# OCHRE DATA SERVICE

MILLER PROSSER

## COLLABORATING FROM A DISTANCE

One of the peculiarities about working in an online, collaborative environment like OCHRE is that when other types of research are put on hold, digital research shifts into overdrive. Our archaeology field projects had amazing, productive seasons excavating in 2019. Looking ahead to the summer of 2020 would require a different strategy. Many of our colleagues have taken this unfortunate moment in history as an opportunity to concentrate efforts on their data. Despite being separated by physical distance and all the attendant risks and challenges of this moment, this past year—and especially the 2020 quarantine period—has been a time of great productivity.

In our previous report, we mentioned the new CEDAR project (funded by the Neubauer Collegium for Culture and Society), which involves researchers in various domain areas (ancient Near East, Divinity School, English) implementing OCHRE as a common text-analysis platform. Over the last year, this project has blossomed to include new partnerships between researchers here in the United States, in the United Kingdom, and in Finland. At the annual Society of Biblical Literature meeting in November 2019, a group of nearly two dozen international researchers agreed to join efforts to work toward the common goal of pursuing digital textual criticism of the Hebrew Bible. Since then, plans have continued to expand. At our annual workshop at the Neubauer Collegium in February 2020, various CEDAR team members presented progress updates and inspiring visions for the future of the project. One prospective expansion involves the inclusion of the Melville Electronic Library in the CEDAR family of projects. Working with Chicago alumnus John L. Bryant, professor emeritus of English at Hofstra University, the CEDAR project hopes to implement for Melville studies the same digital tools as are being used to record and analyze the content of *Gilgamesh*, Shakespeare's *Taming of the Shrew*, and the Hebrew Bible.

Related to our work on text criticism and the CEDAR project, the ODS developed a new visual reconstruction tool to help researchers propose readings where the tablet or manuscript preserves only traces. This new tool uses real examples from the scribe's own handwriting and allows the user to overlay signs on an image to determine possible reconstructions. In the figure below, we compare forms of the Hebrew letter *mem* against an image of a Dead Sea Scroll fragment. By layering real examples from elsewhere in the fragment, the researcher can make a stronger argument for or against a certain reconstruction.[1]

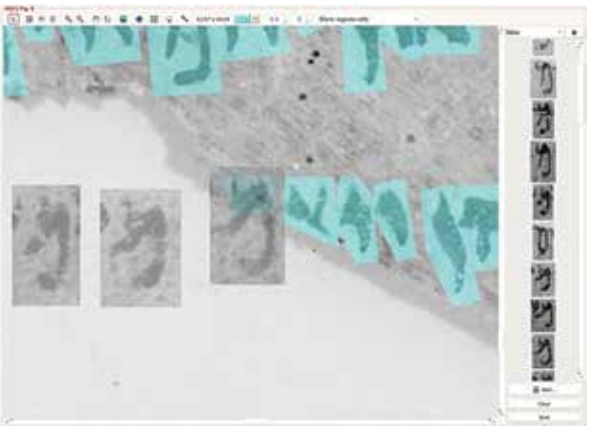

Figure 1. OCHRE text reconstruction tool.

Back when such things were possible, a casual conversation over dinner after a workshop led to a new collaboration between the

OCHRE Data Service and the Computer Science Department at the University of Chicago. When Sandra Schloen and Dr. Sanjay Krishnan struck up this conversation, the OCHRE Data Service had already been collaborating with a highly skilled machine-learning researcher (Eddie Williams) on the first steps toward applying computer vision to reading cuneiform. Together, the newly formed team of Schloen, Krishnan, Prosser, and Susanne Paulus decided to attempt to use machine learning strategies to decipher cuneiform. From these early discussions would emerge the DeepScribe project, an experimental research effort funded by the Center for Data and Computing and the University of Chicago. DeepScribe uses the richly tagged Persepolis Fortification Archive project[2] images as a training set for the machine-learning artificial-intelligence program. The early results are showing great potential for creating a computer program that can read ancient cuneiform. Computers can be trained to recognize modern scripts and even modern handwriting. Can a computer be taught to recognize ancient cuneiform handwriting? And if so, what are the approaches that might make this possible?

The OCHRE database is tailor-made for handling data characterized by features of time and space. As we learned this year, "time" is not limited to the ancient horizon, and "space" is not limited to the Near East. This year we began collaborating with professor Bonnie Clark from the University of Denver on a project investigating the archaeological remains of the World War II internment camp in Amache, Colorado. While we are used to thinking about ceramic vessels and mudbrick walls built by people who speak now-dead languages written in cuneiform, the Amache project has shown that OCHRE is just as useful for reconstructing barracks built of concrete and lumber, and just as helpful for recording data about the spatial distribution of Coca-Cola bottles and Ponds Beauty Cream jars. So far, we have migrated various databases, spreadsheets, photographs, and spatial data files from a variety of sources into the single, unified, and integrated Amache project in OCHRE. We look forward to working with Bonnie, her co–field director April Kamp-Whittaker, and the entire Amache team, once they are able to return to the field for continued research!

This year also found us expanding on the use of OCHRE as a research publication platform. For many years now, OCHRE has been employed as a database for collecting and curating data. It is now time to progress through the data lifecycle to the publication phase. With some recently enhanced features, any research project may now quickly and easily publish all or any part of their data from OCHRE directly to a standard website viewable on a computer or mobile device. With a single click, an entire appendix of data becomes available as a navigable HTML page. This expansion has proven useful for researchers who need a means to publish legacy data or digital projects that are no longer supported on their current platform. Because the OCHRE Data Service works closely

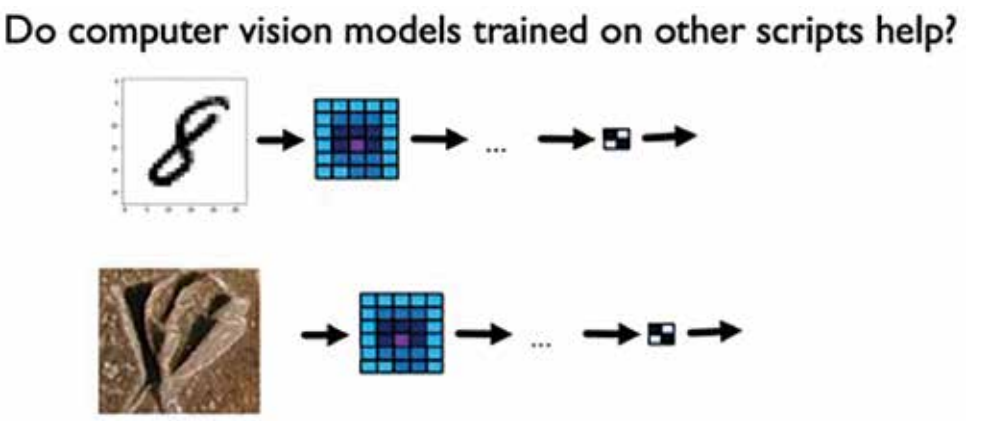


Figure 2. DeepScribe computer vision for cuneiform.

with the Digital Library Development Center at the Regenstein Library, we can offer a long-term home for digital projects. UChicago professor Alan L. Kolata (Bernard E. and Ellen C. Sunny Distinguished Service Professor of Anthropology and of Social Sciences in the College) has published his extensive research on Tiwanaku and Its Hinterland through his OCHRE project. UChicago professor John Lucy (William Benton Professor Emeritus) turned to OCHRE to support his language training website on Mayan Yucatec.[3] Sometimes legacy data is trapped on old media formats like DVD or CD-ROM discs. We discovered this year that the companion disc provided with the printed edition of OIP volume 127 (*Megiddo 3: Final Report on the Stratum VI Excavations*, edited by Timothy P. Harrison) was password protected and that the password had been long forgotten! As an aside, our apologies to those who were unable to access the data on this disc. However, we rescued the data, imported it into the OCHRE database, and published it to the web, now unlocking it for all to access as the Megiddo III Digital Archive. Included in this online presentation are various plates, raw data, and an interactive GIS map.

Figure 3. Megiddo III digital archive, GIS app.

We would be remiss if we failed to mention the continued work of many of our other OCHRE projects. The Persepolis Fortification Archive project continues its work creating text editions, producing digital photographs, and generating an English-Elamite glossary of the archive. The Ras Shamra Tablet Inventory continues adding and curating data on the texts from Ras Shamra-Ugarit. NELC PhD candidate Rhyne King continues editing and analyzing the economics of Late Babylonian Archives. The Chicago-Tübingen Archaeological Project in Sam'al—while not able to excavate as expected this year at the site of Zincirli, Turkey—has launched a many-pronged approach to adding, curating, analyzing, and publishing data already collected.

Looking ahead to the coming year, we can already spy a number of exciting opportunities on the horizon. We are looking forward to working with UChicago linguistics professor Alan C. Yu on his research on the Washo language and UChicago Art History professor Niall Atkinson on his study of the Florentine Catasto of 1427.

The OCHRE Data Service is managed by Sandra R. Schloen and consists of Dr. Miller C. Prosser (database consultant), Dr. David Schloen (professor of Near Eastern archaeology), and Charles Blair (University of Chicago, DLDC).

Visit us at ochre.uchicago.edu.

## *Endnotes*

[1] See more on the reconstruction tool on the OCHRE Wiki: sites.google.com/view/ochrewiki/projects/cedar/reconstruction-tool.

[2] See the Annual Report on the Persepolis Fortification Archive project by professor Matthew Stolper.

[3] We wish to acknowledge the work of John Jung in the DLDC for the development of the LUCY website using OCHRE data.