

# OCHRE DATA SERVICE

MILLER PROSSER and SANDRA SCHLOEN

## WE INTERRUPT OUR REGULAR PROGRAMMING . . .

The Online Cultural and Historical Research Environment (OCHRE), the software platform supported by the [OCHRE Data Service](#), has a long history of development—long, that is, for the technology industry. From its beginnings in 2001 in its current incarnation as a Java-XML application for archaeologists, it has grown to be a full-featured, comprehensive database platform serving fifty plus active projects and over six hundred users in disciplines encompassing not only archaeology and philology—the specialties of the Oriental Institute—but also paleontology, ancient genetics, textual criticism, lexicon building, and language learning, in contexts that span a wide range of space and time. As a stable platform that currently manages over half a million images and documents, and over eight million database items, the OCHRE system is primed to meet the needs of research projects of all kinds. With the active phase of major software programming behind us, our priorities are shifting, allowing us to venture into new territory, both literally and virtually.

## . . . TO TAKE ON NEW ADVENTURES

Since 2016, we have been cultivating a program whereby an OCHRE Data Service research assistant is embedded in an archaeological project onsite to manage the process of “born digital” data collection. Leading up to the beginning of the excavation season, the OCHRE Data Service trains an advanced student in a variety of skills needed to serve as an onsite data manager.

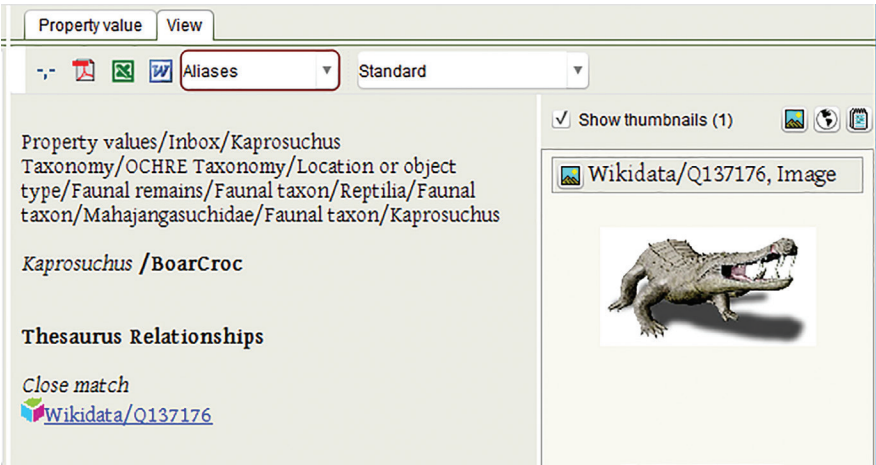
This highly successful arrangement is a win for the student who gains valuable experience in research data management and who is trained on other on-the-ground skills such as aerial photography (using camera-equipped drones) and geo-spatial data processing (using GIS tools). This is also a win for the archaeological field project, which wraps up its season with its data recording completed, photographs linked, specialist data integrated, and all the project’s data available to all project collaborators on the web as they go home to their respective institutions. OCHRE Data Service onsite data managers to date have been Nicholas Schulte, Tel Shimron, Israel (2017); Emma Kerr, Zincirli, Turkey (2018); Abraham Seare, Cor-



University of Chicago  
PhD student Emma  
Kerr, onsite OCHRE data  
manager for the Zincirli  
Expedition, Turkey, 2018.

ral Redondo, Peru (2018, 2019); Stephanie Baumgart, Gobero, Niger (2018); Andrew Wright, at both Nippur, Iraq (2019) and Tel Shimron, Israel (2019); and Alexander Ward, Tell Keisan, Israel (2019). Many of these students present their research experience in posters and presentations at academic conferences, giving them valuable experience in professional contexts.

The increasingly wide geographic spread of our operations continues to expand the breadth of data being recorded in OCHRE. As we began work in Peru, we added descriptors for textile conservation and new species like “llama” not previously encountered in our work, which had been primarily in the Middle East. For work in Niger, we added hitherto unrepresented descriptive values like “fel-site,” “crocodile,” “harpoon,” and “Jurassic,” not to mention “Nigersaurus” and its extinct relatives. In southern Iraq our staff GIS expert, Andrew Wright, spent nine days acquiring aerial footage of the massive and impressive site of Nippur, adding “looter’s pit” to our set of taxonomic descriptors. But whether a skull of a goat in Israel, a skull of a bear in Turkey, a skull of a human in Peru, or a skull of a dinosaur in Niger, a “skull” is a “skull,” and OCHRE allows extensive sharing and reuse of its descriptive taxonomies across projects. OCHRE is not bound to *any* specific time or space, and so it is widely applicable to research projects of all kinds.



“BoarCroc” nicknamed by University of Chicago paleontologist Paul Sereno and cataloged in OCHRE by research assistant Katherine Hodge (University of Chicago BA 2019); illustrated by a Semantic Web link to Wikidata from OCHRE.

## ... TO ADD SOME BELLS AND WHISTLES

As we have expanded our work into new parts of the world and taken on new collaborators, we have become more sensitive to the need for internationalization of the platform. Java provides good support for internationalization, and so the OCHRE *interface* is currently being translated into several languages, beginning with French, Spanish, Hebrew, and Chinese. This reflects our diverse client base and illustrates support for a range of language features including right-to-left writing systems and non-Latin scripts. The OCHRE *data* can be translated by projects into any language of choice. We provide tools to make this easy. Specifically, we have recently integrated the Google Translate API with OCHRE’s string processing features. Simply pick the language to which you wish to translate the current character string, and OCHRE will send it to Google Translate for translation. It can then be edited if you wish to improve on the automatic translation. Once again, we are inspired by our wide range of projects. Translating into French the widely used property “Location or object type” serves our French-Canadian colleagues who worked at Tell Acharneh (Syria), the local museum collaborators in central Niger (where French is the official language), and our own researchers in Chicago who are integrating data about Ras Shamra-Ugarit published originally in French.

The screenshot shows a web interface for the OCHRE Data Service. At the top, there is a 'Name' label. Below it is a horizontal menu with buttons for 'English', 'Chinese', 'French', 'German', 'Italian', 'Spanish', and 'Turkish'. Below the menu is a text input field containing the text 'Type de lieu ou d'un objet'. To the right of the input field is a small icon of a speech bubble with a French flag below it.

Translating  
“Location or  
object type”  
using OCHRE’s  
multilingual  
features.

## ... TO PERFORM RESCUE OPERATIONS

New adventures are always fun, but we are increasingly being called upon to rescue old data. Retiring professors who had been early adopters of technology find themselves saddled with legacy data facing an uncertain future. Data in old formats (do you remember Corel Draw or FoxPro or non-Unicode fonts?) on old hardware (floppy disks anyone?) is doomed to oblivion unless measures are taken to transform it to more stable formats on supported platforms. Modeled as a data warehouse, OCHRE is the perfect resting place for legacy data. Having been designed based on XML—a human-readable, non-proprietary format—OCHRE data is archive-ready. In the care and keeping of the University of Chicago library, OCHRE data will live “forever.”

## ... TO PLAY NICELY IN THE WIDER WORLD OF THE WORLD WIDE WEB

Regardless of its origin, as born-digital or legacy data, OCHRE’s fast-track publication features bring data to life! A user simply uses the “Publish” menu option to activate any OCHRE item’s **Citation URL**, making it uniquely addressable and accessible on the web. The OCHRE API (application programming interface) provides a variety of methods to retrieve OCHRE data dynamically for display in web-based applications. As part of the project “[An Organon for the Information Age](#),” generously funded by the Neubauer Collegium for Culture and Society, and with the assistance of UChicago master’s student Jie Heng, we are working to publish the OCHRE ontology thereby formally documenting the underlying data structures. As part of the project “[Critical Editions for Digital Analysis and Research](#)” (CEDAR), and in conjunction with our other philology projects, we are working to export OCHRE data into TEI/XML (Text Encoding Initiative) and RDF/XML (Resource Description Framework), both of which are familiar formats to those working in Digital Humanities. These efforts will make OCHRE data available to web developers, to archival finding aids, and to the Semantic Web at large.

## ... TO STOP TO THANK OUR COLLABORATORS AND RESEARCH ASSISTANTS

The work of the OCHRE Data Service would not be possible without the long-term technical support of the [Digital Library Development Center](#), led by Charles Blair at the University of Chicago Library, and we greatly value the services of system administrators Peggy Wilkins and Fred Seaton. We also benefit from the support and services of our colleagues at the [Research Computing Center](#) at the University of Chicago. Our project collaborators all over the world keep us inspired by new questions, new challenges, new data, and new problems to solve. Finally, we enjoy and appreciate the steady stream of hard-working students and assistants who cycle through our office to contribute to the many projects we support. Every day is a new adventure!