PetDB, SedDB, EarthChem & SESAR

Resources for CoreWall:

Geochemical Data &

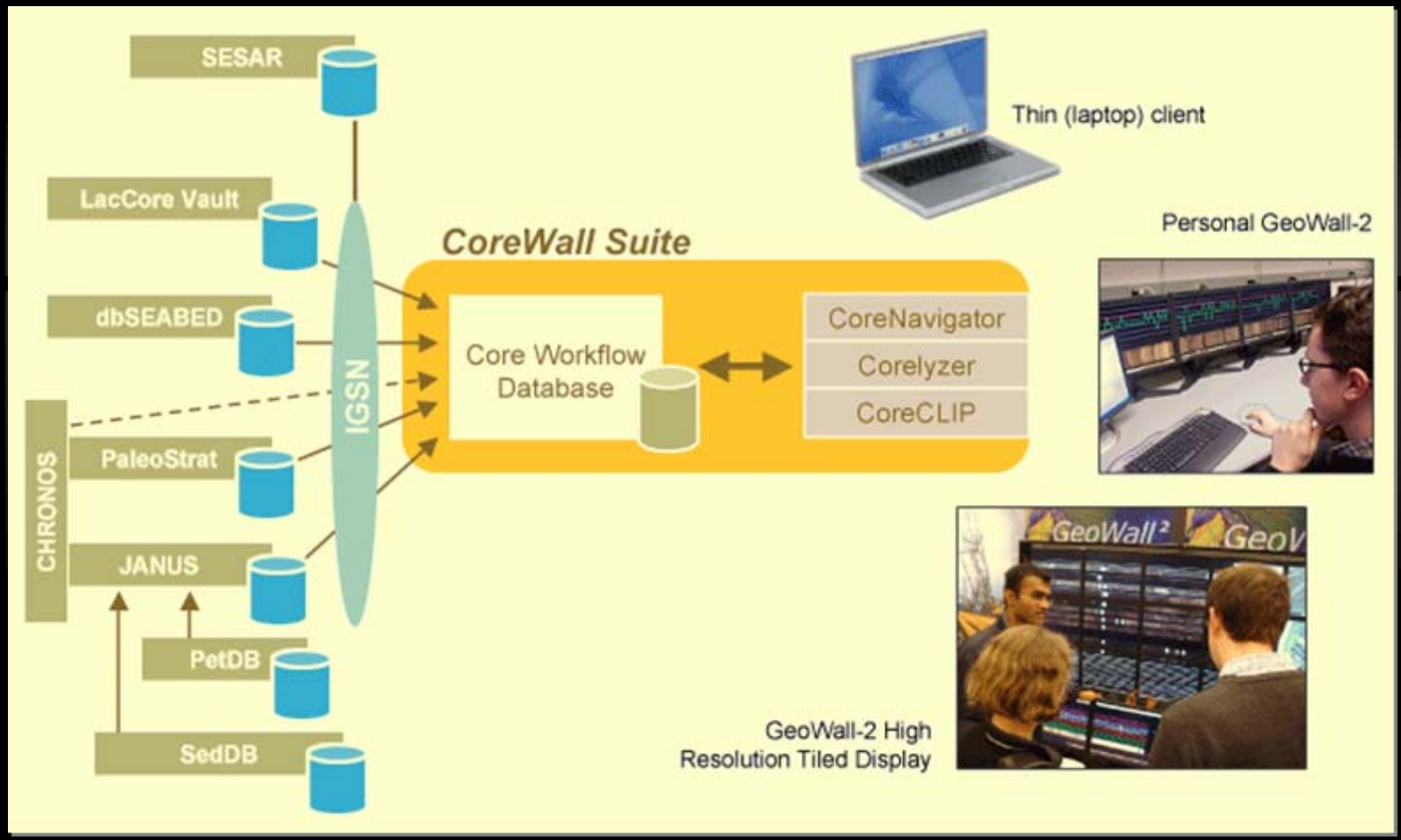
Unique Sample Identifiers

The Team

- Scientific Guidance
  - Steve Goldstein (SedDB, SESAR)
  - Charlie Langmuir (PetDB)
  - Rick Murray (SedDB)
  - Nick Pisias (SedDB)
  - Doug Walker (EarthChem)
  - Mike Mottl (VentDB)
- Data Development & Stewardship
  - Annika Johansson
  - (Karin Block, June 2006)
  - Chris Lenhardt
  - Rusty Lotti Bond
- System Engineering & Operation
  - Sri Vinayagamoorthy
  - Nevila Celo
  - Charles DeGruccio
  - Branko Djapic
  - Artem Fishman
  - Annie Gerard
  - David Strom

The CoreWall Connection I

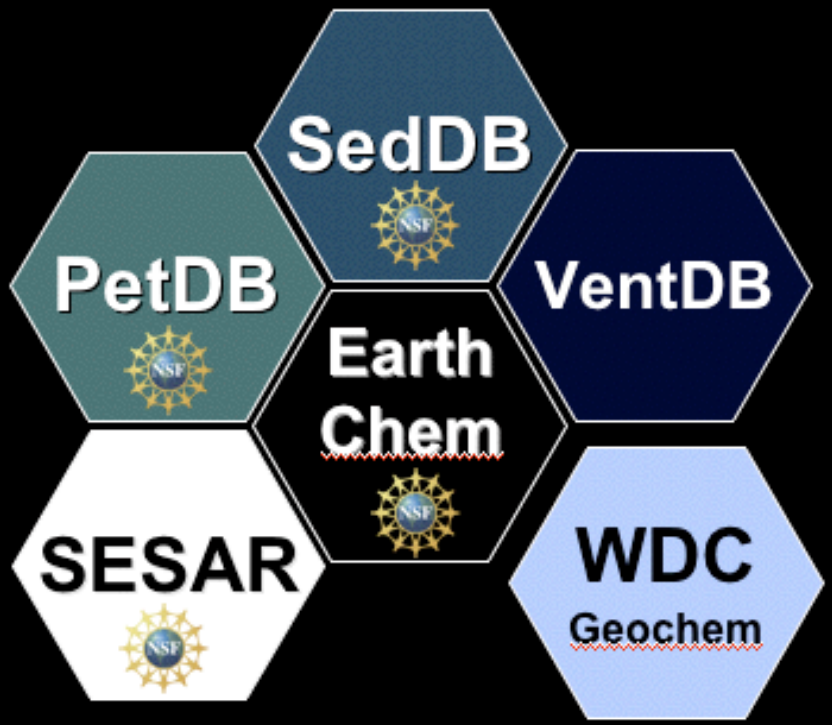
Diagram by Jason Leigh (EVL, U Illinois Chicago)



LAMONT-DOHERTY EARTH OBSERVATORY
THE EARTH INSTITUTE AT COLUMBIA UNIVERSITY

CIESIN Columbia University

Geoinformatics for Geochemistry



PetDB: *Rocks (volcanic, plutonic, mantle rocks) generated at mid-ocean ridges, fracture zones, back-arc basins, young seamounts (+ old oceanic crust)*

SedDB: *Marine sediments & sedimentary rocks*

VentDB: *Seafloor hydrothermal springs*

EarthChem: *Solid earth & beyond*

Digital Data Collections for Geochemistry

Compile and serve ALL ‘raw’ geochemical data

- Major Oxides
- Trace Elements
- Radiogenic Isotopes
- Stable Isotopes
- Noble Gases
- Uranium series
- Analytical Ages

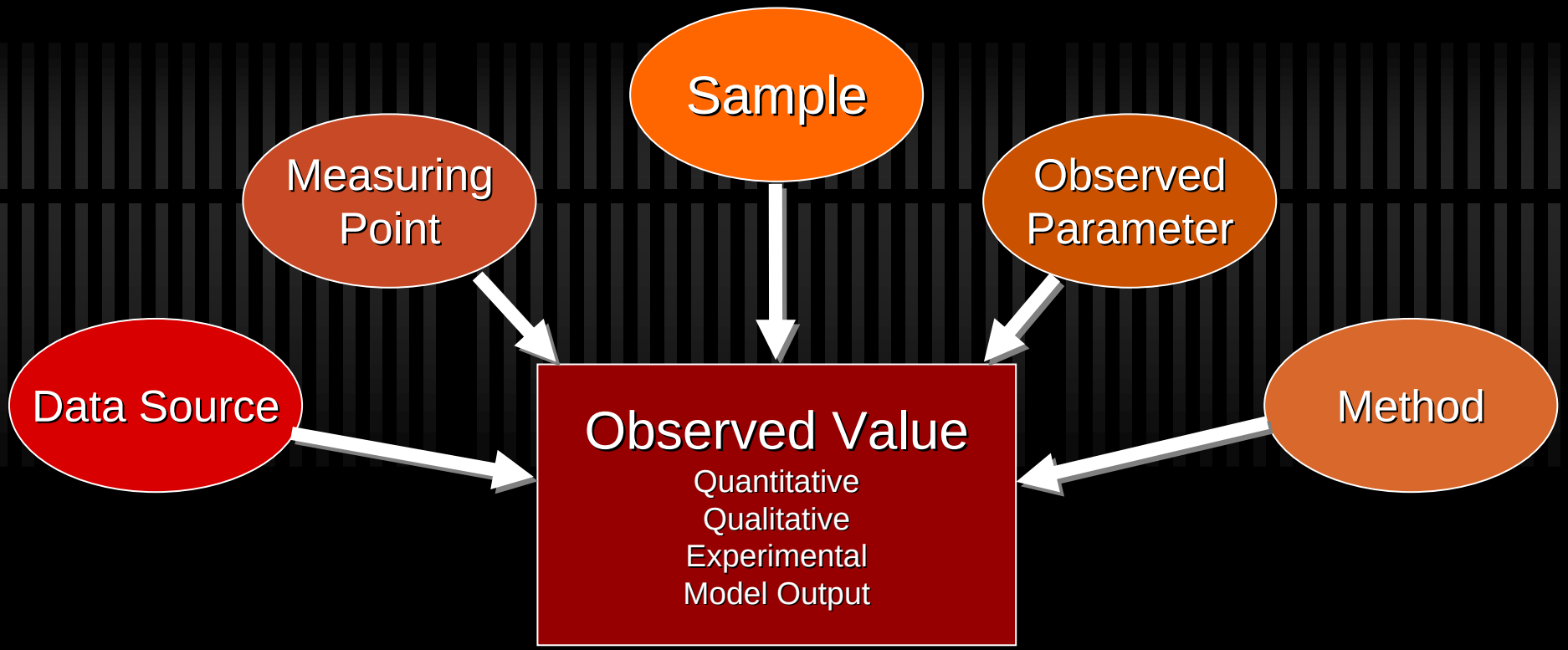
Relational data model

- ✓ Data fully integrated
- ✓ Wide range of sample & analytical metadata
- ✓ All types of chemical data for all materials

Content

- **Legacy data**
  - Data from the literature entered by PetDB personnel
- **New data**
  - Ingest data submitted by PIs
  - Data submission tools, to be built by EarthChem

Generic Data Model for Analytical Data



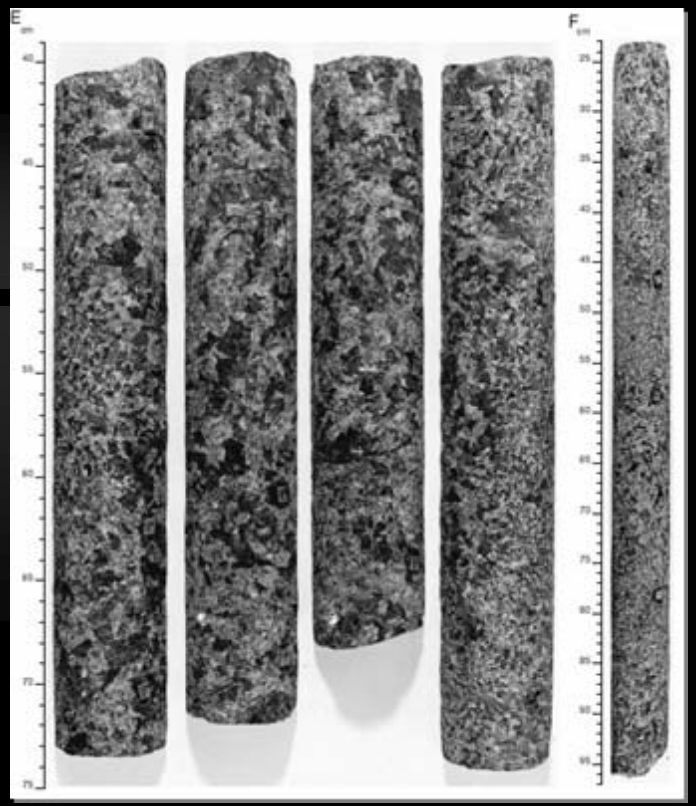
Djapic, Vinayagamoorthy & Lehnert 2006

Unique Features

- Data fully integrated (each individual chemical value is searchable).
- Data for individual samples are linked via unique sample identifiers.
  - ***Sample Profile*** (sample info page) summarizes all available data for a sample.
- Query result integrates data from different sources.
- Data can be filtered by analytical technique.
- Web services allow exchange of information with external systems.

DSDP/ODP in PetDB

	DSDP	ODP
Legs	**56**	**24**
Holes (Stations)	**221**	**75**
Samples	**7887**	**2197**
Chemical Values	**29,523**	**11,251**
Rocks	24305	6011
Minerals	5102	5051
Melt Inclusions	116	132
Rock Modes	123	57
Chemical items	**141**	**113**



http://www.petdb.org

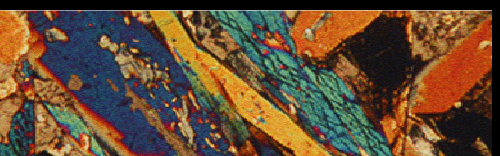
Digital Data Collection for Geochemistry

- Maximize utility of the Geochemical data set in science & education
- Advance scientific discovery
- Allow data integration & visualization across the Geosciences
- Enhance data quality control

EarthChem

- ✓ Consortium founded in 2003
- ✓ based on and expanding the collaboration of the igneous geochemical databases PetDB, GEOROC, and NAVDAT.
- ✓ Goal: Build an integrated data management and information system for solid earth geochemistry,
- ✓ Project funded in 2005.

- ∑ Nurture synergies among projects
- ∑ Minimize duplication of efforts
- ∑ Share tools and approaches



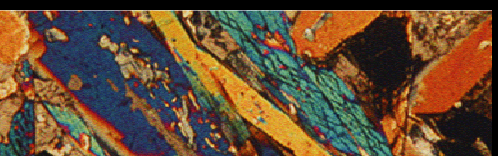
earthchem

ADVANCED DATA MANAGEMENT IN SOLID EARTH GEOCHEMISTRY

The EarthChem Project

"A Community Resource for Geochemistry"

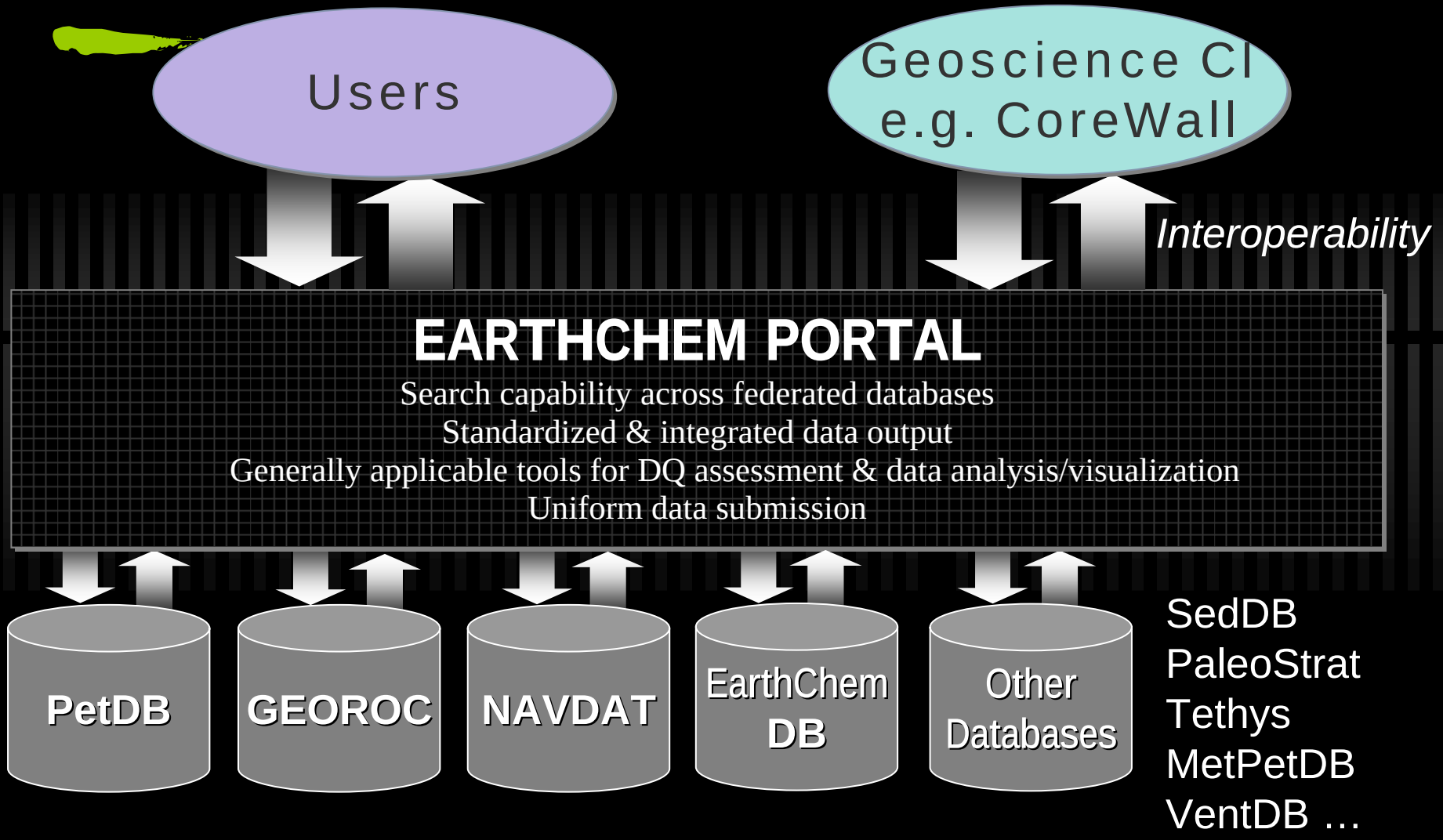
- ✓ Provide a central access point to the broadest range of geochemical data ("One-Stop Shop")
- ✓ Ensure efficient and continuing update and expansion of data holdings
- ✓ Promote & implement standards for data management in Geochemistry



earthchem

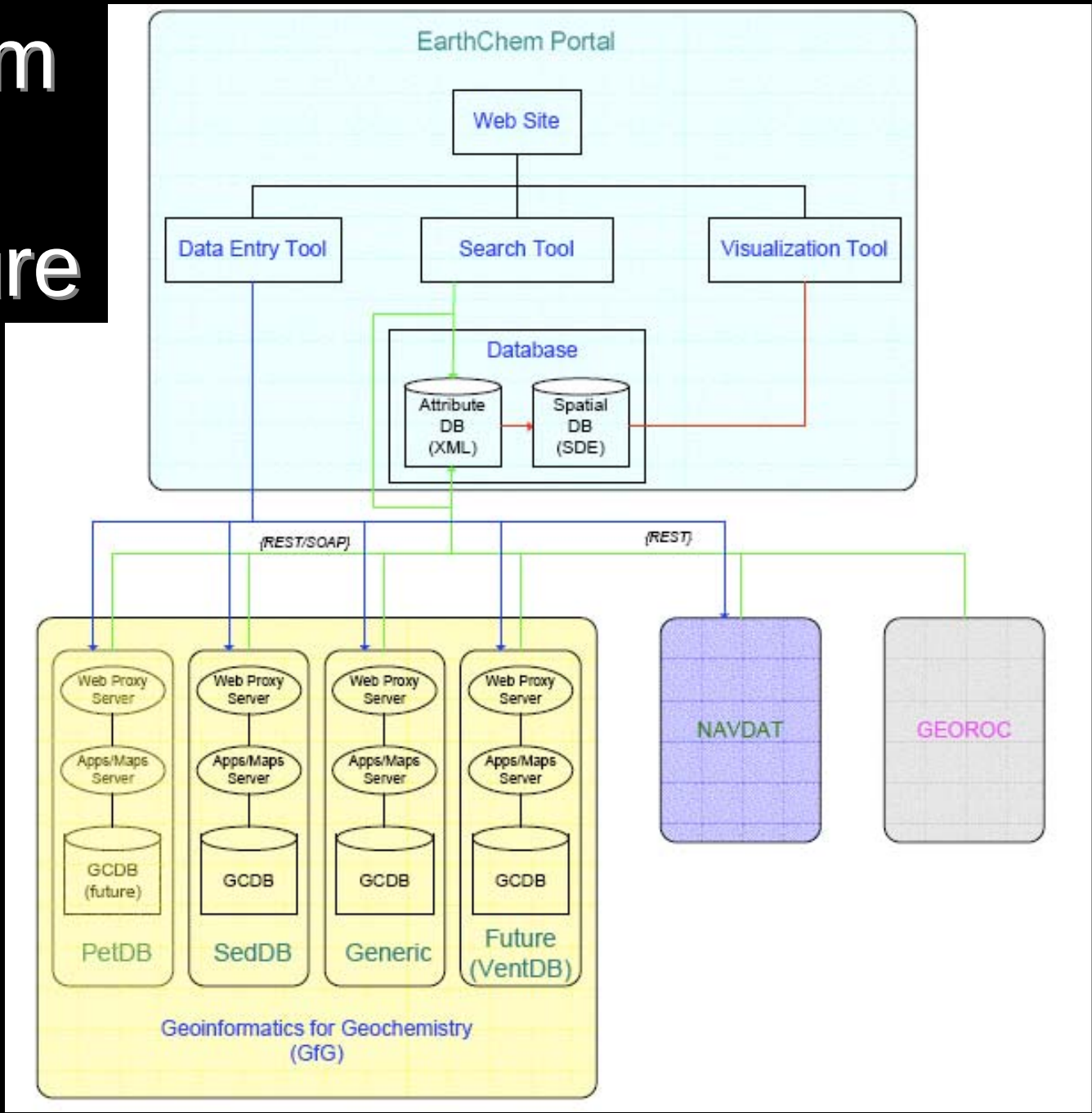
ADVANCED DATA MANAGEMENT
IN SOLID EARTH GEOCHEMISTRY

One-Stop-Shop for Geochemical Data

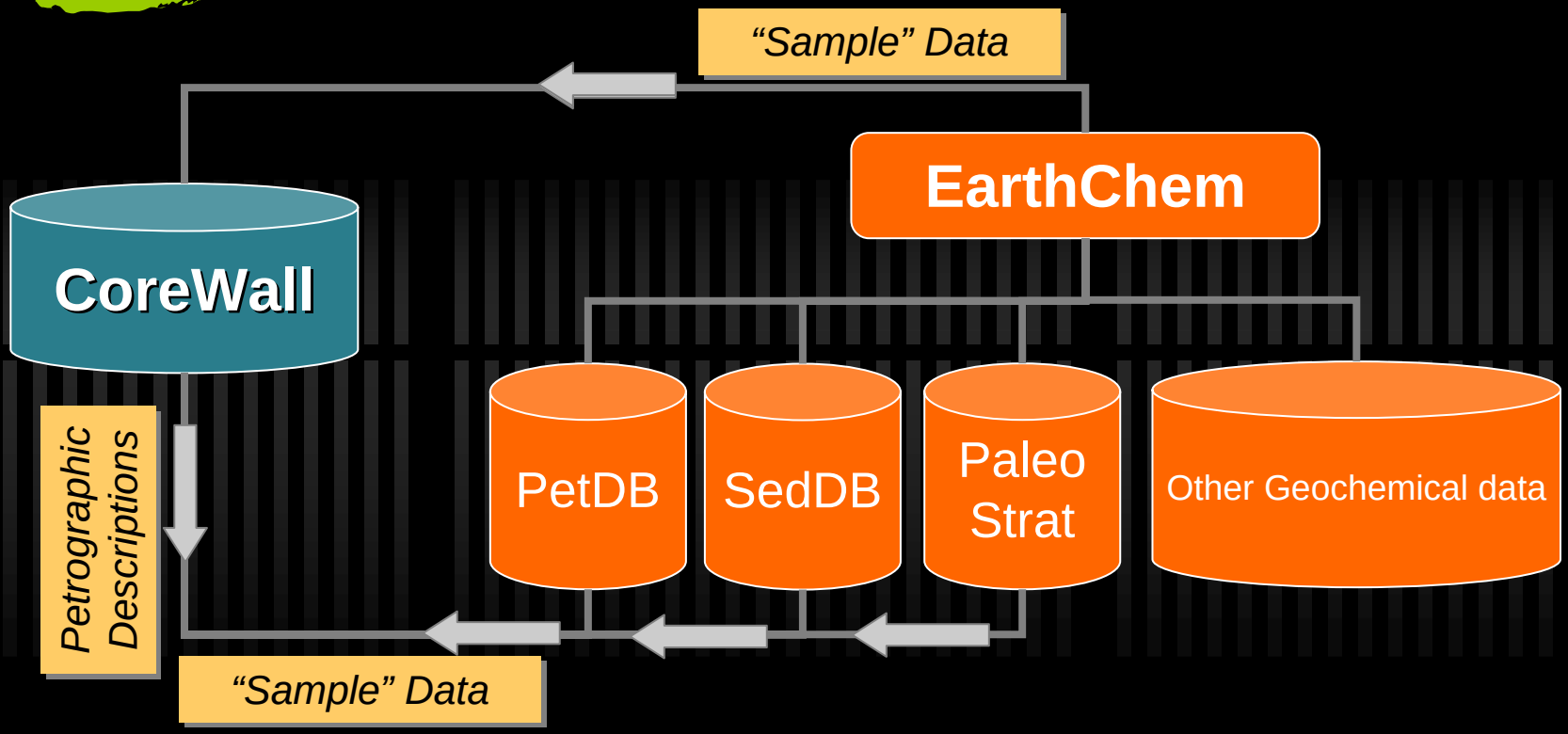


EarthChem Portal Architecture

Beta release anticipated July 1, 2006



The CoreWall Connection II



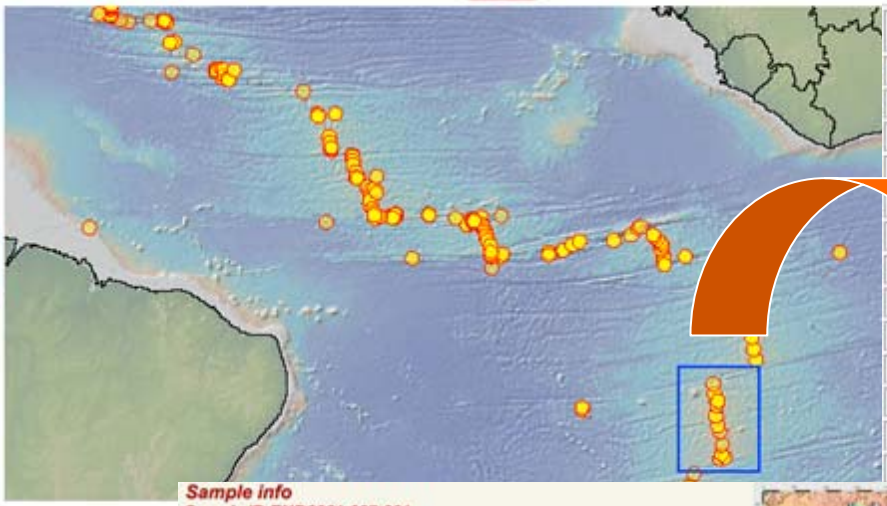
Data exchange via REST/SOAP web services

Interactive map *(beta version)*:

➥ *Mapviewer with base map of bathymetry from mgDMS*

PETDB
Petrological Database of the Ocean Floor

x y GO Select AOI GO Zoom Factor : 2 x



SAMPLE_ID	SAMPLE_URL	GEOM
END0061-008-002A	Sample Profile	Point:(EPSG:4326)-13.398,-8.008
END0061-008-001A	Sample Profile	Point:(EPSG:4326)-13.398,-8.008
END0061-008-001	Sample Profile	Point:(EPSG:4326)-13.398,-8.008
END0061-007-001	Sample Profile	Point:(EPSG:4326)-13.457,-7.507
-032-004	Sample Profile	Point:(EPSG:4326)-13.12,-11.27
32-001	Sample Profile	Point:(EPSG:4326)-13.12,-11.27
1-016	Sample Profile	Point:(EPSG:4326)-12.88,-11.19
-009	Sample Profile	Point:(EPSG:4326)-12.88,-11.19
ABP0005-031-023	Sample Profile	Point:(EPSG:4326)-12.88,-11.19
ABP0005-031-020	Sample Profile	Point:(EPSG:4326)-12.88,-11.19
ABP0005-031-008	Sample Profile	Point:(EPSG:4326)-12.88,-11.19
ABP0005-031-018	Sample Profile	Point:(EPSG:4326)-12.88,-11.19
ABP0005-031-021	Sample Profile	Point:(EPSG:4326)-12.88,-11.19
31-013	Sample Profile	Point:(EPSG:4326)-12.88,-11.19
	Sample Profile	Point:(EPSG:4326)-12.88,-11.19
	Profile	Point:(EPSG:4326)-13.04,-11.05
13-001	Profile	Point:(EPSG:4326)-13.007,-10.553
12-001	ofile	Point:(EPSG:4326)-13.13,-9.93
-011	ofile	Point:(EPSG:4326)-13.23,-9.62
1-001A	Sample Profile	Point:(EPSG:4326)-13.23,-9.62
10-003	Sample Profile	Point:(EPSG:4326)-13.325,-9.107
10-001	Sample Profile	Point:(EPSG:4326)-13.325,-9.107
		SG:4326)-13.325,-9.107
		SG:4326)-13.268,-8.615
		PSG:4326)-13.22,-8.37
		PSG:4326)-13.22,-8.37
		PSG:4326)-13.22,-8.37
44-005	Sample Profile	Point:(EPSG:4326)-13.22,-8.37

Sample info
Sample ID:END0061-007-001

Location:
Latitude: 7.507°S
Longitude: 13.457°W
Elevation: -4030
Tectonic setting: SPREADING_CENTER
Location: MID-ATLANTIC RIDGE:SPREADING_CENTER
Location Comment:

Sample Description:
Rock type:
Classification:
Description:
Alteration:
Age:
Archived at:

Sampl...
Cruise: EN061... Scientist:
(ENDEAVOR, 1981) Technique: Dredge
Date: Station: END0061-007

Rock Data

Reference	Alias	Method	Material	Sr87_Sr86	Nd143_Nd144	Pb206_Pb204	Pb207_Pb204	Pb208_Pb204	He	He3_He4(R/Ra)
HANAN, 1986	EN061 7D-1	MS	Glass			18.888	15.582	38.517		
FONTIGNIE, 1998	EN061 7D-1	MS	Glass	0.702454	0.513159					
GRAHAM, 1992	EN061 7D-1	MS	Glass						6.1E-6	7.88
GRAHAM, 1992	EN061 7D-1	MS	Glass						7.71E-6	7.48

WFS (web feature services) for PetDB stations

Ambiguous Naming of Rock Samples

Name	Location	Publication	Cruise
D3-1	SEIR	*ANDERSON, 1980*	VM3301 (Vema)
D3-1	North Fiji Basin	*EISSEN 1994*	Starmer 1 (Nadir)
D3-1	Shimada Smt	*GRAHAM 1988*	S1-79 (Sea Sounder)
D3-1	Gorda Ridge	*CLAGUE 1984*	KK2-83NP (Kana Keoki)
3-1	Lamont Smts	*BATIZA 1982*	RISE III (New Horizon)

Sample names are duplicated.

Sample names are modified or changed.

Dredge sample 3, Amphitrite Cruise 1963/4	
D3	*Engel 1964*
D-3	*Scheidegger 1981, Schilling 1971*
PD3	*Tatsumoto 1965, 1966*
PD-3	*Hedge 1970, Muehlenbach 1972*
PV D-3	*Engel 1965*
AMPH3D	*Pineau 1976*
AMPH-D3	*MacDougall 1986*
AMPH D-3	*Sun 1980, Schilling 1975*
AMPH 3-PD-3	*Hart 1971*
S-10	*Subbarao 1972*

DSDP Leg 46, Hole 396B, Section 22, Sample 3, 28-33cm	
46396B 22 3,28-38	*Dungan 1978*
396B 22 3,28-38	*Muehlenbach 1979*
249	*Dungan 1978*
DSDP046-0396B-022-003/28-38	*PetDB*

Examples from the PetDB Database

The International Geo Sample Number IGSN

IGSN :SIO001324

Unique user code String of random characters

- Structure
  - 9-digit alphanumeric string
  - First three characters are unique user code (registered with SESAR)
  - Last 5 characters are random (4th digit is contingency)
  - Allows 2,176,782,336 sample identifiers per registrant
- Managed at a central registry (SESAR)
  - Generated by SESAR or by registered users (upon request).
  - Strict compliance with the IGSN structure required.
- Applied in sample curation, data publication, & digital data management.
- Does not replace personal or institutional names.

SESAR System for Earth SAmple Registration

About SESAR || Partners || News || Contact Us

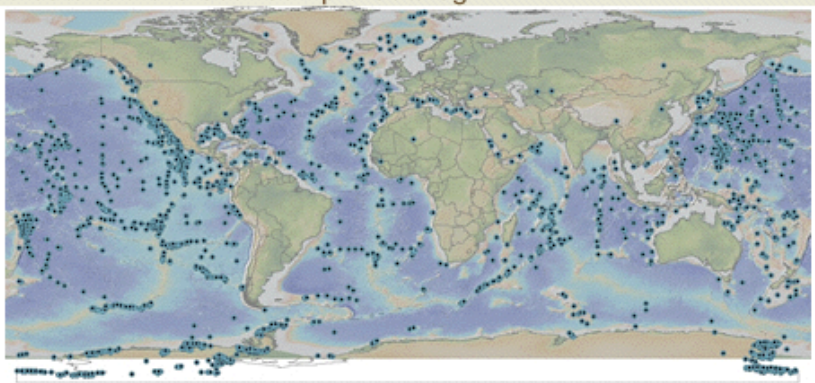
- How to use SESAR
- Logout
- My SESAR
- Query Sample Catalog
- For Developers
- Bulk sample registration forms
- SESAR Home

SESAR is a centralized registry that provides and administers unique identifiers for Geoscience samples - the International Geo Sample Number IGSN.

Use of the IGSN will help to avoid ambiguity, systematize sample designation, and ensure that all information associated with a sample is preserved and accessible on a global scale.

Unique sample identification will aid the preservation and curation of samples, and facilitate sharing of samples and sample-based data. By registering your samples, you automatically contribute to building a global catalog of Geoscience samples.

Contribute to the Global Sample Catalog!



This image
SESAR.

IGSN
International Geo Sample Number

The IGSN as a persistent unique identifier for all types of Geoscience sample, e.g. cores, rocks, minerals, fossils, and fluids, that is used in sample curation, data publication, and digital data management.
More ...

News

The US Polar Rock Repository at the Byrd Polar Research Center of Ohio State University has registered more than 7,000 samples collected in Antarctica since 1964. More .

Provide and manage unique identifiers for samples

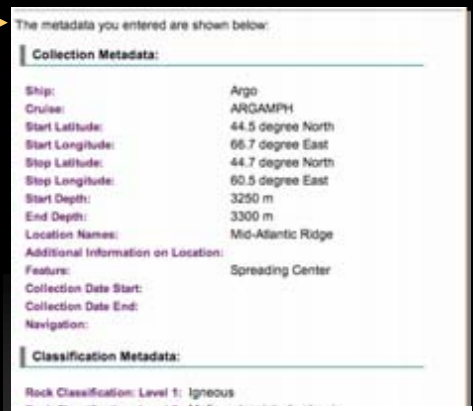
to support global sharing, linking, and integration of information and data about these samples.

www.geosamples.org

SESAR Services

- ✓ Administration of the IGSN
- ✓ Personal/institutional? sample management ("MyGeosamples")
  - Preserve information about your samples
  - Store field notes, photos, etc.
  - Create maps of sample locations
- ✓ Access to a Global Sample Catalog
- ✓ Sample Profiles
  - View all metadata about a sample
  - Discover all online data and publications for a sample

Integrating Sample Data and Metadata via IGSN in a DOI Framework in a Distributed Geoscience System Environment


The metadata you entered are shown below:

Collection Metadata:

Ship:	Argo
Cruise:	ARGAMPH
Start Latitude:	44.5 degree North
Start Longitude:	66.7 degree East
Stop Latitude:	44.7 degree North
Stop Longitude:	60.5 degree East
Start Depth:	3250 m
End Depth:	3300 m
Location Names:	Mid-Atlantic Ridge
Additional Information on Location:	
Feature:	Spreading Center
Collection Date Start:	
Collection Date End:	
Navigation:	

Classification Metadata:

Rock Classification: Level 1: Igneous

Data for this sample can be found at
doi: 10.xxxx/….1
doi: 10.xxxx/….2
doi: 10.xxxx/….3
doi: 10.xxxx/….4
doi: 10.xxxx/….5

doi: 10.xxxx/IGSN.KL1000005

Data for IGSN.KL100005

Data for IGSN.KL100005

Data for IGSN.KL100005

Data for IGSN.KL100005

Data for IGSN.KL100005

doi: 10.xxxx/..1

doi: 10.xxxx/..2

doi: 10.xxxx/..3

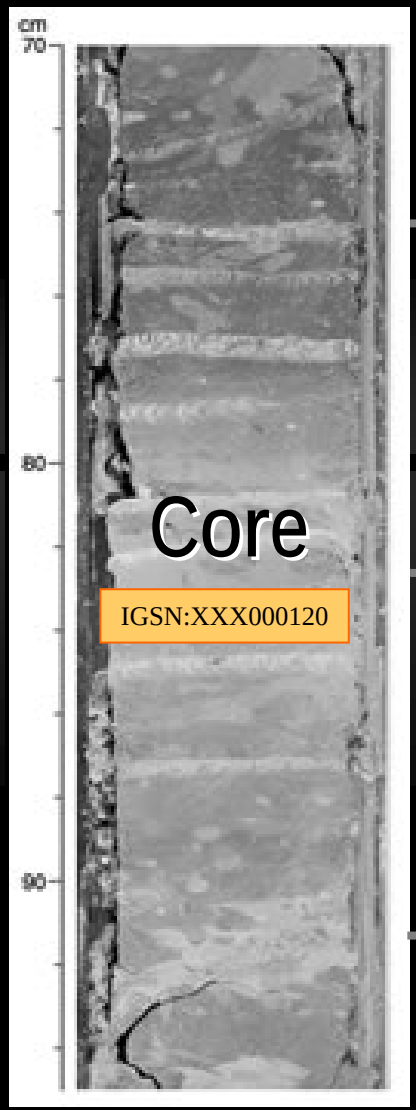
doi: 10.xxxx/..4

doi: 10.xxxx/..5

Sample Types & Granularity

- Any type of Geoscience ‘Object’:
  - “Sampling events” such as holes, wells, dredges, stratigraphic sections
  - “Individual samples”: rocks, minerals, fossils, fluid samples, precipitates, synthetic material, etc.
  - “Sub-samples” of any of above: mineral or fossil separates, leachates, pore fluids, etc.

Granularity & Relations of Registered Samples



Parent ⟷ Child

Parent ⟷ Child

Parent ⟷ Child

- Core Section 1 (IGSN:XXX0065B3)
  - Sample 1
    - Fossil separate
    - Microprobe mount
  - Sample 2
- Core Section 2 (IGSN:XXX07ST4K)
  - Sample 1
  - IGSN:ABC0L53NW
    - Rock powder (IGSN:ABC0L653X)
    - Mineral conc. (IGSN:ABC078HGB)
    - Leachate (IGSN:ABC0L98SW)
  - Sample 3
- Core Section 3 (IGSN:XXX9K23G6)
  - Sample 1
  - Sample 2
  - Sample 3 (IGSN:XYZ0G693M)

Implementation Challenges

- Diversity of applications
  - Sample types
  - Sample relations
  - Legacy vs new samples
  - Sample management
- Diversity of users
  - individual investigators
  - structured science & field programs
  - museums & repositories
  - IT

- *Who should register: e.g. curators, collectors, investigators, data management systems?*
- *When should a sample be registered: in the field, during curation, at time of data publication?*
- *At what granularity should samples be registered?*
- *How far does SESAR need to support sample management?*

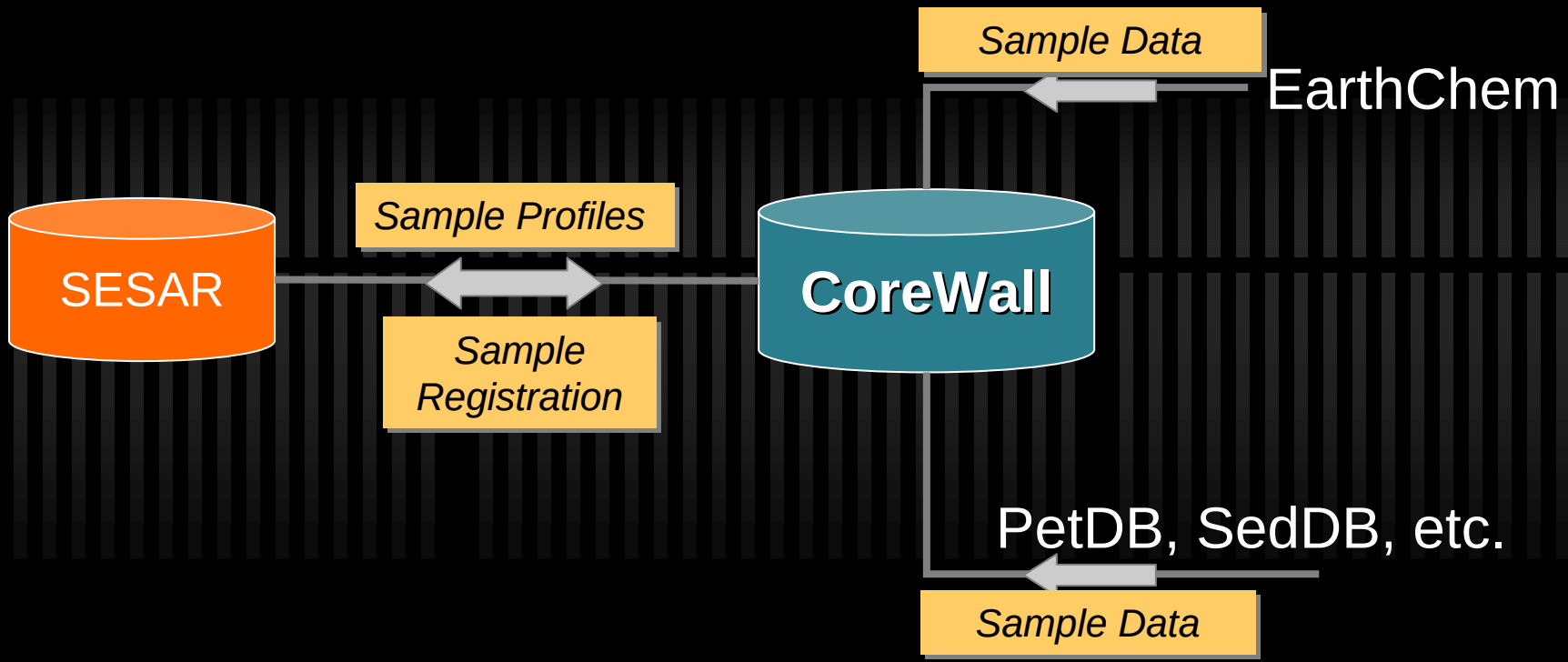
Collection Registrations

- ✓ Scripps dredge & core collection
- ✓ US Polar Rock Repository
- ✓ DSDP/ODP*
- ✓ LDEO core & dredge collection*
- ✓ WHOI core & dredge collection
- ✓ Newark Basin cores*
- ✓ ARF Core collection*
- ✓ Harvard rock & mineral collection
- ✓ LacCore Repository

- ✓ State Geological Surveys KS, AZ, DE
- ✓ Bureau of Economic Geology, TX
- ✓ Energy Research Center, UT
- ✓ GEOMAR, MARUM (Germany)

* Sample metadata submitted

The CoreWall Connection III



Data exchange via REST/SOAP web services