

Tephra-EPMA Use Case

Stephen C. Kuehn

EPMA and Tephra Lab
Physical and Environmental Sciences
Concord University

Wednesday, July 26, 2023

FAIR Facilities and
Instruments workshop,
Boulder, Colorado

Personal Background and Interests in FAIR and Open Data

- Faculty at small PUI
 - Teaching, mostly
 - But lots of other things too
- EPMA lab manager
 - Instruments, people, projects, samples, data
 - Method and reference material development
 - Collaborative resource for other people's work
- Research in tephra and applications
 - Integrates both pyroclastic deposits at source volcanoes and far-reaching volcanic ash deposits
 - Eruption history, physical volcanology, tephrochronology, and applications to paleolakes, paleoseismology, etc

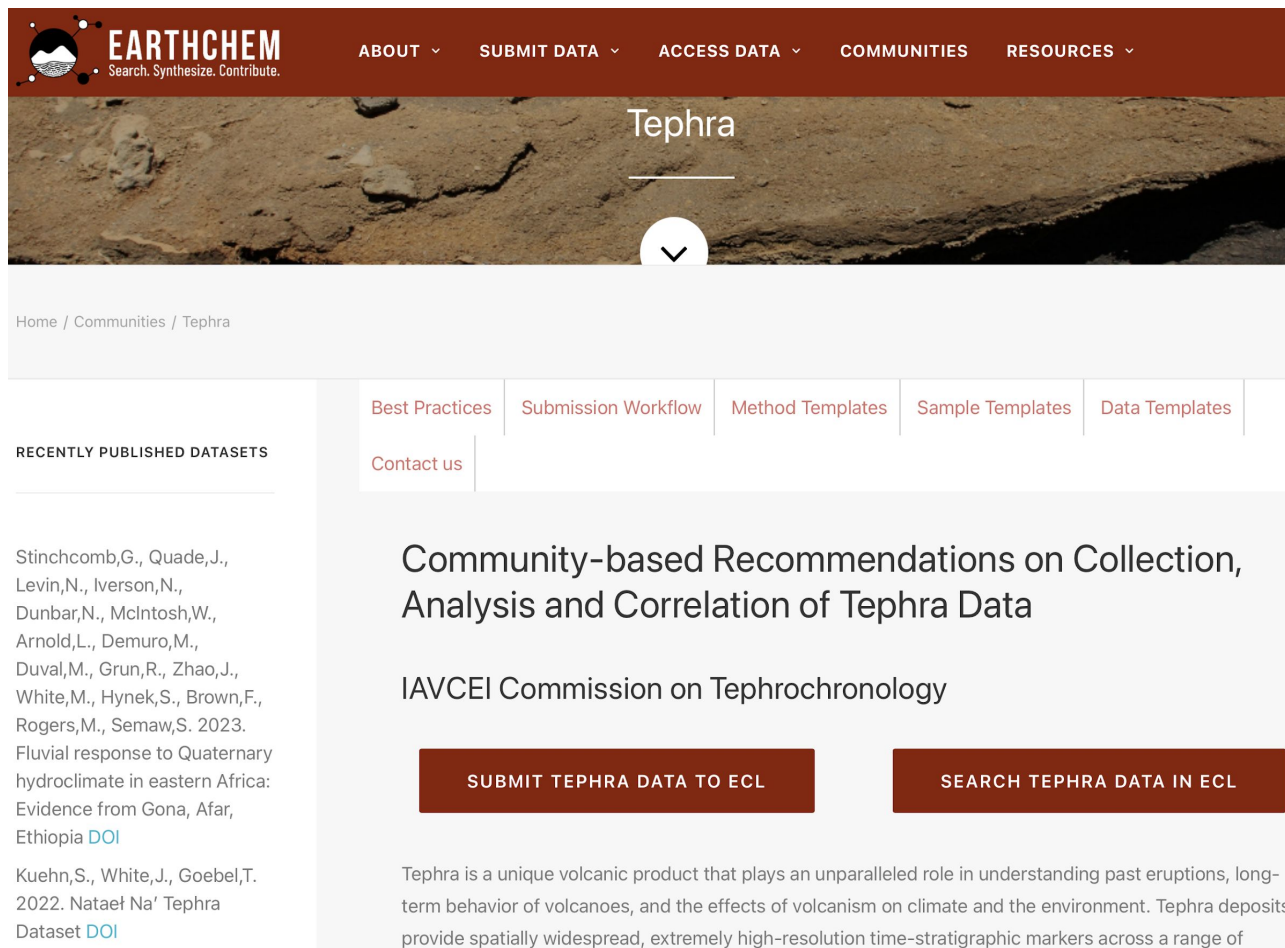


Background and Interests in FAIR and Open Data - Use Case

- Work on FAIR data standards and recommended practices with the tephra community for about 10 years (+ NSF EarthCube community as well)
- Led to community recommendations / standards
 - Wallace et al., Community established best practice recommendations for tephra studies—from collection through analysis, Scientific Data 9 (2022), <https://dx.doi.org/10.1038/s41597-022-01515-y>
 - Abbott et al., Community Established Best Practice Recommendations for Tephra Studies-from Collection through Analysis, Zenodo (2022), <https://dx.doi.org/10.5281/zenodo.3866266>
- Collaboration with StraboSpot, SESAR, EarthChem, Sparrow, and Throughput
 - On initial implementations – e.g. field app, community portal, metadata templates
 - Key components for a future data ecosystem spanning the field, lab, and data repository

Background and Interests in FAIR and Open Data - Use Case

Community portal



The screenshot shows the EarthChem website's Tephra community portal. The header features the EarthChem logo and navigation links: ABOUT, SUBMIT DATA, ACCESS DATA, COMMUNITIES, and RESOURCES. The main content area is titled 'Tephra' and includes a breadcrumb trail: Home / Communities / Tephra. A navigation bar contains links for Best Practices, Submission Workflow, Method Templates, Sample Templates, Data Templates, and Contact us. The 'RECENTLY PUBLISHED DATASETS' section lists two publications: one by Stinchcomb et al. (2023) on fluvial response to Quaternary hydroclimate in eastern Africa, and another by Kuehn et al. (2022) on the Natae' Na' Tephra Dataset. A large featured article titled 'Community-based Recommendations on Collection, Analysis and Correlation of Tephra Data' by the IAVCEI Commission on Tephrochronology is highlighted with two buttons: 'SUBMIT TEPHRA DATA TO ECL' and 'SEARCH TEPHRA DATA IN ECL'. The article text begins with 'Tephra is a unique volcanic product that plays an unparalleled role in understanding past eruptions, long-term behavior of volcanoes, and the effects of volcanism on climate and the environment.'

EARTHCHEM
Search. Synthesize. Contribute.

ABOUT ▾ SUBMIT DATA ▾ ACCESS DATA ▾ COMMUNITIES RESOURCES ▾

Tephra

Home / Communities / Tephra

Best Practices Submission Workflow Method Templates Sample Templates Data Templates

Contact us

RECENTLY PUBLISHED DATASETS

Stinchcomb,G., Quade,J., Levin,N., Iverson,N., Dunbar,N., McIntosh,W., Arnold,L., Demuro,M., Duval,M., Grun,R., Zhao,J., White,M., Hynek,S., Brown,F., Rogers,M., Semaw,S. 2023. Fluvial response to Quaternary hydroclimate in eastern Africa: Evidence from Gona, Afar, Ethiopia [DOI](#)

Kuehn,S., White,J., Goebel,T. 2022. Natae' Na' Tephra Dataset [DOI](#)

Community-based Recommendations on Collection, Analysis and Correlation of Tephra Data

IAVCEI Commission on Tephrochronology

SUBMIT TEPHRA DATA TO ECL **SEARCH TEPHRA DATA IN ECL**

Tephra is a unique volcanic product that plays an unparalleled role in understanding past eruptions, long-term behavior of volcanoes, and the effects of volcanism on climate and the environment. Tephra deposits provide spatially widespread, extremely high-resolution time-stratigraphic markers across a range of

Background and Interests in FAIR and Open Data - Use Case

SESAR
metadata
template built
on community
standards

This works and fills an
important gap, but
remains cumbersome
to complete – More
digital from birth
metadata and
automation could help

Partial support in
database, but can
archive complete
record in file repository

Metadata fields in columns

Category	*station ID or name of sampling station	core ID name - long	*stratum or stratigraphic interval ID	*sample ID	parent IGSN (Core ID from a core, Site ID if a grab sample)	sample IGSN	other sample ID	sample labelling convention	purpose
Tephra Best Practice: Explanation/Definition/Example	must use the same station ID from the "Station (Site)" tab	give the name of the core that the sample was collected	must use the same stratum or stratigraphic interval ID from the "Stratum or Stratigraphic Interval" tab	must be unique. can be helpful to include station ID in sample ID, e.g. 09KWRD-1 (year, collector initial, volcano or location, sample 1)	list the IGSN number for the site or core that the sample was collected from	give International Geo Sample Number (IGSN) if applicable (strongly encouraged to register all samples in the System for Earth Sample Registration)	include other sample ID's assigned to this sample (e.g. lab-assigned ID's)	explain how a sample ID is generated (year, initials, volcano name, area name etc.), IGSN (international sample numbering)	why are you sampling? correlation etc.
SESAR: Explanation (Example)			(Coconino Sandstone; Fig Tree Formation)	Sample Name: Collector's sample name. Mandatory. (TR-PROW)		IGSN: Leave blank if you want SESAR to assign the IGSN (IGMEG001)	Other name(s): Other name(s) used for the sample. (TRPOW; T-PROW)		Purpose: Free text collection purpose (Stratigraphic or Geochronology)
Metadata Entry Example	Summer Lake 2006 Trench 1		Composite Section Stratum 9				CUI071	site code (FF) and sequential number; a,b or -1, -2 indicate subsamples of a tephra layer or sample splits	tephrochronology
Input your Metadata Here	Summer Lake 2006 Trench 1		Composite Section Stratum 9	FF1			CUI071	site code (FF) and sequential number; a,b or -1, -2 indicate subsamples of a tephra layer or sample splits	tephrochronology
	Summer Lake 2006 Trench 1		Composite Section Stratum 11	FF2			CUI072	site code (FF) and sequential number; a,b or -1, -2 indicate subsamples of a tephra layer or sample splits	tephrochronology
	Summer Lake 2006 Trench 1		Composite Section Stratum 23	FF3			CUI073	site code (FF) and sequential number; a,b or -1, -2 indicate subsamples of a tephra layer or sample splits	tephrochronology
	Summer Lake 2006 Trench 1		Composite Section Stratum 24	FF4			CUI074	site code (FF) and sequential number; a,b or -1, -2 indicate subsamples of a tephra layer or sample splits	tephrochronology
			Composite Section Stratum 28	FF5a			CUI075a	site code (FF) and sequential number; a,b or -1, -2 indicate subsamples of a tephra layer or sample splits	tephrochronology
			Composite Section Stratum 28	FF5b			CUI075b	site code (FF) and sequential number; a,b or -1, -2 indicate subsamples of a tephra layer or sample splits	tephrochronology
	Summer Lake 2006 Trench 1		Composite Section Stratum 29	FF6			CUI076	site code (FF) and sequential number; a,b or -1, -2 indicate subsamples of a tephra layer or sample splits	tephrochronology
	Summer Lake 2006 Trench 2		Trench 2 lined with erosion	EE7.1			CUI077.1	site code (FF) and sequential number; a,b or -1, -2 indicate subsamples of a tephra layer or sample splits	tephrochronology

SESAR and tephra community explanation rows

Metadata example row

Put your metadata here, as many lines as needed

Tabs for introduction and different metadata frameworks

Background and Interests in FAIR and Open Data - Use Case

Analytical Method Descriptors

Re-usable with DOI so only needs to be done once (e.g. by the method expert)

Concord EPMA_Glass_METHOD_6-0 (final)

Home Insert Draw Page Layout Formulas Data Review View Automate Developer XLSTAT Tell Me

A2

Primary Analytical Metadata

BOLD, CAPITAL headings indicate MANDATORY fields

This spreadsheet is designed for entry of metadata related to a single EPMA or SEM-EDS analytical method. Use a separate spreadsheet for each additional method.

IDENTIFICATION	TECHNIQUE	INSTRUMENT	LABORATORY	Laboratory ID	Additional method references	Funding	ANALYTICAL PROCEDURE			
Metadata Field	Method DOI	METHOD NAME	Method start date	TECHNIQUE	INSTRUMENT	LABORATORY	Laboratory ID	Additional method references	Funding	ANALYTICAL PROCEDURE
Metadata Field Description	DOI or unique identifier for this method (generated after submission)	Short descriptive name for the method, with version number	First date when method was used in YYYYMMDD format (format as plain text)	Select either EPMA or SEM-EDS	Instrument manufacturer and model - e.g. CAMECA SX50	Name of laboratory or institution - e.g. Woods Hole Oceanographic Institute	add ROR if available (https://ror.org)	DOI or URL, for any other document that describes this method in further detail; include an evaluation of accuracy and precision. Ideally it will also be open access (e.g. Zenodo)	Grants and other funding that supported instrument purchase, major upgrades, key supporting equipment, and personnel. Include grant IDs where applicable	SOFTWARE analytical and automation software, including versions
Enter only metadata on this line		CU routine lepta glass version 6	2019-03-26	EPMA	ARL SEMQ	Concord University, Athens, West Virginia, USA		This method is a derivative of method 1 described in this 2016 conference poster: https://doi.org/10.5281/zenodo.6271693	2019 West Virginia Higher Education Policy Commission Division of Science and Research, grant number HEPC.dir.16.10.2015 West Virginia Higher Education Policy Commission Division of Science and Research, grant number HEPC.dir.16.11.2013 West Virginia Higher Education Policy Commission Division of Science and Research, grant number HEPC.dir.16.11.2013	Probe for EPMA 9.6.4 and Bruker Esprit 1.9.4.3448

EPMA-SEM Part 1 EPMA-SEM Part 2 +

EPMA and SEM-EDS done
Working on LA-ICP-MS, others

Concord EPMA_Glass_METHOD_6-0 (final)

Home Insert Draw Page Layout Formulas Data Review View Automate Developer XLSTAT Tell Me

A1

Acquisition and Post-Processing Metadata

OPERATION	CALIBRATION	PROCEDURAL BLANK															
PARAMETER	BEAM CURRENT	Spectrometer	Sequence	WDS Diffracting Crystal	Detector Type	X-ray Line	Peak Acquisition Time (seconds)	Background Acquisition Method	WDS Background Acquisition Time	WDS PHA Setting	EDS Dead Time	Primary Standard Name	Primary Standard ID	Interfering Elements	Inference Correction Standards	Blank Method	
measured element or oxide - e.g. SiO ₂ , Ba	beam current measured in nA	type and number - e.g. WDS 1, WDS 2, EDS	order of analysis on spectrometer - e.g. 1, 2, 3	WDS diffracting crystal - e.g. PET, LIF, TAP, RAP, etc.	e.g. xenon, P-10, P-10 low pressure, P-10 high pressure, SDD silicon drift detector, Si(Li) detector	e.g. Ka, La	report seconds, numerical value only	e.g. two backgrounds, high side only; multi-point, MAN (mean atomic number mode); Separate background method is not applicable for EDS	total for background(s) - e.g. 30 sec (not applicable if using MAN)	e.g. integral, differential	percent dead time reported by the EDS spectrometer	reference material used for element standardization (calibration)	IGSN, GeoRAM ID, Smithsonian number, or similar	list any elements for which a spectral interference correction was used (e.g. Ba is interfered by Ti ka)	reference materials used to calibrate the inference conditions	e.g. single sample, multi-sample weighted mean	
Use one row for each measured element	SiO2	10	EDS	1	SDD	Ka	8				10-12%	BHVO-2g glass, NKT-1g glass, or Lipari obsidian ID3506		Ba	Sandomite	integral	offline, multi-standard, weighted mean blank correction computed and applied in Excel
	TiO2	10	WDS 6	1	PET	P-10	Ka	180	MAN	Integral		NKT-1g glass		Ba	Sandomite	integral	offline, multi-standard, weighted mean blank correction computed and applied in Excel

EPMA-SEM Part 1 EPMA-SEM Part 2 +

Add PIDs for labs and instruments?

Background and Interests in FAIR and Open Data - Use Case

Micro-Analytical Data and Metadata Template at EarthChem

Report both sample data and reference material (secondary standard) data together linked by analytical session

Supports many in-situ point analyses of the same target, mean, stdev, grain ID, etc.

Supports Smithsonian GVP volcano and eruption IDs

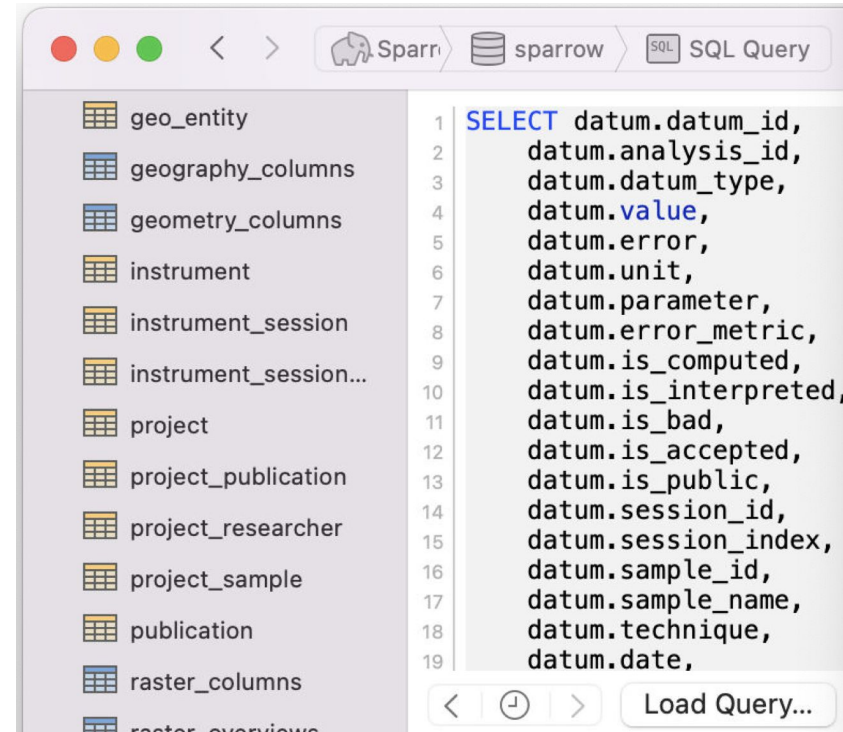
PID links to method DOIs, IGSNs, ORCIDs

	A	B	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	
1	DATA TABLE NAME	Kuehn-White-G																						
2	Analytical Method Name	CU routine teph																						
3	ANALYTICAL METHOD DOI	https://doi.org/10.26434/chemrxiv-2022-08-02																						
4	TECHNIQUE (EPMA or SEM)	EPMA																						
5	INSTRUMENT	ARL SEMQ																						
6	LABORATORY	Concord Univers																						
7																								
8	Sample, grain, and population information			Additional analytical metadata										For glass: REPORT NORMALIZED ANALYTICAL DATA (totals to 100%) AND THE ORIGINAL DATA. For minerals: Report unnormalized data and the original totals; For Fe-Ti Oxides, report both total Fe and Ti. REPORT OXIDE WEIGHT PERCENTAGES (m/m%) FOR ALL ELEMENTS EXCEPT H. REPORT ALL DATA - Data must not be censored (removed or replaced with zero). Data lines with large negative values should be set aside as these have potential problems with background subtraction. Recommended: To avoid rounding error, report at least one extra significant figure (typical for Fe, Ti, and Mn). Recommended: If an explicit correction for time-changing count rates is utilized (e.g., TDC), report the corrected data. Analyzed parameters (e.g. SiO2, TiO2, Total, Na-TDI%); Report normalized data.										
9	SAMPLE ID	IGSN	GVP Event or Episode ID	Analyst	ANALYSIS DATE and TIME	ANALYTICAL SESSION ID	Method Name	Method DOI	Beam Diameter	NUMBER OF ANALYSES	DATA LINE TYPE			SiO2	TiO2	Al2O3	FeOT	MnO	MgO	CaO	Na2O	K2O	P2O5	
82	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 7:30 PM	2020-01-01-A			5		single			68.12	1.10	13.71	4.59	0.078	1.32	3.39	4.34	2.91	0.268	
83	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 7:44 PM	2020-01-01-A			5		single			66.33	1.19	14.44	5.16	0.073	1.38	3.69	4.51	2.79	0.296	
84	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 7:58 PM	2020-01-01-A			5		single			66.17	1.15	14.77	5.08	0.077	1.45	3.72	4.35	2.80	0.290	
85	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 8:01 PM	2020-01-01-A			5		single			65.99	1.13	14.60	5.21	0.080	1.60	3.79	4.30	2.85	0.261	
86	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 8:05 PM	2020-01-01-A			5		single			66.97	1.14	14.65	4.84	0.087	1.53	3.57	4.23	2.90	0.275	
87	WRST 24075-tephra (CU19: BOF00000C			0000-0000	2020/01/07	2020-01-01-A			5	17	Mean			66.24	1.16	14.50	5.13	0.080	1.53	3.73	4.37	2.83	0.285	
88	WRST 24075-tephra (CU19: BOF00000C			0000-0000	2020/01/07	2020-01-01-A			5	17	StDev			0.91	0.07	0.50	0.29	0.016	0.16	0.34	0.13	0.22	0.019	
89																								
90	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 7:37 PM	2020-01-01-A			5		single			72.65	0.432	14.40	2.27	0.078	0.441	1.48	4.09	3.91	0.093	
91	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 7:40 PM	2020-01-01-A			5		single			71.99	0.424	14.70	2.22	0.072	0.441	1.46	4.78	3.75	0.079	
92	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 7:47 PM	2020-01-01-A			5		single			64.47	1.12	14.58	5.63						0.287	
93	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 7:51 PM	2020-01-01-A			5		single			63.78	1.06	15.24	5.67						0.261	
94	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 6:58 PM	2020-01-01-A			5		single			62.05	1.073	15.19	6.93						0.349	
95	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 7:33 PM	2020-01-01-A			5		single			61.17	0.382	27.73	2.03	0.081	0.413	1.58	2.92	3.36	0.094	
96	WRST 24075-tephra (CU19: BOF00000C			0000-0000	1/7/20 7:54 PM	2020-01-01-A			5		single			67.27	0.744	16.48	3.74	0.161	0.958	2.12	5.20	2.93	0.200	
97																								
98	BHVO-2G					2020-01-01-A			8		single			50.23	2.80	13.81	11.20	0.140	7.32	11.47	2.21	0.528	0.286	
100	BHVO-2G					2020-01-01-A			8		single			50.23	2.79	13.68	11.17	0.173	7.38	11.51	2.24	0.559	0.265	
101	BHVO-2G					2020-01-01-A			8		single			50.25	2.79	13.63	11.04	0.200	7.41	11.56	2.25	0.551	0.278	
102	BHVO-2G					2020-01-01-A			8		single			50.05	2.82	13.61	11.35	0.158	7.45	11.46	2.28	0.524	0.273	
103	BHVO-2G				0000-0000	1/7/20 11:17 PM	2020-01-01-A		5		single			50.62	2.84	13.54	10.95	0.216	7.34	11.52	2.21	0.493	0.263	

Add PIDs for labs and instruments?

Background and Interests in FAIR and Open Data - Use Case

- This work has relatively comprehensive data and metadata coverage.
 - Field data in StraboSpot
 - Sites, samples, and cores in SESAR
 - Methods, geochemistry, and additional metadata at EarthChem
- Developing a schema for Sparrow to help capture most data and metadata during the progress of laboratory work flows

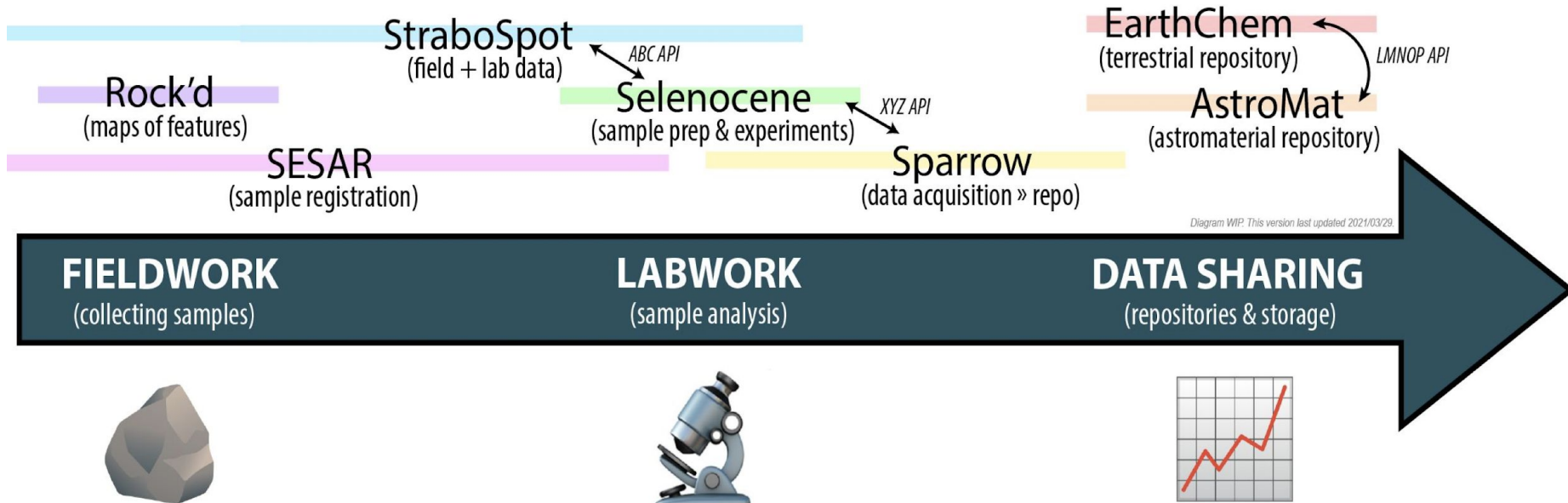


Background and Interests in FAIR and Open Data - Use Case

- Developing a schema for Sparrow lab data system with comprehensive coverage of tephra data and metadata – adaptable to e.g. min/pet/geochron too
 - Projects, researchers, funding, and publications
 - Field locations, strata, and cores
 - Samples, sample curation, and geochemical reference materials
 - Sample preparation and resulting fractions, grain mounts, and thin sections
 - Analytical instruments, methods, and sessions
 - Geochemistry with full association of sample and reference material/calibration (QC) data
 - Optical and electron microscopy images (see also StraboMicro)
 - Grain size - sieve data and discrete measurements
 - Grain type (componentry, mineralogy) and grain density
 - Volcanic source assignment using Smithsonian GVP IDs
 - Data access rights and embargos
 - ORCIDs, IGSNs, DOIs, grant IDs, etc. captured as available
- Add PIDs for labs and instruments?*

Background and Interests in FAIR and Open Data - Use Case

- Building on multiple components for an inter-connected ecosystem spanning the field, lab, and data repository
- PIDs key for linking related information within and across data systems



Approach / Philosophy / Needs

- I'm a big fan of the databased from birth approach – e.g. in StraboSpot for the field
- I'm a fan of integrated databases wherever possible instead of just heterogeneous file-based repositories
- We need more use of the above, especially for lab workflows and geochemistry.
- By capturing the data and metadata on the fly, FAIR archiving of research outputs in *integrative open databases* becomes much easier. And, we'll get more comprehensive metadata too for better data provenance and improved re-usability.
- Right now, data and metadata archiving is still much too labor intensive and thus a huge barrier for most researchers.
- The more readily we can capture key information as it is generated, the more people will do it and the better all of this FAIR and open data stuff will work out.
- PIDs are really important for a connected FAIR data-metadata ecosystem. Adding

Thoughts about Instrument-Facility PIDs

- Improved data provenance – better access to information about analytical facilities where data is generated
- Better linking of instruments, labs, technicians, grants, methods, data, and publications
- Better discovery of similar / related datasets from the same lab or discovery of potential collaborators
- Better tracking of impacts of equipment/facility funding (makes funders happy)

Thoughts about Instrument-Facility PIDs

- Could have instrument/facility landing pages similar to ORCID records
 - Connections to all of the previous
 - funding inputs – e.g. grants
 - research outputs – e.g. datasets, publications
 - etc.
 - Instrument history – initial installation, major changes/upgrades, relocation to different institution
 - Laboratory staff history – potential to better credit technicians, link to ORCIDs
 - Instrument capabilities ?
 - Lab specializations ?