

9. Data Management Plan

Data Sharing

For each research project, typically an excavation at a particular site, DRC collaborators enter context and artifact data into the DRC Application and its PostgreSQL backend. The primary means by which data are shared with scholars and the general public is the DAACS website (www.daacs.org), which is built in WordPress. The Principle Investigator (PI) for each project or site works with DAACS staff to create a “home page” for each project and link to it discursive background historical information, archaeological chronologies, stratigraphic summaries, and overall digital site plans, sections, and photographs (*e.g.*, <https://www.daacs.org/sites/sugarloaf/#home>). After final checks are complete and the site’s PI has signed off, data for the project in the PostgreSQL backend are made available to the query module on the DAACS website (<https://www.daacs.org/query-the-database/>).

The query module gives users access to a point-and click interface that generates custom SQL queries against the PostgreSQL backend. Users can query on chosen sites and artifact classes and chosen variables that describe them and have the data returned at different levels of aggregation, ranging from individual excavation contexts to entire sites.

Based on user choices, the query module can deliver fine-grained data on artifacts and their excavation contexts from related tables, in which each record represents a different attribute value on an artifact, and single artifacts, identified by a unique DAACS artifact ID number, span multiple records. This is useful, for example, in the study of correlations among decorative motifs on ceramics. The query module can also deliver coarse-grained data as well. For example, site-level assemblage summaries, in which each record represents a broad artifact class (*e.g.*, cut nails and its total count). The new “DRC Silver” and “DRC Bronze” cataloging modes will require catalogers to record values for smaller subsets of the complete set of “DRC Gold” fields. The DRC Advisory Committee will help us identify the fields that will be required for each mode. Finer grained data will not be available for sites catalogued in Silver and Bronze modes. But the benefit of the tradeoff is the ability to catalog larger assemblages on tighter budgets.

Data Types

Below is a summary of the kinds of data types cataloged into, and managed, by the DRC Application and the DAACS website.

Archaeological Field Records: DRC-certified archaeologists enter all information found on the original archaeological field records, including sediment descriptions for layers, and the stratigraphic relationships among those layers, into the DAACS database’s context tables and related tables. This process creates easy to search records that are linked to each artifact from that context. DRC partners link photos and scanned paper field records to their corresponding context records. All context data are available for download in .html format through the *Query the Database* section of the DAACS website.

Artifact Records: The structure of the DAACS database allows for recording of detailed information about individual artifacts. DAACS classification and measurement protocols are described in [online manuals](#). Detailed guidelines in written form ensure consistency among catalogers and provide researchers with an opportunity to understand how the data they seek

to use were generated. DRC catalogers and analysts undergo specialized training and testing in these protocols to ensure that they are implemented accurately across all collections. All artifact data are available for download in .html format through the *Query the Database* section of the DAACS website (<https://www.daacs.org/query-the-database/>).

Artifact Images: All unique, illustrative, or diagnostic artifacts and all artifacts exhibiting any sort of post-manufacture modification are digitally imaged. A set of DAACS cataloging protocols specifies how these artifacts should be imaged, named, and stored. Depending on size, select artifacts are either scanned or photographed using a professional camera setup. Images are uploaded and image records are linked to the appropriate artifact and/or object record. They are served to the public through the Query-the-Database section of the DAACS website.

Context and Other Excavation Images: Existing photographs and slides are scanned at 300 dpi, saved as archival .tiffs and .jpegs, and are linked to the appropriate project and context records in the database. They are served to the public through the Query-the-Database section of the DAACS website.

Site Maps, Excavation Plans and Profiles: Hand-drawn site maps and excavation plans and profiles are scanned following established DAACS protocols at 300 dpi and saved as archival .tiffs and .jpegs. Plan and sections of individual contexts are linked in the DAACS database to the appropriate context records. Composite site plans are then digitized in vector format using a CAD program from the mosaics of the scans, following standard DAACS protocols to ensure consistency in the depiction of particular subsurface features, above-ground, extant architectural elements, and excavation areas across all sites launched online. Vector plans are saved for delivery on the DAACS website in .dgn (Microstation) and .dxf (AutoCad) formats. Site plans in .pdf (Adobe) also are made available online for download.

Metadata Resources: DAACS maintains detailed metadata on the DAACS data structures and the process for creating the archaeological data in DAACS, such as cataloging manuals, and tutorials on how to use the data. These are shared on the DAACS website (<https://www.daacs.org/about-the-database/daacs-cataloging-manual/>)

Source Code

As detailed in the Final Product and Dissemination section of the Narrative, all source code for the DRC interfaces and Open API will be documented and made freely available for download, reuse, and modification in a public repository at GitHub.

Storage, Maintenance, and Protection of Data

Currently, the DAACS database backend and all linked artifact and context images are housed on a dedicated PostgreSQL server located at the University of Virginia's Institute for Advanced Technology in the Humanities. The entire server is backed up daily on two Quantum VS1 DLT tapes. The DAACS PostgreSQL database is not directly connected to the Web. The DRC Application website (www.daacsrc.org) and PostgreSQL database are tarred and gzipped to another server not accessible to the public after every update. This process adds a third layer of backup and allows previous versions of the website to be retained for historical purposes. Finally, the DAACS PostgreSQL database is also manually backed up prior to any database maintenance or updates. This

process of data storage and maintenance has worked seamlessly since 2014. We will continue using this process for all software and data produced as a result of the *Expanding DRC* project. The DAACS website is built in WordPress, an open-source content management system. It presents, in HTML, discursive sites summaries and the items listed above.