ESS-DIVE Package Service Updates

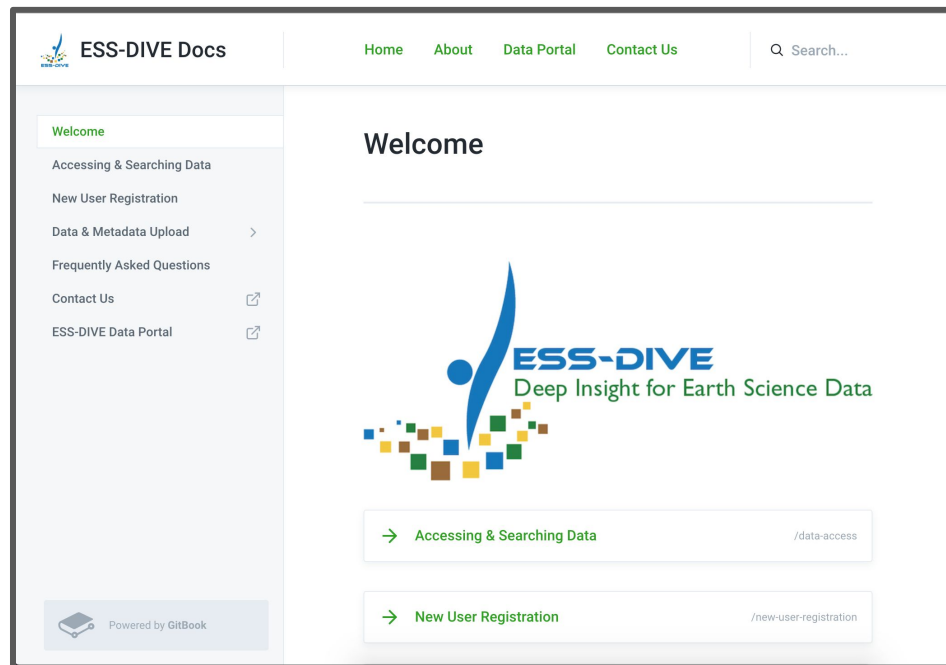**Fianna O'Brien**, Val Hendrix, Hesham Elbashandy

What is it?

*Through **ESS-DIVE Package Service**, organizations can **write code to store data packages** and then **reuse** the code to upload other data packages in the same or different repositories.*

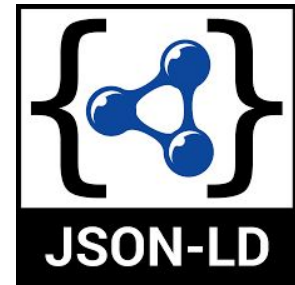# What's new in v1.2.0

NEW

- Update existing data packages
  - Update metadata
  - Add new data files
  - Replace existing data files
- Improved package display
  - Full metadata in single package search, including data files, with easy links to website
- New & improved documentation website
  - Easy to navigate tutorials
  - Troubleshooting guides
  - Helpful tips

JSON for Linked Data (JSON-LD)

JSON-LD (JavaScript Object Notation for Linked Data), is a method of encoding Linked Data using JSON (see https://json-ld.org/)

- The ESS-DIVE metadata schema is a restricted subset of https://schema.org/Dataset specification
- This covers all of the fields that ESS-DIVE collects from users (see ESS-DIVE JSON-LD Schema Proposal)
- JSON-LD is recommended by DataCite for package submission.
- JSON-LD has broad tool support and can be embedded in landing pages for harvesting by DataCite and indexing by Google.

Metadata Crosswalk example

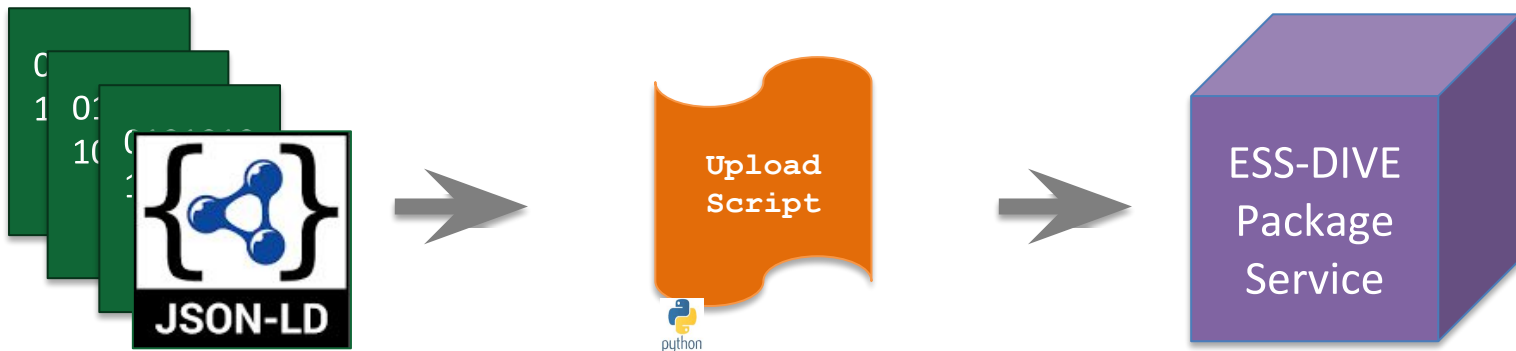| ESS-DIVE Field	JSON-LD	DataCite 4.1
Title	name	title
Alternative Identifiers	alternateName	alternateIdentifiers
Abstract	description	description[@type=Abstract]
Keywords	keywords	subjects
Data Variables	variablesMeasured	subjects
Publication Date	datePublished	publicationYear

bit.ly/ess-dive-crosswalk

REST API

The ESS-DIVE package service is a REST API

- REST (**Re**presentational **S**tate **T**ransfer) is an architecture style of designing internet-scale systems (scalable, fault tolerant, extensible)
- REST takes advantage of the standard web protocols
  - Uses well-defined status codes for error handling
  - Makes use of web protocol verbs for managing data
- Some aspects of a REST API are:
  - Data is received in a convenient format (e.g. JSON)
  - There are operations for manipulating data (create, update, delete)

How they all work together?

Upload & update data package *metadata + data* to ESS-DIVE.
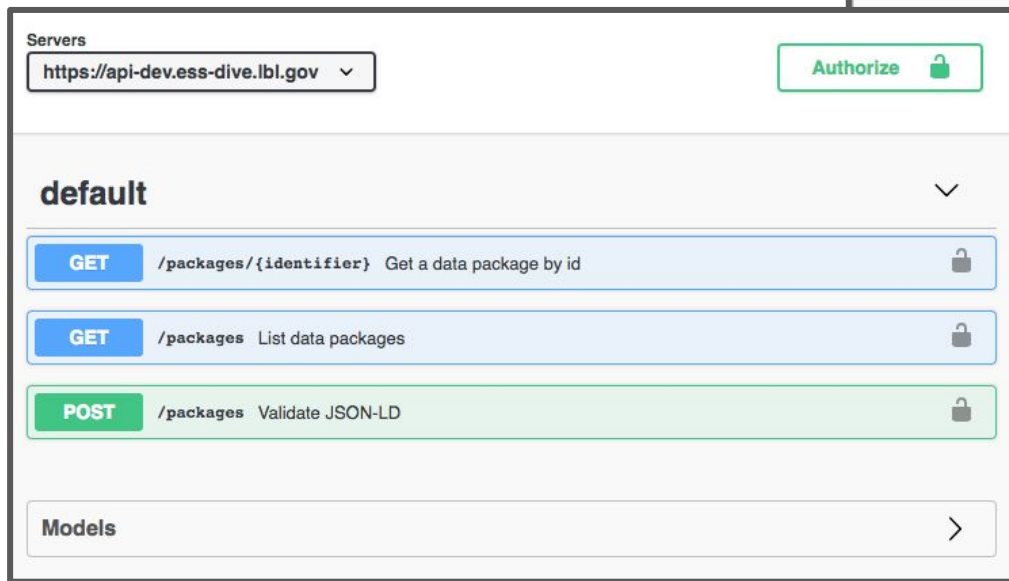
- ✔ Authenticate with a token
- ✔ Validate metadata with detailed error messages provided for invalid content
- ✔ Upload up to 1GB of data (metadata+data)
- ✔ Retrieve a list data packages submitted by you.
- ✔ Learn how to use the API from detailed documentation and examples

API Documentation

The ESS-DIVE API is described using the *OpenAPI Specification* description format.

We use *Swagger UI* to transform the OpenAPI description into rich interactive documentation of the API

- Each *operation* is defined in detail with examples
- *Structure of JSON* expected and produced is defined.

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 *1 GB* as a total size limit.

Validation Error Response

```
{
    "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"
    ]
}
```

Successful Upload

```
{
    "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
            proident, sunt in culpa qui officia deserunt mollit anim id est laborum.",
        "datePublished": "2019",
        "keywords": [
            "EARTH SCIENCE > BIOSPHERE > VEGETATION",
            "Climate Change"
        ],
        "editor": {
            "givenName": "Valerie",
            "familyName": "Hendrix",
            "email": "vchendrix@lbl.gov"
        },
        "creator": {
```

Update existing data package

NEW

https://api.ess-dive.lbl.gov/packages

- ***Metadata Updates***
  - Submit only the fields you want to update using JSON-LD
- ***Data File Updates***
  - *Upload additional files to an existing data package*
- For all actions, you must ***authenticate*** with a token

```
{
    "id": "ess-dive-186cc163880256f-20190927T203722674073",
    "viewUrl": "https://data-sandbox.ess-dive.lbl.gov/view/
        ess-dive-186cc163880256f-20190927T203722674073",
    "detail": "Data Package updated successfully.",
    "errors": null,
    "dataset": {
        "@context": "http://schema.org/",
        "@type": "Dataset",
        "@id": "ess-dive-186cc163880256f-20190927T203722674073",
        "name": "ESS-DIVE Webinar Demo [UPDATED]",
        "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 proident, sunt in culpa qui officia deserunt mollit anim id est
                laborum."
        ],
        "creator": [
            {
                "@type": "Person",
                "givenName": "Danielle",
                "familyName": "Christianson",
```

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
- For all actions, you must ***authenticate*** with a token.

https://api.ess-dive.lbl.gov/packages

```
"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
                enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi
```

Retrieve single data package metadata

- ***Retrieve a single data package*** *submitted by you with complete metadata*
  - *View complete JSON-LD record*
  - *View list of associated uploaded data files*
- For all actions, you must ***authenticate*** with a token.

https://api.ess-dive.lbl.gov/packages

```
"id": "ess-dive-3850023d22077eb-20190904T192506206816",
"viewUrl": "https://data-sandbox.ess-dive.lbl.gov/view/
    ess-dive-3850023d22077eb-20190904T192506206816",
"dateUploaded": "2019-09-04T19:25:06.615Z",
"dateModified": "2019-09-04T19:25:07.722Z",
"isPublic": false,
"citation": "Hanson P J; Riggs J (2015): Updated dataset title. SPRUCE.
    ess-dive-3850023d22077eb-20190904T192506206816",
"dataset": {
    "@context": "http://schema.org/",
    "@type": "Dataset",
    "@id": "ess-dive-3850023d22077eb-20190904T192506206816",
    "name": "Updated dataset title",
    "description": [
        "This data set reports selected ambient environmental monitoring data from the
            S1 bog in Minnesota for the period June 2010 through December 2016.
            Measurements of the environmental conditions at these stations will serve
            as a pre-treatment baseline for experimental treatments and provide driver
            data for future modeling activities.",
        "The site is the S1 bog, a Picea mariana [black spruce] – Sphagnum spp. bog
            forest in northern Minnesota, 40 km north of Grand Rapids, in the USDA
            Forest Service Marcell Experimental Forest (MEF). There are/were three
            monitoring sites located in the bog: Stations 1 and 2 are co-located at the
            southern end of the bog and Station 3 is located north central and adjacent
```

How do we put this into practice?

ESS-DIVE
Documentation Portal

ess-dive.gitbook.io/docs/

GitHub Examples

github.com/ess-dive/essdive-package-service-examples

Questions?