PAC & MAC • Iowa City, Iowa

From Segregated to Integrated: Rethinking Big Data with GRSLE

Charles Orngard[a,b], Lawrence C. Todd[b], Daniel Dalmas[c,b], David Rapson[d]

[a]Iowa State University; [b]GRSLE Inc.; [c]University of Utah; [d]University of Wyoming

Archaeology increasingly engages with "big data," yet most datasets remain fragmented, site-based, artifact-class specific, or built from incompatible systems. The GRSLE Project offers an alternative: a "little big data" approach developed through continuous high-elevation fieldwork in northwest Wyoming since 2002. Rather than segregated artifact classes, GRSLE employs a unified coding system that records every object (lithics, faunal remains, features, historic materials, and modern traces) at the individual level. Over 250,000 artifacts have been cataloged, producing scale without losing context and enabling interpretations that move beyond traditional site boundaries. The system balances detail with accessibility: refined enough to capture variability yet simple enough for new students to apply consistently. This poster introduces GRSLE's integrated big data framework and evaluates its broader utility. We argue that cumulative, context-rich datasets open avenues of research not possible in fragmented ones, and that integration, not aggregation, is what makes Big Data transformative for archaeology.

Figure 1. Examples of GRSLE's multifaceted accumulation of data: (a) on-site, catch and release chipped stone recording, (b) systematic pedestrian transect surveys, (c) faunal excavation.

I. The Problem: Fragmented Archaeological Data

Archaeology is increasingly engaging with "big data," but too often it remains fragmented. Simply aggregating large swaths of information, isolated by site, class, or incompatible database design, offers fuzzy insight in making inferences about past human behavior.

When datasets are built from systems that cannot communicate with each other, their potential for large-scale insight is lost. Information stays locked in silos, and connections across materials, contexts, and time are obscured (Figure 2).

Aggregation ≠ Integration

Simply merging datasets doesn't take full advantage of different types of relationships; only integrated systems preserve the context and comparability that make large archaeological datasets powerful.

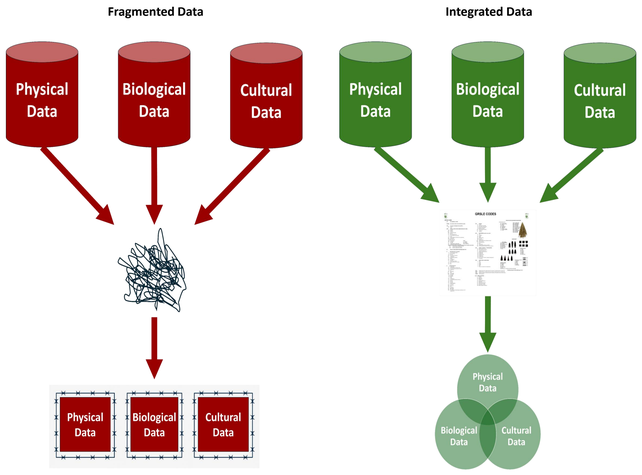

Figure 2. Fragmented data (left) versus integrated data through the GRSLE system (right).

II. The GRSLE Approach: "Little Big Data"

The GRSLE Project takes a distinct approach to big data in archaeology. Rather than aggregating disconnected datasets, GRSLE builds scale through integration at the point of collection. Every object from lithics, faunal remains, historic materials, and even modern traces is recorded within a single unified coding system (Figure 4).

This approach has been refined through continuous high-elevation fieldwork in northwest Wyoming since 2002. Over 250,000 artifacts have been individually catalogued, resulting in a dataset that continues to grow cumulatively while maintaining context and comparability.

The result is a form of "little big data", large in scope but grounded in detail. The system is refined enough to capture variability, yet simple enough for new students to apply consistently, striking a balance between accessibility and analytical depth.

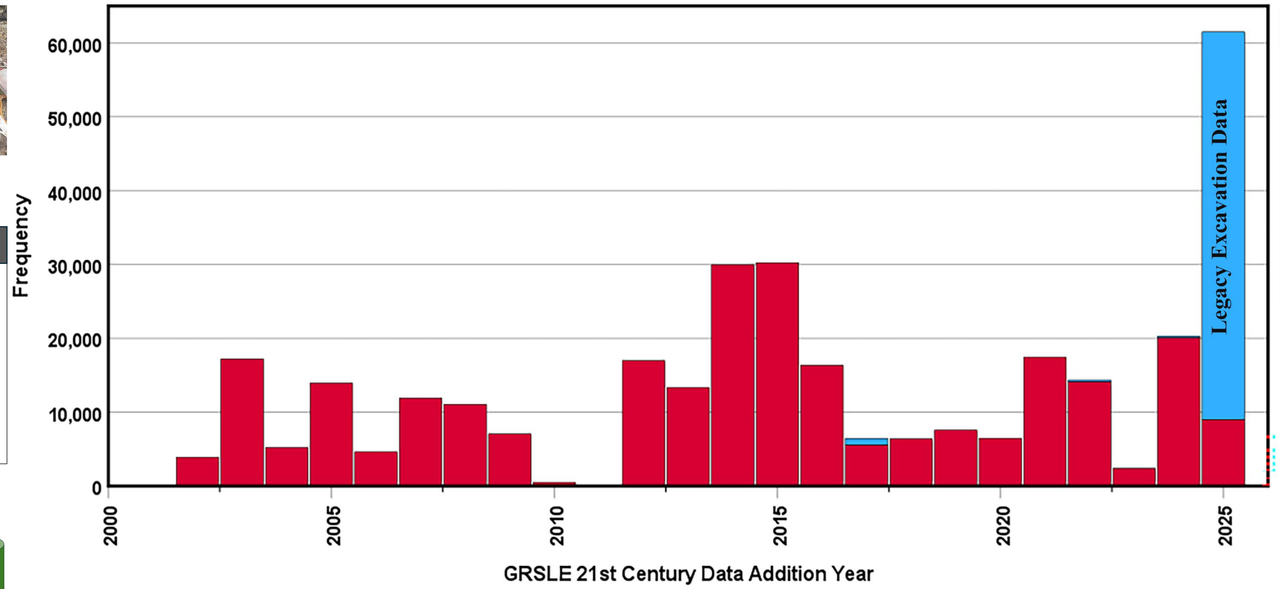

Figure 5. Growth of the GRSLE dataset over time, including integration of legacy excavation data in 2024.

III. The GRSLE System

GRSLE's framework is a shared attribute structure that applies to every recorded object. Regardless of class, each data entry is documented using the same hierarchy of provenience, material, form, portion, and contextual data.

This unified structure ensures that information is both comparable and cumulative. A chipped stone flake, bone fragment, or glass shard can be analyzed together as they all follow the same descriptive logic, which links spatial, technological, and environmental variables without translation between datasets.

Rather than dividing data into parallel databases (Figure 2, left), GRSLE treats every entry as part of a single, interconnected record of human activity (Figure 2, right). This relational design transforms individual observations into a network of interpretable relationships, bridging artifact types, context, and temporal boundaries within a single analytical framework.

Figure 3. Examples of different types of data recorded in the GRSLE database: (a) chipped stone (e.g, projectile point), (b) modern refuse (e.g., bullet cartridge), (c) historic artifacts (e.g., barbed wire), (d) recreation features (e.g, recreational hearth), (e) local flora and fauna (e.g., ground squirrel), (f) archaeological features (e.g., prehistoric cairn).

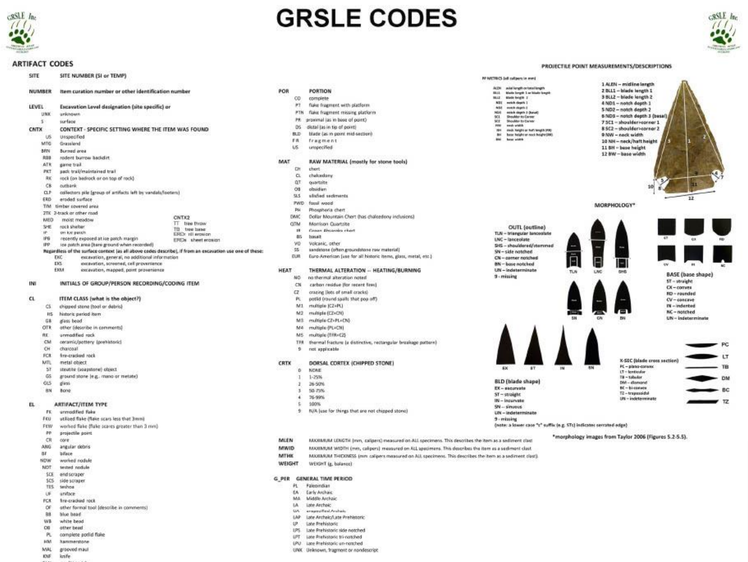

Figure 4. Code sheet of standardized GRSLE codes used both on-site and post-field.

IV. Bugas-Holding: The GRSLE Approach's Foundation

The incorporation of a site in the Absaroka Mountains excavated in the mid-1980s (Bugas-Holding, 48PA563) demonstrates the adaptability of the GRSLE framework. In 2025, over 51,000 records from the Bugas-Holding site were translated into the existing dataset using GRSLE's standardized attribute logic (Figure 5).

This process shows that GRSLE's structure is not limited to new field collections; it can also restructure and harmonize preexisting datasets. Data originally recorded under different conventions was systematically aligned with current attributes, allowing it to join the broader analytical network without loss of contextual integrity.

A key example of this adaptability lies in artifact numbering. While many projects restart numbering at each site or unit, GRSLE assigns a unique identifier to every recorded object, tied to its spatial unit (e.g., G1), year (e.g., 25), and sequential position. This system, adapted from practices at Bugas-Holding, ensures that every item, past or present, can be distinctly referenced within a single, cumulative dataset.

With now more than 300,000 linked records, the result demonstrates how a consistent recording logic can be applied across projects and generations, transforming legacy data into an active component of an evolving, regional-scale dataset.

Figure 6. Bugas-Holding excavation area in Sunlight Basin, northwest Wyoming.

Figure 7. Excavation and documentation at the Bugas-Holding site, illustrating detailed provenience recording and artifact mapping practices foundational to GRSLE's integrated framework.

V. Broader Implications

The GRSLE framework demonstrates that archaeological "big data" becomes transformative only when it is integrated through shared structure, not simply aggregated. By linking data across materials, sites, and time, the system enables questions that exceed the limits of traditional, site-based analysis.

Its design is replicable and scalable, offering a model for how regional and legacy datasets can be made interoperable without sacrificing contextual detail. The result is a cumulative, evolving record of human activity that is ever-growing without compromising consistency or detail.

Integration at this scale creates new opportunities for comparative research, long-term landscape analysis, and collaborative synthesis, providing a pathway toward truly interconnected archaeological data.

Fieldwork for this project was conducted as part of the Greybull River Sustainable Landscape Ecology (GRSLE) initiative. Special thanks to Dr. Lawrence C. Todd and Daniel Dalmas for their guidance and collaboration, and to Dr. David Rapson for his excellent work integrating the legacy Bugas-Holding data into the GRSLE database. Gratitude is also extended to the Shoshone National Forest U.S. Forest Service for access and support in the Shoshone National Forest.

Charles Orngard
Iowa State University
Anthropology | World Languages and Cultures
orngard@iastate.edu