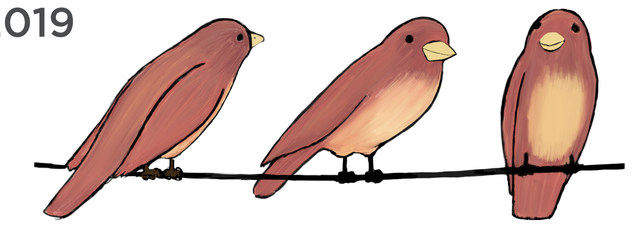
Sparrow: a cyberinfrastructure component supporting data management and sharing in geochronology laboratories

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



Data archival software for geochemical labs

- Organizes lab analytical data in a flexible, modern, industry-standard database application
- Augments current lab analytical processes
- Tools for curating and showcasing metadata

Interfaces with data users and community archives through a standard API

Does not dictate specific data-analysis tooling

Keeps labs in control of the canonical version of data

https://sparrow-data.org
Project information, documentation, and examples

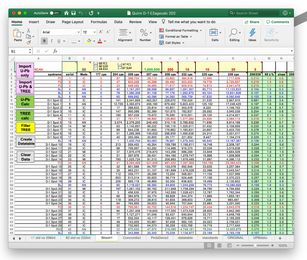
https://github.com/EarthCubeGeochron/Sparrow
Source code for core software

Linking metadata

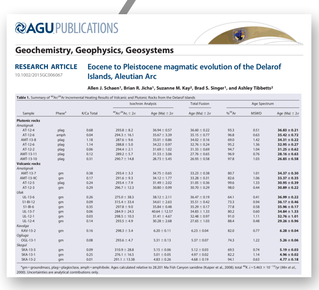


Geological context

Machine-readable
Cross-referenced
Available for synthesis



Analytical data



Published science result

Design principles

Extensible

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

Modular

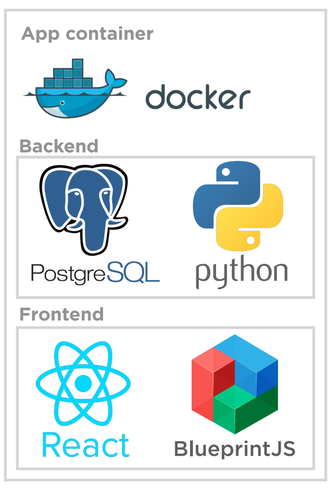
- Based on standard, widely used tools
- Decoupled API, frontend, and database
- Importers and exporters interface with other lab systems

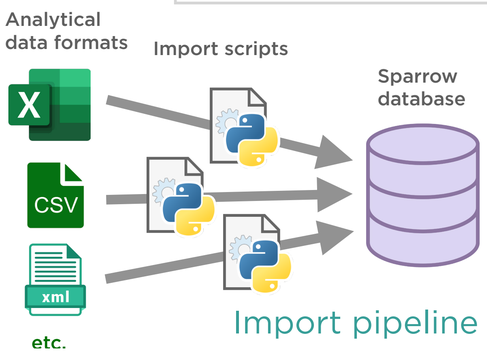
Lightweight

- Small footprint at implementing labs
- Abstracts the complexity of underlying software systems

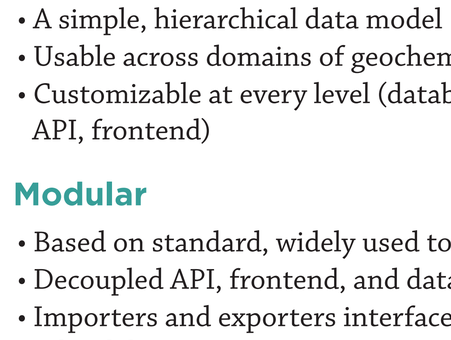
Data import scripts are implemented outside the core software, helping Sparrow stay minimal and focused.

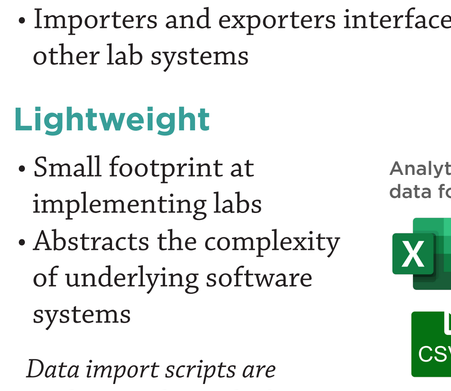
Software stack





Metadata management interface





Sparrow's main interface supports viewing analytical data and managing projects, sample metadata, and embargo.

The value of interoperable data

Reproducibility

- Adjustment and intercalibration
- Archival for re-analysis
- Audit by peer facilities

Audience: Expert workers within the geochronology community.
Data requirements: Detailed analytical information

Aggregation and synthesis

- Discovery of relevant geochemical data
- Integration with other data types (stratigraphic, ecological, etc.) for synthesis studies

Audience: out-of-domain scientists
Data requirements: Interpreted ages, simplified uncertainties, context

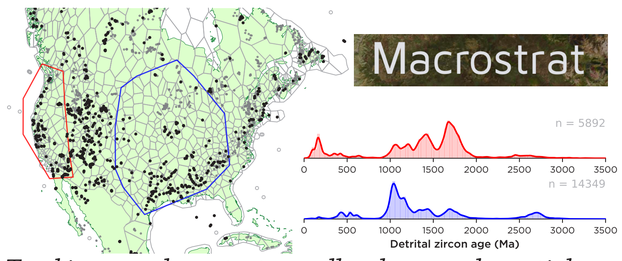
Obstacles

- Community data standards are often inflexible, do not account for subtle methodological differences in cutting-edge work
- Curating and archiving data demands significant researcher time, breaks focus

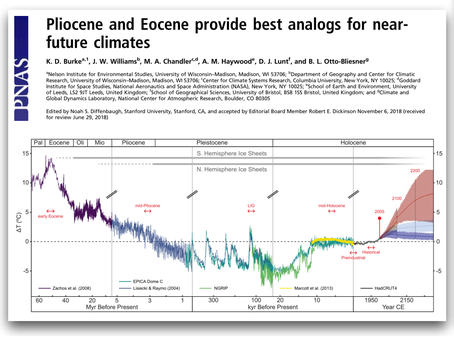
EarthChem
EARTHTIME ACCELERATING ADVANCES IN GEOCHRONOLOGY SINCE 2003

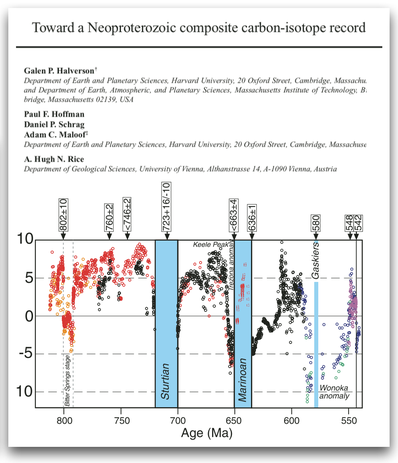
Interoperability has been a crucial aspect of initiatives focused on intercalibration and re-analysis.

Important composite time series in the geosciences are anchored by a collection of geochronology measurements, emphasizing the need for well-contextualized age data.

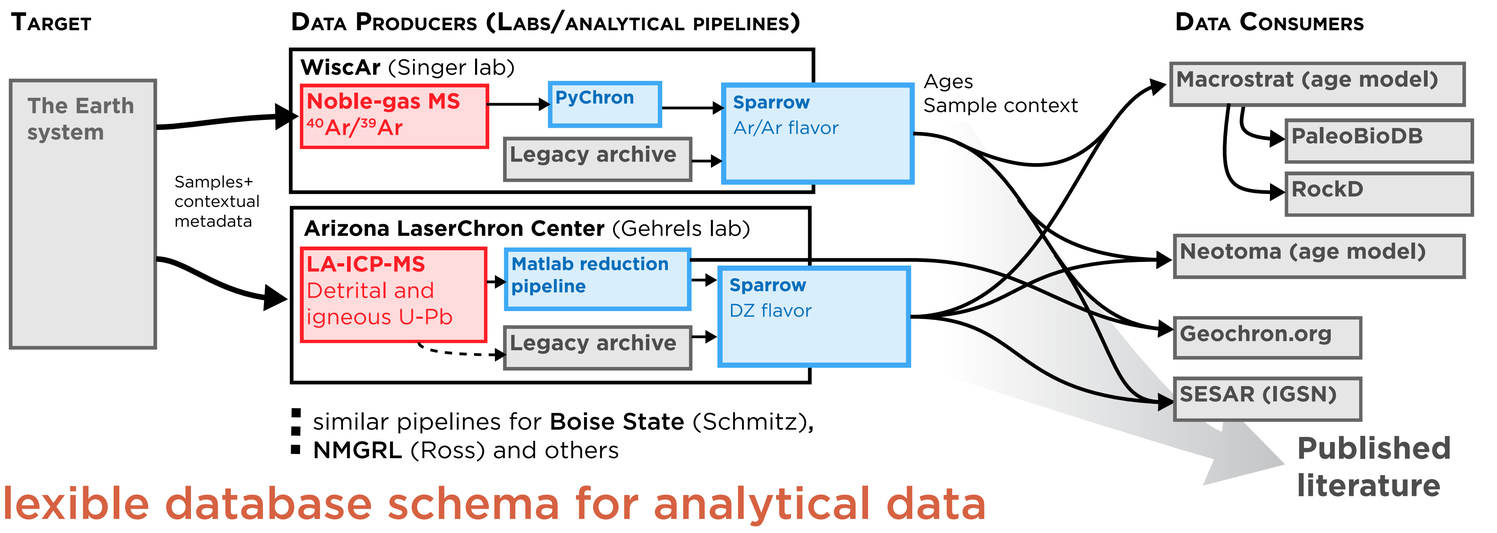


Tracking sample context can allow large-scale spatial and temporal subsets of age data to be compared with other datasets (here, Macrostrat *stratigraphic columns)*





Automated connections to external data consumers

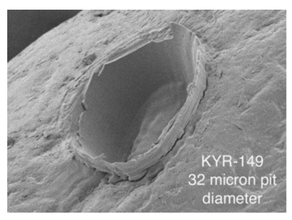


Flexible database schema for analytical data

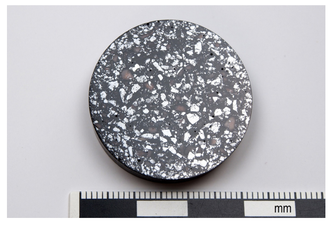
Single measured values

H2O content:	5.34 ±0.534(1s) %
De:	21.6 ±0.900(1s) Gy
OD:	24.0 %
luminescence age:	5.36 ±0.342(1s) ka
total dose rate:	4.03 ±0.195(1s) Gy/ka

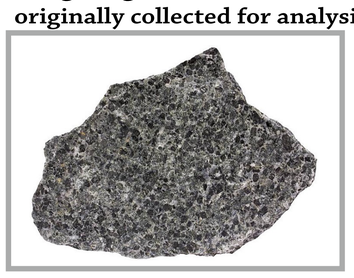
Single instance of space- or time-resolved data collection



Grouped measurements in a single instrument run



The geologic material that was originally collected for analysis



Datum → **Analysis** → **Session** → **Sample**

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

Analysis-focused vocabulary

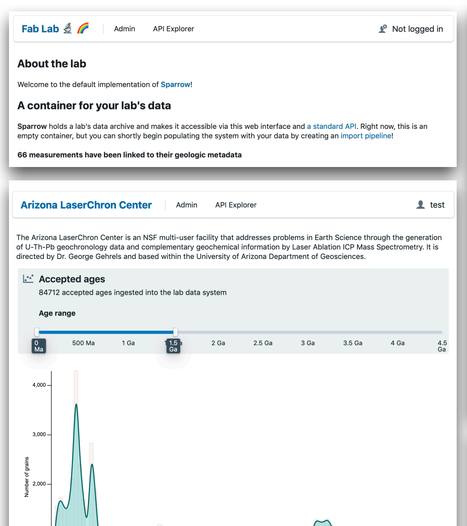
ratio of uranium isotope U238 to lead isotope Pb204	EarthChem	*To community archives (e.g.* **Geochron.org**) *or peer facilities.*
ratio of uranium isotope U238 to lead isotope Pb206	EarthChem	
ratio of uranium isotope U238 to thorium isotope Th230	EarthChem	
ratio of uranium isotope U238 to thorium 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	

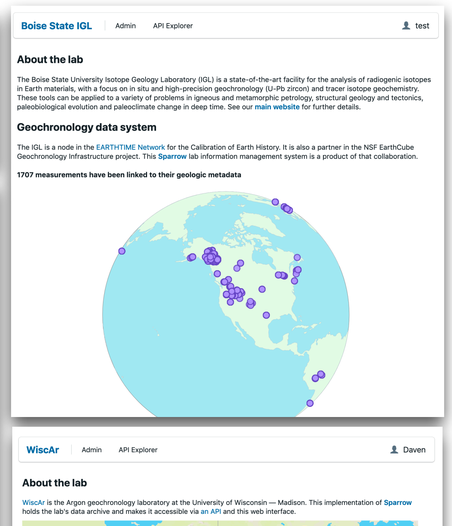
Interpretation-focused vocabulary

detrital grain best age	Sparrow	*To end users (e.g. the* **Macrostrat** *Earth model,* **Neatoma** *ecology database)*
igneous formation age	Sparrow	
cooling age	Sparrow	
exposure age	Sparrow	

Allows labs to immediately capture the full richness and fidelity of the measured data, without having to fit its description to a community standard.

Customizable lab website





Implementation progress

Arizona LaserChron Center (U-Pb)

~1500 samples (80,000 zircon grains) imported (~10% of archive)
Challenges: large scale of data, inconsistently linked metadata.

Boise State Isotope Geology Laboratory (U-Pb)

Archive of 1200 IGSN-tracked samples linked.
ETRedux XML importer created
~200 ETRedux XML files imported.
Challenges: representing rich sample/project metadata

WiscAr (Ar-Ar)

Entire post-2012 archive from MAP spectrometer (~1200 samples) imported.
Working on linking analyses to publication data.
Challenges: building metadata editing components

New Mexico Geochron. Research Laboratory (Ar-Ar)

PyChron importer completed
Initial test import successful

UW Madison Cosmogenic Nuclides Lab

Archive of ~1000 lab and literature samples ingested; work beginning on process management components

Test instances for other techniques: chemo-stratigraphy, electron microprobe, optical-ly-stimulated luminescence

Early-stage integrations: British Geological Survey and others