

# Leveraging the Globus Platform in your Web Applications

GlobusWorld – April 26, 2018

Greg Nawrocki

[greg@globus.org](mailto:greg@globus.org)



globus

# Topics and Goals

- **Platform Overview – Why expose the APIs**
- **A quick touch of the Globus Auth APIs**
- **Globus Transfer APIs**
  - Jupyter Notebook walkthrough
- **Integrating Globus into Science Gateways and Portals**
  - Modern Research Data Portal (MRDP) demo
- **Overall Goal - Build the foundation for what you will see in subsequent talks**



How can I integrate Globus into my research workflows?

Specifically applications I've developed or am planning to develop.

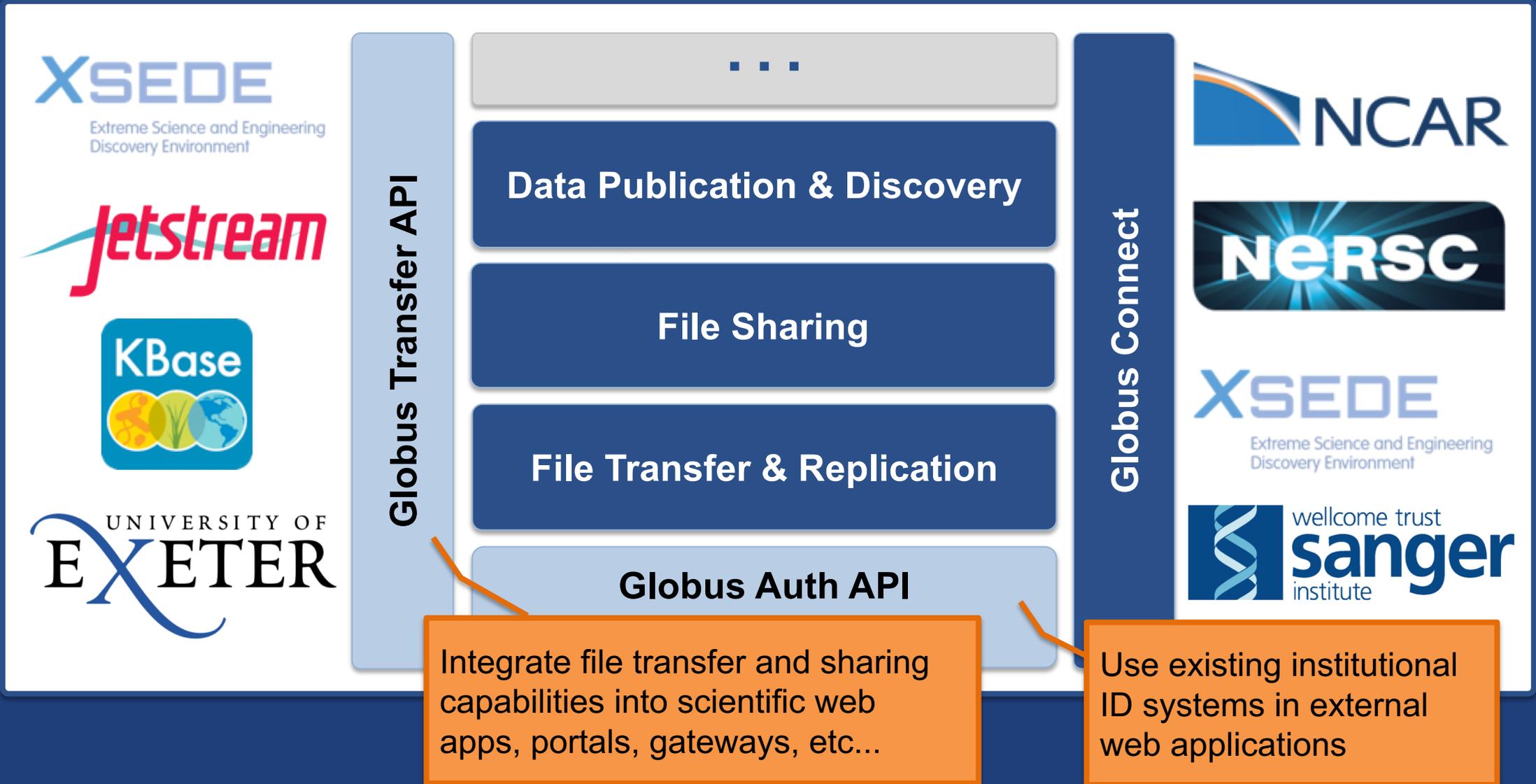


Globus serves as...

A platform for building science gateways, portals and other web applications in support of research and education



# Globus Platform-as-a-Service

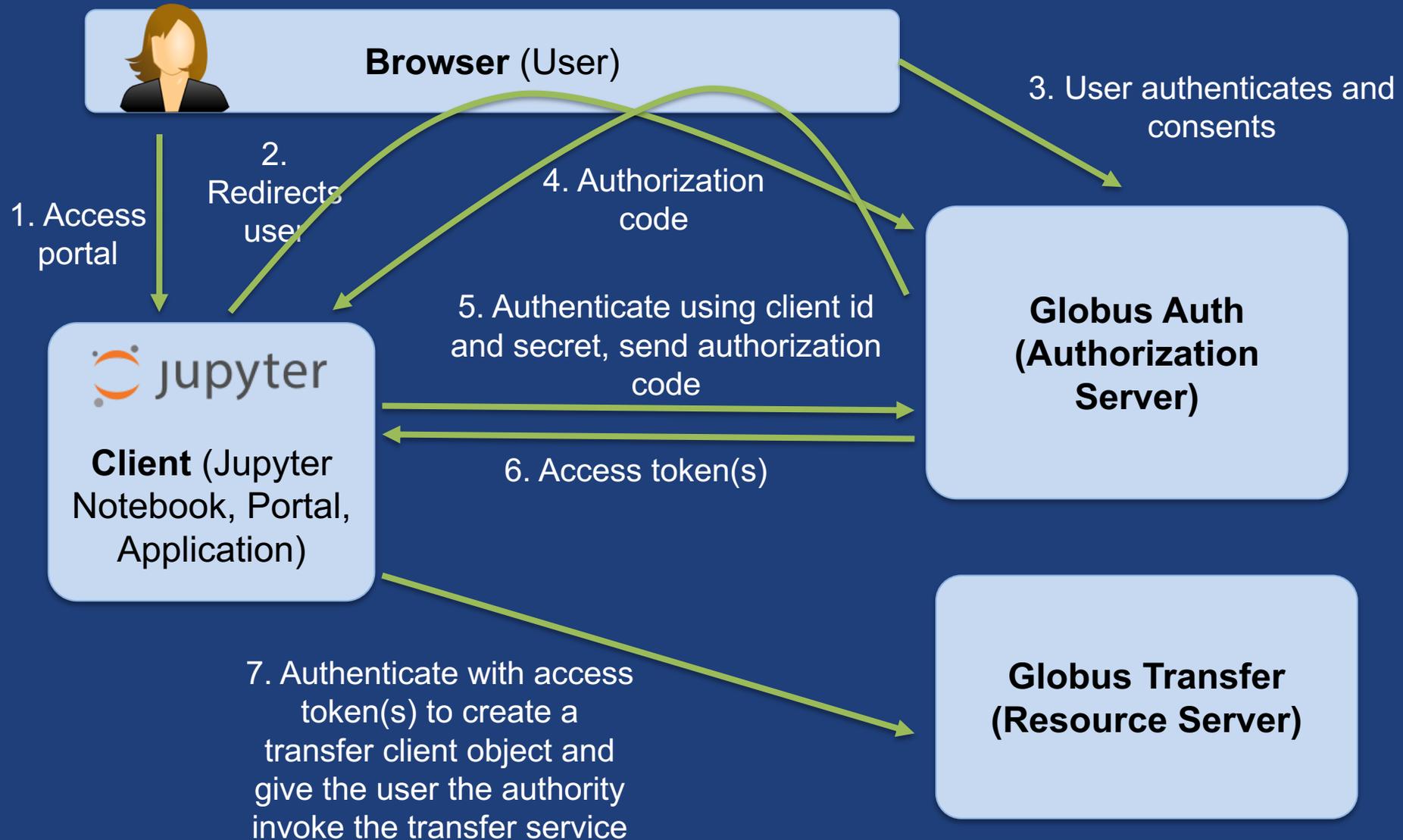


# PaaS Security Challenges – Globus Auth

- **How to provide:**
  - Login to apps
    - Web apps (Jupyter Notebook, Portals), Mobile, Desktop, Command line
  - Protect all REST API communications
    - App → Globus service (Jupyter Notebook, MRDP)
    - App → non-Globus service (MRDP)
    - Service → service (MRDP)
- **While:**
  - Not introducing even more identities
    - Providing a platform to consolidate those identities
  - Providing least privileges security model (consents)
  - Being agnostic to programming language and framework
  - Being web friendly
  - Making it easy for users and developers



# Authorization Code Grant





# Globus Platform **Transfer API**



# Globus Transfer API

- **Globus Web App consumes public Transfer API**
- **Resource named by URL (standard REST approach)**
  - Query params allow refinement (e.g., subset of fields)
- **Globus APIs use JSON for documents and resource representations**
- **Requests authorized via OAuth2 access token**
  - Authorization: Bearer asdfkqhafsdafeawk

**[docs.globus.org/api/transfer](https://docs.globus.org/api/transfer)**

# Globus Python SDK

- **Python client library for the Globus Auth and Transfer REST APIs**
- **`globus_sdk.TransferClient` class handles connection management, security, framing, marshaling**

```
from globus_sdk import TransferClient  
tc = TransferClient()
```

**[globus.github.io/globus-sdk-python](https://globus.github.io/globus-sdk-python)**



# TransferClient low-level calls

- **Thin wrapper around REST API**

- `post()`, `get()`, `update()`, `delete()`

`get(path, params=None, headers=None, auth=None, response_class=None)`

- `path` – path for the request, with or without leading slash
  - `params` – dict to be encoded as a query string
  - `headers` – dict of HTTP headers to add to the request
  - `response_class` – class response object, overrides the client's `default_response_class`
  - Returns: `GlobusHTTPResponse` object

# TransferClient higher-level calls

- **One method for each API resource and HTTP verb**
- **Largely direct mapping to REST API**

```
endpoint_search(filter_fulltext=None,  
                filter_scope=None,  
                num_results=25,  
                **params)
```



# Walkthrough Jupyter Notebook

- <https://jupyter.demo.globus.org>
  - Sign in with Globus
  - Verify the consents
  - Start My Server (this will take about a minute)
  - Open folder: globus-jupyter-notebooks
  - Open folder: globusWorld2018
  - Run SDK\_autoAuth.ipynb
- **If you mess it up and want to “go back to the beginning”**
  - Back down to the root folder
  - Run NotebookPuller.ipynb



# Endpoint Search

- **Plain text search for endpoint**
  - Searches owner, display name, keywords, description, organization, department
  - Full word and prefix match
- **Limit search to pre-defined scopes**
  - all, my-endpoints, recently-used, in-use, shared-by-me, shared-with-me
- **Returns: List of endpoint documents**

# Endpoint Management

- **Get endpoint (by id)**
- **Update endpoint**
- **Create & delete (shared) endpoints**
- **Manage endpoint servers**

# Endpoint Activation

- **Activating endpoint means binding a credential to an endpoint for login**
- **Globus Connect Server endpoint that have MyProxy or MyProxy OAuth identity provider require login via web**
- **Auto-activate**
  - *Globus Connect Personal* and *Shared* endpoints use Globus-provided credential
- **Must auto-activate before any API calls to endpoints**



# File operations

- **List directory contents (ls)**
- **Make directory (mkdir)**
- **Rename**
- **Note:**
  - Path encoding & UTF gotchas
  - Don't forget to auto-activate first



# Task submission

- **Asynchronous operations**
  - Transfer
    - Sync level option
  - Delete
- **Get submission\_id, followed by submit**
  - Once and only once submission

# Task management

- **Get task by id**
- **Get task\_list**
- **Update task by id (label, deadline)**
- **Cancel task by id**
- **Get event list for task**
- **Get task pause info**



# Bookmarks

- **Get list of bookmarks**
- **Create bookmark**
- **Get bookmark by id**
- **Update bookmark**
- **Delete bookmark by id**
  
- **Cannot perform other operations directly on bookmarks**
  - Requires client-side resolution

# Shared endpoints and access rules (ACLs)

- **Shared Endpoint – create / delete / get info / get list**
- **Administrator role required to delegate access managers**
- **Access manager role required to manage permission/ACLs**
- **Operations:**
  - Get list of access rules
  - Get access rule by id
  - Create access rule
  - Update access rule
  - Delete access rule

# Management API

- **Allow endpoint administrators to monitor and manage all tasks with endpoint**
  - Task API is essentially the same as for users
  - Information limited to what they could see locally
- **Cancel tasks**
- **Pause rules**



How can I do this in my  
[science gateway, data  
portal, web app, ...]?

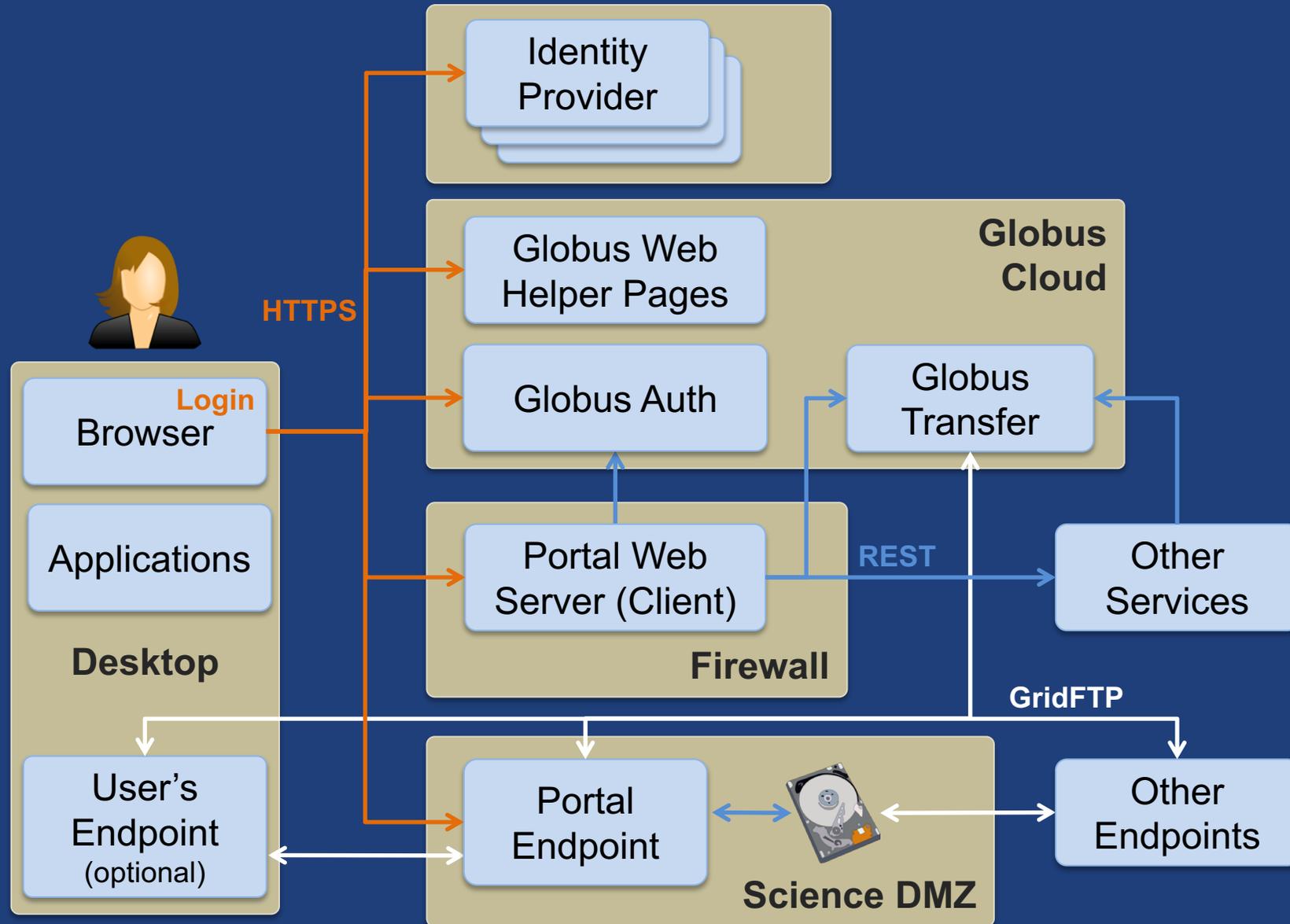


# Demonstration Modern Research Data Portal

<https://docs.globus.org/modern-research-data-portal/>



# Prototypical research data portal





# Data Distribution: ARM Climate Research Facility

### Data Selection Summary

Signed in as ANANTHAKRISHNANR1.

mergesonde1mace c1 @ fkb M1 [Generate Citation](#) 274 file(s) // 6014 MB

Order Complete Datastream  Extract Specific Measurements

Note: All variables will be delivered for this datastream.

**Measurement** : Atmospheric temperature  
**Variable** : Temperature // temp

2007-04-01

2007-12-31

Combine files by datastream

File format

Remove data flagged by Data Quality Reports (DQR) of type  Incorrect  Suspect

**Data Delivery Options**

- FTP
- Globus
- THREDDS
- Dropbox

Extraction options only apply when "Extract Specific Measurements" is selected.

Original files will be delivered as part of all orders.



# Analysis Workflow Integration: Wellcome Sanger

Sanger Imputation Service **Beta**

Home About Instructions ▾ Resources Status

## Sanger Imputation Service

This is a free genotype **imputation** and **phasing** service provided by the [Wellcome Trust Sanger Institute](#). You can upload GWAS data in VCF or 23andMe format and receive imputed and phased genomes back. Click [here](#) to learn more and [follow us on Twitter](#).

### Before you start

Be sure to [read through the instructions](#).

You will need to set up a free account with [Globus](#) and have [Globus Connect](#) running at your institute or on your computer to transfer files to and from the service.

### Ready to start?

If you are ready to upload your data, please fill in the details below to **register an imputation and/or phasing job**. If you need more information, see the [about](#) page.

Full name

Organisation

Email address

What is this [?](#)

Globus user identity

[➔ Next](#)

### News [@sangerimpute](#)

**11/05/2016**  
Thanks to [EAGLE](#), we can now return **phased data**. The HRC panel has been updated to r1.1 to fix a [known issue](#). See [ChangeLog](#) for more details.

**15/02/2016**  
Globus API changed, please see [updated instructions](#).

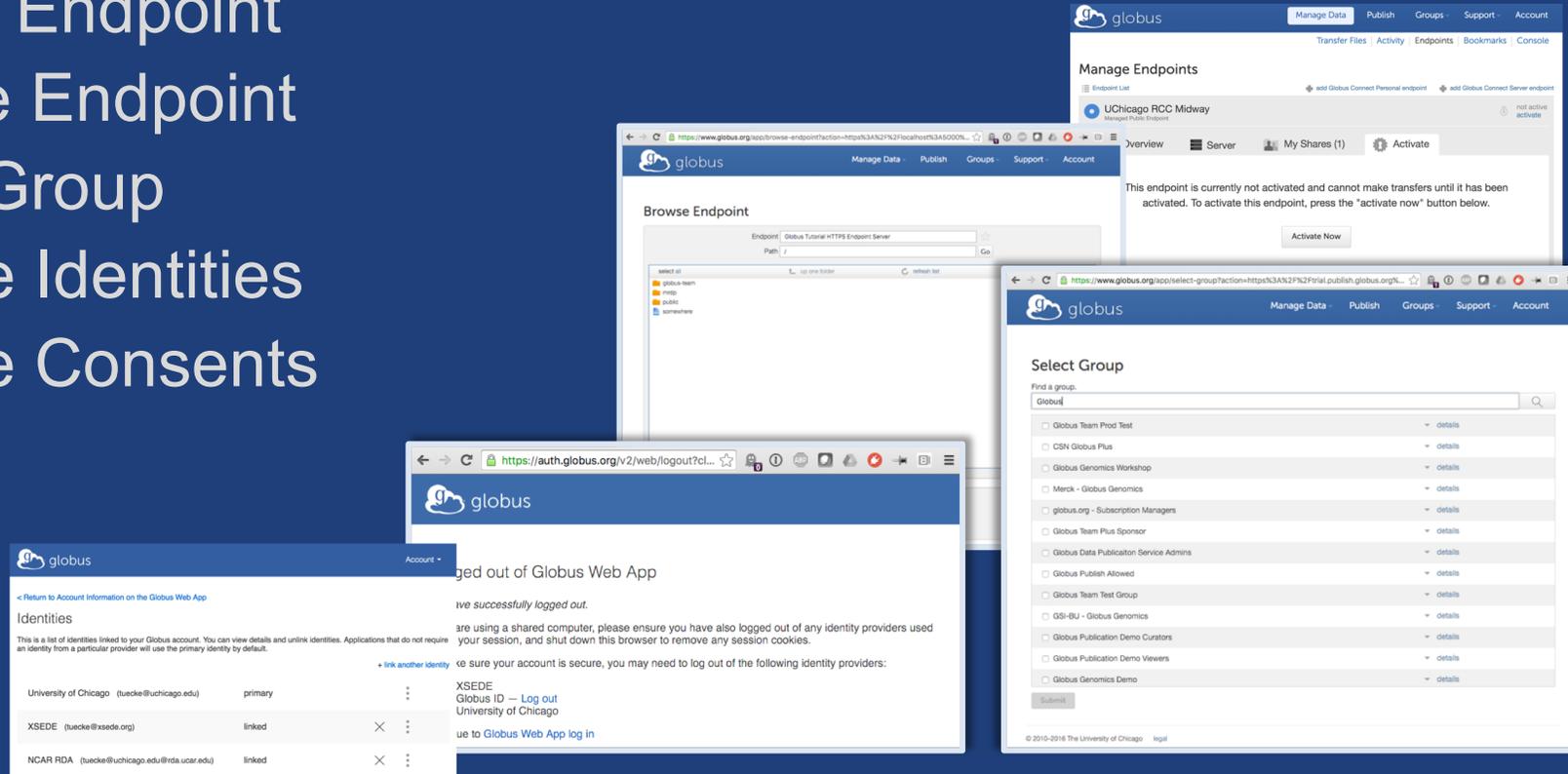
**17/12/2015**  
New status page and reworked internals. See [ChangeLog](#).

**09/11/2015**  
Pipeline updated to add some features requested by users. See [ChangeLog](#).

[📄 See older news...](#)

# Globus Helper Pages

- **Globus pages designed for use by your web apps**
  - Browse Endpoint
  - Activate Endpoint
  - Select Group
  - Manage Identities
  - Manage Consents
  - Logout



[docs.globus.org/api/helper-pages](https://docs.globus.org/api/helper-pages)



# Globus PaaS developer resources

globus-sdk-python 0.2.5 documentation »

Table Of Contents

- Globus SDK for Python (Beta)
- Installation
- Basic Usage
- API Documentation
- License

## Python SDK

### Installation

The Globus SDK requires Python 2.6+ or 3.2+. If a user does not have Python installed, they should first install Python.

The simplest way to install the Globus SDK is using the following instructions:

```
pip install globus-sdk
```

This will install the Globus SDK and its dependencies.

Bleeding edge versions of the Globus SDK can be installed using the following instructions:

```
git checkout https://github.com/globus/globus-sdk-python
cd globus-sdk-python
python setup.py install
```

### Basic Usage

Modern Research Data Portal

# Modern Research Data Portal

It's how research data management is done!

LOGIN | SIGN UP

## Requirements

- You need to be in the tutorial users group for sharing: <https://www.globus.org/app/groups/50b6a29c-63ac-11e4-8062-22000ab68755>
- Installed Globus Python SDK

## Jupyter Notebook

```
In [15]: from __future__ import print_function
tutorial_endpoint_1 = "ddb59ae1-0004-11e5-ba46-22000b92c6ec" # endpoint "Globus"
tutorial_endpoint_2 = "ddb59af0-6d04-11e5-ba46-22000b92c6ec" # endpoint "Globus"
tutorial_users_group = "50b6a29c-63ac-11e4-8062-22000ab68755" # group "Tutorial Users"
```

## Configuration

First you will need to configure the client with an OAuth2 access token. For the purpose of this tutorial, you can use the token from the Globus website. Click the "Jupyter Notebook" option and copy the resulting text below, or click on "Globus CLI" and

```
In [16]: transfer_token = None # if None, tries to get token from ~/.globus.cfg file
```

## Sample Application

[docs.globus.org/api](https://docs.globus.org/api)

[github.com/globus](https://github.com/globus)

# Support resources

- **Globus documentation: [docs.globus.org](https://docs.globus.org)**
- **Helpdesk and issue escalation: [support@globus.org](mailto:support@globus.org)**
- **Customer engagement team**
- **Globus professional services team**
  - Assist with portal/gateway/app architecture and design
  - Develop custom applications that leverage the Globus platform
  - Advise on customized deployment and integration scenarios