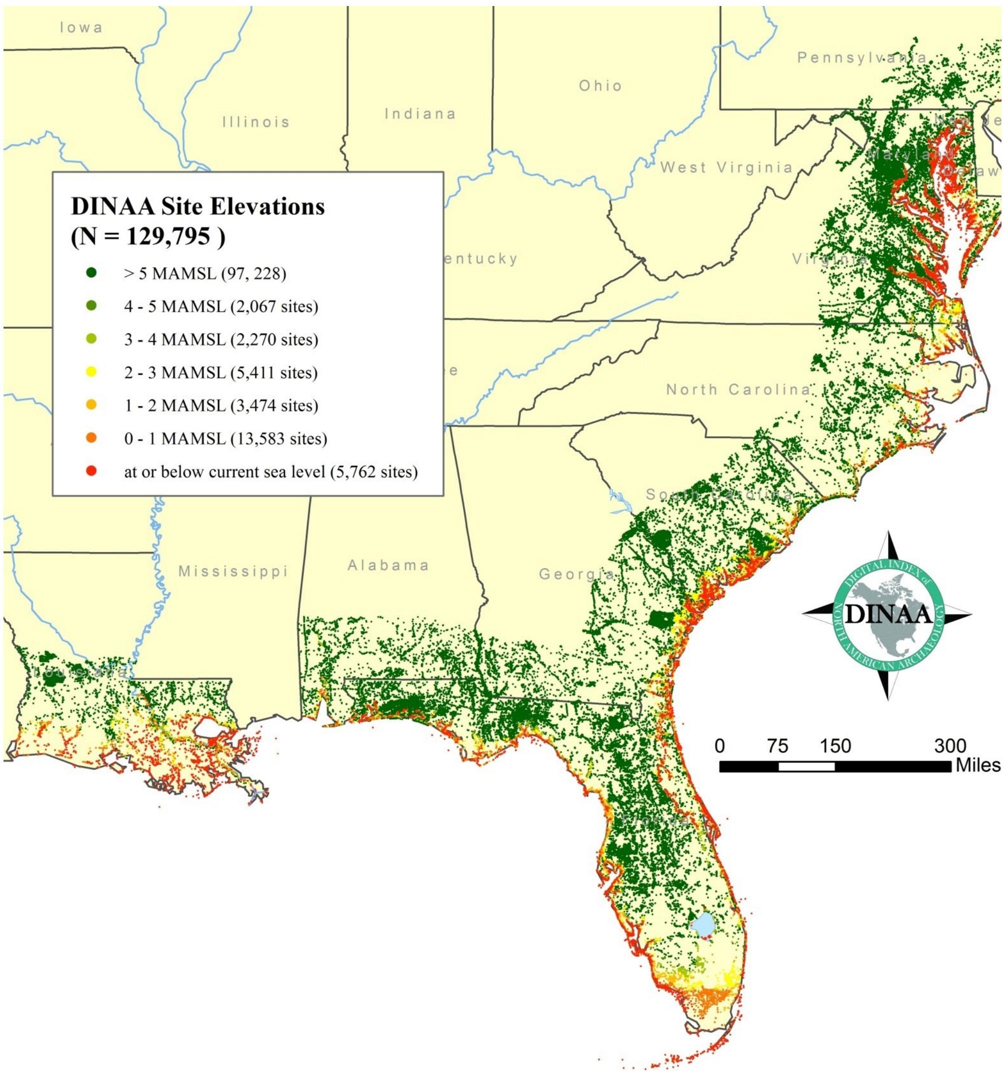


Figure 5. DINAA permits links to information in a wide range of online data repositories, using site numbers as the common denominator. DINAA directs users to these outlets, but access and content control remains on their systems.

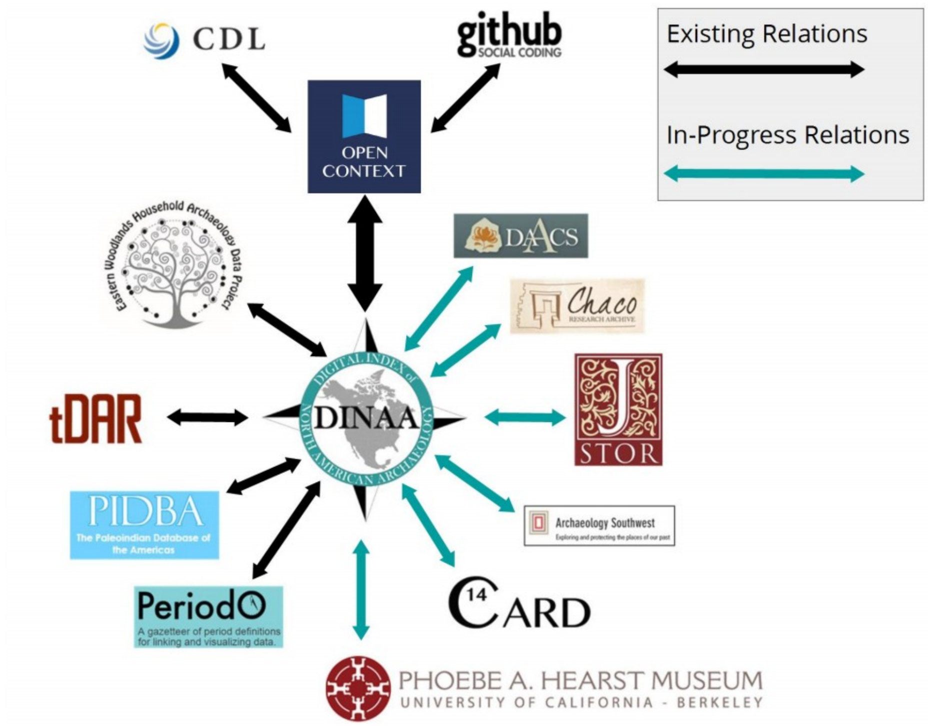
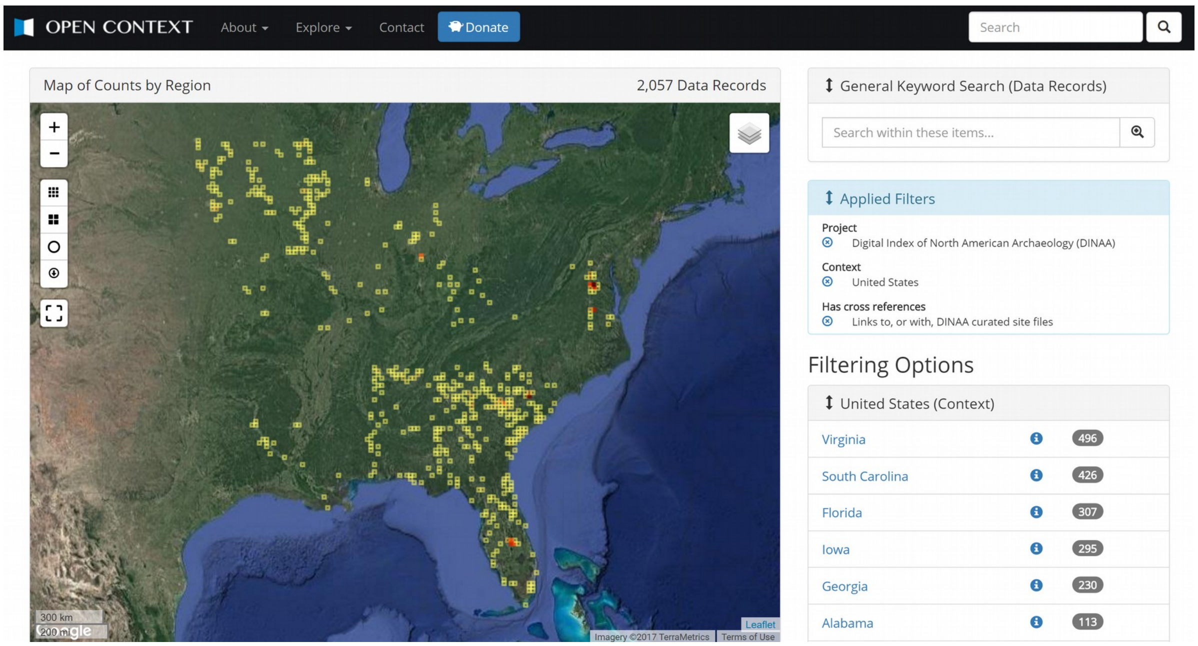


Figure 6. Sites in DINAA that have Linked Data relationships with reports or other data in external data repositories. The coverage presented here reflects states with site file data in tDAR, the Eastern Woodlands Household Archaeology Database (White 2014), the Paleoindian Database of the Americas (Anderson and Miller 2017; Anderson et al. 2010), and the Federal Register.



TABLES

Table 1. Archaeological site and National Register of Historic Places eligible property loss in the southeastern United States, given sea-level rise in 1-meter increments.

	No El. Data	<0	0 to 1 m	>1 to 2 m	>2 to 3 m	>3 to 4 m	>4 to 5m	Totals
SITE TOTALS	331	5762	13583	3474	5411	2270	2067	32898
Cumulative Site Totals	331	6093	19676	23150	28561	30831	32898	
NRHP PROPERTY TOTALS	0	283	1035	393	440	187	134	2472
Cumulative NRHP Property Totals	0	283	1318	1711	2151	2338	2472	

*No archaeological site data currently available from Mississippi

SUPPLEMENTARY MATERIAL

None included.