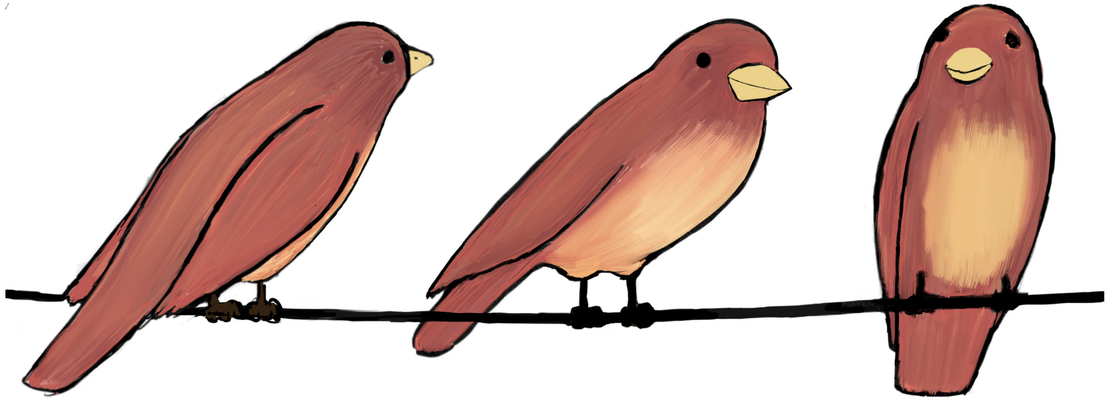
Sparrow: a cyberinfrastructure component supporting data management and sharing in geochronology laboratories

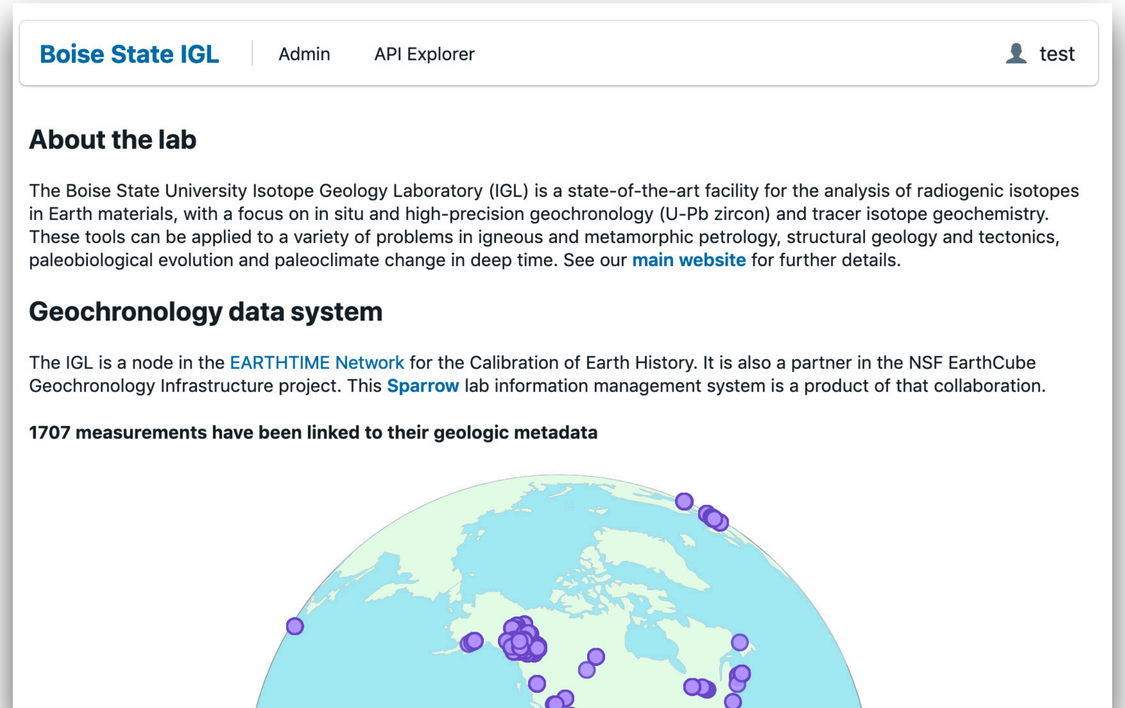
Daven P. Quinn[1], A.J. Schaen[1], K. Sundell[2], G. Gehrels[2], S.E. Peters[1], J.I. Ross[3], M. Schmitz[4], and B.S. Singer[1]

1. *University of Wisconsin – Madison*
2. *University of Arizona*
3. *New Mexico Bureau of Geology & Mineral Resources*
4. *Boise State University*



Goldschmidt 2019, Barcelona, España

https://github.com/EarthCubeGeochron/Sparrow
https://sparrow-data.org



Supported by
NSF EarthCube

Reconstructions of geologic time underpin synthesis studies across Earth science domains

always ultimately driven by detailed geochronology data

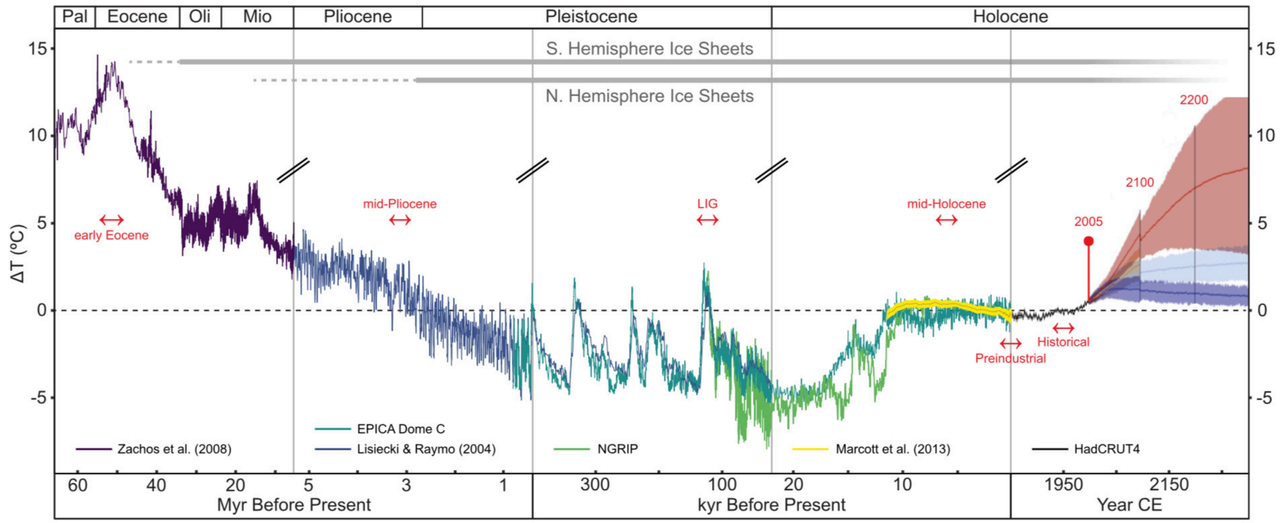
PNAS

Pliocene and Eocene provide best analogs for near-future climates

K. D. Burke[a,1], J. W. Williams[b], M. A. Chandler[c,d], A. M. Haywood[e], D. J. Lunt[f], and B. L. Otto-Bliesner[g]

[a]Nelson Institute for Environmental Studies, University of Wisconsin–Madison, Madison, WI 53706; [b]Department of Geography and Center for Climatic Research, University of Wisconsin–Madison, Madison, WI 53706; [c]Center for Climate Systems Research, Columbia University, New York, NY 10025; [d]Goddard Institute for Space Studies, National Aeronautics and Space Administration (NASA), New York, NY 10025; [e]School of Earth and Environment, University of Leeds, LS2 9JT Leeds, United Kingdom; [f]School of Geographical Sciences, University of Bristol, BS8 1SS Bristol, United Kingdom; and [g]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO 80305

Edited by Noah S. Diffenbaugh, Stanford University, Stanford, CA, and accepted by Editorial Board Member Robert E. Dickinson November 6, 2018 (received for review June 29, 2018)

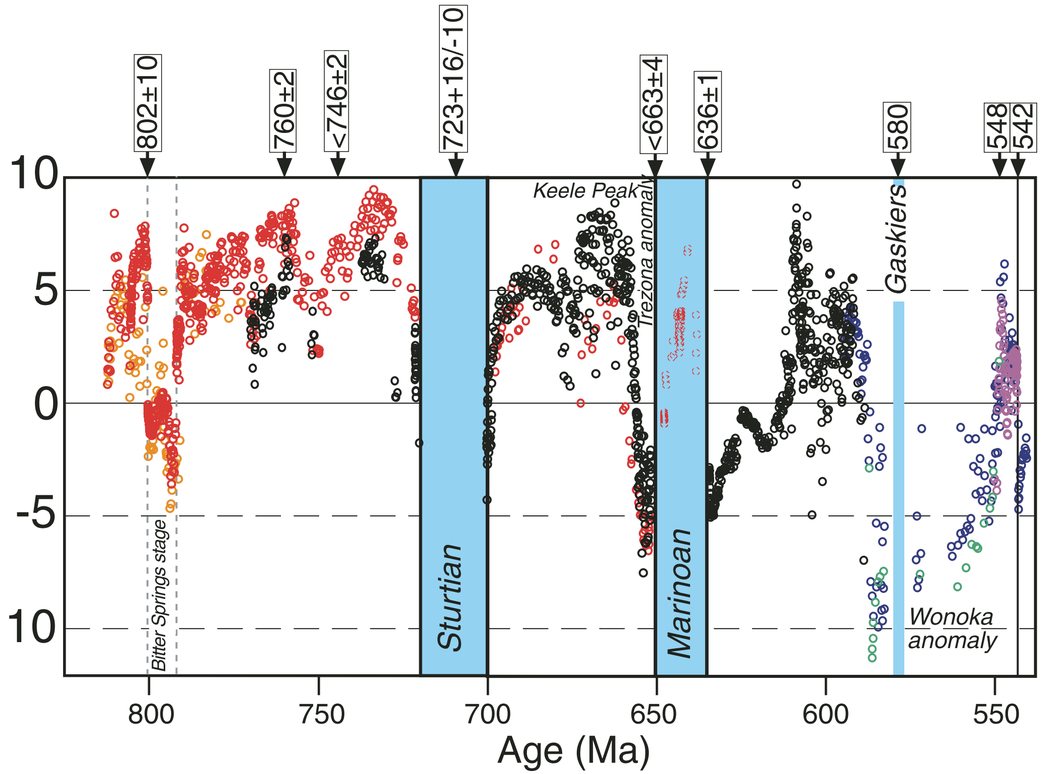


Toward a Neoproterozoic composite carbon-isotope record

Galen P. Halverson[†]
Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, Massachusetts 0213 and Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Building 54 bridge, Massachusetts 02139, USA

Paul F. Hoffman
Daniel P. Schrag
Adam C. Maloof[‡]
Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, Massachusetts 02138

A. Hugh N. Rice
Department of Geological Sciences, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria



Interoperable data for geochronology

A long-recognized need







A lot of work!

How can labs take on this responsibility while maintaining their focus on geochronology?

Benefits

Reproducibility and intercalibration

- Re-standardization and calibration
- Archival for re-analysis
- Audience: Expert workers within the geochronology community
- Data requirements: Detailed analytical data

Aggregation and synthesis

- Discovery of relevant geochronology measurements
- Integration with other types of data (stratigraphic, ecological, etc.)
- Audience: scientists in other domains
- Data requirements: Interpreted ages, simplified uncertainties

Geochronology data workflow
and standardization challenges

Rich contextual metadata

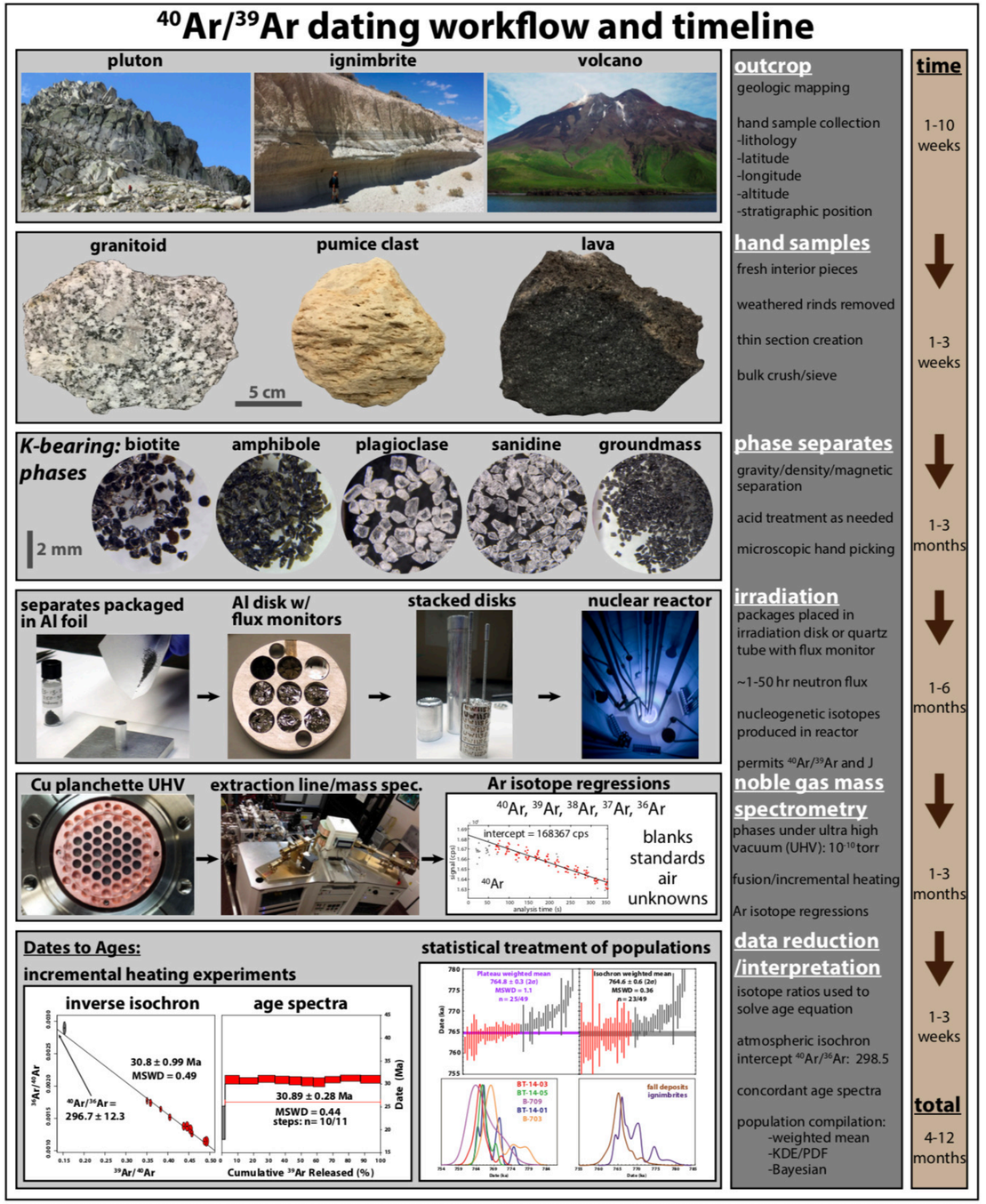
often not tracked systematically

Analytical process data

differently structured between isotope systems, analytical techniques, *and* instruments

Interpreted age data

encodes expert intervention that can be challenging to justify and reproduce



Building a standardized lab data pipeline

Analytical processes differ among geochronology labs

Less diverse *metadata*

- Sample information
- Geological context
- Project/publication metadata

Similar archival challenges

- Heterogeneous data files
- Poorly structured legacy data
- Manual recalibration
- Not a good use of lab/user time

EarthCube Geochronology Frontiers consortium

WiscAr

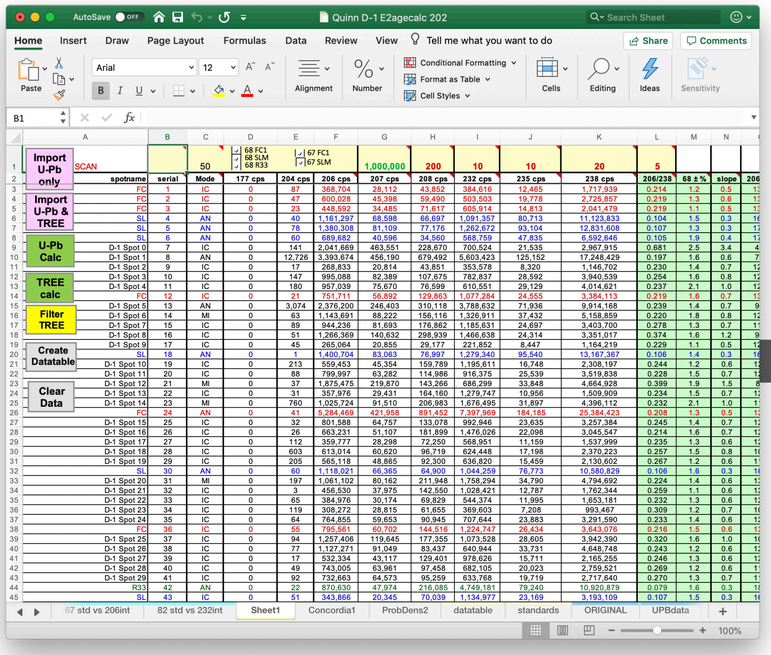
ARIZONA LASERCHRON CENTER
Department of Geosciences
University of Arizona

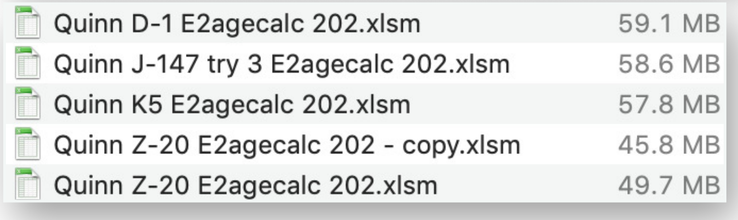
New Mexico Geochronology Research Laboratory

ISOTOPE GEOLOGY LABORATORY
BOISE STATE UNIVERSITY

Goal

Analytical data archives

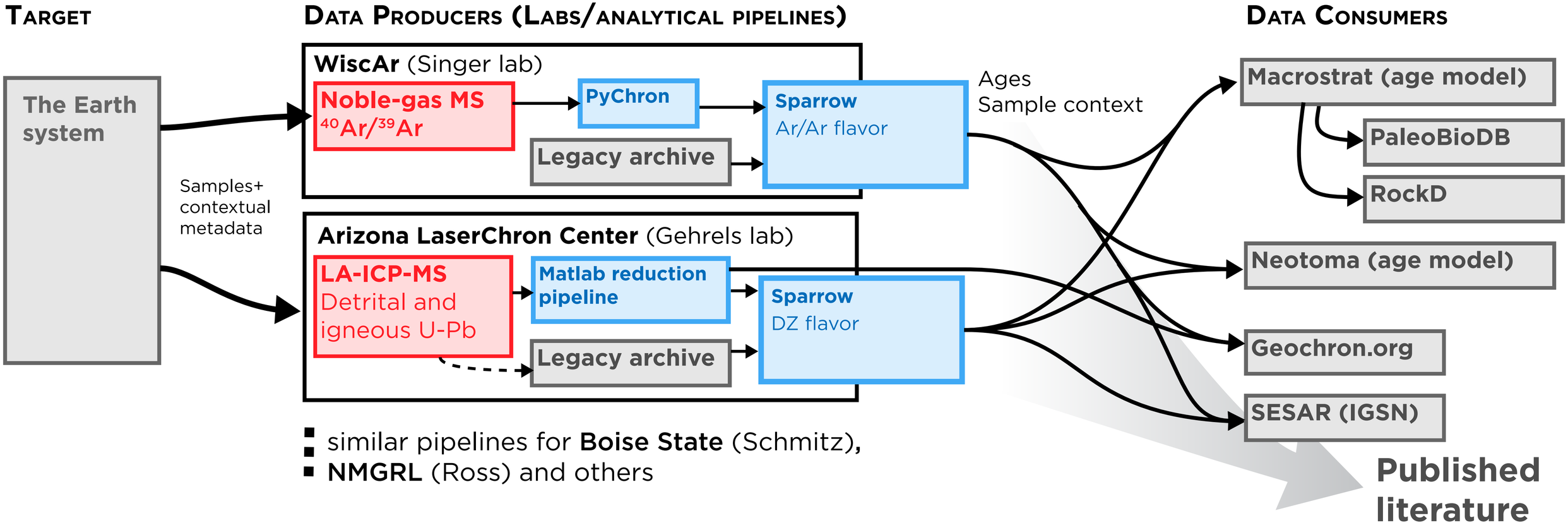




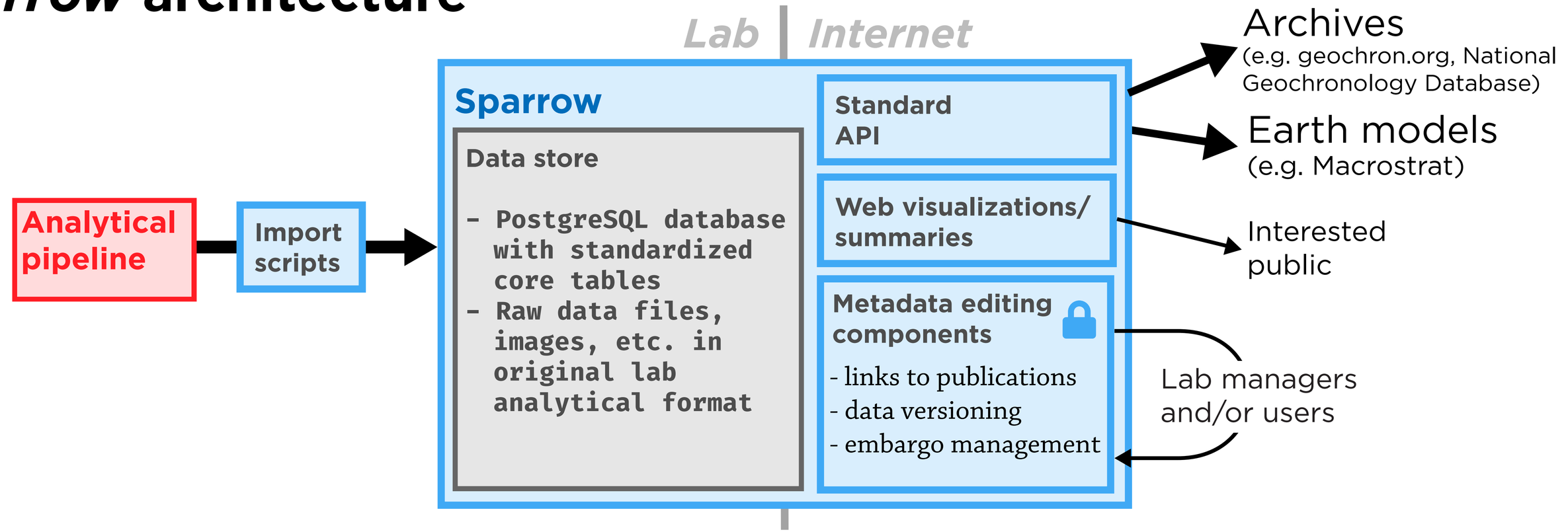
- Machine-readable
- Community-accessible
- Contextualized
- Ready for synthesis

Sparrow: a software interface between a lab and the world

- Each lab maintains control of its own internal process
- Harmonized endpoint enables easy data sharing



Sparrow architecture



Structured data store

- Analytical data in original file format
- Geologic context and project metadata
- Structured subset of the data for external use

Interface to lab data

- Standardized API providing access to the lab's analytical data
- Data-driven public-facing website for the lab
- Web application to manage metadata (e.g. data embargo, sample/geological info, projects/researchers, links to publications)

Sparrow design principles

Adds to (rather than replacing) current lab processes

Does not include or dictate analytical pipeline

Extensible

- A simple, hierarchical data model
- Usable across domains of geochemistry
- Customizable at every level (database, API, frontend)

Modular

- Based on standard, widely used tools
- Decoupled API, frontend, and database
- Importers and exporters interface with other lab systems

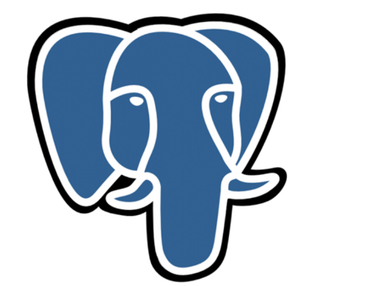
Lightweight

- Small footprint at implementing labs
- Abstracts the complexity of underlying systems (e.g. Docker, PostgreSQL)

Technology stack

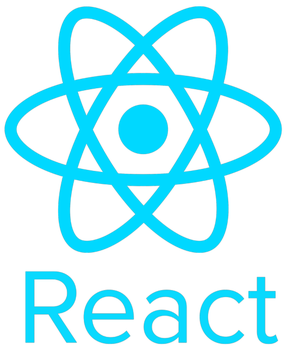


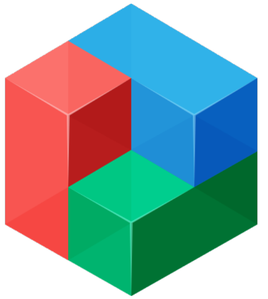
Backend





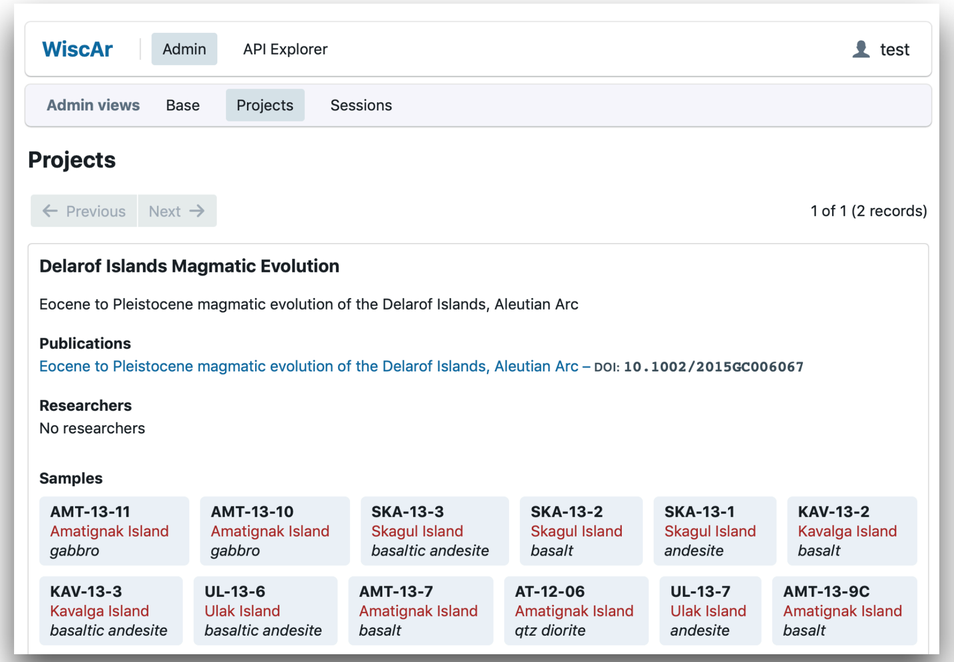
Frontend

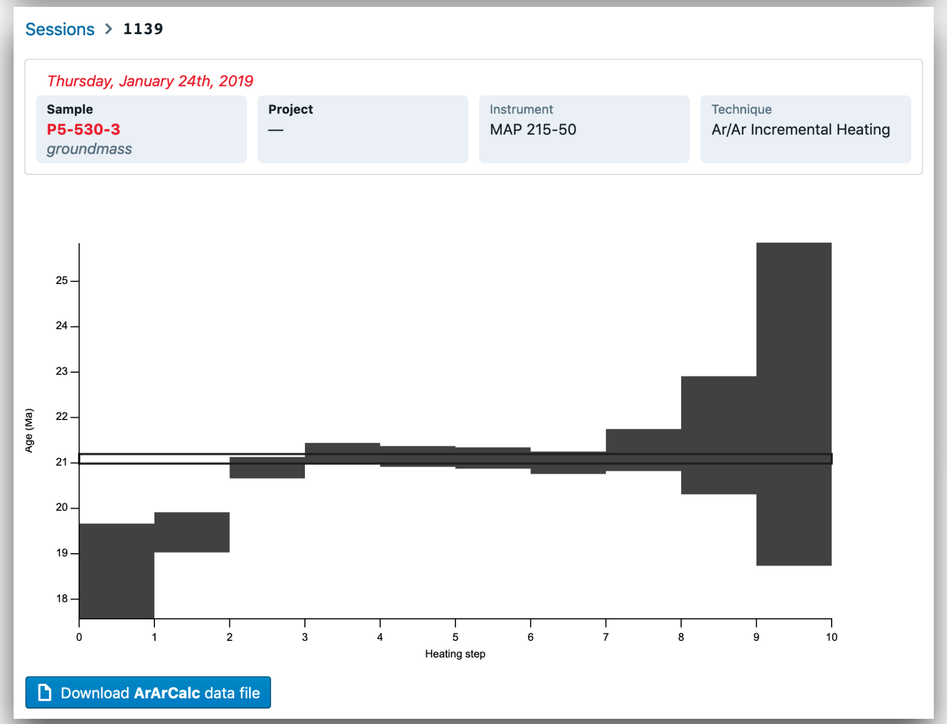


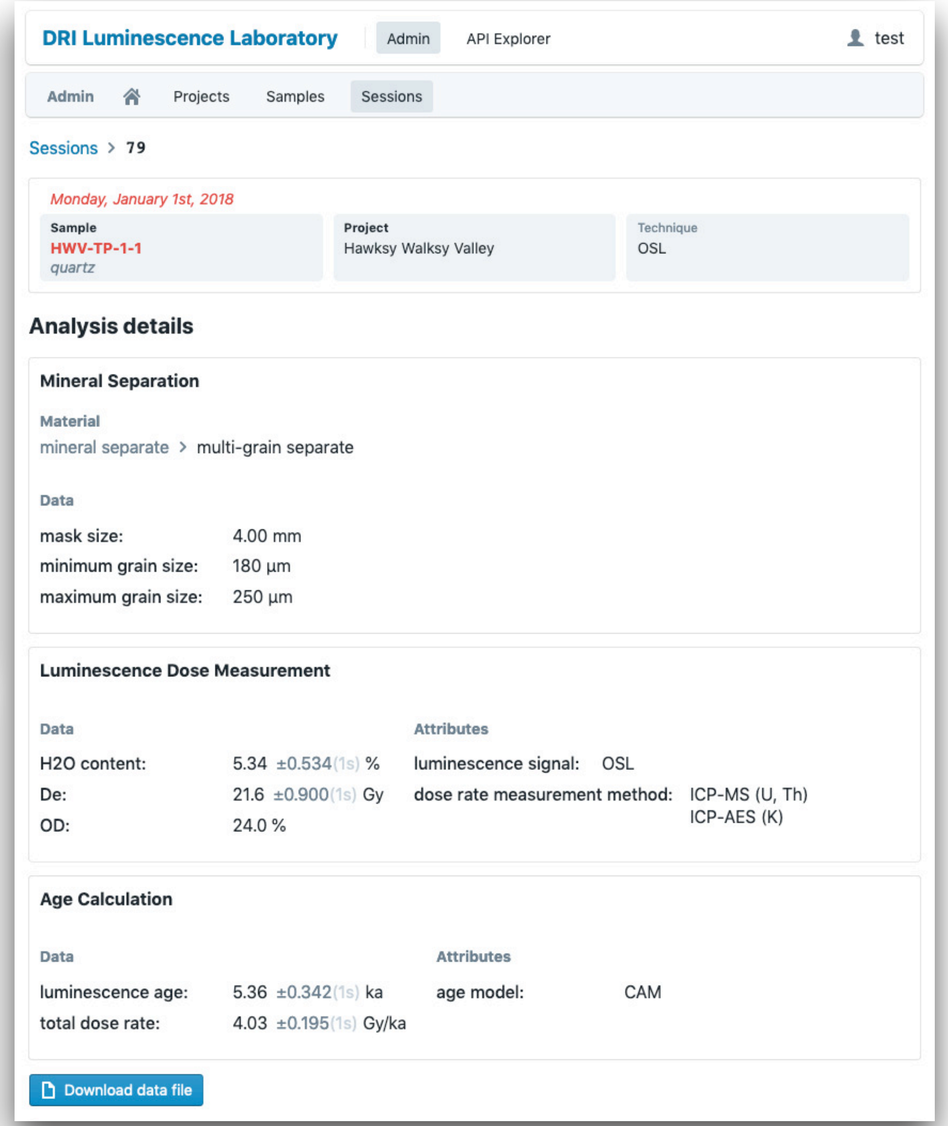


Web-based management frontend

- Access detailed analytical data and original data files
- Import and export data files
- **Manage project- and sample-level metadata**
- Link data to publications
- Request and link IGSNs *(optional; planned)*





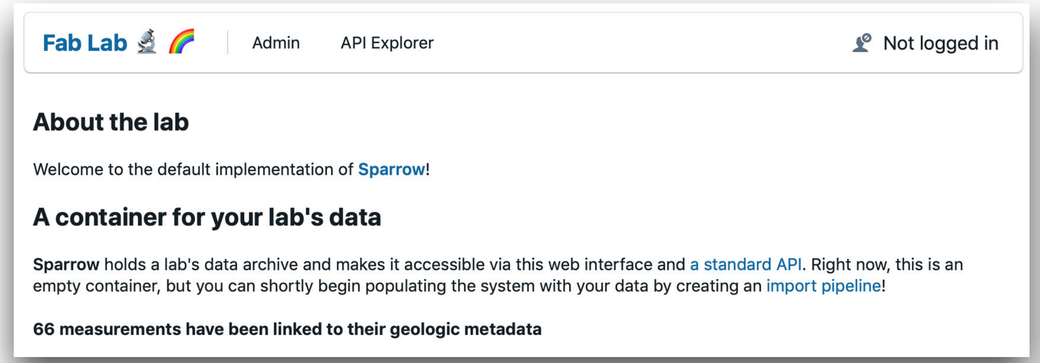


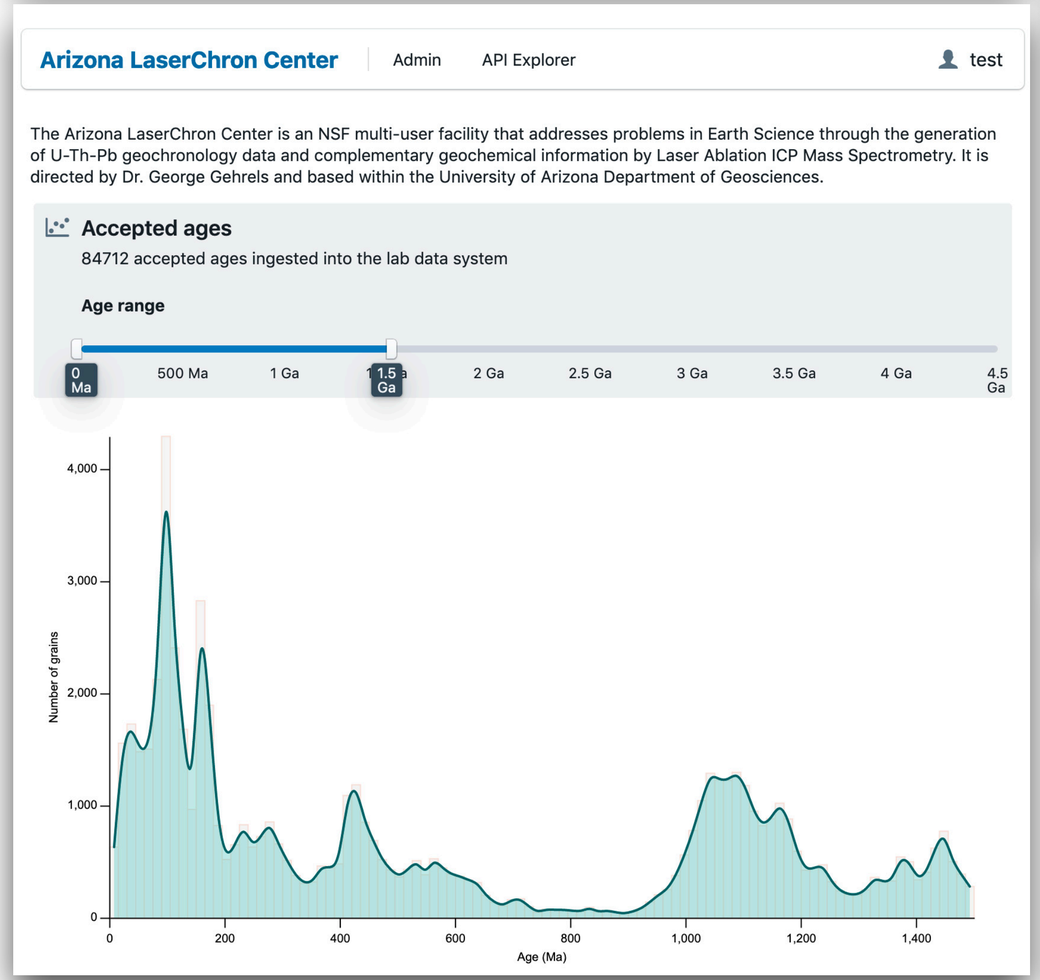
Extensible architecture

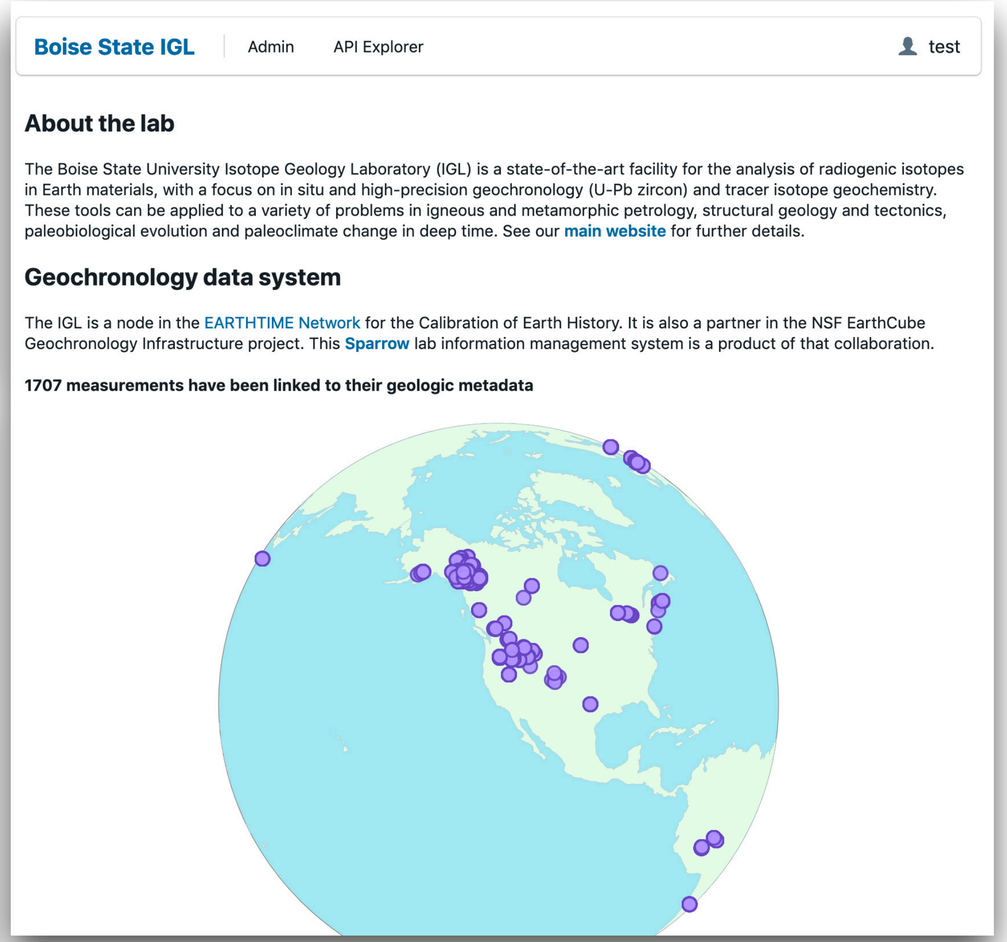
All parts of application are customizable by individual labs.

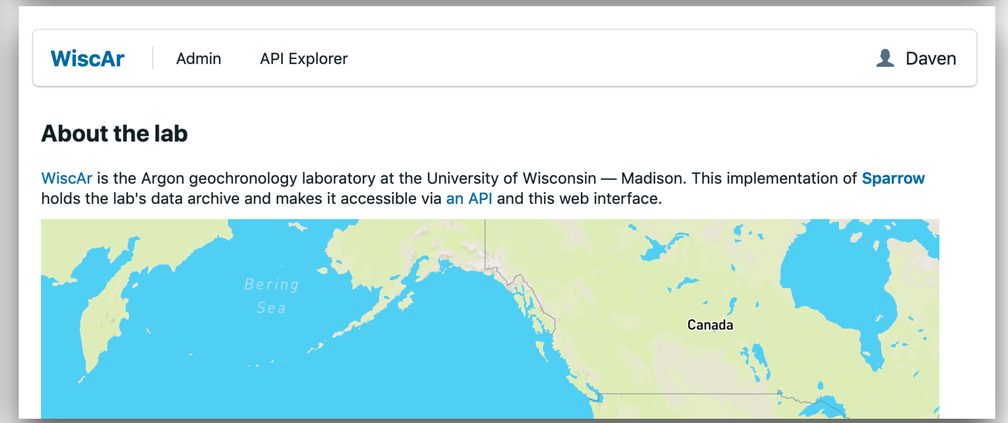
All features are optional!

WordPress *for geochemistry data! (+ an API)*









Interoperability and flexibility
enabled by structured vocabularies

Lab-specific vocabulary

Ratio of 206Pb*/238U	Arizona LaserChron Center
Ratio of 207Pb*/235U*	Arizona LaserChron Center
206Pb*/207Pb* age	Arizona LaserChron Center
206Pb*/238U* age	Arizona LaserChron Center
207Pb*/235U* age	Arizona LaserChron Center
Concordance between U/Pb age systems	Arizona LaserChron Center

configurable mappings

Shared vocabulary (analytical focus)

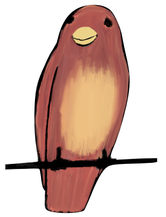
For export to, e.g. Geochron.org

ratio of uranium isotope U238 to lead isotope Pb204	EarthChem
ratio of uranium isotope U238 to lead isotope Pb206	EarthChem
ratio of uranium isotope U238 to thorium isotope Th230	EarthChem
ratio of uranium isotope U238 to thorium isotope Th230, re...	EarthChem
ratio of uranium isotope U238 to thorium isotope Th232	EarthChem
ratio of lead isotope Pb207 to uranium isotope U235	EarthChem
ulvospinel	EarthChem

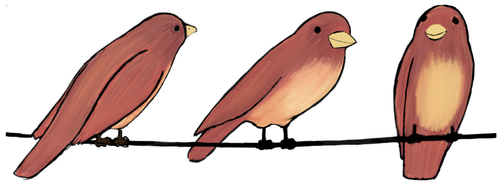
Shared vocabulary (interpretive focus)

For Macrostrat, Neatoma, end-users

detrital grain best age	Sparrow
igneous formation age	Sparrow
cooling age	Sparrow
exposure age	Sparrow



- Lab-specific data types with precise meanings
- Community-accepted standard vocabularies for sharing and reuse
- Importers and exporters for common file types (e.g. ETRedux XML)



Project status

https://github.com/EarthCubeGeochron/Sparrow
https://sparrow-data.org

Implementations

U-Pb facilities

- Arizona LaserChron Center
- Boise State Isotope Geology Laboratory

Argon-Argon facilities

- WiscAr
- New Mexico Geochronology Research Laboratory

Other techniques (test instances)

- Chemostratigraphy
- Optically-stimulated luminescence
- Cosmogenic nuclides

Importers for ETRedux, ArArCalc, PyChron, and other standard formats completed!

Core software development

winter 2019–present

- database and API design
- interactive UI components
- documentation
- design evaluation

Community consultation

- EarthCube workshop in May 2019: ~25 geochronologists discussed design and implementation challenges
- Beginning to train others to contribute code