Paleo Core

A Platform for data integration in paleoanthropology

Denné Reed
Dept. of Anthropology University of Texas at Austin

TEXAS
The University of Texas at Austin

NSF

TACC
TEXAS ADVANCED COMPUTING CENTER

JET BRAINS

DbVisualizer
The Universal Database Tool

Overview

https://paleocore.org

- **Background** - What is Paleo Core?

- **Orientation** - Where does Paleo Core fit in the data management ecosystem?

- **Roadmap** - Where is it heading in the near future?

Data Deluge

Every year paleobiologist, archaeologists and geologists excavate thousands of artifacts and generate large amounts of physical and digital data.

In the past 35 years:
1,000,000 field studies
1,000,000,000 items curated

Data Silos

Independent collection
results in data silos

1. **Data Integrity and Preservation**
   Data are vulnerable to loss,
   corruption and obsolescence.

2. **Data Accessibility and Mobility**
   Data are difficult to discover,
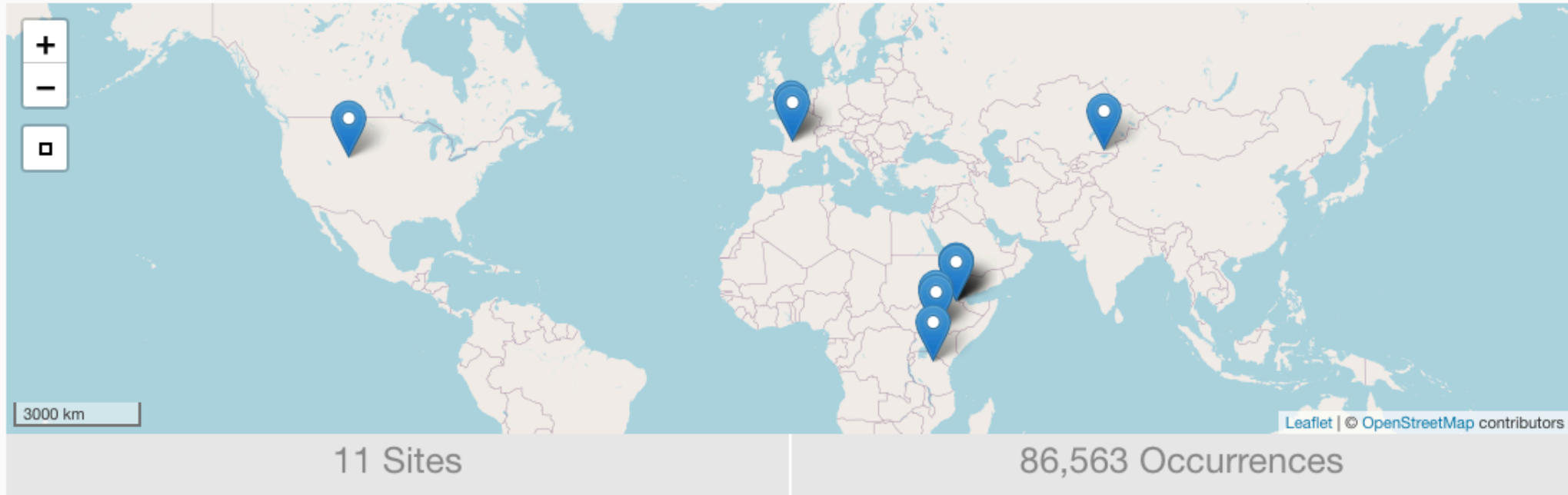   access and integrate with other
   data.

Better Data

New information does not necessarily bring clarity or insight.

Addressing challenging questions requires systematic integration of diverse data from multiple sources.

Paleo Core Projects

11 Sites	86,563 Occurrences

Combe-Capelle Bas (cc)

A Pleistocene archaeological site in France.

🔓 Private   19064 records

Great Divide Basin (gdb)

A Paleogene site in central North America.

🔓 Private   12516 records

Omo-Mursi Research Project (omo_mursi)

omo-mursi

🔓 Private   709 records

Dikika Research Project (drp)

A Pliocene Site in Ethiopia

🔓 Private   1979 records

Hadar Research Project (hrp)

A Plio-Pleistocene site in the Afar region of northeaster Ethiopia.

🔓 Private   9969 records

Paleo Silk Road (psr)

A survey of Pleistocene sites in central Asia.

🔓 Private

Eyasi Plateau Paleontology Project (eppe)

Laetoli, a Plio-Pleistocene site in northern Tanzania.

🔓 Public   13716 records

Ledi-Geraru Research Project (lgrp)

A Plio-Pleistocene site in the Afar region of northeaster Ethiopia.

🔓 Private   5241 records

West Turkana Archaeology Project (wtap)

A Plio-Pleistocene project in West Turkana, Kenya.

🔓 Private

Fontechevade (fc)

A Pleistocene archaeological site in France.

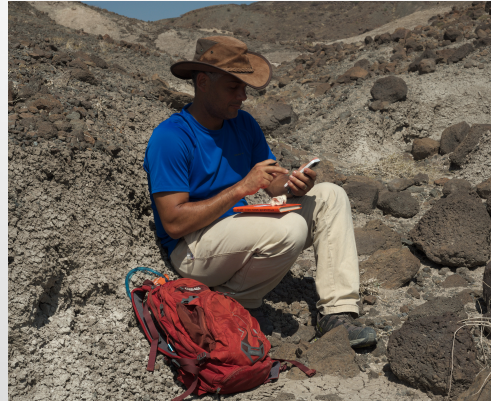🔓 Private   20329 records

Mille-Logya Project (mlp)

A Plio-Pleistocene site in the Afar region of northeaster Ethiopia.

🔓 Private   3040 records

Paleo Core

Digital Data Collection Tools

Reed D et al. (2015) Digital Data Collection in Paleoanthropology.

Paleo Core

Digital Data
Collection Tools

Standards
Best Practices

Collaborative Data
Management Platform

Reed D et al. (2015) Digital Data Collection in
Paleoanthropology.

Reed D, et al. (2018) PaleoCore: an open-source
platform for geospatial data integration in
paleoanthropology.

- Paleo Core is geared towards **Paleoanthropology** broadly defined, including paleobiology, archaeology and geology.

- Paleo Core is focused on efficiently procuring and managing **new data** collected by researchers rather than legacy data already in museums.

- Paleo Core emphasizes free, open-source geospatial solutions (**FOSS4G**) in a Spatial/Scientific Data Infrastructure (**SDI**).

- Paleo Core emphasizes **data standards** and best practices to promote **integration** across projects.

- Paleo Core emphasizes ontologies and **linked open data** (LOD) to promote knowledge integration with the semantic web.

Mobile Devices

CentOS Linux

Remote Clients

Django
Web Application Server

PaleoCoreR
API for R Statistics

FOSS4G SDI

GeoServer
Web Feature/Map Server

PostgreSQL
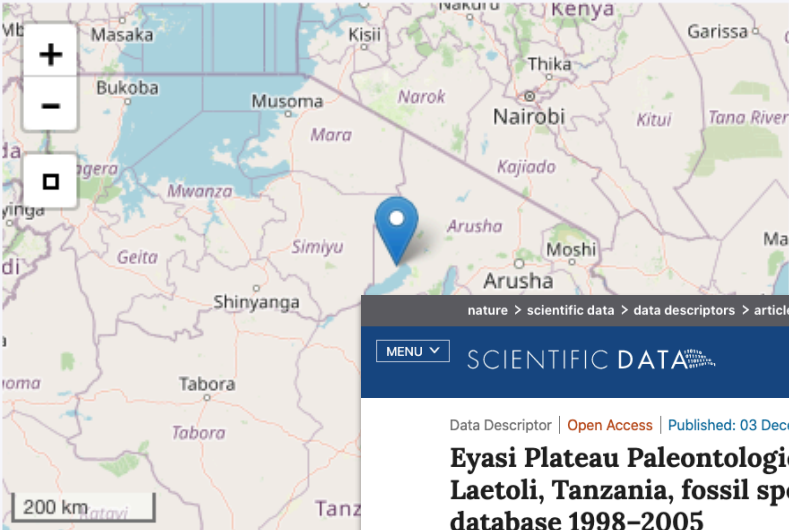PostGIS

Paleo Core
Spatial
Database

django

Orientation

Eyasi Plateau Paleontological Expedition (EPPE)

+
−
□

Masaka Kisii Kenya Garissa
Bukoba Musoma Narok Nairobi Kitui Tana River
Mwanza Mara Kajiado
Geita Simiyu Arusha Moshi Ma
Shinyanga Arusha
Tabora
Tabora Tanz

200 km

| Description | Datasets |

The EPPE Laeotoli specimen database contains 13713 records of plant an
specimens) collected by EPPE field teams working at Laetoli, Tanzania. Th
of the original hard-copy specimen catalog and documents the discovery,
taxonomic diversity of Pliocene and Pleistocene fauna and flora in norther
between 3.7 Ma to 200 ka. The collections include important lectotypes fo
the first record of *Paranthropus aethiopicus* outside Kenya and early recor
sapiens. The taxonomic identifications appearing in this dataset are the or
are provisional. Any taxonomic analysis employing this dataset should refe
published by specialists in the appropriate taxonomic domain (e.g. Harriso

Eyasi Plateau Paleontological Expedition, Laetoli, Tanzania, fossil specimen database 1998–2005

Denné Reed ✉, Terry Harrison & Amandus Kwekason

Abstract

The Eyasi Plateau Paleontological Expedition (EPPE) Laetoli specimen database contains 13716 records of plant and animal fossils (ca. 28248 specimens) collected by EPPE field teams working at Laetoli, Tanzania between 1998 and 2005. This dataset is a digital version of the original hard-copy specimen catalog, and it documents the discovery, stratigraphic provenience and taxonomic diversity of Plio-Pleistocene fauna and flora in northern Tanzania between 4.4 Ma and >200 ka. Laetoli is renowned for the discovery of important hominin fossils, including the lectotype for *Australopithecus afarensis*, one of our early hominin ancestors, the first record of *Paranthropus aethiopicus* outside Kenya-Ethiopia, and an early record of our own species *Homo sapiens*. This database is one of the few publicly available palaeoanthropological fossil datasets and serves as an example for expanding open access to primary fossil occurrence data in palaeoanthropology. The taxonomic identifications appearing in this dataset are the original field identifications and are provisional. Any taxonomic analysis employing this dataset should refer to updated taxonomic identifications published by specialists.
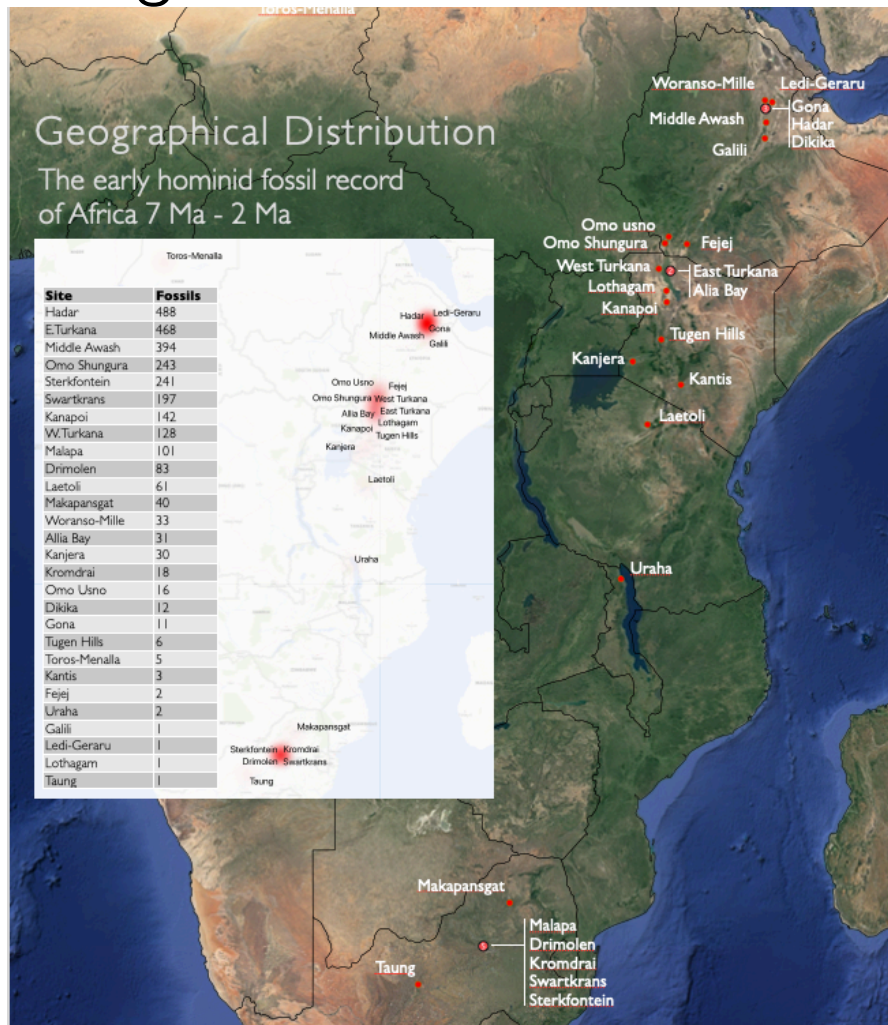
Future

African Rift Valley Research Consortium

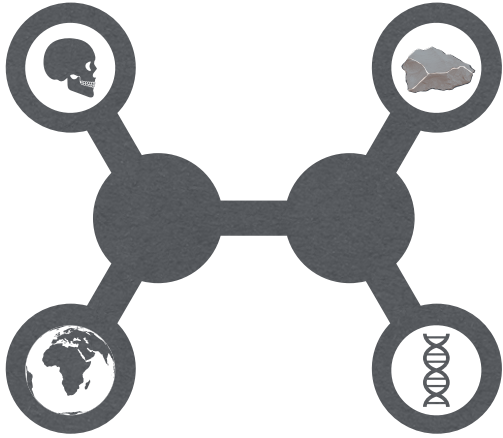# Summary

Insights

Collaboration

Access

Paleo Core Workshop Nairobi

Acknowledgements

Contributors and Developers

Tomislav Urba, TACC

John Kappelman, U Texas, Austin

W Andrew Barr, George Washington U.

Shannon McPherron, Max Planck

References

Reed DN, Barr WA, McPherron S, Bobe R, Geraads D, Wynn JG, Alemseged Z. (2015) Digital Data Collection in Paleoanthropology. Evolutionary Anthropology. 24:238-249.

Reed DN, Barr WA, Kappelman J. (2018) PaleoCore: an open-source platform for geospatial data integration in paleoanthropology. In: Robert Anemone and Glenn Conroy (eds.) "New Geospatial Approaches in Anthropology." University of New Mexico Press.

Reed D, Harrison T, Kwekason A. (2019) Eyasi Plateau Paleontological Expedition, Laetoli, Tanzania, fossil specimen database 1998-2005. *Scientific Data* **6**(304): 1-11. Doi: 10.1038/s41597-019-0304-2.