

CONFERENCE ON ELECTRONIC PUBLICATION OF ANCIENT NEAR EASTERN TEXTS

David Schloen

The following report offers a summary of the conference, *Electronic Publication of Ancient Near Eastern Texts*, sponsored by the Oriental Institute and the Franke Institute for the Humanities, to discuss the prospect of using the World-Wide Web for electronic publication of ancient Near Eastern texts. The conference was held at the Franke Institute, 8–9 October 1999.

The organizing committee (Gene Gragg, Charles Jones, John Sanders, David Schloen, and Matthew Stolper) was pleased by the enthusiastic response to the announcement of a conference. Nearly thirty colleagues from out of town attended the two-day meeting, including several from Europe and Israel. They were joined by roughly the same number from the University of Chicago and neighboring institutions. During some sessions the number in attendance reached seventy, and it never fell below forty. More important than numbers, however, is the high level of discussion and debate that was achieved during the conference, as all who participated will attest. This is due to the energy and acumen both of those who made formal presentations and those who spoke up during the discussion sessions. We are grateful to all of the participants for the time and effort they invested in coming to Chicago and in contributing to the success of this conference.

A major goal of the conference was to assess the prospects for establishing a formal international standards organization charged with setting technical standards for the interchange of Near Eastern data in digital form. Both the conference and the establishment of such an organization are timely in light of the recent development of Internet-oriented data standards and software that now provide a common ground for cooperation among diverse philological and archaeological projects, which have heretofore adopted quite idiosyncratic approaches. This common ground, not just for academic research but in all areas of information exchange, is created by the Extensible Markup Language (XML) and a growing array of software tools that make use of XML to disseminate information on the Internet.

As noted in the original announcement of the conference, XML is a non-proprietary “open” or public standardized data format that provides a simple and extremely flexible “tag”-based syntax for both representing complex information as a stream of ordinary text and delivering it over the World-Wide Web (for more details, see <http://www.oasis-open.org/cover>). Furthermore, XML is based on a proven approach because it is a subset of the International Standards Organization-ratified Standard Generalized Markup Language (SGML), which has been used for electronic publication world-wide for more than a decade. XML therefore makes possible

powerful and efficient forms of electronic publication via the Internet, including academic publication of philological and archaeological data. But XML itself is merely a starting point because its simplicity and flexibility, which ensure its widespread adoption, require the development of specific XML tagging schemes or “markup languages” appropriate to each domain of research. Such a tagging scheme expresses the abstract logical structure of a particular kind of data in a rigorous and consistent fashion. Thus, for example, chemists have already created a “chemical markup language” using XML to express the structure of molecules and chemical reactions, so that the data they work with can be easily shared and searched on the web. Likewise, NASA has created an “astronomical instrument markup language,” biologists have created a “biological markup language,” and so on. Once such tagging schemes exist, various kinds of software can then be developed to present different views of logically structured data for different purposes, or to create new sets of data structured in a particular way, with the assurance that these data structures can be created and viewed on any computer anywhere without special conversions or translations.

The consensus of the conference participants is that XML should be used as the basis for future electronic publication of Near Eastern data. The establishment of a formal working group for Near Eastern text markup was also endorsed as a vehicle for the collaborative development and dissemination of suitable XML tagging schemes and associated software. The name and scope of the new standards organization remain to be decided. A number of conference participants emphasized the importance of including Near Eastern languages and texts of all periods within the scope of the text markup group, rather than arbitrarily limiting it to ancient Near Eastern texts, because the same issues arise in dealing with non-European scripts and languages regardless of their date. Similarly, several people expressed what seemed to be a generally held desire to find ways to include electronically published archaeological data within the standards-setting effort. This would ensure maximum interoperability of textual and archaeological datasets, so that it would be easy to obtain information about the spatial provenience and the material-cultural context of excavated or monumentally inscribed texts, and conversely so that it would be easy to obtain philological information about texts viewed as artifacts from an archaeological perspective.

In the opinion of the conference organizing committee, therefore, a new standards organization could develop XML markup for Near Eastern information of all kinds, including both primary data (philological, archaeological, and geographical) and relevant secondary literature. Even restricting the scope to “Near Eastern” information is rather arbitrary from a technical standpoint, but this mirrors the scope of the existing academic infrastructure of Near and Middle Eastern departments, institutes, and centers to which members of this organization would belong. Note that the mission is not just to devise XML tagging schemes but also to facilitate the development of well-documented web browser-based software, which could be widely shared among Near Eastern projects, and to coordinate training and professional development for researchers who want to learn how to use these tagging schemes and software. Thus at some point it might also be desirable to create a formal task force for training and professional development within this organization.

In the aftermath of the conference, discussion is underway concerning these details, including the name and the precise scope and mode of operation of a new international organization, as well as a schedule of future meetings. Decisions will be announced in the future, but it is clear already that there is a widespread desire to make this organization as broadly based as possible so that it can facilitate the cooperative development of effective and widely accepted technical standards. The Oriental Institute will continue to sponsor this effort and to support it with its

PHILOLOGY

reputation and resources, in collaboration with the University of Chicago's Department of Near Eastern Languages and Civilizations, Center for Middle Eastern Studies, and Committee on the Ancient Mediterranean World.

The conference program and abstracts, a list of attendees, and summary is available on the Oriental Institute website at http://www-oi.uchicago.edu/OI/INFO/XML_Conference_1999.html.

A full account of the conference was published as "Electronic Publication of Ancient Near Eastern Texts" by Charles E. Jones and David Schloen in *Ariadne* 22 (December 1999), which is available on-line at <http://www.ariadne.ac.uk/issue22/epanet/>
