

ESS-DIVE Advanced Bootcamp

November 30th, 2020 10:00 - 11:00 am PDT



Which project or group are you representing today?

“Watershed Function SFA”

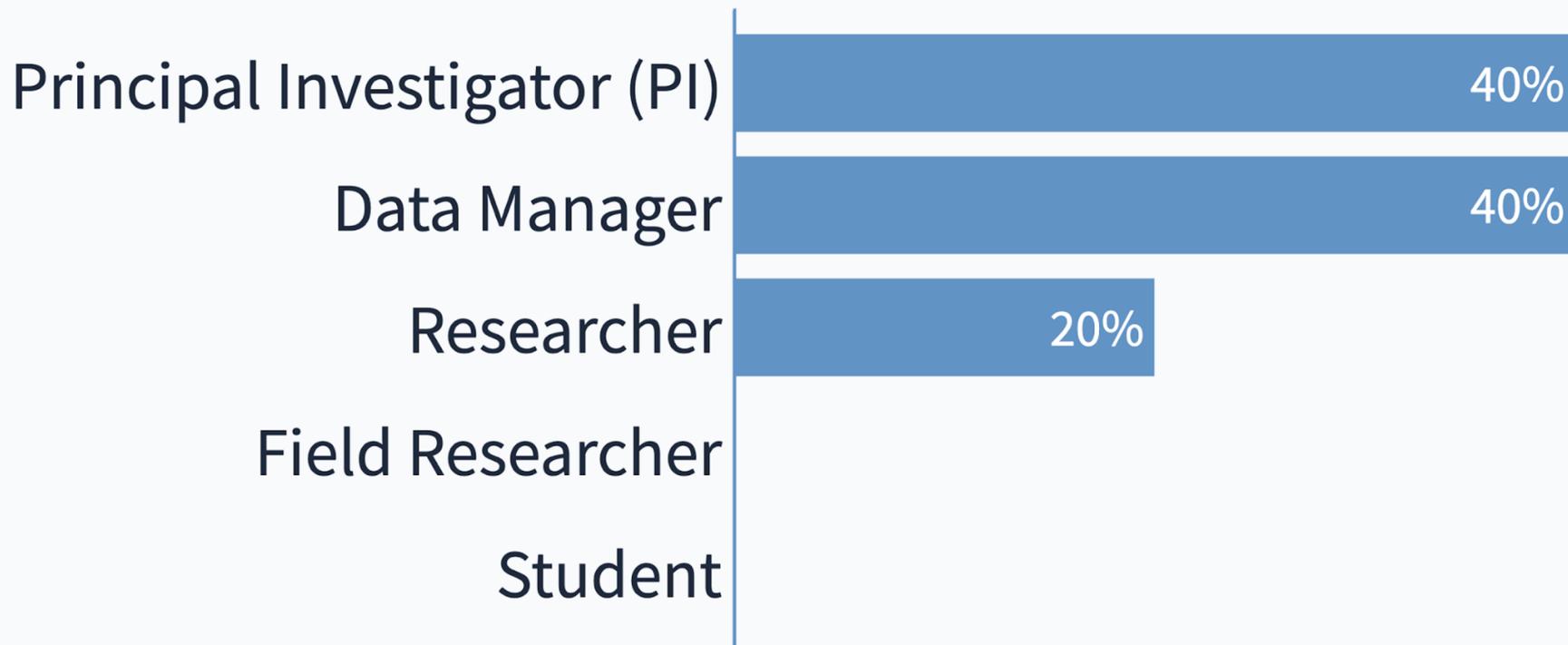
“NGEE-Arctic”

“ESS-DIVE”

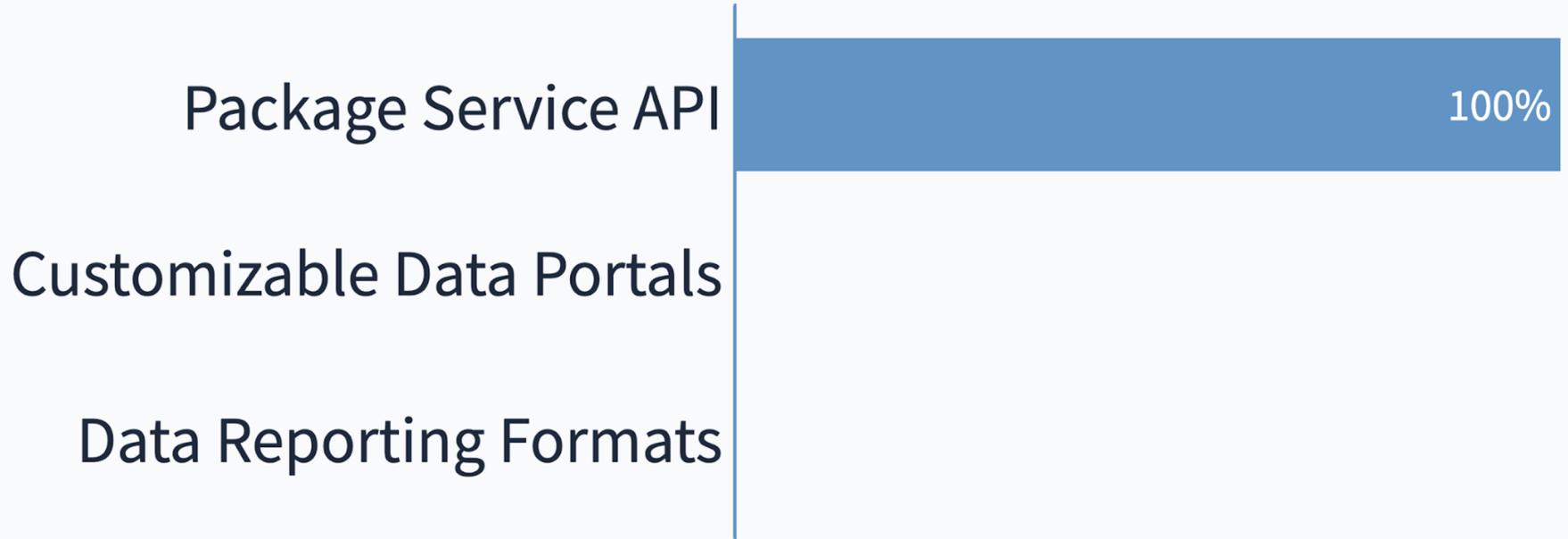
“Brookhaven National Lab (NGEE projects)”

“Various: TES at PNNL, ESS CF Partners project, ...”

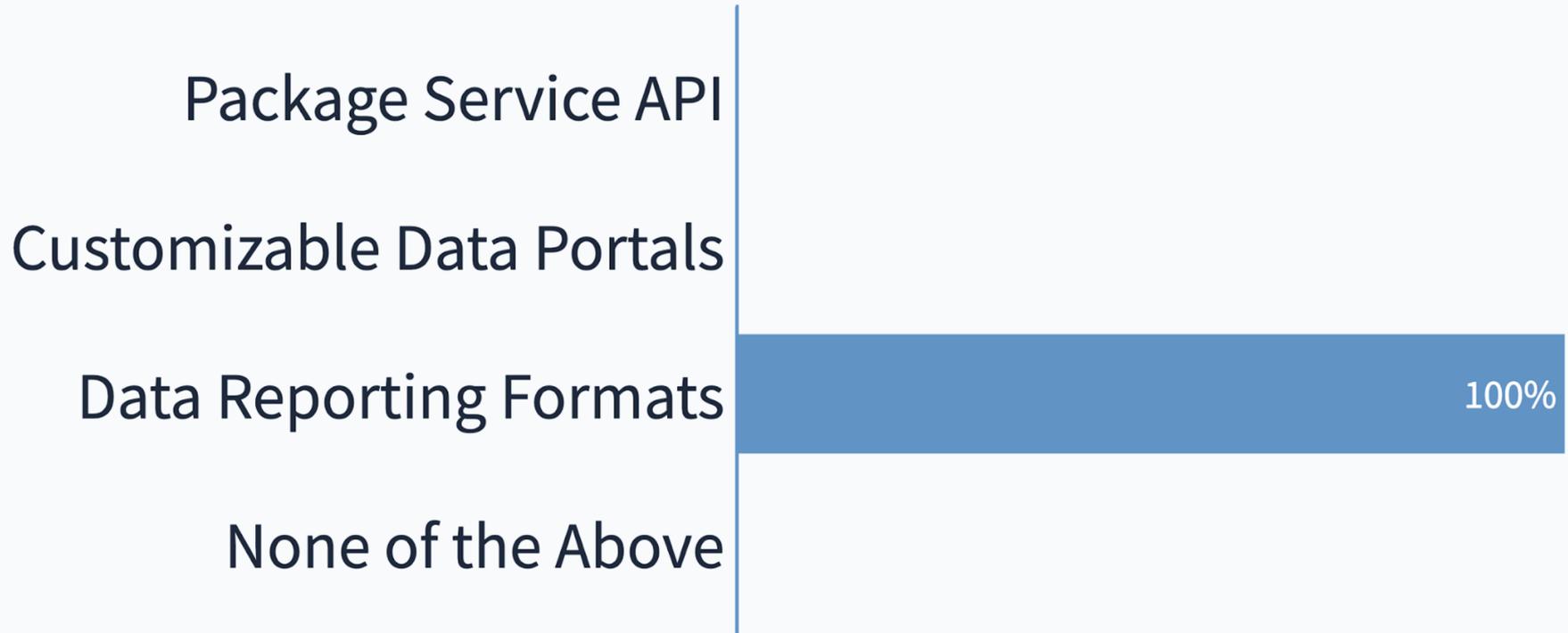
What is your role in your group?



Which session are you most interested to learn more about today?



Which of these topics are you familiar with?



ESS-DIVE Advanced Bootcamp

November 30th, 2020 10:00 - 11:00 am PDT

Madison Burrus, Rob Crystal-Ornelas, Fianna O'Brien

Zarine Kakalia, Emily Robles, Deb Agarwal, Charuleka Varadharajan, Shreyas Cholia, Valerie Hendrix, Joan Damerow, Hesham Elbashandy, Mario Melara, Makayla Shepherd, Maegen Simmonds, Cory Snively, and Karen Whitenack



ESS-DIVE November 2020 Webinar



Advanced Bootcamp Overview

- **Programmatic** submissions and updates to data packages using the Package Service API
- **Customizable data portals** to organize your data collections
- How to use the ESS-DIVE **community github** for reviewing and commenting on the ESS-DIVE proposed metadata and data **reporting formats**

This is an interactive webinar! We have demos, polls, and question breaks prepared.

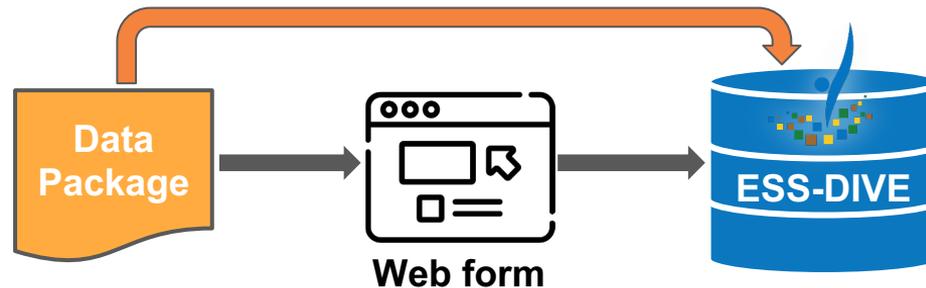
ESS-DIVE Package Service API

What is it?



*We've discussed upload via the web interface.
What if that becomes cumbersome?*

Through **ESS-DIVE Package Service API**, organizations can **write code to store, modify, and search data packages** on ESS-DIVE.



Package Service Use Cases

When to use this

- Uploading **large data**
- **Programmatic** upload
- **Bulk data** upload
- **Reusing information** across data packages, such as location data.

Who should use this?

- Projects with a large number of data packages
- Teams wanting to duplicate package metadata
- Projects with consistent metadata standards

What is your preferred coding language?

Activate 

Show new 

Lock 

Clear responses 

Full screen 

Top

| 1 | 

R

| 0 |

Bash shell

| 0 |

I'm not familiar with coding

| 0 |

R

| 0 |

C++

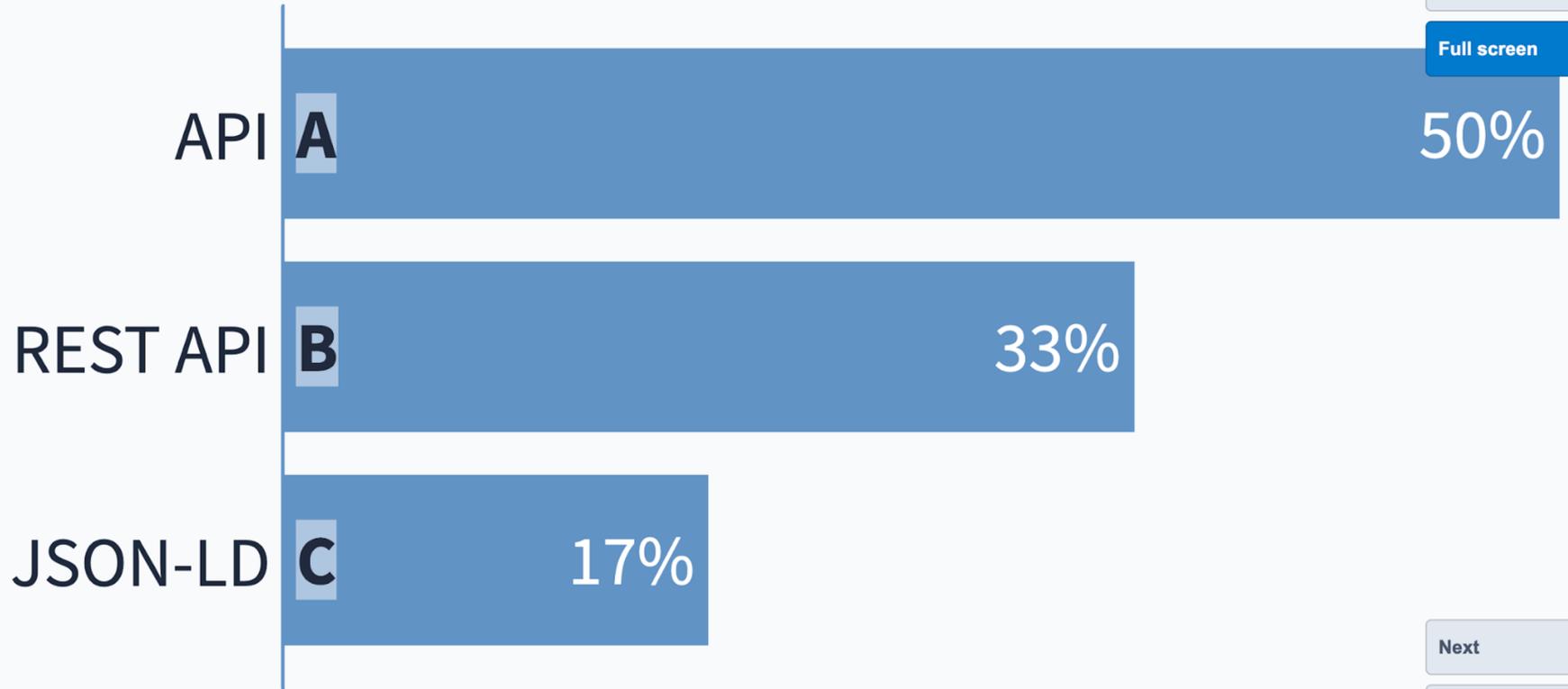
Which of these terms are you familiar with?

Show responses 

Lock 

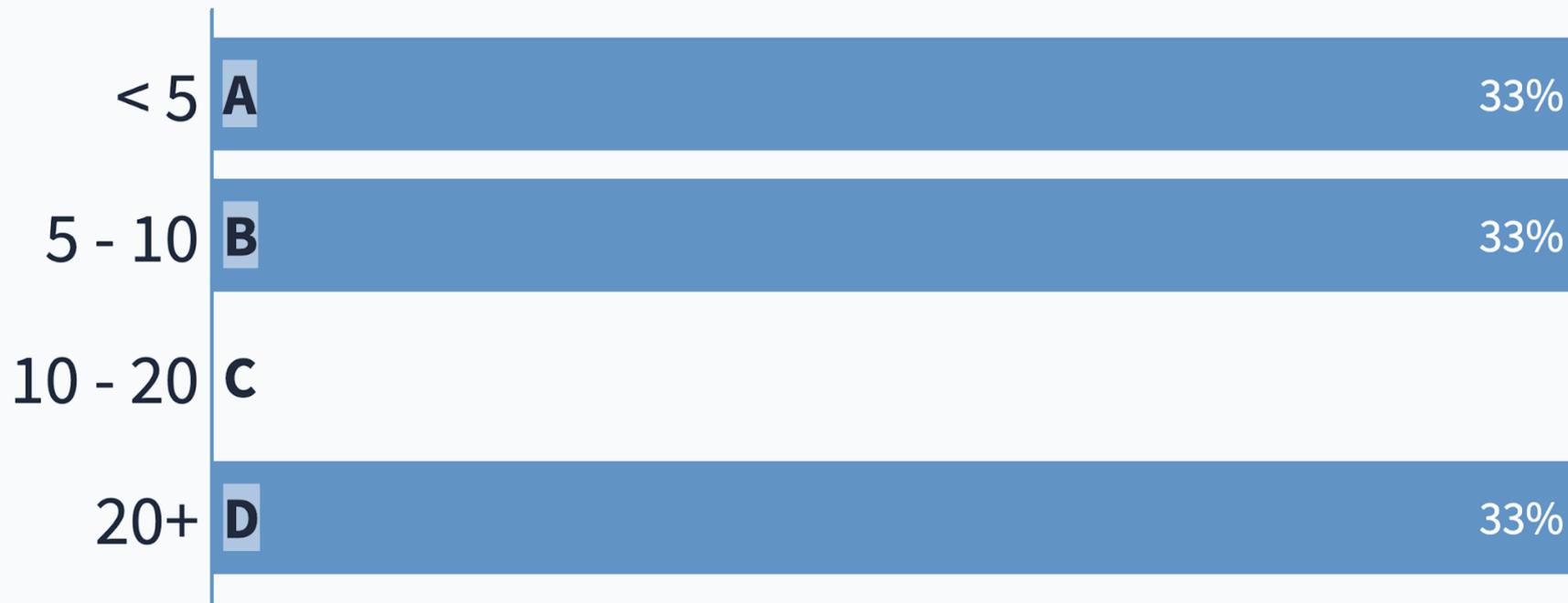
Clear responses 

Full screen 



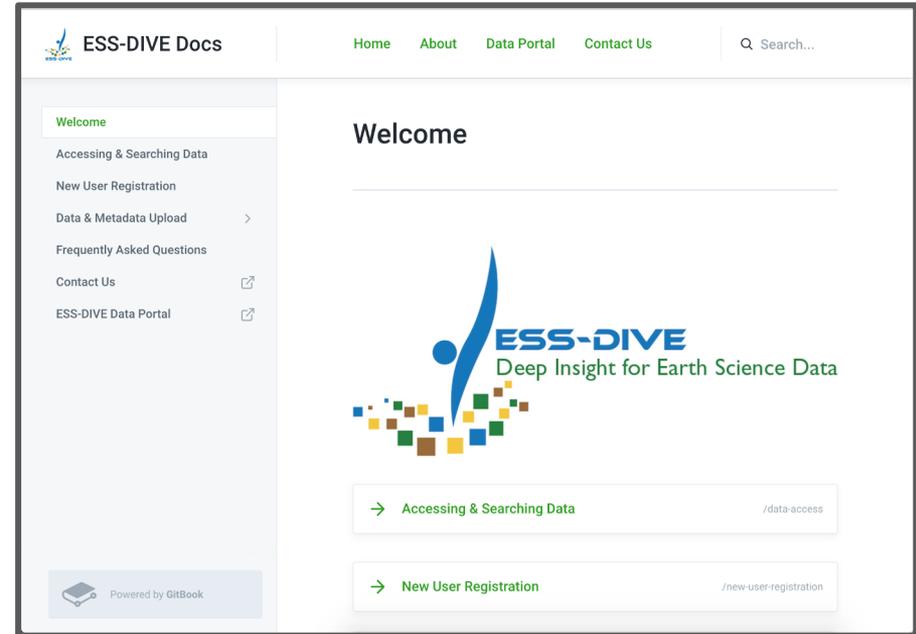
Next 

How many data packages do anticipate submitting to ESS-DIVE in total by the end of 2021?



Package Service API Capabilities

- **Create & update data packages**
 - Update metadata
 - Add data files
 - Replace existing data files
- **Upload Large Data Files**
- **Explore existing package metadata**
 - Full metadata in single package search, including data files, with easy links to website
- **Package Service Documentation**
 - Easy to navigate tutorials
 - Troubleshooting guides
 - Helpful tips



docs.ess-dive.lbl.gov

Package Service API Upload Process



Review Documentation

Get Upload
Token

Create
Metadata

Upload Metadata
& Data

Review tutorial documentation for Package Service API

Get authentication token from ESS-DIVE

Create package metadata in JSON-LD

Success! View package landing page on ESS-DIVE

Errors? Fix any errors found by validation using documentation

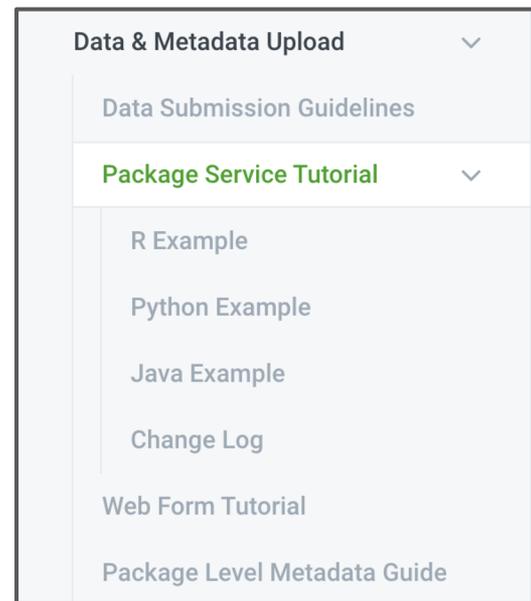
API Documentation Tutorial



To get started with the Package Service API, we recommend using ESS-DIVE's [API Tutorial](#).

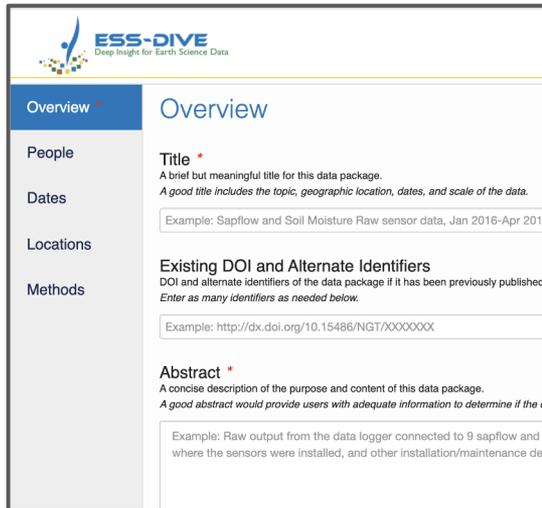
This tutorial includes:

- All **prerequisites** needed to use the API
- **Code examples** of how to submit, update, and search data packages
- Complete examples in three **different coding languages**:
 - R
 - Python
 - Java



Accessible Coding

- Tutorial documentation gives complete walkthrough of API submission
- API metadata fields are equivalent to the fields in web upload form
- Quick learning curve for users of varying experience



ESS-DIVE
Deep Insight for Earth Science Data

Overview

People

Dates

Locations

Methods

Title *
A brief but meaningful title for this data package.
A good title includes the topic, geographic location, dates, and scale of the data.
Example: Sapflow and Soil Moisture Raw sensor data, Jan 2016-Apr 2016

Existing DOI and Alternate Identifiers
DOI and alternate identifiers of the data package if it has been previously published.
Enter as many identifiers as needed below.
Example: <http://dx.doi.org/10.15486/NGT/XXXXXXX>

Abstract *
A concise description of the purpose and content of this data package.
A good abstract would provide users with adequate information to determine if the c
Example: Raw output from the data logger connected to 9 sapflow and ... where the sensors were installed, and other installation/maintenance de



```
{
  "@context": "http://schema.org/",
  "@type": "Dataset",
  "@id": "http://dx.doi.org/10.3334/CDIAC/spruce.001",
  "name": "Correct Package with a one liner description",
  "description": ["This is a very long paragraph This is a very"],
  "creator": [
    {
      "@id": "http://orcid.org/0000-0001-7293-3561",
      "givenName": "Paul J",
      "familyName": "Hanson",
      "affiliation": "Oak Ridge National Laboratory",
      "email": "hansonpj@ornl.gov"
    },
    {
      "givenName": "Jeffrey",
      "familyName": "Riggs",
      "affiliation": "Oak Ridge National Laboratory"
    }
  ],
  {
    "givenName": "C",
    "familyName": "Nettles",
    "affiliation": "Oak Ridge National Laboratory"
  }
}
```

Metadata Crosswalk example

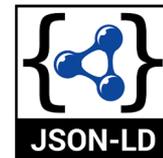
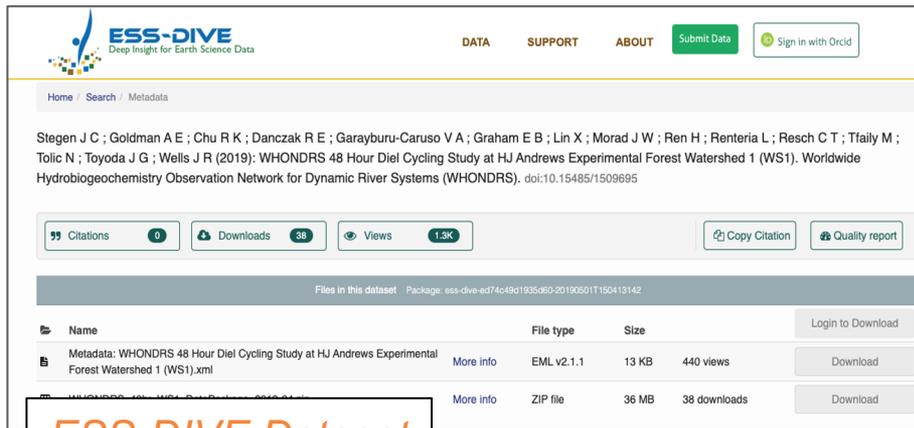


ESS-DIVE Field	JSON-LD	DataCite 4.1
Title	name	title
Alternative Identifiers	alternateName	alternateIdentifiers
Abstract	description	description[@type=Abstract]
Keywords	keywords	subjects
Data Variables	variablesMeasured	subjects
Publication Date	datePublished	publicationYear

bit.ly/ess-dive-crosswalk

JSON-LD Metadata Standard

- **JSON object notation is easier to read, understand & manipulate** for humans than many other object notation styles
- JSON-LD embedded in data package landing page allows **harvesting by Google**

ESS-DIVE
Deep Insight for Earth Science Data

DATA SUPPORT ABOUT [Submit Data](#) [Sign in with Orcid](#)

Home / Search / Metadata

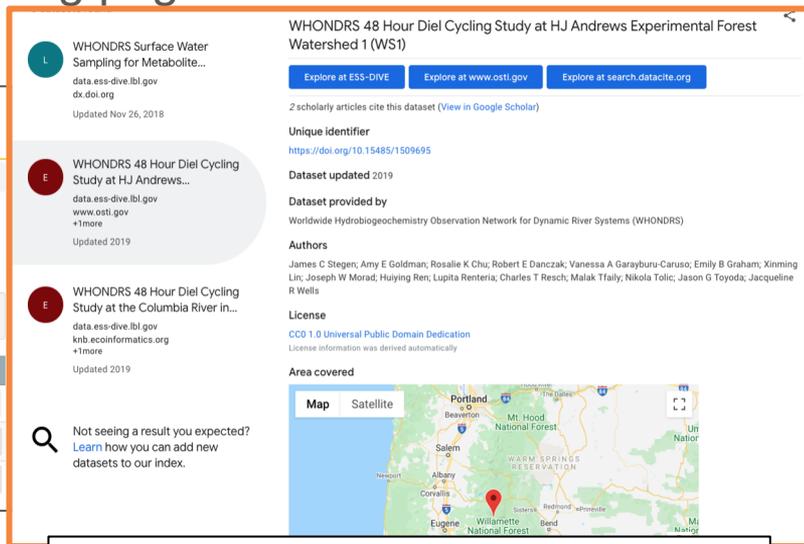
Stegen J C ; Goldman A E ; Chu R K ; Danczak R E ; Garayburu-Caruso V A ; Graham E B ; Lin X ; Morad J W ; Ren H ; Renteria L ; Resch C T ; Tfaily M ; Tolic N ; Toyoda J G ; Wells J R (2019): WHONDORS 48 Hour Diel Cycling Study at HJ Andrews Experimental Forest Watershed 1 (WS1). Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems (WHONDORS). doi:10.15485/1509695

Citations 0 Downloads 38 Views 1.3K [Copy Citation](#) [Quality report](#)

Files in this dataset Package: ess-dive-ed74c49d1935d60-201905011150413142

Name	File type	Size	Views	Download
Metadata: WHONDORS 48 Hour Diel Cycling Study at HJ Andrews Experimental Forest Watershed 1 (WS1).xml	EML v2.1.1	13 KB	440 views	Download
WHONDORS 48 Hour Diel Cycling Study at HJ Andrews Experimental Forest Watershed 1 (WS1).zip	ZIP file	36 MB	38 downloads	Download

*ESS-DIVE Dataset
Landing Page*



WHONDORS Surface Water Sampling for Metabolite...
data.ess-dive.lbl.gov
dx.doi.org
Updated Nov 26, 2018

WHONDORS 48 Hour Diel Cycling Study at HJ Andrews...
data.ess-dive.lbl.gov
www.osti.gov
+1more
Updated 2019

WHONDORS 48 Hour Diel Cycling Study at the Columbia River in...
data.ess-dive.lbl.gov
knbc.econinformatics.org
+1more
Updated 2019

WHONDORS 48 Hour Diel Cycling Study at HJ Andrews Experimental Forest Watershed 1 (WS1)
[Explore at ESS-DIVE](#) [Explore at www.osti.gov](#) [Explore at search.datacite.org](#)

2 scholarly articles cite this dataset ([View in Google Scholar](#))

Unique identifier
<https://doi.org/10.15485/1509695>

Dataset updated by
2019

Dataset provided by
Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems (WHONDORS)

Authors
James C Stegen; Amy E Goldman; Rosalie K Chu; Robert E Danczak; Vanessa A Garayburu-Caruso; Emily B Graham; Xinming Lin; Joseph W Morad; Huiying Ren; Lupita Renteria; Charles T Resch; Nikola Tfaily; Nikoia Tolic; Jason G Toyoda; Jacqueline R Wells

License
CC0 1.0 Universal Public Domain Dedication
License information was derived automatically

Area covered

Map Satellite

Not seeing a result you expected?
[Learn](#) how you can add new datasets to our index.

Google Dataset Search

Create new data package

- You will be able to **validate and submit JSON-LD** against the interface.
- If the JSON-LD is invalid, details about the errors will be given
- You may upload files up to **100 GB**.

```
{
  "@context": "http://schema.org/",
  "@type": "Dataset",
  "@id": "http://dx.doi.org/10.3334/CDIAC/spruce.001",
  "name": "Correct Package with a one liner description",
  "description": ["This is a very long paragraph This is a very"],
  "creator": [
    {
      "@id": "http://orcid.org/0000-0001-7293-3561",
      "givenName": "Paul J",
      "familyName": "Hanson",
      "affiliation": "Oak Ridge National Laboratory",
      "email": "hansonpj@ornl.gov"
    },
    {
      "givenName": "Jeffrey",
      "familyName": "Riggs",
      "affiliation": "Oak Ridge National Laboratory"
    },
    {
      "givenName": "C",
      "familyName": "Nettles",
      "affiliation": "Oak Ridge National Laboratory"
    }
  ]
}
```

API Submission Responses

```
{
  "detail": "One or more fields raised validation errors.",
  "errors": [
    "'datePublished' is a required property",
    "'provider' is a required property",
    "editor 'email' is a required property"
  ]
}
```



```
{
  "id": "ess-dive-c5ab2ee757aebc4-20190927T203522703923",
  "viewUrl": "https://data-sandbox.ess-dive.lbl.gov/view/ess-dive-c5ab2ee757aebc4-20190927T203522703923",
  "detail": "Data Package created successfully.",
  "errors": null,
  "dataset": {
    "name": "ESS-DIVE Webinar Demo",
    "description": "Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non"
  }
}
```



Update existing data package

- **Metadata Updates**
 - Submit only the fields you want to update using JSON-LD
- **Data File Updates**
 - *Upload additional files to an existing data package*

```
{
  "spatialCoverage": [
    {
      "description": "Site ID: S1 Bog Site name: S1 Bog, Marcell",
      "geo": [
        {
          "name": "Northwest",
          "latitude": 47.50285,
          "longitude": -93.48283
        },
        {
          "name": "Southeast",
          "latitude": 47.50285,
          "longitude": -93.48283
        }
      ]
    }
  ]
}
```

Retrieve list of data package metadata

- **Retrieve a complete list** of data packages *submitted by you* with important metadata, including:
 - Title
 - URL
 - Public Visibility
 - Citation

```
"total": 146,  
"user": "http://orcid.org/0000-0001-9061-8952",  
"page_size": 25,  
"row_start": 1,  
"result": [  
  { ...  
},  
  {  
    "id": "ess-dive-b29ca461452a7d3-20190924T000453220359",  
    "viewUrl": "https://data-sandbox.ess-dive.lbl.gov/view/  
    ess-dive-b29ca461452a7d3-20190924T000453220359",  
    "dateUploaded": "2019-09-24T00:04:53.612Z",  
    "dateModified": "2019-09-24T00:04:53.612Z",  
    "isPublic": false,  
    "citation": "Christianson D (2019): My First Data Package. My project.  
    ess-dive-b29ca461452a7d3-20190924T000453220359",  
    "dataset": {  
      "@context": "http://schema.org/",  
      "@type": "Dataset",  
      "@id": null,  
      "name": "My First Data Package",  
      "description": "Lorem ipsum dolor sit amet, consectetur adipiscing elit,  
      sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut  
      anim ad minim veniam, quis nostrud exercitation ullamco laboris nisi
```

Common Errors

- **JSON-LD validation errors**
 - Use tutorials, technical documentation, or packages examples to understand JSON-LD schema expectations.
- **Out of date token**
 - Tokens expire daily. Make sure token is up to date

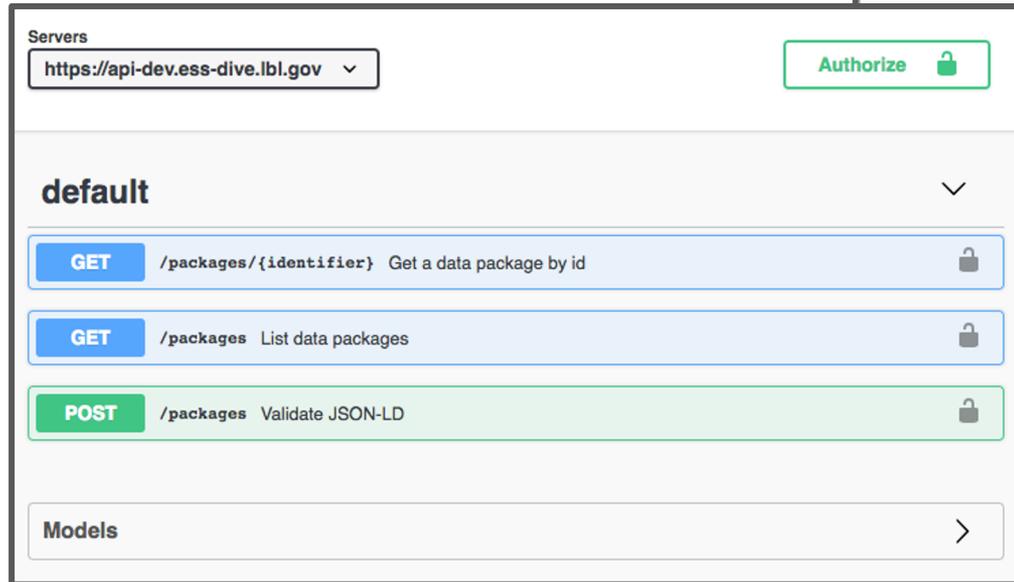
```
{
  "detail": "One or more fields raised validation errors.",
  "errors": [
    "'datePublished' is a required property",
    "'provider' is a required property",
    "editor 'email' is a required property"
  ]
}
```

```
{"detail":"You do not have authorized access"}
```

API Technical Documentation



For technical details about ESS-DIVE's Package Service API, visit our [extended documentation](#).



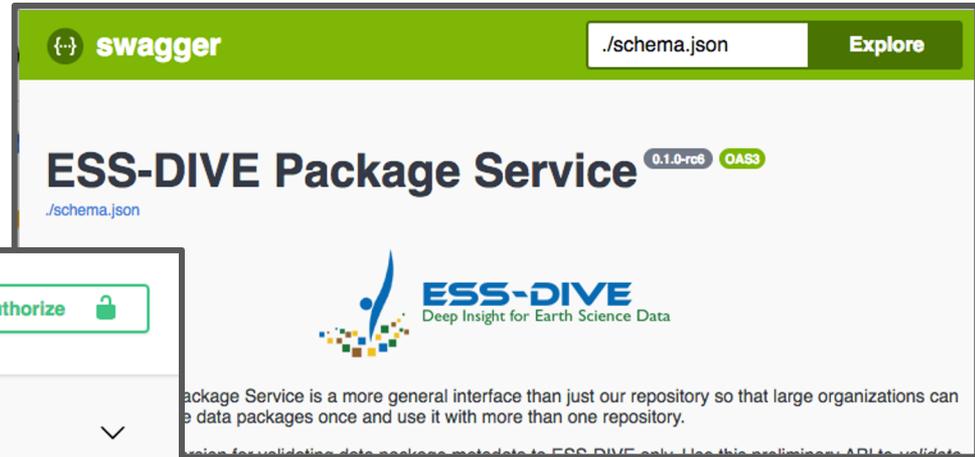
Servers
https://api-dev.ess-dive.lbl.gov

Authorize

default

- GET /packages/{identifier} Get a data package by id
- GET /packages List data packages
- POST /packages Validate JSON-LD

Models



swagger .schema.json Explore

ESS-DIVE Package Service

0.1.0-rc8 OAS3

./schema.json



ESS-DIVE
Deep Insight for Earth Science Data

Package Service is a more general interface than just our repository so that large organizations can use the same data packages once and use it with more than one repository.

- This is hosted on a separate webpage from other ESS-DIVE documentation
- It's best to use this resource when **troubleshooting errors**



Package Service API Summary

- **Programmatically upload, update, and search** data packages
- **Reuse code for duplicated information** across packages
- Upload **large data files** & perform **bulk uploads**
- Use **tutorial documentation** to get comfortable

Tutorial documentation: docs.ess-dive.lbl.gov

API documentation: api.ess-dive.lbl.gov

Customized Data Portals

What is a data portal?



*A data portal is a **collection** of any ESS-DIVE **data packages**. Researchers, projects, labs, and even individuals can easily **highlight and share** their datasets and **research topics** using a data portal.*

ESS-DIVE Data Search Interface



ESS-DIVE's Data Search Interface can also be considered a data portal. In this **data view**, we can see **every data package** ever published in ESS-DIVE.

The screenshot displays the ESS-DIVE Data Search Interface. At the top left is the ESS-DIVE logo and tagline "Deep Insight for Earth Science Data". Navigation links for "DATA", "SUPPORT", and "ABOUT" are visible, along with a "Submit Data" button and a "Sign in with Orcid" button.

The main content area is divided into two columns. The left column contains a search bar with the placeholder "Search phrase" and a search icon. Below it are filter options under "Filter by:", including "Identifier", "Region description", "Creator", and "Year".

The right column shows a list of datasets, titled "DATASETS 1 TO 25 OF 441". The list includes:

- Dataset 1: Chadwick K D ; Brodrick P ; Grant K ; Henderson A ; Bill M ; Breckheimer I ; Williams C F R ; Goulden T ; Falco N ; McCormick M ; Musinsky J ; Pierce S ; Hastings Porro M ; Scott A ; Brodie E ; Hancher M ; Steltzer H ; Wainwright H ; Williams K ; Maher K (2020): **NEON AOP foliar trait maps, maps of model uncertainty estimates, and conifer map, East River, CO 2018.** A Multiscale Approach to Modeling Carbon and Nitrogen Cycling within a High Elevation Watershed. doi:10.15485/1618133
- Dataset 2: Damerow J ; Varadharajan C ; Boye K ; Brodie E ; Burrus M ; Chadwick D ; Cholia S ; Crystal-Ornelas R ; Elbashandy H ; Eloy Alves R ; Ely K ; Goldman A ; Hendrix V ; Jones C ; Jones M ; Kakalia Z ; Kemner K ; Kersting A ; Maher K ; Merino N ; O'Brien F ; Perzan Z ; Robles E ; Snavey C ; Sorensen P ; Stegen J ; Weisenhorn P ; Whitenack K ; Zavarin M ; Agarwal D (2020): **ESS-DIVE Global Sample Numbers and Metadata Reporting Format for Environmental Systems Science (IGSN-ESS).** Environmental Systems Science Data Infrastructure for a Virtual Ecosystem (ESS-DIVE). doi:10.15485/1660470
- Dataset 3: Dong W ; Beutler C ; Brown W ; Newman A ; Versteeg R ; Williams K H (2020): **Cation Data for the East River Watershed, Colorado.** Watershed Function SFA. doi:10.15485/1668055
- Dataset 4: Gilfillan D ; Marland G ; Boden T ; Andres R (2020): **Global, Regional, and National Fossil-Fuel CO2 Emissions: 1751-2017.** CDIAC-FF, Research Institute for Environment, Energy, and Economics, Appalachian State University. doi:10.15485/1712447

Each dataset entry includes a small icon for information, a document icon, and a view count (e.g., "1" or "3").

On the right side of the interface is a map of the United States and parts of Canada and Mexico. The map is overlaid with a grid of blue squares, each containing a number representing the number of datasets for that geographic area. The map is titled "Hide Map" and includes a "Satellite" button and a "Terrain" button. A scale bar at the bottom right indicates "Map data ©2020 1000 km" and "Terms of Use".

At the bottom of the page, there are logos for "BERKELEY LAB", "NERSC", "NCEAS", and "U.S. DEPARTMENT OF ENERGY Office of Science".

Key Portal Features



- Portals can be **branded**
 - All portals have **customizable** banners, logos, and URLs
- Provide a space for **data and research context**
- Contain portal specific **data metrics**
- Easily **share** your portals

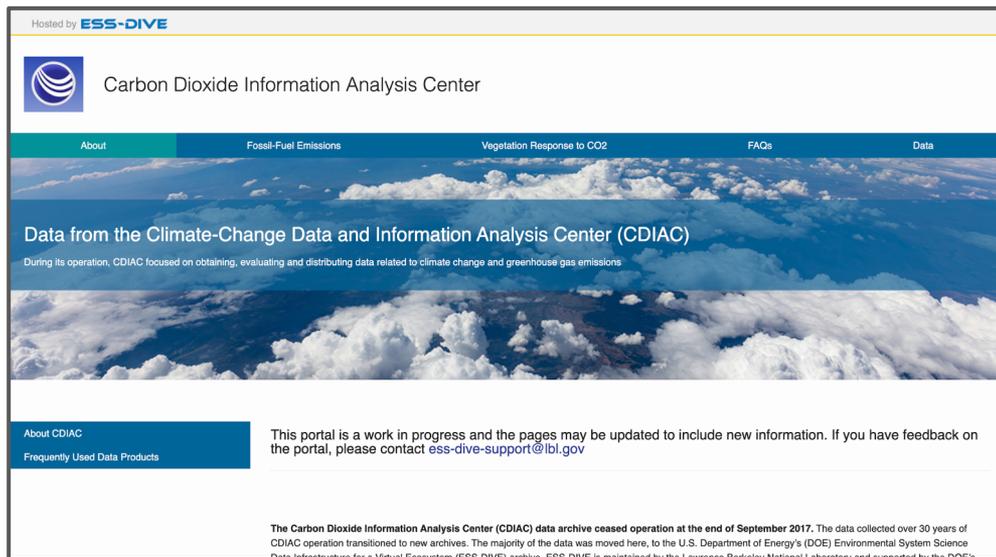
<https://data.ess-dive.lbl.gov/portals/CDIAC>

A screenshot of the Carbon Dioxide Information Analysis Center (CDIAC) portal page. The page is hosted by ESS-DIVE. It features a navigation menu with links for "About", "Fossil-Fuel Emissions", "Vegetation Response to CO2", "FAQs", and "Data". The main content area displays a banner image of a cloudy sky with the text "Data from the Climate-Change Data and Information Analysis Center (CDIAC)" and a sub-header "During its operation, CDIAC focused on obtaining, evaluating and distributing data related to climate change and greenhouse gas emissions". A footer section contains a "Frequently Used Data Products" link and a notice stating that the CDIAC data archive ceased operation at the end of September 2017, with data moved to the U.S. Department of Energy's (DOE) Environmental System Science Data Infrastructure for a Mutual Ecosystem (ESS-DIVE) archive. The notice also mentions that ESS-DIVE is maintained by the Lawrence Berkeley National Laboratory and supported by the DOE's Office of Biological and Environmental Research.

Key Portal Features

- You can imagine a portal as a personalized website for your research that's hosted by ESS-DIVE
- If you already have a website for your research, the portal link can be inserted into your website to allow easy access to your data

<https://data.ess-dive.lbl.gov/portals/CDIAC>

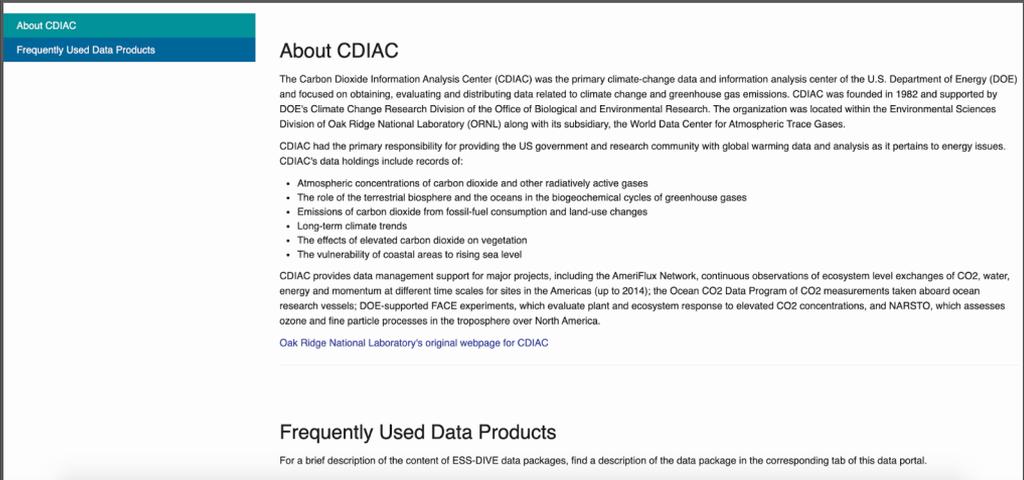
A screenshot of the Carbon Dioxide Information Analysis Center (CDIAC) portal. The page is hosted by ESS-DIVE. It features a navigation menu with links for "About", "Fossil-Fuel Emissions", "Vegetation Response to CO2", "FAQs", and "Data". The main content area displays the title "Data from the Climate-Change Data and Information Analysis Center (CDIAC)" and a sub-header "During its operation, CDIAC focused on obtaining, evaluating and distributing data related to climate change and greenhouse gas emissions". Below this is a large image of a cloudy sky. At the bottom, there is a section for "About CDIAC" and "Frequently Used Data Products", along with a notice that the portal is a work in progress and contact information for feedback: ess-dive-support@lbl.gov. A footer note states that the CDIAC data archive ceased operation at the end of September 2017 and was transitioned to new archives at the U.S. Department of Energy's (DOE) Environmental System Science Data Infrastructure for a Mutual Ecosystem (ESS-DIVE) archive.

Portals are a powerful tool for connecting related datasets together

Portals give projects and groups an identity in ESS-DIVE



- Portals allow you to give your datasets scientific context beyond what's capable with just your data package alone.
- **Your portal can describe...**
 - **Your research group**
 - **The study site**
 - **The research question your addressing**

A screenshot of the Carbon Dioxide Information Analysis Center (CDIAC) website. The page has a blue header with two tabs: "About CDIAC" (selected) and "Frequently Used Data Products". The main content area is white and contains the following text:

About CDIAC

The Carbon Dioxide Information Analysis Center (CDIAC) was the primary climate-change data and information analysis center of the U.S. Department of Energy (DOE) and focused on obtaining, evaluating and distributing data related to climate change and greenhouse gas emissions. CDIAC was founded in 1982 and supported by DOE's Climate Change Research Division of the Office of Biological and Environmental Research. The organization was located within the Environmental Sciences Division of Oak Ridge National Laboratory (ORNL) along with its subsidiary, the World Data Center for Atmospheric Trace Gases.

CDIAC had the primary responsibility for providing the US government and research community with global warming data and analysis as it pertains to energy issues. CDIAC's data holdings include records of:

- Atmospheric concentrations of carbon dioxide and other radiatively active gases
- The role of the terrestrial biosphere and the oceans in the biogeochemical cycles of greenhouse gases
- Emissions of carbon dioxide from fossil-fuel consumption and land-use changes
- Long-term climate trends
- The effects of elevated carbon dioxide on vegetation
- The vulnerability of coastal areas to rising sea level

CDIAC provides data management support for major projects, including the AmeriFlux Network, continuous observations of ecosystem level exchanges of CO₂, water, energy and momentum at different time scales for sites in the Americas (up to 2014); the Ocean CO₂ Data Program of CO₂ measurements taken aboard ocean research vessels; DOE-supported FACE experiments, which evaluate plant and ecosystem response to elevated CO₂ concentrations, and NARSTO, which assesses ozone and fine particle processes in the troposphere over North America.

[Oak Ridge National Laboratory's original webpage for CDIAC](#)

Frequently Used Data Products

For a brief description of the content of ESS-DIVE data packages, find a description of the data package in the corresponding tab of this data portal.

Portals give projects and groups an identity in ESS-DIVE



Hosted by **ESS-DIVE**

 Carbon Dioxide Information Analysis Center

About **Fossil-Fuel Emissions** **Vegetation Response to CO2** **FAQs** **Data**

Fossil-Fuel CO2 Emissions

CDIAC's most well-cited dataset collection



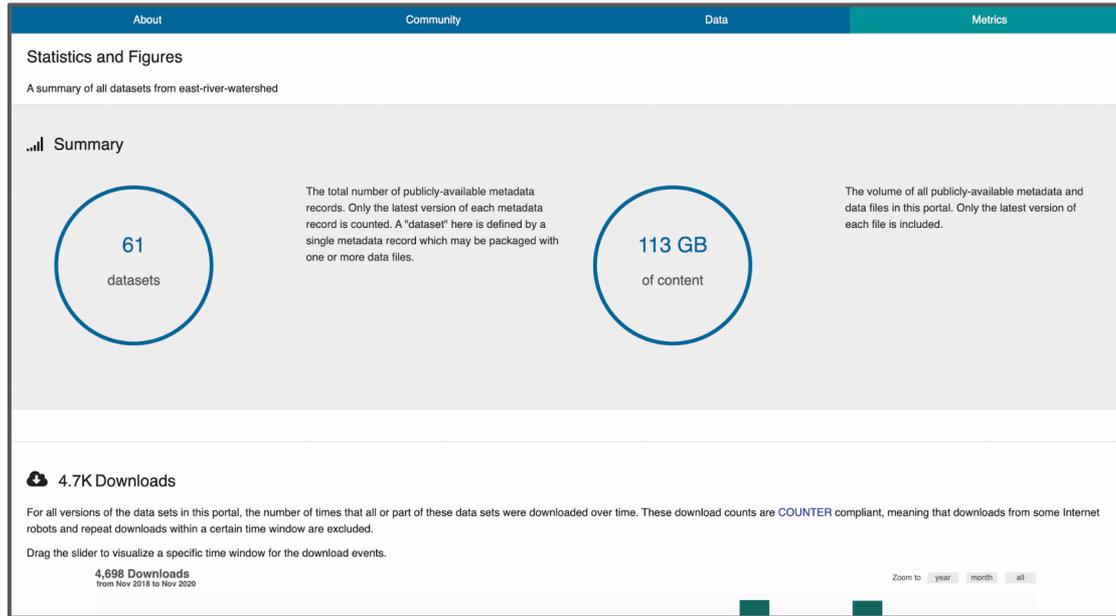
Fossil-Fuel CO2 Emissions
Global
Regional
National
Full Global Carbon Budget
Historical Global CO2 Estimates

This page contains multiple datasets. Tab through the table of contents on the left hand side to skip to a particular dataset of interest. Click the header of any dataset to view the data on the ESS-DIVE archive.

This portal is a work in progress and the pages may be updated to include new information. If you have feedback on the portal, please contact ess-dive-support@lbl.gov

In the CDIAC portal, we focus on the scientific purpose of this data collection

Data portal specific metrics

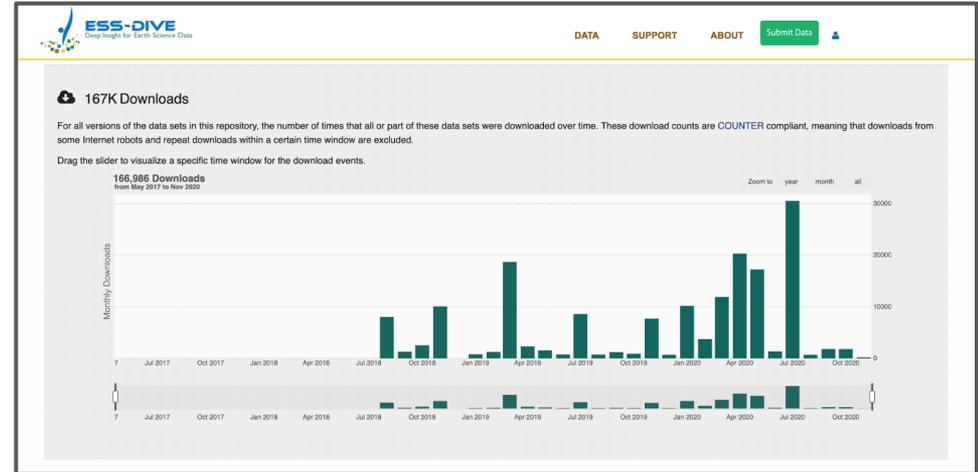


This metrics page is personalized to the data packages that you choose to add to your portal

Data portal specific metrics



- Some additional metrics are:
 - Number of times your portal data has been cited
 - Number of times the portal data has been downloaded
 - Number of times your data package landing pages have been viewed



*Visit ESS-DIVE's repository summary
to see the full metric capabilities
available to all portals*

<https://data.ess-dive.lbl.gov/profile>

Creating Portals



- Any registered **Data Contributor** can create a portal!
- After you **login** to ESS-DIVE with your ORCID and you can create a new portal from your **settings page**

A screenshot of the ESS-DIVE web application interface. The top navigation bar includes "DATA", "SUPPORT", "ABOUT", "Submit Data", and a user profile "MADISON". A search bar is on the left. The main content area displays a list of datasets, with the first one highlighted: "Chadwick K D ; Brodrick P ; Grant K ; Henderson A ; Bill M ; Breckheimer I ; Williams C F R ; Goulden T ; Falco N ; McCormick M ; Musinsky J ; Pierce S ; Hastings Porro M ; Scott A ; Brodie E ; Hancher M ; Stelzer H ; Wainwright H ; Williams K ; Maher K (2020): NEON AOP foliar trait maps, maps of model uncertainty estimates, and conifer map, East River, CO 2018. A Multiscale Approach to Modeling Carbon and Nitrogen Cycling within a High Elevation Watershed. doi:10.15485/1616133". To the right of the dataset list is a map of North America with numbered data points. An orange arrow points to a "MY SETTINGS" menu in the top right corner of the map area. The menu includes options for "MY PROFILE", "MY DATA PACKAGES", and "SIGN OUT".

Creating Portals



- Any registered **Data Contributor** can create a portal!
- After you **login** to ESS-DIVE with your ORCID and you can create a new portal from your **settings page**

A screenshot of the ESS-DIVE web application interface. The top navigation bar includes the ESS-DIVE logo, the text "Deep Insight for Earth Science Data", and links for "DATA", "SUPPORT", "ABOUT", "Submit Data", and a user profile for "MADISON". Below the navigation bar, there are tabs for "My Data" and "Settings". The "Settings" tab is active, showing a dropdown menu with options: "Settings", "My account", "My portals ★ Alpha", "Groups", and "Authentication Token". An orange arrow points to the "My portals" option. To the right, the "My Portals" section displays a table of existing portals. The table has columns for "Logo", "ID", and "Title". Each row includes an "Edit" button. A green "+ New portal" button is located at the top right of the table.

Logo	ID	Title	
	east-river-watershed	Data from the East River Watershed, Colorado	Edit
	PNNLRiverCorridorSFA	PNNL River Corridor SFA	Edit
	WHONDRS	Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems	Edit
	portal-tutorial	Portal Tutorial	Edit

Creating Portals



- ESS-DIVE has extensive documentation on how to develop your portal
- Our DataONE partners at the Arctic Data Center have tutorial videos available

A screenshot of the ESS-DIVE Docs website. The page title is "ESS-DIVE Docs". The navigation menu includes "Home", "About", "Data Portal", and "Contact Us". The left sidebar contains a table of contents with the following items: "Welcome", "Accessing & Searching Data", "New User Registration", "Data & Metadata Upload" (with a right arrow), "Portals" (highlighted in green with a down arrow), "Portal Main Components", "Creating Portals", "How to Publish Portals", "Sharing Portals", "Tutorial Videos", "Frequently Asked Questions", "Contact Us" (with an external link icon), and "ESS-DIVE Data Portal" (with an external link icon). At the bottom of the sidebar is a "Powered by GitBook" logo. The main content area has the heading "Portals" and a sub-heading "What is a Portal?". The text under "What is a Portal?" explains that a portal is a collection of private or public ESS-DIVE data packages on a unique webpage, and that it provides a means to preserve information regarding the projects' objectives, scopes, and organization. It also mentions that portals leverage ESS-DIVE's metric features to create statistics describing the project's data packages. The text is partially cut off at the bottom of the screenshot.

Creating Portals

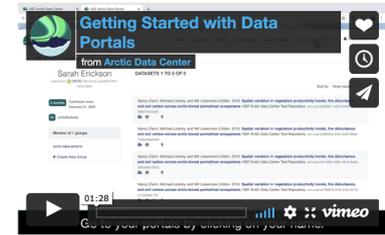


- ESS-DIVE has extensive documentation on how to develop your portal
- Our DataONE partners at the Arctic Data Center have tutorial videos available

Introduction to Portals



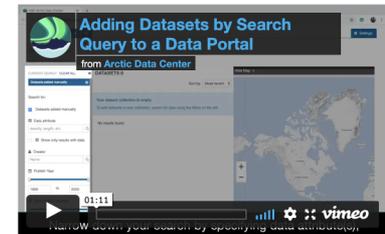
Getting Started



Creating a Freeform Text Page



Adding Datasets by Search Query



Creating Portals

- ESS-DIVE has extensive documentation on how to develop your portal
- Our DataONE partners at the Arctic Data Center have tutorial videos available

Give it a try!

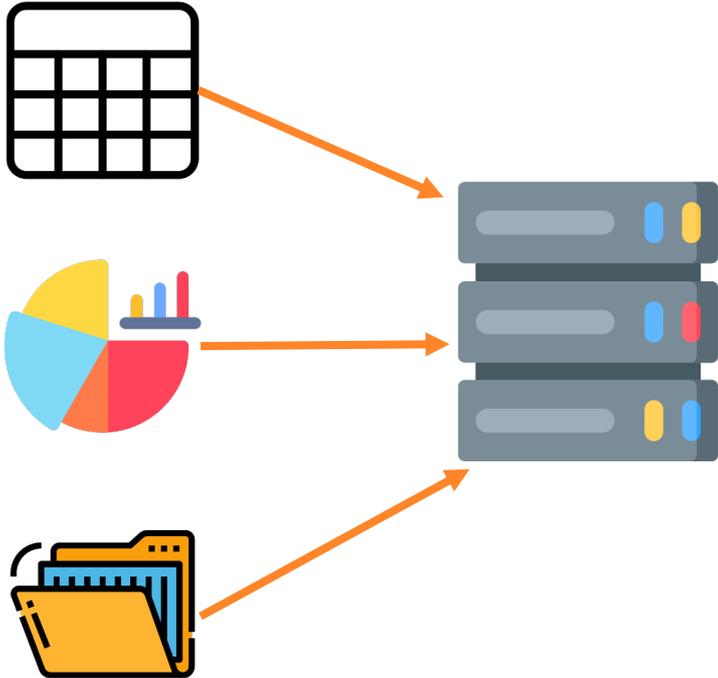
A collage of four video thumbnails from the Arctic Data Center, arranged in a 2x2 grid. Each thumbnail shows a different tutorial video. The top-left video is titled 'Introduction to Portals' and features a landscape of ice and water. The top-right video is titled 'Getting Started' and shows a user interface with a search bar and a list of datasets. The bottom-left video is titled 'Creating a Freeform Text Page' and shows a user interface with a text input field and a 'Save' button. The bottom-right video is titled 'Adding Datasets by Search Query' and shows a user interface with a search bar and a map of the Arctic region. Each video thumbnail includes a play button icon and a VIMEO logo at the bottom.

Are you or your group interested in creating a portal for your published data packages?



ESS-DIVE data and metadata reporting formats

Data stored in repositories should be FAIR



60% of publicly archived data is unusable.

Findable

Accessible

Interoperable

Reusable

Clarifying terminology: Data standards and reporting formats



- **Data Standards** - Decades of development, international community of governance



Darwin Core

- **Reporting Formats** - In development, though still enable data harmonization and synthesis

Non-standardized data are difficult to reuse and synthesize



idNumber	material	temperature
3928	soil	23.2
3234	groundwater	9.02



sampleNum	substance	temp
8765	dirt	21.1
2312	ground liquid	7.0

Standardized data are more readily reused and synthesized



idNumber	material	temperature
3928	soil	23.2
3234	groundwater	9.02

idNumber	material	temperature
8765	soil	21.1
2312	groundwater	7.0

Community-led reporting format development to make diverse ESS data FAIR



**Package Level
Metadata/JSON-LD**

Agarwal, Hendrix (LBNL)



**Package Metadata
Quality**

Damerow (LBNL)



**Sample
IDs/Metadata**

Damerow (LBNL)

ESS-DIVE team developed **high priority, base data formats**



**File-Level/
csv Metadata**

Velliquette, Heinz,
Devarakonda (ORNL)



Soil Respiration

Bond-Lamberty,
Pennington (PNNL)



Leaf Physiology

Rogers, Ely (BNL)



**Hydrologic
Monitoring**

Goldman (PNNL)



**Water/Soil
Chemistry**

Boye (SLAC)



**Amplicon
Sequencing**

Weisenhorn (ANL)

Developing **reporting formats for select data types** with \$1M from DOE's Biological Environmental Research (BER) program

ESS-DIVE's GitHub Community Space is for reporting format, code, tools



ESS-DIVE Community Space

<https://github.com/ess-dive-community>

A workspace for code, tools, reporting formats and other products related to the ESS-DIVE repository. Please message ess-dive-support@lbl.gov to contribute

Berkeley, CA USA <http://ess-dive.lbl.gov> @ESSDIVE ess-dive-support@lbl.gov

Repositories 6 Packages People 4 Teams Projects Settings

Pinned repositories

[Customize pinned repositories](#)

[essdive-community-space-guide](#)

A guide to starting or contributing to ESS-DIVE community repositories

R



Joan Damerow (LBNL)

ess-dive-community / **essdive-sample-id-metadata**

<> Code Issues 2 Pull requests Actions Projects Wiki

master Go to file Add file Code

JEDamerow Update guide.md ✓ 15 days ago 237

terms	Update sampleMetadata_sources.md	2 months ago
README.md	Update README.md	15 days ago
contribute.md	Update contribute.md	2 months ago
guide.md	Update guide.md	15 days ago
instructions.md	Update instructions.md	2 months ago
sampleTemplate.csv	Add files via upload	last month
sampleTemplate.xls	Add files via upload	last month



Joan Damerow (LBNL)

ess-dive-community / **essdive-sample-id-metadata**

<> Code Issues 2 Pull requests Actions Projects Wiki

master Go to file Add file Code

JEDamerow	Update guide.md	15 days ago	237
terms	Update sampleMetadata_sources.md	2 months ago	
README.md	Update README.md	15 days ago	
contribute.md	Update contribute.md	2 months ago	
guide.md	Update guide.md	15 days ago	
instructions.md	Update instructions.md	2 months ago	
sampleTemplate.csv	Add files via upload	last month	
sampleTemplate.xls	Add files via upload	last month	

essdive-sample-id-metadata / guide.md

Sample Description:

Material* | Field name informal classification | Sample Description |

Material	
Proposed ESS-DIVE Element	material Required
Definition	Material that the sample consists of.
Format	SESAR controlled list . See ESS-DIVE's proposed material terms from Environment Ontology (ENVO)
Additional Instructions	Use a semi-colon to delimit multiple materials where needed. ESS-DIVE is requesting additional terms for organisms, organic material, and water samples. Please provide feedback on any other terms needed.
Examples	soil; sediment; surface water ENVO:00002042 ; groundwater ENVO:01001004

GitHub “Issues” enable us to track suggestions to reporting format from user community



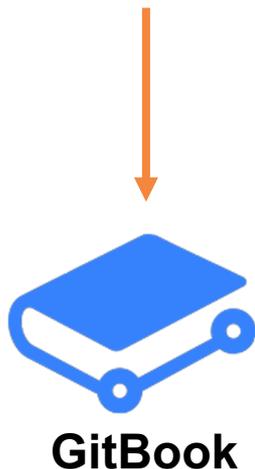
<> Code **! Issues 4** Pull requests Actions Projects

Title

Write Preview H B I

****Term to change****: Exact name of term you are suggesting change to.
****Submitter****: Provide your name.
****Justification****: Why is this change necessary?

Updates to GitHub instantly pushed to the GitBook



<> Edit file

👁 Preview changes

```
1 # ESS-DIVE Sample ID and Metadata Reporting
2
3 ESS-DIVE recommends registering samples for [Global Sample Numbers
```



Sample ID and
Metadata



ESS-DIVE Sample ID and Metadata
Reporting

ESS-DIVE Sample ID and Metadata Reporting

Easy translation of GitHub content to user-friendly documentation



A screenshot of a web page titled "ESS-DIVE Sample ID and Metadata Reporting". The page has a light gray header with a logo and the text "Sample ID and Metadata". A search icon is in the top right. The main content area has a large heading "ESS-DIVE Sample ID and Metadata Reporting" followed by a paragraph: "ESS-DIVE recommends registering samples for Global Sample Numbers (IGSNs) through the System for Earth Sample Registration (SESAR). IGSNs are associated with standardized metadata to characterize a variety of different samples and their collection details. These sample identifiers facilitate sample discovery, tracking, and reuse; they are especially useful when sample data is shared with collaborators, sent to different labs or user facilities for analyses, or distributed in different data files, datasets, and/or publications." A sidebar on the left contains a search bar and a list of links: "ESS-DIVE Sample ID and Metadata Reporting" (highlighted with an orange box), "ESS-DIVE Instructions to Register Samples for IGSNs", "terms" with a right arrow, "ESS-DIVE Sample ID and Metadata Guide", and "Guidelines for contributing".

<https://ess-dive.gitbook.io/sample-id-and-metadata/>

Learn more and contribute to reporting formats on our GitHub Community Space



<https://github.com/ess-dive-community>

File-Level
Metadata

CSV Files

Soil
Respiration

Sample ID
Metadata

Leaf
Physiology

Hydrologic
Monitoring

Water & Soil
Chemistry

Amplicon
sequencing



Questions?

Detailed guides are available at docs.ess-dive.lbl.gov

ESS-DIVE Community Space on GitHub <https://github.com/ess-dive-community>

Contact us at ess-dive-support@lbl.gov



ESS-DIVE Glossary

List of common ESS-DIVE terms :

- **DataONE** - The Data Observation Network for Earth (DataONE) is a distributed framework and sustainable cyberinfrastructure that provides open and secure access to Earth observational data. ESS-DIVE is a DataONE member.
- **DOE** - The U.S. Department of Energy (DOE) is a Cabinet-level department of the United States whose mission is to ensure America's security and prosperity by addressing its energy, environmental and nuclear challenges through transformative science and technology solutions.

ESS-DIVE Glossary (*cont.*)



- **ESGF** - The Earth System Grid Federation (ESGF) is a collaboration that develops, deploys and maintains software infrastructure for the management, dissemination, and analysis of model output and observational data. It is an interagency and international effort led by the U.S. Department of Energy with various co-funding agencies.
- **ESS** - Environmental Systems Science (ESS) is a U.S. Department of Energy Office of Science program under the Biological and Environmental Research Program seeking to advance a robust predictive understanding of terrestrial surface and subsurface ecosystems.



ESS-DIVE Glossary (*cont.*)

- **ESS-DIVE** - Environmental System Science Data Infrastructure for a Virtual Ecosystem (ESS-DIVE) is a U.S. Department of Energy repository for earth and environmental science data, models and software generated from research on terrestrial and subsurface environments.
- **NCEAS** - The National Center for Ecological Analysis and Synthesis (NCEAS) at UC Santa Barbara partners with ESS-DIVE on data preservation. NCEAS is a DataONE member and a recognized expert in ecological data, digital libraries, and standards for data format.

ESS-DIVE Glossary (*cont.*)



- **NERSC** - The National Energy Research Supercomputing Center (NERSC) is the primary scientific computing facility for the Office of Science in the Department of Energy, and is the primary data storage center for ESS-DIVE.
- **ORCID** - The Open Researcher and Contributor Identifier provides anyone a persistent digital identifier (an ORCID iD) that distinguishes researchers from one another and provides a record that supports automatic links among all professional activities.