### ESS Publication Workflow

Gareth Murphy Experimental Control and Data Curation 2018-07-03

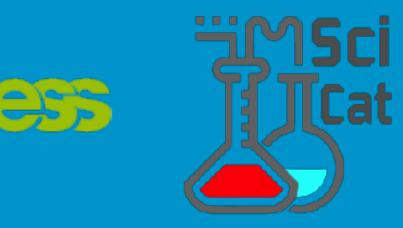








EUROPEAN SPALLATION SOURCE



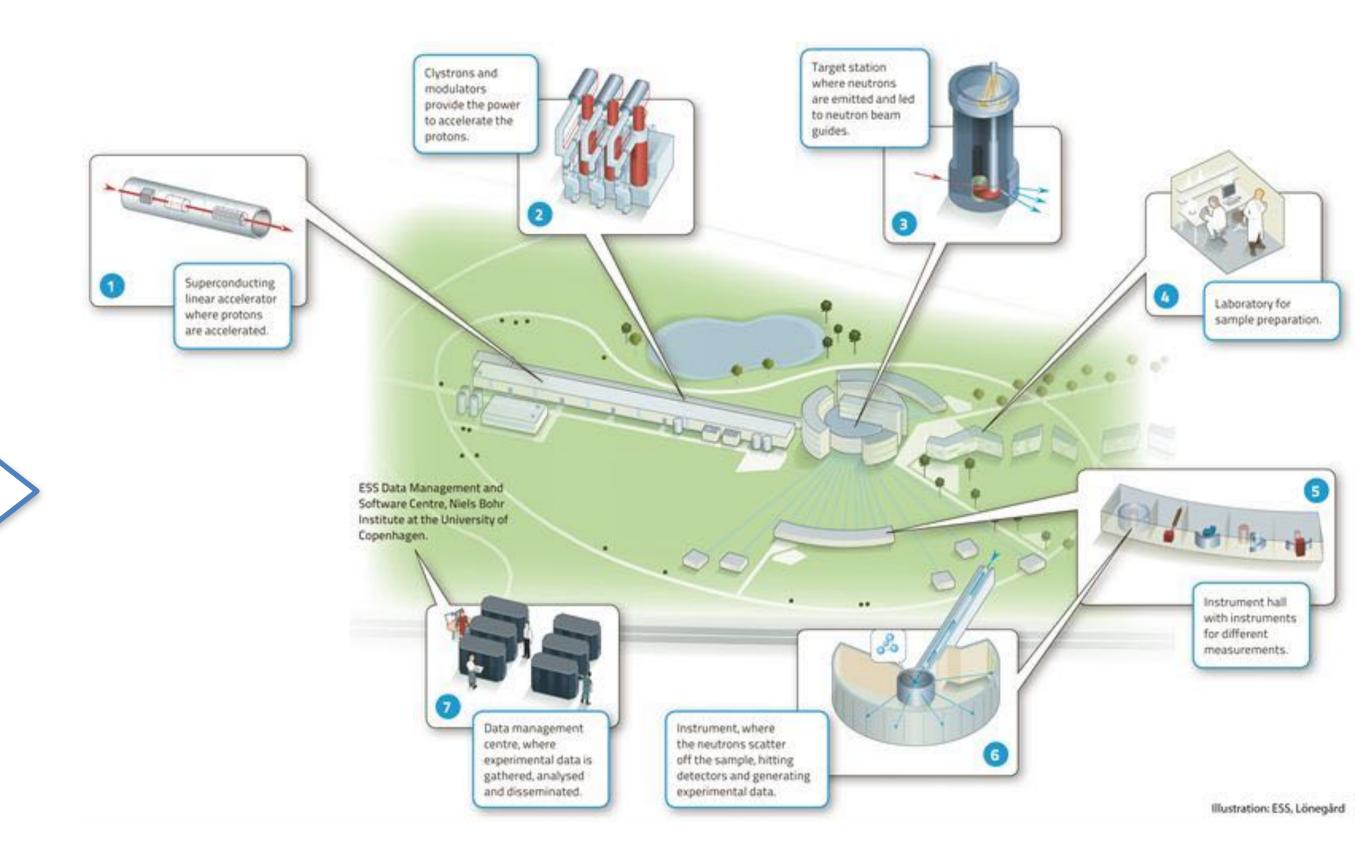
### Powerful pulsed neutron source

- 17 Partner countries
- Construction work in progress





EUROPEAN SPALLATION SOURCE





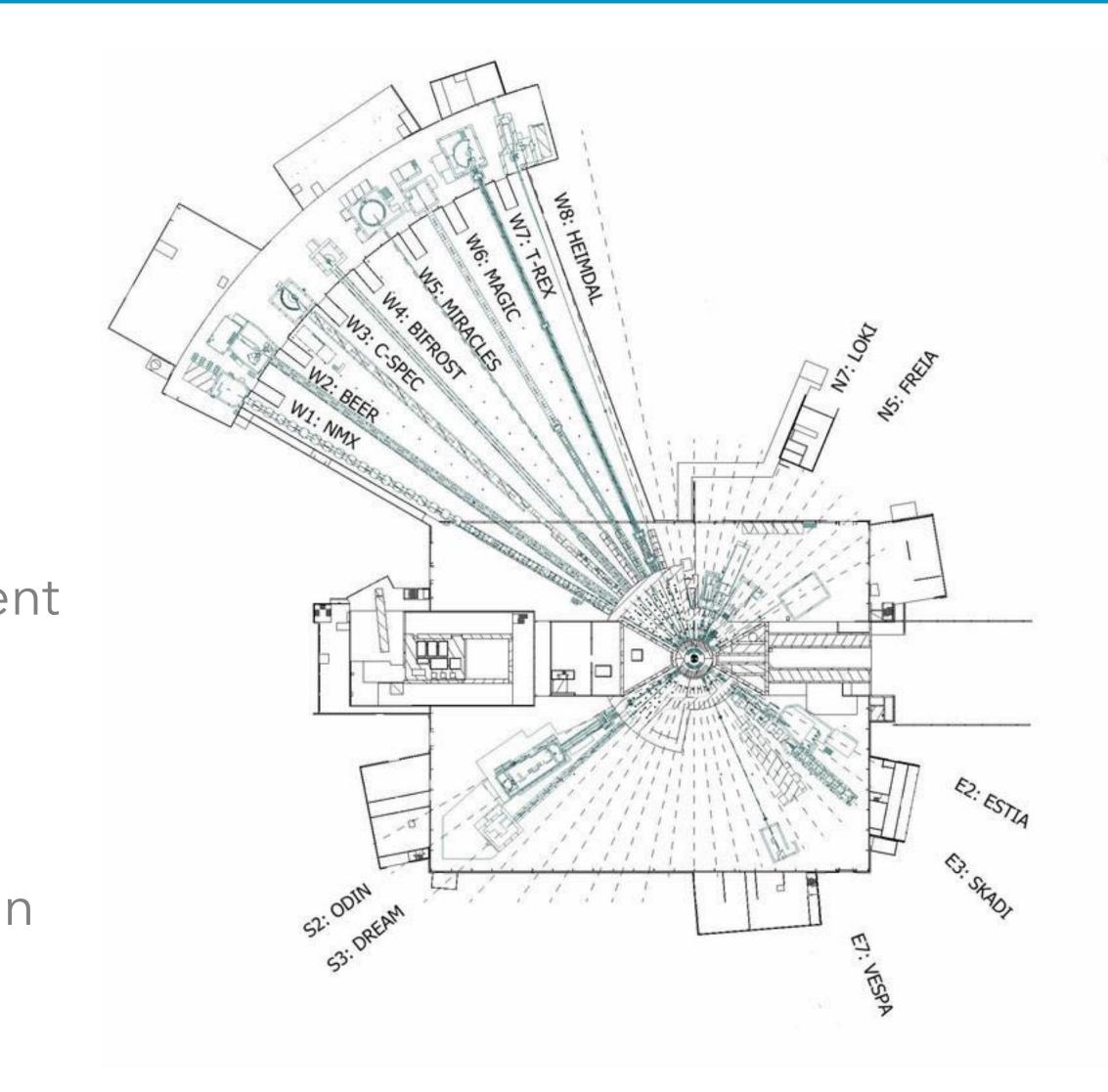


# **European Spallation Source**

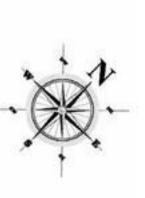
- European Spallation Source Scandinavia
- 15 instruments/beamlines
- Imaging, spectroscopy, diffraction
- Each instrument has different data requirements
- Traditionally, communities have had different data types, formats, analysis and reduction methods, standards - problem for data management
- By standardizing across instruments, we can make this process simpler and quicker



**EUROPEAN SPALLATION** SOURCE

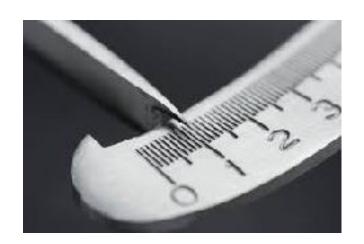


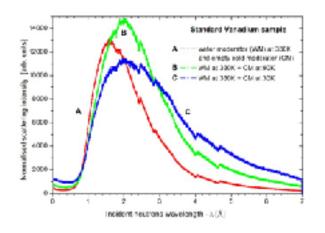




# Raw, reduced and derived data







- Raw data unprocessed data at full resolution, with communications artifacts removed (e.g. frame headers)
- Reduced transformed and corrected from instrument units to physical units,
- Derived data images, plots, statistics
- NASA define several processing levels raw = level 0, reduced = level 1, derived = level 2



**EUROPEAN** SPALLATION SOURCE





# Data Management & Software Centre (DMSC)



- DMSC one team to rule the data
- analysis
- catalogue

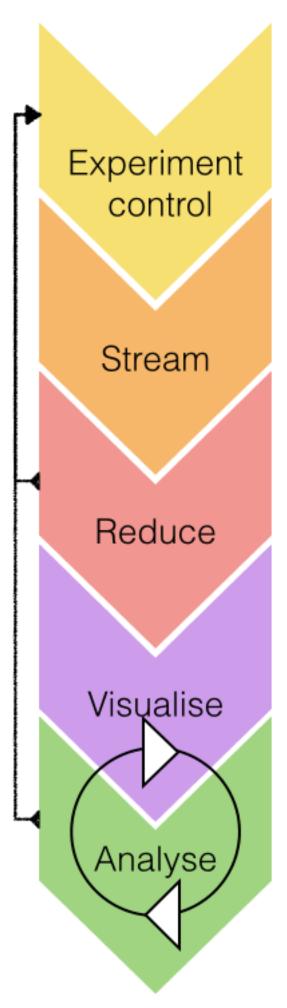


EUROPEAN SPALLATION SOURCE

 Create uniform file writer for every beam line Connect data acquisition to data reduction and

Create/acquire metadata and send to data

• Owner + ORCiD, time, wavelength, license, type





# brightness

- BrightnESS is a Horizon2020 program to support ESS
- Years of archives of legacy brightnESS data need to be supported
- 250,000 files, many different formats and types
- Metadata need to be preserved and made accessible
- Test case for SciCat



EUROPEAN SPALLATION SOURCE



### ESS Data Policy

data and results.

- RAW AND META DATA COLLECTED BY ESS 3.
  - Raw data and associated meta data 3.1.

    - the point of creation of said container.
  - Security, curation, archival of raw and meta data 3.2.
    - 3.2.1.
    - 3.2.2.
    - 3.2.3.
    - cycle when the raw data was taken.
  - 3.3.

    - collected.
    - 3.3.4.
    - 3.3.5. data.
  - Open Access to Raw and meta data 3.4.



**EUROPEAN SPALLATION** SOURCE

3.1.1. DMSC will act as the custodian of raw data and associated meta data.

3.1.2. ESS shall provide a persistent identifier for containers of raw data and meta data at

Raw data will be read-only for the duration of their life time.

Raw and meta data will be stored in a defined format.

### ESS will curate a means to read the raw data and meta data.

3.2.4. Two (2) copies of the raw data and meta data will be kept in physically different locations for at least first five (5) years commencing at the end of an ESS operations

3.2.5. ESS will move raw data and meta data to archival facilities for long term curation for a minimum of 10 years after the restricted access period.

### Access to Raw data and associated meta data by ESS/PT

3.3.1. Access to the raw data and meta data shall be restricted to the original PT and ESS staff for the first three (3) years after the end of the ESS operations cycle when the raw data was taken. After this period, the raw and meta data will be open access.

3.3.2. ESS staff shall respect the confidentiality of the scientific research data for the restricted access period but may exercise the right to make public thereafter.

3.3.3. The PT shall be able to obtain an electronic copy of the scientific research data they

It is the responsibility of the PT to curate their electronic copy.

Publications including data collected at ESS will cite the persistent identifier of the



Organizati Document Date		ESS-0081403 10/05/2017
	3.4.1.	Three (3) years a generated, ESS s an open on-line
	3.4.2.	The PT may required by submitting a vertex of scient advisory commit
	3.4.3.	The PT may requ access period.
	3.4.4.	The online index sample names/c runs, volume of temperature ran the PI).
	3.4.5.	Access to the on with the PT, prov access as defined data policy.
	3.4.6.	Publications or o data from ESS w proposal numbe
4. Res	ULTS	
4.1	0	a archin of all rocu

4.1.

EUROPEAN SPALLATION SOURCE

after the end of an ESS operations cycle when the raw data was shall make public an index of the raw data and meta data collected in catalogue and notify PT of this event.

uest an extension of the restricted access period of up to three years, written request specifying the reasons for the extension to the ESS nce who will decide following the advice of the chair of the science ittee.

uest their data be made open before the expiration of the restricted

x shall catalogue raw data and associated meta data for example: compositions, instrument names and basic parameters, number of f data, sample environment equipment used, sample data itself, nges, etc., and the title from the original proposal with the name of

nline catalogue of open data will be given to a user, not associated oviding said user registers with ESS, fulfils the ESS requirements for ed in the ESS Access Policy, and accepts the terms of this scientific

other generated intellectual property resulting from access to open vill cite the original persistent identifier of the data, the PT and the er.

Ownership of all results (intellectual property) derived from the analysis of the raw data shall be determined by the specific user agreement, subject to the European Spallation Source ERIC Intellectual Property Rights and Inventions Policy.



# Landing pages

- accessible landing page, which displays metadata
- Need landing page server
- and acquire a DOI and landing page



EUROPEAN SPALLATION SOURCE

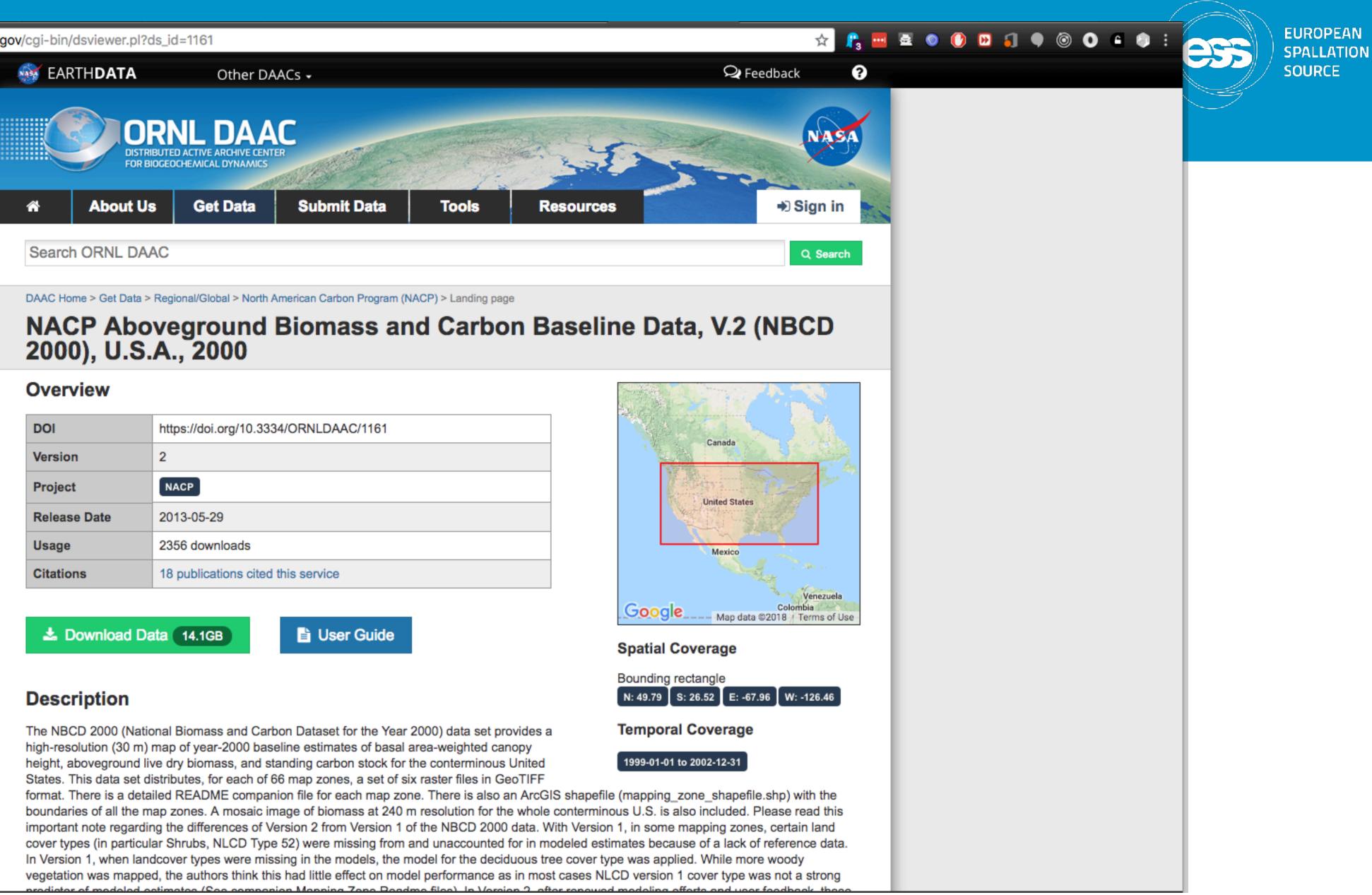
# Digital Object Identifier (DOI) must connect to

### <u>https://doi.org/10.17199/BRIGHTNESS.D5.1</u>

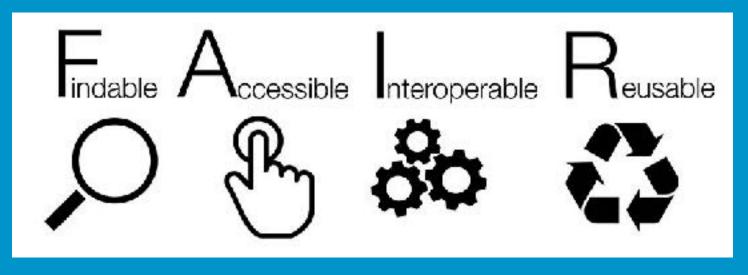
• Users should be able to make their data public



C ☆ Secure | https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds\_id=1161



DOI	https://doi.org/10.3334/ORNLDAAC/1161
Version	2
Project	NACP
Release Date	2013-05-29
Usage	2356 downloads
Citations	18 publications cited this service





### **To be Findable:**

- F2. data are described with rich metadata.
- F4. metadata <u>specify</u> the data identifier.

### **TO BE ACCESSIBLE:**

### **TO BE INTEROPERABLE:**

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (meta)data use vocabularies that follow FAIR principles. I3. (meta)data include <u>qualified references</u> to other (meta)data.

### **TO BE RE-USABLE:**

- R1.2. (meta)data are associated with their provenance.



**EUROPEAN** SOURCE

F1. (meta)data are assigned a globally unique and eternally persistent identifier. F3. (meta)data are registered or indexed in a searchable resource.

A1 (meta)data are retrievable by their identifier using a standardized communications protocol. A1.1 the protocol is open, free, and universally implementable.

A1.2 the protocol allows for an authentication and authorization procedure, where necessary. A2 metadata are accessible, even when the data are no longer available.

R1. meta(data) have a plurality of accurate and relevant attributes. R1.1. (meta)data are released with a clear and accessible data usage license. R1.3. (meta)data meet domain-relevant community standards



# Photon and Neutron Open Science Cloud (PANOSC)

- FAIR PaNOSC will comply with the FAIR principles in the following ways:
- Findable all data will have a DOI, rich metadata, common api for federated search
- Accessible api will support open protocol, metadata accessible even without data
- Inter-operable metadata to follow community standards (Nexus), register metadata
- Reusable follow community standardise metadata, clear licence (CC-BY)



**EUROPEAN** SPALLATION SOURCE



# Researcher persistent identifier

### ORCID

- Can uniquely identify researcher using instruments
- Can follow data use and citations
- Data creator/steward can be identified uniquely



D ORCID 🧜 🕎 🗟 📢 🜔 🖻 🌖 🌒 🎯 💽 🖴 🌒 C ☆ Secure | https://orcid.org/orcid-search/quick-search/?searchQuery="european%20spalla... ☆ English QQ Search ORCI FOR RESEARCHERS FOR ORGANIZATIONS ABOUT HELP SIGN IN SIGN IN REGISTER FOR AN ORCID ID LEARN MORE Connecting Research

4,963,023 ORCID iDs and counting. See more...

### Showing 40 of 93 results

and Researchers

.....

ORCID ID	First/given name	Last/family name	Other names	Affiliations
https://orcid.org/0000-0002-0078-0372	Nikolaos	Gazis		European Spallation Source
https://orcid.org/0000-0003-3797-0476	Konstantin	Batkov		European Spallation Source
https://orcid.org/0000-0001-5371-9199	Emanuele	Laface		European Spallation Source
https://orcid.org/0000-0003-0175-179X	Javier	Cereijo Garcia		European Spallation Source, European Spallation Source University of Vigo, University of A Coruña, University of Santiago de Compostela
https://orcid.org/0000-0002-0206-0387	Susan	Everett	S. M. Everett	Oak Ridge National Laboratory, European Spallation So ERIC, University of Tennessee
https://orcid.org/0000-0002-2109-1226	Mads	Bertelsen		
https://orcid.org/0000-0003-1875-4700	Chung- Chuan	Lai		European Spallation Source ESS AB, Linköping Universit National Tsing-Hua University
https://orcid.org/0000-0001-8688-4238	Masatoshi	Arai		European Spallation Source ESS AB, Tohoku University
https://orcid.org/0000-0001-5434-3728	Morten	Sales		
https://orcid.org/0000-0002-7015-1053	Mats	Lindroos		Lunds Universitet, European Spallation Source ESS AB, European Organization for Nuclear Research, Chalmers tekniska hogskola
https://orcid.org/0000-0001-8287-0269	Zoe	Fisher		European Spallation Source ESS AB, Los Alamos Nationa Laboratory, University of Florida, University of Stellenbo





nal

### Researcher logbooks

Position 19, 14, 13.5, 17(--) Reflected. A\_Reflected Beaus\_No: Heasurements on No. magnutic frige) 10 min. each Direct Beam ( Position: -16) B- black Beam : Measurements without any isagle Aurin Lis (43 file) (167. 15/ 2:80) Current / Total (yA) 35 min START (160 (244) STOP B: Direct Bean Scattering -t mastrements sport 2 core or Duct Beam C-Direct Beam : 5 mm Al plate in front of the window. 2 minutes (167 / 2.82) START (166.6 / 25.31) STOP to min (5 Hles) C: Tay Thickness Al window (166.05/1.63) START (166 /25.18) STOP domin (5 files)

← → C  localhast:000/group/ AS all prime services	prime-subscriber-delivery-missions ☆
CHANNELS 4	_
A general .	AS all.prime.services 14:40 @lee.manager: Drone 1 on mission AB wa
DIRECT MESSAGES +	FM fleet.manager 140fl (edited) (caliphines.ervices: Verfying services
PRVATE GROUPS +	AS al.ptime.services 14.41 glee.manager: Yeah please hurry up. Ort
E prime-subscribe-delivery-mi	FM fleet.manager 14.42 (clinel) (caliprime.services) Otrange! That :
	AS all.prime.services 14.44 giver.manager: You think it is a faulty care
	FM fleetmanager 14:45 (edited) (cali prime services) Standard open
	AS al.prime.services 1445 Pr@#/Elt is mission A3 waypoint 3 - ne way
	FM fleet.manager 1448 (calippine services): Just get OK for
	AS al.prime.services 14:0 gftee.manager: 50 what?Weblow it up or
	FM fleet.manager 14:51 (edited) (cali.prime.services :: We've got som
	AS al.prime.services 14:5 gnec.manager: granditied their theorem
	3dr.dromes 14:50 (call prime services): Yes 318
	AS al.prime.services 14:27 @3dr.drones: Dur drone 1 . model SoldWay
	3D 2dr.drones 14.18 (adited) (Caliprime services) WTEI That bird
	AS all prime services 14:59 all prime services 14:59 all prime services 14:59
DCKET.CHAT	de Morrage



**EUROPEAN** SOURCE

# How to capture researcher logs Use a web chat application like Rocket Chat to capture collaboration

	승 =
very-missions J	Drene 1: mission A3 waypoint 2 fuel 130%. Drene 2: mission A7 waypoint 2 fuel 130%.
D0160er17,2015	M mentys
ypdint 7 locks out of fuel. How can that be?	
ce records - piease hold.	When With Tween
ly Ji/h fael left, doesn't look good.	Dufulo WebWires
aird went missioning with 100% fuel only 2 minutes ago.	
sorgaupe? Hen much can ve twet 1/?	
ting procedule says we must trust the instrumentations. Call it back to base!	Google Man data 00015 Encode Terros of Use Report a man error
r in hell it is going to make it bank with 30% fuel!	ALTUAN TO BASE CANCEL
contingency services from HQ on Drone 1 mission A3 waypoint 7.	
t over a bush? Tankit into rotention.pond? Or crush and it for civilian recovery?	
ne MF inserance coverage! Even indirect forest fire up to 200 acres. De whatever you want!	
a. CK. Will confer with engineering. VD, this is a new proce from stor grad. drones	
TecLorwy, serial 229838 has a taulty fuel grage please advise of disposal.	
is still in our labs. It has not been FAA cleared for civilian airspace. How can you guys be flying it?	
ufacturing therr for you in Shenzhes, remember?	
"bold" "tatus, valuta» ("init	





### Summary

- SciCat already provides PIDs
- We still need DOI integration and landing pages
- SciCat publication workflow



EUROPEAN SPALLATION SOURCE

# ESS requires Open Access metadata and data

BrightnESS data can be a good test case for







### Open Access data

- Persistent Identifiers (PIDs)
- Digital Object Identifiers (DOIs)
- Landing page



SOURCE



C 🗋 🗎 Secure 🛛	https://daac.ornl.gov/cgi-bin/dsviewer.pl?	'ds_id=1161	🖈 👫 🔤 🛎 💿 🕐 🖸 (
	🚳 EARTH <b>DATA</b>	Other DAACs 🗸	Q Feedback ?
			NASA
		BIOGEOCHEMICAL DYNAMICS	
	About U	s Get Data Submit Data Tool	s Resources Sign in
	Search ORNL DA	AC	Q Search
	DAAC Home > Cot Date >	> Regional/Global > North American Carbon Program (NACP) > Land	
	2000), U.S		rbon Baseline Data, V.2 (NBCD
	Overview		
	DOI	https://doi.org/10.3334/ORNLDAAC/1161	THE REPORT OF A LOCAL
	Version	2	Canada
	Project	NACP	
	Release Date	2013-05-29	United States
	Usage	2356 downloads	Mexico
	Citations	18 publications cited this service	
			Venezuela Colombia
	📩 Download D	ata 14.1GB 🕒 🖹 User Guide	Colombia Map data ©2018 / Terms of Use
			Spatial Coverage
	D		Bounding rectangle
	Description		N: 49.79 S: 26.52 E: -67.96 W: -126.46
		ional Biomass and Carbon Dataset for the Year 2000) data map of year-2000 baseline estimates of basal area-weight	
	height, aboveground l	ive dry biomass, and standing carbon stock for the conterm	inous United 1999-01-01 to 2002-12-31
		distributes, for each of 66 map zones, a set of six raster file ailed README companion file for each map zone. There is	s in Geo I IFF also an ArcGIS shapefile (mapping_zone_shapefile.shp) with the
			for the whole conterminous U.S. is also included. Please read this 0 2000 data. With Version 1, in some mapping zones, certain land
			unted for in modeled estimates because of a lack of reference data.
		dcover types were missing in the models, the model for the	





