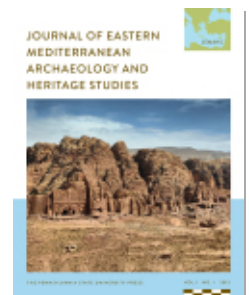# Additional Thoughts on Sustaining and Promoting Open Data in Archaeology

Eric C. Kansa, Sarah Whitcher Kansa

publications that rely on that data, thus increasing the value of both the data sets and the interpretive works. If one of the biggest barriers to the publication of data is the lack of formal credit within the academy, it strikes me that this could be a good place to start chipping away at the wall, by linking the published data to the things that do, in fact, count.

The section of Kansa and Whitcher Kansas's analysis that I would like to expand upon is the sustainability discussion. Who does pay for data publication? Right now, grant funding is available for data initiatives like those described, but presumably those granting agencies are trying to seed the path toward data publication with the belief that a sustainable model will eventually be found. For a born-digital entity like Open Context or other archiving services like tDar/Digital Antiquity, an open access model makes a lot of sense, but what happens if/when the funding dries up? Who pays to keep the service alive? The authors of the datasets (whether directly or through their project-specific research grants)? Author fees may limit participation and create a less egalitarian result than that proffered here. Do the institutions that provide the hosting for the various platforms simply eat the cost? If institutions pay, then the free rider problem takes center stage, unless participation is limited to that institution's faculty (not exactly an "open" publishing model). Plus, if the institution is willing to foot the bill now and does not care about freeloaders in the system, will they continue to feel the same in two years if their budgets are squeezed? From a university press point of view, this is a critical question. As university presses have learned the hard way, what was once a funding priority for the university might seem less important in leaner years. Perhaps most significantly, how does the financial equation change if data publication becomes the norm, as the authors of this article hope? Will the existing staff and technological resources at data publishing services be sufficient? How much additional revenue will be required and how does that alter the "who pays" equation? Currently, most data services like those described in this article have a relatively high level of support compared to the level of submissions. If successful, that is almost certainly going to flip.

I do not want to suggest that sustainability of an open data model is not possible, but I do think it is worth noting that commercial and scholarly publishers are all trying to wrap their heads around the sustainability of various forms of open access publishing, and nobody has found *the* answer. There have been successes, such as *PLOSOne*, there are some intriguing new models being explored, such as *PeerJ*, and more are in development. But there have also been numerous failures and disappointments as well. The sustainability piece of the equation is absolutely critical. Just because something is of value to the scholarly community does not mean that the financial support will automatically follow. I do hope, though, that services like Open Context and Digital Antiquity find that elusive balance.

REJOINDER

Additional Thoughts on Sustaining and Promoting Open Data in Archaeology

ERIC C. KANSA AND SARAH WHITCHER KANSA

We are grateful to the respondents for their thoughtful contributions to this forum. Many of their comments focus on the issue of the profitability of open access/open data models for publication (see Pratt, Alexander, and Hall). Our own livelihood depends on sustainability, so we are naturally concerned about this issue. In our work with Open Context, we have benefited from public and private granting agencies for funding, libraries for data management services, and contracts for helping to build similar systems in other domains. This mix of funding will no doubt change as time goes on and we adapt to new circumstances. Digital Antiquity and the Archaeology Data Service, both large disciplinary repositories, by necessity have greater infrastructure and other costs, and will need other business models. Similarly, university-based repositories, such as the California Digital Library, rely on other institutional funding streams. We think this diversity of business models and organizational forms is

healthy. Some solutions will fail, but some will succeed. Since open data encourages the free flow of information, it represents an important strategy of resiliency. If Open Context fails, its data, archived in the California Digital Library and other repositories, will be available for scholars and open for incorporation in future information systems. Policy needs to promote the health of the overall information ecosystem that sustains research so experimentation and turnover can happen without loss of irreplaceable research content.

We are encouraged that the respondents all agree that open data needs to play a greater role in archaeology. For the time being, we face the great challenge of bridging two realities—moving from currently entrenched practices to a future of more open and diverse scholarly outputs. The growth of new, data-intensive research methods will lead to increased pressure for open access and open data. Hall rightly highlighted how data (and text) mining are only feasible with open access. Though not addressed in this venue, Kansa (2012) further explores open access with respect to text-mining in archaeology.

Today's young scholars may be more accustomed to sharing many aspects of their lives via the Web. Young scholars also largely expect fully digital workflows in their research. However, habituation with Facebook and Excel does not necessarily lead to greater data professionalism and openness. Academia's highly competitive job market makes any deviation from established patterns of success extremely risky, replicating cultural norms that work against openness. Porter's comments about "data as symbolic capital" can give needed conceptual tools to explore data's role in scholarly culture. He rightly highlights data's roles in building prestige and recognition. Currently, data may not routinely be publicly shared, but it does circulate privately among networks of researchers to reinforce collaborative ties and alliances. Better understanding of data as symbolic capital is needed to encourage researchers to move such data exchanges toward public channels, where we can better document provenance, improve quality, and preserve data with professionally managed repositories.

Instead of reinforcing a dysfunctional status quo, established leaders in the field can be drivers of reform. Such individuals have the job security to take risks and innovate (supposedly the purpose of tenure!). They can use their prestige to make open access and open data respectable and respected. A hopeful sign of progress, Digital Antiquity boasts a board that includes past and current presidents of the Society for American Archaeology. However, we don't want to set the bar too low. We need to do more than simply add data archiving to existing and highly constrained publishing practices. To fully realize the value of archaeological scholarship, we need innovation in what constitutes publication, innovation in new ways of promoting quality, and innovations that better promote accessibility and interoperability of archaeology on the open Web.