



# TOWARDS A WEB-ENABLED GEO-SAMPLE WEB: AN OPEN SOURCE RESOURCE REGISTRATION AND MANAGEMENT SYSTEM FOR CONNECTING GEO-SAMPLES TO THE WEB

Anusuriya Devaraju<sup>1</sup>, Jens Klump<sup>1</sup>, Victor Tey<sup>1</sup>, Simon Cox<sup>2</sup> and Ryan Fraser<sup>1</sup>

International Conference for Free and Open Source Software for Geospatial (FOSS4G 2017), August 14-19, 2017, Boston, MA.

<sup>1</sup>MINERAL RESOURCES, <sup>2</sup>LAND AND WATER

[www.csiro.au](http://www.csiro.au)



# Presentation Outline

- Challenges in sample identification and discovery
- International Geo Sample Number (IGSN)
- IGSN Implementation in CSIRO
- Applications
- Lessons Learned and Conclusions

# Challenges

- Inconsistent sample cataloguing practices.
- The lack of online catalogues for discovering physical samples.

The screenshot displays the EarthChem data portal. On the left, a world map shows sample locations marked with colored dots. The right panel provides detailed information for a specific sample. The 'Other Names' field is circled in red, showing a list of identifiers: '140-504B-214R-1, 73-76(ALT, 1996)', '140-504B-214R-1, 73-76(BACH, 2003)', '504B-214R1, 73-76(PEUCKER-EHRENBRINK, 2003)', '140-504B-214R-01, 73-76(ZULEGER, 1995)', and '140-504B-214R-1, 73-76(ALT, 1995)'. The 'Identification' field shows the PetDB Identifier: ODP0140-0504B-214R-001/073-076. The 'Sample Description' field lists the rock type as 'igneous volcanic mafic' and the classification as 'DIABASE(BACH, 2003)'. The 'Sampling Information' field shows the field program/cruise as 'ODP-140' and the station as 'ODP0140-0504B'. The 'Location' field shows the latitude as 1.22695°N, longitude as 83.73045°W, elevation as -3474, and tectonic setting as 'OLD\_OCEANIC\_CRUST'.

The EarthChem data portal returns 71 samples called 'A-1'.

Different names representing the same rock sample on the PetDB (Petrological Database)

# Examples

## ◀ Date: Three Weeks Ago

[REDACTED] (... There are 2 bags of Pulp Samples addressed to 'Blackham ... Wed 12/07/... 19 KB ☐  
Administration Support Officer Property (WA, SA, NT & Irymple) CSIRO Business &

## ◀ Date: Last Month

[REDACTED] Is anyone expecting 2 x roller racks 50kg? - please come to... Wed 14/06/... 16 KB ☐  
mann Administration Support Officer Property (WA, SA, NT & Irymple) | CSIRO Business &

## ◀ Date: Older

[REDACTED] ... Looking for crude oil samples Thu 27/04/2... 10 KB ☐  
Hi All, I have a project for which I need a small batch of crude oil samples.. I need roughly 6 ranging

[REDACTED]  
[REDACTED] (... A pallet of Mineral Samples in core trays, has been put in B... Thu 13/04/2... 520 KB ☐  
Administration Support Officer Property (WA, SA, NT & Irymple) CSIRO Business &

[REDACTED] ... Missing sample Wed 22/02/... 94 KB ☐  
MESSAGE FROM: [REDACTED] Hi folks, I am missing a pretty important sample. It is a zoned

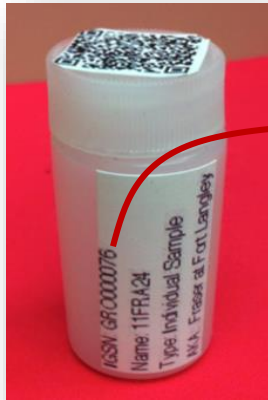


# Presentation Outline

- Challenges in sample identification and discovery
- International Geo Sample Number (IGSN)
- IGSN Implementation in CSIRO
- Applications
- Lessons Learned and Conclusions

# International Geo Sample Number (IGSN)

- Globally unique, persistent identifiers for physical samples (e.g., rocks, water, bio specimens), sample collections, and sampling features (e.g., boreholes).



Fraser River water sample  
(image source: @igsn\_info)

IGSN:GRO000076  
<http://hdl.handle.net/10273/GRO000076>

IGSN ensures  
preservation and access  
of sample metadata

A screenshot of the SESAR Sample Profile page for IGSN: GRO000076. The page shows a primary image of the sample, a QR code, and a table of metadata.

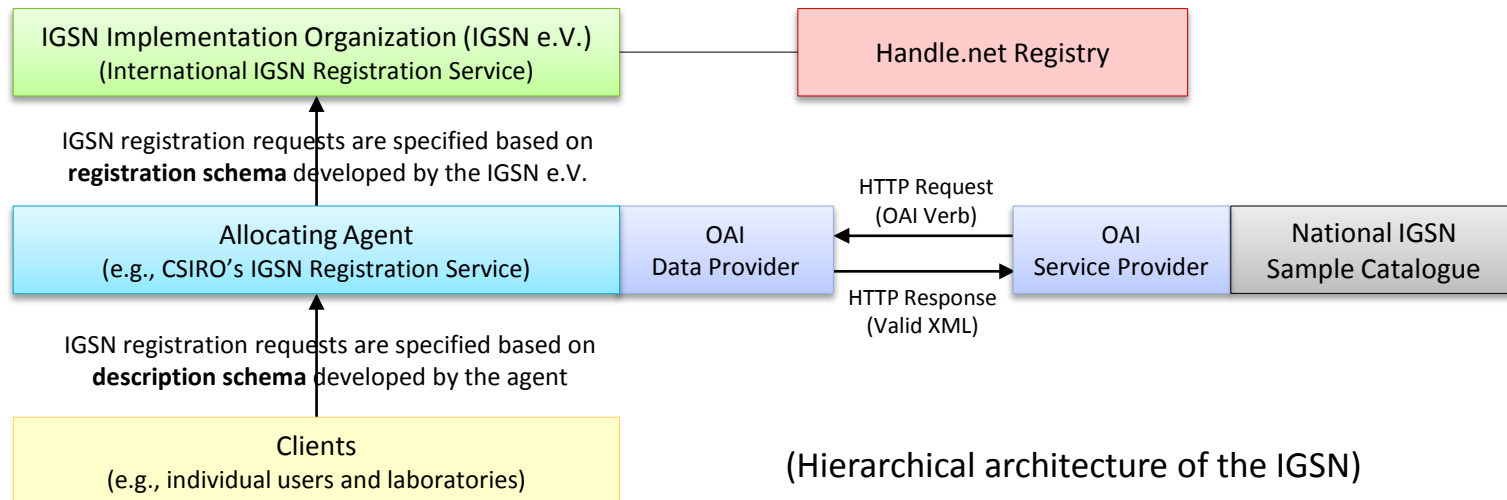
Description	
Material:	Liquid>aqueous
Classification:	Not Provided
Field Name:	Not Provided
Description:	Not Provided
Age (min):	Not Provided
Age (max):	Not Provided
Collection Method:	Not Provided
Collection Method Description:	Not Provided
Size:	Not Provided
Geological Age:	Not Provided
Geological Unit:	Not Provided
Comment:	Not Provided
Purpose:	Not Provided

Geolocation	
Latitude (WGS84):	49.18008
Longitude (WGS84):	-122.56723



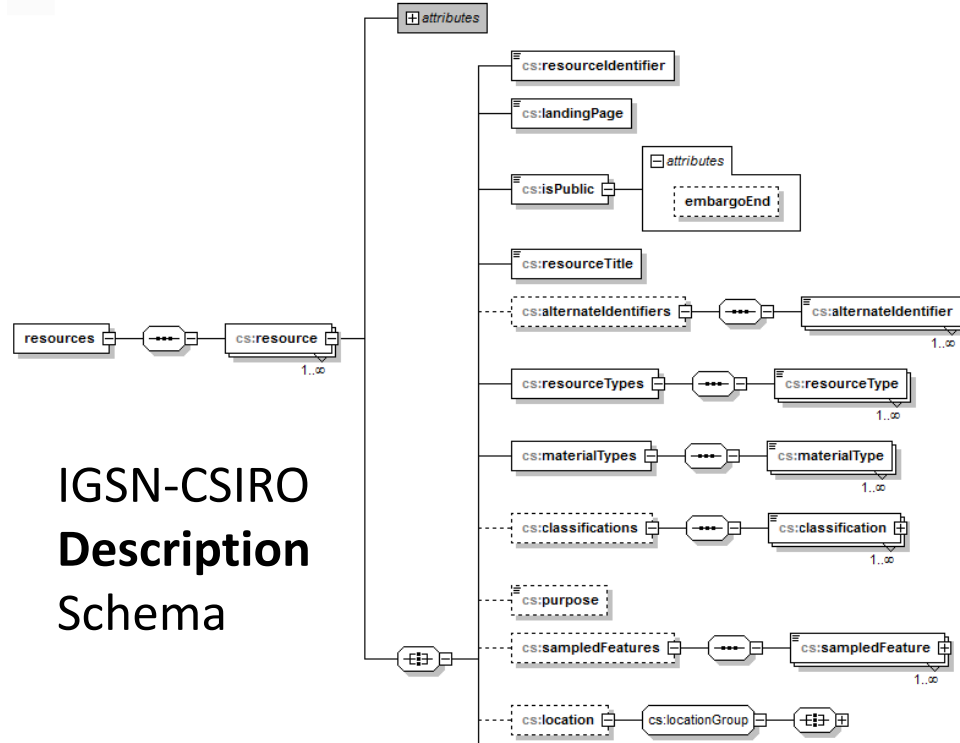
# IGSN System Overview

- Clients send sample registrations to an allocating agent; the allocating agent forwards the registrations to the IGSN e.V.
- Australian IGSN allocating agents are CSIRO, Geoscience Australia, Curtin Uni.

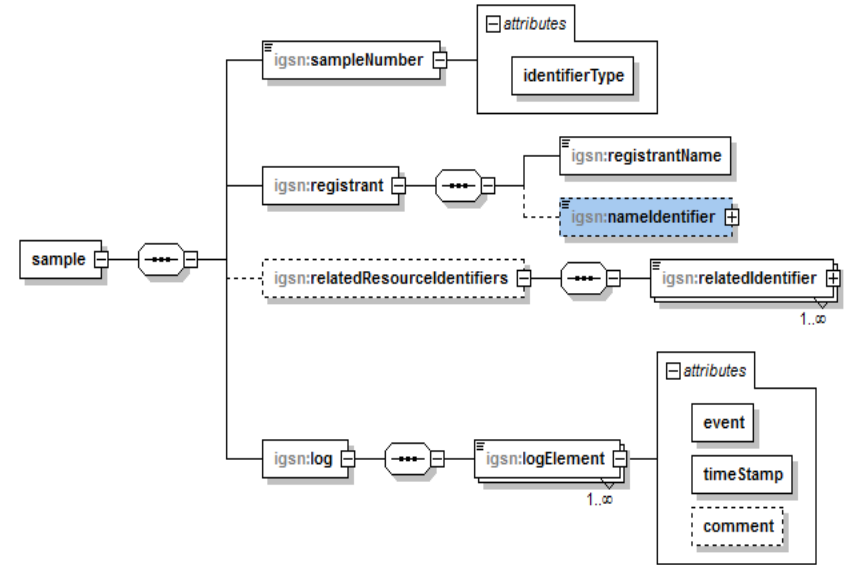


# Metadata Schemas

## IGSN-CSIRO Description Schema



## IGSN Registration Schema





# Presentation Outline

- Challenges in sample identification and discovery
- International Geo Sample Number (IGSN)
- IGSN Implementation in CSIRO
- Applications
- Lessons Learned and Conclusions

# IGSN Implementation in CSIRO

- CSIRO became a member of IGSN in 2013, and the use of IGSN was initiated by CSIRO Mineral Resources.
- IGSN are currently used for:

Sample Data Repositories	Material Types	Registrations (as at 31.05.2017)
Repository of the Australian Resources Research Centre(ARRC)	rock, mineral, soil	25652
Capricorn Distal Footprints	rock, vegetation, water, regolith	4232
Reflectance Spectra Reference Libraries	mineral, rock, synthetic material	94

- Future use cases are National Soil Archive and Australian National Insect Collection.

# Identifier Governance in CSIRO

- Different namespace strategies may be used by other allocating agents.
- Following the IGSN Technical Specification, we established rules for allocating IGSNs in different parts of the organisation.

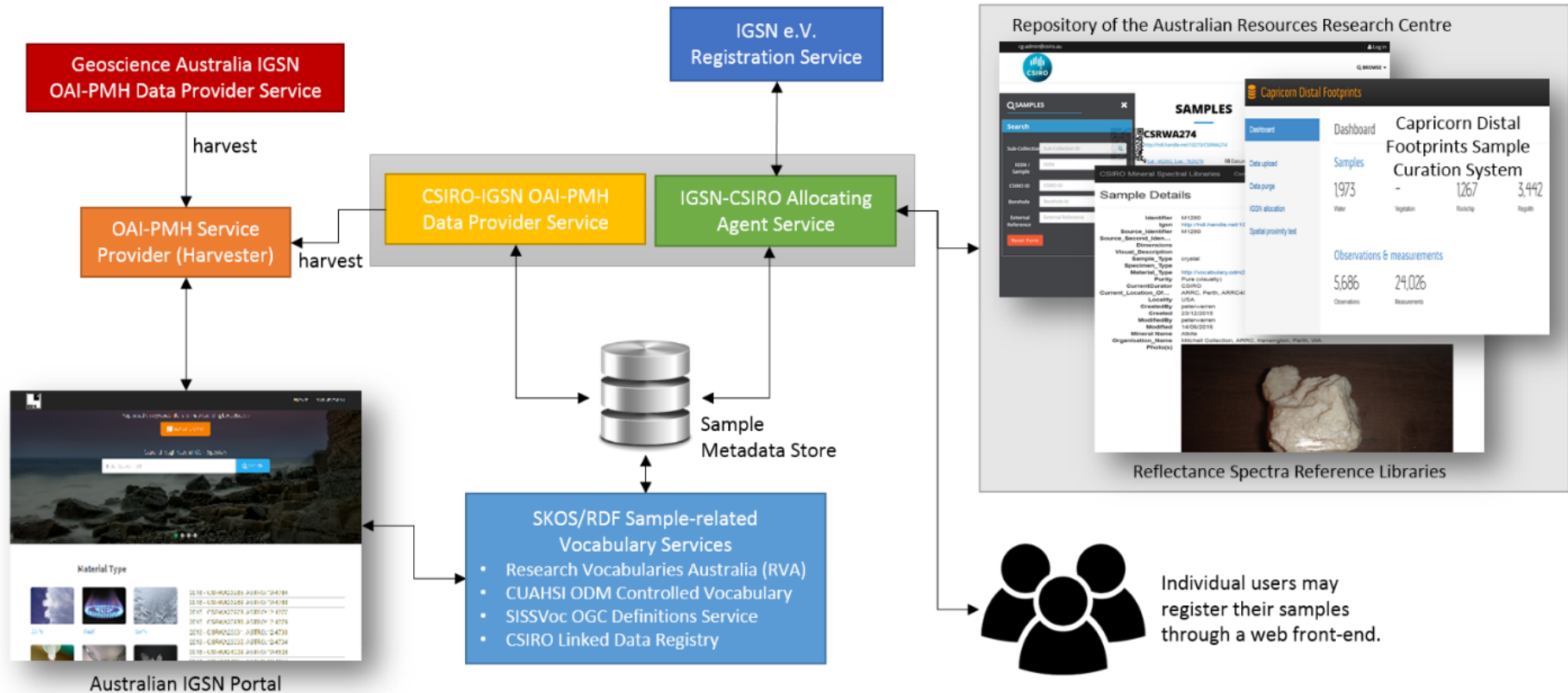


- 1 Allocating agent namespace ('CS' represents CSIRO)
- 2 Sub-namespace of a client
- 3 Sample code specified by the client

IGSN : CSRWASCO0630



# System Implementation



# A Web Front-End for Sample Registrations

Individual users may register their samples through the GUI.

The screenshot displays the CSIRO sample registration web front-end. The top navigation bar includes the CSIRO logo, a user profile icon, and a 'Log out Anusuriya' link. A 'Log out Victor' link is also visible. The main content area is divided into two sections: 'Metadata Information' and 'Registrants'.

**Metadata Information**

**Primary Info**

- Geosample:** A dropdown menu with a 'Tick Auto-gen to randomly generate ids' checkbox and a 'Use default' button.
- Registered Object Type:** A dropdown menu with a 'Select registered object type' label.
- Landing Page:** A text input field with a 'Use default' button and a 'Provide a title for message' label.
- isPublic:** A checkbox with a 'Select public access' label.
- Resource Type:** A dropdown menu with a 'Provide a title for message' label.

**Location**

- Locality:** A text input field with a 'Locality uri' label.
- Vertical Datum:** A text input field with a 'Geometry uri' label.
- Well Known Text:** A text input field with a 'Spatial Reference System Identifier(SRID)' label.
- Location well known text:** A text input field with a 'Well known text, eg POINT(100 100)' label.

**Curation Details**

- Curator:** A text input field with a 'Curator' label.
- Curation Date:** A date picker with a 'Curation Date' label.
- Curation Location:** A text input field with a 'Curation Location' label.

**Registrants**

Name	Email	Active
Victor	victor@csiro.au	Yes
Anusuriya	anusuriya@csiro.au	No
Victor	victor@csiro.au	No

**Registrants List**

- Victor:** Status: **Active** (button: **ADD SAMPLE**)
- Anusuriya:** Status: **Active** (button: **ADD SAMPLE**)
- Victor:** Status: **Active** (button: **ADD SAMPLE**)
- Anusuriya:** Status: **Active** (button: **ADD SAMPLE**)
- Victor:** Status: **Active** (button: **ADD SAMPLE**)
- Anusuriya:** Status: **Active** (button: **ADD SAMPLE**)
- Victor:** Status: **Active** (button: **ADD SAMPLE**)
- Anusuriya:** Status: **Active** (button: **ADD SAMPLE**)

Prefix and user registration can be performed via the user interface from the Registrant menu option.

# CSIRO-IGSN Description Metadata Schema

- Built upon the DataCite Metadata Schema (v4.0).
- An XML Schema which represents the core concepts of physical samples:
  - Sample identification, e.g., identifier, alternate identifiers, title, visibility, material, specimen type..
  - Sampling activity, e.g., location, time, purpose, method...
  - Sample curation, e.g., curator, curation location and time...
  - Related resources, e.g. publications, datasets...
- Supports batch sample registrations and has minimal restrictions on mandatory metadata elements.
- Captures the provenance of sample curation, and the relations between samples and related resources.
- Flexible in terms of representing spatial and temporal information.

# SKOS-based Controlled Vocabularies

Vocabularies	Examples	Service	Provider
Material types	soil, rock, vegetation	Master Controlled Vocabulary Registry for ODM2	CUAHSI
Specimen types	thin section, grab, dredge, cuttings	Master Controlled Vocabulary Registry for ODM2	CUAHSI
Nil-reason types	missing, unknown, withheld.	OGC definitions of nil reasons CSIRO	CSIRO
Contributor types	originator, custodian, point of contact.	Linked Data Registry	CSIRO
*Registration types	physical sample, sample collection, sampling features.	ANDS Research Vocabularies Australia	CSIRO & Geoscience Australia
*Identifier types	DOI, IGSN, LSID, ORCID	ANDS Research Vocabularies Australia	CSIRO & Geoscience Australia
*Relation types	isDerivedFrom, hasDocument, hasDigitalRepresentation.	ANDS Research Vocabularies Australia	CSIRO & Geoscience Australia

A list of existing and newly developed SKOS vocabularies. New vocabularies are indicated by asterisks\*.



# CSIRO-IGSN Allocating Agent Service

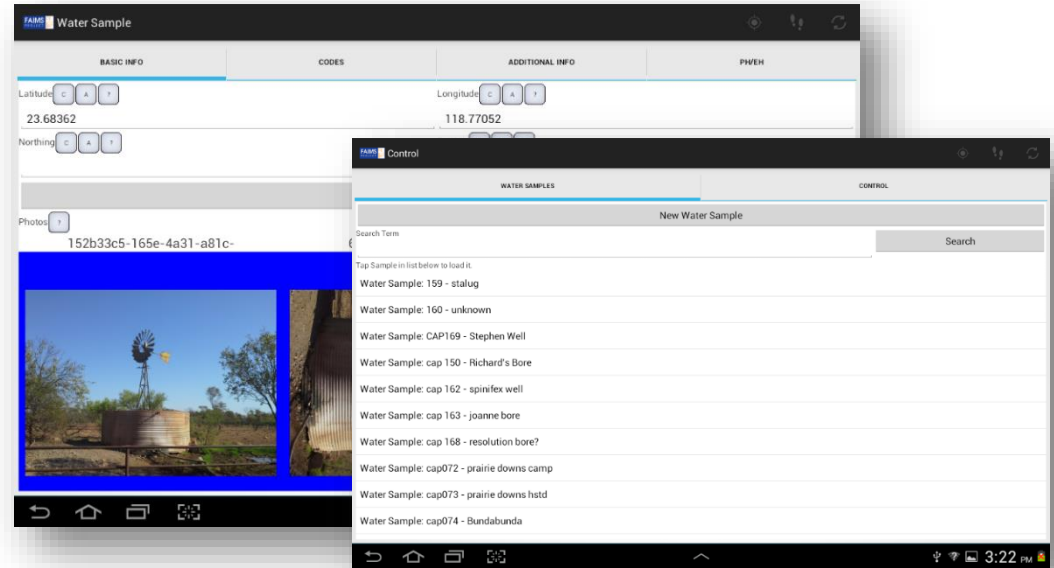
- Sample registrations must be specified based on the CSIRO-IGSN Description Schema.
- A RESTful web service that enable clients to
  - Register a sub-namespace
  - Get all registered sub-namespaces
  - Register samples
  - Get metadata of a sample or a list of samples
  - Mark a sample metadata as deprecated.

# Presentation Outline

- Challenges in sample identification and discovery
- International Geo Sample Number (IGSN)
- IGSN Implementation in CSIRO
- Applications
- Lessons Learned and Conclusions

# Applications

Tracking samples from the field to the sample repository.



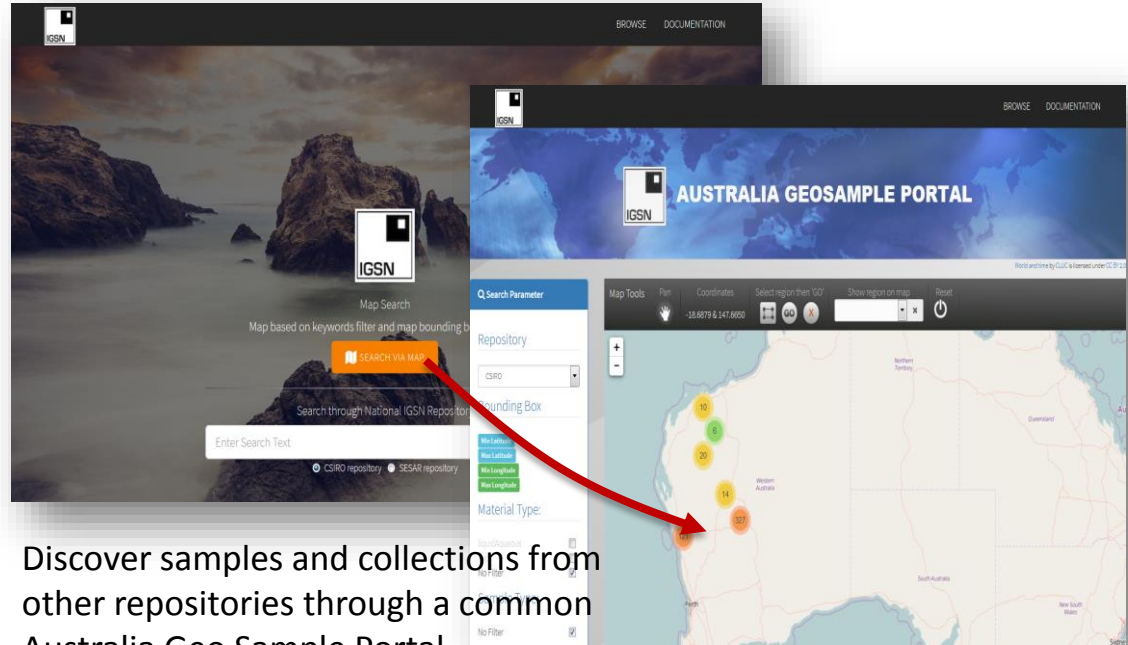
IGSN pre-allocation in mobile field data acquisition systems.

Images courtesy of Reid, N. (2015), Golodoniuc, P. (2017)

# IGSN Applications



Smart labels for sample inventory management.



Discover samples and collections from other repositories through a common Australia Geo Sample Portal.

# IGSN Applications

Reference samples by using IGSNs in publications, e.g., reports, articles, grant proposals.

ScienceDirect Journals Books

Download PDF Export Search ScienceDirect Advanced search

Article outline Show full outline

Abstract Abbreviations 1. Introduction 2. Background 3. Methods 4. Results 5. Discussion 6. Conclusions Acknowledgements Appendix A. Appendix B. Appendix C. Appendix D1. Appendix D2. Appendix E. Supplementary data References

Geochimica et Cosmochimica Acta  
Volume 122, 1 December 2013, Pages 101–126

ELSEVIER

Climate dependence of feldspar weathering in shale soils along a latitudinal gradient

Ashlee L. Dero<sup>a</sup>, Timothy S. White<sup>a,b</sup>, Richard H. April<sup>c</sup>, Brian Reynolds<sup>d</sup>, Thomas E. Miller<sup>e</sup>, Elizabeth P. Knapp<sup>f</sup>, Larry D. McKay<sup>g</sup>, Susan L. Brantley<sup>a,h</sup>

Show more

doi:10.1016/j.gca.2013.08.001

Get rights and content

Publication URL :

<http://dx.doi.org/10.1016/j.gca.2013.08.001>

Download PDF Export Search ScienceDirect Advanced search

Article outline Show full outline

Abstract Abbreviations 1. Introduction 2. Background 3. Methods 4. Results 5. Discussion 6. Conclusions Acknowledgements Appendix A. Appendix B. Appendix C. Appendix D1. Appendix D2. Appendix E. Supplementary data References

uncertainty in the stratigraphic position of SSHO, a rock sample from the bottom of our sampled site at Whipple Dam was used as parent material. Rock samples collected in Puerto Rico include shale from nearby road cuts and stream beds ( $n = 21$ ). At each outcrop, we used a Jacob's staff and an Abney Level to measure off the top of the Tuscarora Sandstone, a prominent stratigraphic marker and resistant landscape element, to estimate the location of soil pits within the stratigraphy of the Rose Hill Shale. Appendix C summarizes detailed rock sampling locations.

Table 2  
Major elemental chemistry of shale collected across the transect and corresponding depth of sample (d) where applicable. All rock samples were collected at local outcrops with the exception of PlymQ-RF and ALD-10-158, which were recovered from the bottom of soil pits and ALD-10-432, which is a weathered shale chip recovered from the bottom of the augered core.

Sample Site	name	IGSN <sup>a</sup>	d (m)	Al (%)	Ca (%)	Fe (%)	K (%)	Mg (%)	Mn (%)	Na (%)	P (%)	Σ (%)
Wales	PlymQ-RF	SSH000GG	0.35	12.0	0.04	6.77	3.15	1.41	0.43	0.57	0.04	2
	ALD-10-01	SSH000STR	–	11.8	0.03	6.40	2.96	1.36	0.33	0.68	0.04	2
	ALD-10-02	SSH000STS	–	11.0	0.01	5.77	2.88	1.23	0.21	0.61	0.04	2
	ALD-10-03	SSH000STT	–	11.8	0.01	6.73	2.87	1.47	0.23	0.62	0.05	2
	ALD-10-04	SSH000STU	–	11.6	0.05	6.45	2.66	1.63	0.13	0.77	0.06	2
	ALD-10-06	SSH000STW	–	11.9	0.05	6.17	2.91	1.53	0.21	0.76	0.06	2
	ALD-10-07	SSH000STX	–	11.7	0.01	6.45	2.93	1.33	0.29	0.62	0.05	2
	ALD-10-08	SSH000STY	–	11.9	0.02	6.82	2.94	1.50	0.62	0.62	0.06	2
	ALD-10-09	SSH000STZ	–	11.2	0.07	6.51	2.77	1.41	0.27	0.64	0.06	2
	ALD-10-33	SSH000SU0	–	11.7	0.04	6.32	2.97	1.41	0.21	0.64	0.06	2



# Presentation Outline

- Challenges in sample identification and discovery
- International Geo Sample Number (IGSN)
- IGSN Implementation in CSIRO
- Applications
- Lessons Learned and Conclusions



# Lessons Learned

- We ensure the global uniqueness of registered identifiers within the organization through the IGSN namespace governance.
- Integration of identifiers into new and existing systems
  - New sampling campaigns – adopt IGSN at the early stage of the sampling activity.
  - Existing sample curation systems – prepend the IGSN namespace to the local sample identifiers.
- What can be identified with IGSNs? We used IGSNs to identify not only physical samples but also sample collections and sampling features..
- Outreach, e.g., talks, meetings, and workshops.
- Documentations & public repositories.
- National collaboration, i.e., a joint project with the other allocating agents in Australia (e.g., Geoscience Australia and Curtin University).

# Conclusions

- We described the implementation of IGSN to identify and publish physical samples and sample collections in CSIRO.
- This serves as a working example for implementing IGSN in a large organization.
- Benefits of implementing IGSN in CSIRO
  - Identify samples unambiguously and discover them easily.
  - Avoid duplicate sampling activities.
  - Promote re-use of the samples for new purposes.
- The use of IGSN in CSIRO may be expanded to other collections and operated as a central service by Information Management and Technology (IM&T) as part of CSIRO's National Facilities and Collections.

# Related Links

Components	Link
CSIRO Allocating Agent Service	<a href="https://igsn.csiro.au/igsn30/api">https://igsn.csiro.au/igsn30/api</a>
Description Metadata Schema	<a href="https://igsn.csiro.au/schemas/3.0/">https://igsn.csiro.au/schemas/3.0/</a>
Metadata Store (source)	<a href="https://github.com/AuScope/igsn30/tree/master/sql">https://github.com/AuScope/igsn30/tree/master/sql</a>
CSIRO-IGSN OAI-PMH Data Provider	<a href="https://igsn.csiro.au/igsn30/api/service/30/oai">https://igsn.csiro.au/igsn30/api/service/30/oai</a>
National IGSN Web Portal	<a href="http://igsn.org.au">http://igsn.org.au</a>
OAI-PMH Harvester and National IGSN Web Portal (source repositories)	<a href="https://github.com/AuScope/NatPortalIGSN">https://github.com/AuScope/NatPortalIGSN</a>
Sample Registration and Management GUI	<a href="https://igsn.csiro.au/igsn30">https://igsn.csiro.au/igsn30</a>

A list of the components developed and their links.

## Mineral Resources

Anusuriya Devaraju  
Postdoctoral Fellow  
e anusuriya.devaraju@csiro.au

## Land and Water

Simon Cox  
Chief Research Scientist  
e simon.cox@csiro.au

## Mineral Resources

Jens Klump  
Science Leader Earth Science Informatics  
E jens.klump@csiro.au

## Mineral Resources

Ryan Fraser  
Research Manager  
e ryan.fraser@csiro.au

## Mineral Resources

Victor Tey  
Senior Software Engineer  
E victor.tey@csiro.au

MINERAL RESOURCES  
[www.csiro.au](http://www.csiro.au)



PROVIDING AUSTRALIAN  
RESEARCHERS WITH WORLD-CLASS  
HIGH-END COMPUTING SERVICES

