Joshua Wells is an associate professor of anthropology and social informatics, and director of CERES, the Center for Excellence in Research and Scholarship, at Indiana University South Bend. He works on big data issues in heritage management and archaeological science, science-technology-society interactions, and affordances and limitations of technological strategies in evolving disciplines. He is especially interested in the ethical and community power of “free” and “open” software and data strategies, including geospatial resources. Wells is a member of the editorial board for the primary archaeological data publishing group, Open Context. From 2010–2018 he chaired the Digital Data Interest Group for the Society for American Archaeology. He has reviewed grant proposals for numerous research agencies, including NSF and NEH. He has consulted on big data, open science and technology access with groups as small as county libraries and as large as the European Union.

**Perspective Statement**

As an anthropological archaeologist, I am interested in the ways that geographic location provides an entry point to contextualize information about the past across a variety of potentially intersecting sociocultural and physical science vectors. Within anthropology and allied social sciences, the understanding that cultures and societies are bounded spatiotemporal phenomena provides the potential to categorize the scope of a particular belief system or social order in time and space and to illuminate processes of change as elements of group behavior accumulate change to form seemingly distinct horizons of human experiences. With this in mind, geographic location becomes key to understanding the longitudinal persistence or change in particular sociocultural behaviors, through time in a particular area (whether from the scale of a village or city-state to a geographic region). In a complementary fashion, geographic location becomes key to understanding transmission of sociocultural behaviors from one place to another through combinations of communication, trade, syncretism, or population movement. Thus, in archaeology, the sociocultural phenomena described to define past peoples have salient geographical aspects that range from local points to large polygonal areas of interest, and often including overlapping areas of combined definitional sets. Furthermore, such phenomena are often deduced from physical science investigations in discrete locations, including areal surveys, excavations, and proveniences of particular materials used to derive information based in astronomy, biology, chemistry, geology, and physics. Subsequently, the communication of location is important, but the definition of that information may have to consider a mixture of possibilities, including how that
location was represented by past peoples, how that location was understood by archaeological investigators in a research agenda, and how that location is defined in a modern geopolitical context where jurisdictions from the local to the international have salient characteristics.

In practice, enabling discovery of research data in archaeology, per the geographic qualities described above, may seem thorny. However it is important to remember that all of those qualities may be represented in some fashion on the face of the Earth, and the solution involves vocabulary control that defines where particular concepts are important and where they overlap. For instance, about 1200 CE, in Evansville, Indiana, on the Ohio River, was the Angel Site, the fourth largest indigenous city-state of the Mississippian cultural tradition, a walled town with thousands of inhabitants who lived in ethnic neighborhoods among monumental earthworks. The Angel Site may be known as part of the Mississippian phenomenon, in which case key terms and geographic definitions of the extent of Mississippian peoples from Minnesota to Florida become important. The site may be understood as the capital of a polity about 50 kilometers in radius, which would have its own spatial boundaries that fluctuated in time in southern Indiana and northern Kentucky. The site may be understood in terms of the archaeological investigations conducted there for about 75 years, with changing geographic contexts as scientific sampling strategies (units, trenches, levels, horizons, features, samples, etc.) changed in the discipline. Finally, the site may be understood as a protected heritage resource which is a National Historic Landmark and an Indiana State Park, with a Smithsonian Trinomial Number of 12VG1 (an alphanumeric intelligent key, 12 refers to Indiana, VG to Vanderburgh County where the site is located, and 1 designates that the site was the first recorded to the government roster in that county), which will be used to denote information related to the Angel Site in government documents, in many research journals, and in many museum and archival collections.

Spatialization techniques that can enable data discovery within the realm of archaeology must then be able to account for a variety of overlapping and complementary concepts in space. They may also need to take into account the need for periodization, to recognize that a particular set of concepts which may have applied to geographic area during one time period may not be appropriate or useful for all time periods. Archaeology as a discipline is currently engaged in a massive effort to wrangle ontologies and vocabularies in regions, nations, and across international consortia. Many best practices currently focus on tagging and accumulating definitions in areas, rather than providing rigid vocabulary controls that may exclude nuance. More particular temporal and investigative strategy tools may then be applied to refine and isolate particular instances of information.