Figshare API Workshop

Open Repositories 2023 Stellenbosch, South Africa Twitter: #openrepos2023



Quick Activity

Photo by iorni.com on Unsplash





Instructors



Adrian-Tudor Pănescu Technical Team Leader



Andrew Mckenna-Foster Product Specialist (Assistant instructor)

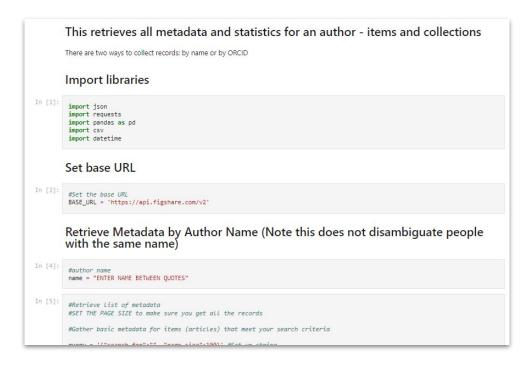




This workshop is an opportunity to create something useful to you

A personal statistics report

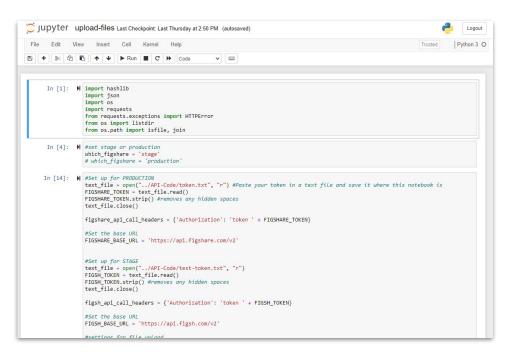
Try this in Google Colab.





Manage repository records

Wesleyan University uses the API to manage records and upload large file collections





Harvest records to create a data catalog

Either from another repository to Figshare or from Figshare to another repository

Macquarie University built a tool to harvest metadata from Dryad and deposit in their repository as linked file records

Dryad to Figshare harvest

Python 3 script to harvest metadata from a selected organisation's Dryad data records and generate matching metadata-only records in an institutional Figshare repository.

Votes		
	Browse Search on Macquarie Univer Q. Log	in
 This script only harvests metad dryad-based datasets from an i 		
 The existing (dryad-generated) generated Figshare record. 		
• The 'link file' option is used on		
 This is 'one-time only' script - t script to be an ongoing harvers 	File(s) stored somewhere else	
hananticitas	https://doi.org/10.5061/dryad.ct121	
Prerequisites	Please note: Linked content is NOT stored on Macquarie University and we can't guarantee its availability, quality, security or accept any i	liability.
o use this script, you will need to c epository account under which rec		
# SET TOKEN		
	Data from: Developmental stress increases reproductive success in male zebra finches	
	Cite Share + Collect	
	Dataset posted on 2022-06-10, 21:12 authored by Ondi L. Crino, Colin T. Prather, Stephanie USAGE METRICS (2) C. Driscoll, Jeffrey M. Good, Creagh W. Breuner 19 0 0	
	views downloads citations There is increasing evidence that exposure to stress during development can have	

https://github.com/mq-eresearch/dryad_to_figshare/tree/v1.0.0



Build a custom web application

University of Sheffield built a custom search interface: <u>https://orda.shef.ac.uk/</u>

Browse ORDA	Home / Search res	ult					
≡ All items			Discove	er our reseau	ch		
Faculties	Search:	All fields -	data	i our resear	cii.		Q
E Categories	Search.	, ar noido	untu				
😩 Conferences							
					Fi	Iter: All types 🔻	Sort: Date published
E Statistics		Workflow	tanate of Oman by Matteo Bormetti on 06 Jun 202 0 🗧 LIL				
Workflow Recording protocol for the identification and study of' by Mattoe Bornetti on 06/06/2023 0+03 PM Report Enabling Research Participation		Humbe Report by	ng Research Participation r Habiba Aminu et al on 05 Jun 202 19 📕 🛄		nderserve	d Populations ir	Yorkshire and
Power in Underserved Populations in' by Habiba Aminu et al on 05/06/2023 12:39 PM Software Yanalysis code associated with project The effect of' by Inge Kensbergen et al on 05/06/2023 12:20 PM Dataset		Analysis code associated with project "The effect of proportional pricing on hypothetical alcohol purchasing in two online experiments" Software by Inge Kersbergen et al on 05 Jun 2023 12:20:41					
Data associated with project The effect of proportional' by inge Kersbergen et al on 05/06/2023 12:16 PM		purcha (EMBARGI	sociated with project "Th sing in two online experi DED) Dataset by Inge Kersbergen e	ments"		al pricing on hyp	oothetical alcohol



Please note down 1-2 projects you'd like to work on

As we progress through the workshop apply what you learn to your project

You may not finish the project today, but should leave with a clear roadmap

Photo by No Revisions on Unsplash





Workshop Website

https://amckennafoster.github.io/figshareapi-workshop.github.io/index.html

- Schedule
- Links
- Resources





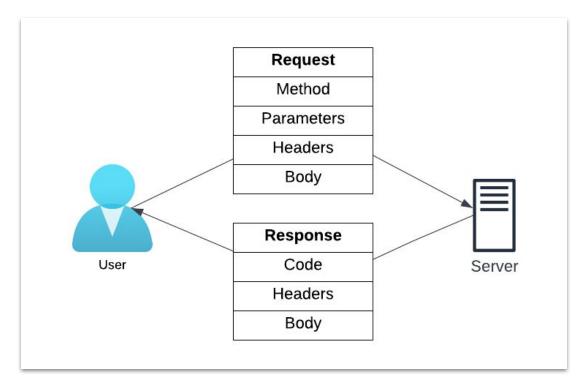
Workshop Outline

- Introduction to HTTP requests and APIs
- Introduction to Figshare's API(s)
- Accessing Figshare's REST API using Postman
- Building an application over Figshare's REST API



HTTP Requests and APIs

What is a HTTP request?





What is a HTTP request?

POST /v2/articles/search?offset=10&limit=1 HTTP/1.1

Host: api.figshare.com

```
Content-Type: application/json
```

Authorization: Bearer b2be49036a3158c5edd5a0553ae9

```
{"search for": ":tags: test"}
```



HTTP is stateless

- Requests are independent from one another
- Each request needs to contain enough information in order for the server to be able to respond to it
- This is why, for example, you need to send authentication information with each HTTP request

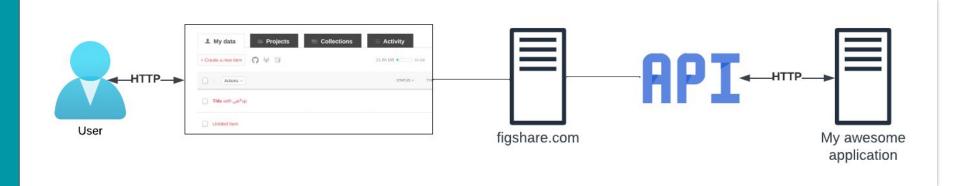


What is an API?

- Abbreviation for **Application Programming Interface**
- It's an interface which allows other applications to use the functionality of a software system



What is an API?





What is REST?

- Abbreviation for REpresentational State Transfer
- A set of principles on how web APIs should be built
- Servers should always respond with the representation of a *resource*



Example of a REST resource

```
"files": [
        "id": 25133441,
        "name": "Panescu Advances in Digital Repositories PhD Thesis.pdf",
        "size": 1411929,
        "is link only": false,
        "download url": "https://ndownloader.figshare.com/files/25133441",
        "supplied md5": "87f39b4da1a371572523b59bf83bb32e",
        "computed md5": "87f39b4da1a371572523b59bf83bb32e"
],
"custom fields": [],
"authors": [
        "id": 1402906.
        "full name": "Adrian-Tudor Panescu".
        "is active": true,
        "url name": "Adrian-Tudor Panescu",
        "orcid id": "0000-0002-8940-898X"
1.
"figshare url": "https://figshare.com/articles/thesis/Advances in Digital Repositories/8247626",
"description": "This thesis presents the evolution of research repositories, from collections of peer-reviewed outputs, such as
   journal articles or monographs, to collections holding a wide array of materials, including data sets, preprints, scientific
    software or protocols, based on the new realities of the scientific research endeavour. It also discusses a number of novel
   practical solutions to repository issues such as licencing, bibliographic record modelling, dissemination and linking, or
    record management and migration, by employing new technologies such as linked data, blockchain, and extract, transform and
    load frameworks.".
"funding": null,
"funding list": [],
"version": 2,
"status": "public".
```

What is JSON?

- Abbreviation for JavaScript Object Notation
- Human-readable data interchange format
- Consists of key-value pairs, lists and other basic types (numbers, strings)
- For example: { "name": "Jane Doe", "someList": [1, "two", 3] }



REST and request methods (verbs)

- GET: retrieve a resource
- POST: add a new resource
- PUT: update a new resource
- PATCH: partially update a new resource
- DELETE: remove a resource



HTTP status codes

- Standard (and easiest) way of understanding what happened with your requests
- Some common status codes:
 - 200 OK
 - 400 Bad request
 - 403 Forbidden
 - 404 Not Found
 - 500 Internal Server Error
- Full list: <u>https://developer.mozilla.org/en-US/docs/Web/HTTP/Status</u>



HTTP request body

- The actual data sent to the server
- It's any sequence of bytes (JSON, XML, an image file) but...
- We need to tell the server what we are sending:
 - Content-Type: application/json
- The server response can also contain any sequence of bytes but...
- We can tell the server what we accept:

• Accept: */*

The request/response bodies are optional



HTTPS

- Secure HTTP
- A secure channel is created between the client and the server
- Request details (URL, headers, body) are not visible on the open network
- Important if sensitive data (e.g., authentication credentials) is exchanged



Tea Break 30 min

The Figshare API

Figshare API(s)

- Figshare actually has 4 different APIs:
 - REST API: <u>https://api.figshare.com/v2</u>
 - Stats API: <u>https://stats.figshare.com/</u>
 - OAI-PMH: <u>https://api.figshare.com/v2/oai</u>
 - ResourceSync: <u>https://scholardata.sun.ac.za/.well-known/resourcesync</u>



Figshare API(s)

- REST API:
 - v1: not very RESTful, deprecated in 2021
 - v2: current version, centered around resources, backwards-compatible
 - v3: will better expose all UI functionality, improve documentation, better integrate stats
- Stats API:
 - Distinct due to the fact that stats are held in a different database and processed by different services than metadata
- OAI-PMH:
 - implemented in order to allow harvesting all Figshare items
- ResourceSync:
 - Another standard from OAI allowing synchronization between systems
 - Used for implementing sitemaps



ACTIVITY

Try GET request

Instructions to retrieve full metadata and retrieve views are here: <u>https://amckennafoster.github.io/figshare-api-workshop</u>. .github.io/workshop/workshop-api-basics.html

OpenAPI and Swagger

- OpenAPI
 - A specification language for REST APIs
 - Allows prototyping and generating documentation, client code and test cases
- Swagger:
 - OpenAPI implementation
 - Editor: <u>https://editor.swagger.io/</u>
 - Swagger UI: https://swagger.io/tools/swagger-ui/



Figshare and OpenAPI

- Documentation publicly available at <u>https://docs.figshare.com</u>
- Source code for documentation: <u>https://github.com/figshare/user_documentation</u>
- Swagger file: <u>https://github.com/figshare/user_documentation/blob/master/swagger_documentation/swagger.json</u>



API authentication

REST API - personal token:

https://help.figshare.com/article/how-to-get-a-personal-token

- OAuth2 create applications which require users to log into their own Figshare account
- Stats API user/password, contact support@figshare.com for an account
- Credentials are sent to the API server via the Authorization header
- Make sure credentials are stored in a secure location!



ACTIVITY

Authenticate and try a POST request

https://amckennafoster.github.io/figshare-api-workshop .github.io/workshop/workshop-api-basics.html

Sandbox accounts

global.user@figsh.com.bk	global.user2@figsh.com.bk	global.user3@figsh.com.bk	
Bongani Jwara	Andisiwe Magocoba	Samuel Simango	
Charl Roberts	Bubele Bido	Songezo Mpikashe	
Nambitha Manqola	Janina Van der Westhuizen	Sizwe Ngcobo	
Nkululeko Magwaza	Yan Han	J	

https://global.figsh.com

global.user4@figsh.com.bk	global.user5@figsh.com.bk
Regina Sikhosana	Eddie Mathiba
Richard Nobebe	Martin Dreyer
Xabiso Xesi	Motlanalo Hlophe
	Tshinakaho Malesa



DEMO

A quick overview of the Figshare REST API resources

Live repository: <u>https://docs.figshare.com/</u> Sandbox repository: <u>https://docs.figsh.com/</u>

Q&A & PROJECT TIME

Use this time to ask questions and lay out the requirements for your project(s):

- 1. What is your objective
- 2. What information will you need to retrieve or send
- 3. What endpoints will you need to use

Lunch 1hr

Performing Figshare API requests

Postman

- Tool for performing API requests
- Can be downloaded from https://www.postman.com/downloads/
 - Web version available at <u>https://web.postman.co/</u>
- All demo requests available as a Postman collection: <u>https://www.postman.com/tudor/workspace/public-workspace/collection/3296</u> <u>25-1a755e1d-811a-48b8-8471-860e3f09f0ec</u>



DEMO

Using Postman to perform requests

https://amckennafoster.github.io/figshare-api-workshop.gith ub.io/workshop/postman-use-api.html

Retrieve item metadata

OR23	OR23 Workshop / Retrieve item metadata									
GET × https://api.figsh.com/v2/articles/8418909 Send ~										
Parar	Params Authorization Headers (5) Body Pre-request Script Tests Settings Cookies									
Quer	y Params									
	Кеу	Value	Description •••• Bulk Edit							
	Кеу	Value	Description							
Body Pret	Cookies Headers (12) Test Results ty Raw Preview Visualize JSON ~	Status: 200 OK	Time: 448 ms Size: 4.04 KB 🖺 Save as Example 🚥							
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17	<pre>"files": [</pre>	4007cd431f376101a",	1							



Create > Add metadata > Publish



Create item

OR23 Workshop / Create item	🖺 Save 🗸 👓 🥖 🗐
POST v https://api.figsh.com/v2/account/articles	Send ×
Params Authorization • Headers (8) Body • Pre-request Script Tests Settings	Cookies
🜑 none 🌑 form-data 🌑 x-www-form-urlencoded 💿 raw 🜑 binary 🜑 GraphQL JSON 🗸	Beautify
<pre>1 { 2 ····"title": "My test item" 3 }</pre>	I
Body Cookies Headers (11) Test Results 🕀 Status: 201 Created Time: 149 ms	Size: 615 B 🖺 Save as Example 👓
Body Cookies Headers (11) Test Results Image: 149 ms Image: 149 ms	Size: 615 B 🖺 Save as Example 👓



Retrieve categories

OR23 Workshop / Get categories	🖺 Save 🗸 👓
GET v https://api.figsh.com/v2/categories	Send ×
Params Authorization • Headers (6) Body Pre-request Script Tests • Settings	Cookies
<pre>pm.environment.set("categoryId", pm.response.json()[1].id);</pre>	Test scripts are written in JavaScript, and are run after the response is received. Learn more about tests scripts ↗ > Snippets Get a variable Get a collection variable Set an environment variable Set a global variable Set a global variable
ody Cookies Headers (12) Test Results	ie: 542 ms Size: 431.88 KB 🖺 Save as Example 🚥
Pretty Raw Preview Visualize JSON ~ =	



Update item

OR23 V	Vorkshop / Up	odate item							🖺 Save	~	000	0	Ē
PUT	~ ht	tps://api.figsh.com/v2/a	ccount/articles/8	508592								Send	
Params	Authorizatio	on • Headers (8)	Body • Pre-r	equest Script	Tests Setti	ngs						Co	ookies
none	form-data	a 🔵 x-www-form-urle	ncoded 🛛 🦲 raw	linary	GraphQL J	SON	~					Be	eautify
1 2 3 4 5 6	descrip catego	: "My test item", ption": "Lorem ipsu ries": [{{categoryI ["test"]											T
Body C	cookies Heade	ers (11) Test Results			(A	Status: 205 Reset Content	Time: 233 ms	Size: 597 B	B) Save a	as Examp	le •••
Pretty	Raw	Preview Visualize	JSON 🗸	<u>-</u>								ſ	Q
1 2 3 4	{ "locatio "warning }	on": " <u>https://api.f</u> gs": []	igsh.com/v2/ad	count/articles	s/8508592",								I



Publish item

OR23 Workshop / Publish item	🖺 Save 🗸 👓 🥖 🗐
POST v https://api.figsh.com/v2/account/articles/8508592/publish	Send v
Params Authorization • Headers (7) Body Pre-request Script Tests Settings	Cookies
Type Bearer Token V Token c91a0bcfda12aef3a877508b64fat	fe5cb4e2{
The authorization header will be automatically generated when you send the request. Learn more about <u>authorization</u> A	
Body Cookies Headers (11) Test Results	ms Size: 560 B 🖺 Save as Example 👓
Pretty Raw Preview Visualize JSON V	G Q
<pre>1 { 2 "location": "https://api.figsh.com/v2/articles/8508592" 3 }</pre>	I



Retrieve views

GET	https://stats.figsh.com/total/views/article/8460124							
Params	Authorization • Headers (6)	Body Pre-request Script Tests	Settings		Cookies			
Туре	Basic Auth \sim	Username		global_stats				
Body Coo	kies Headers (8) Test Results			C Status: 200 OK Time: 710 ms Size: 291 B	🖺 Save as Example 🛛 👓			
Pretty	Raw Preview Visualize	JSON V 🔁			r Q			
1 { 2 3 }	"totals": 5				T			



Retrieve monthly views breakdown

OR23 Workshop / Get monthly stats	🖺 Save 🗸 👓
GET v https://stats.figsh.com/global/breakdown/month/views/article/8418909	Send ~
Params Authorization • Headers (6) Body Pre-request Script Tests Settings	Cookies
💿 none 💿 form-data 💿 x-www-form-urlencoded 💿 raw 💿 binary 💿 GraphQL	
This request does not have a body	
Body Cookies Headers (10) Test Results 🕀 Status: 200 OK Time: 485 ms	Size: 496 B 🖺 Save as Example 👓
Pretty Raw Preview Visualize JSON ~ -e	r _d Q
<pre> "breakdown": { "2023-05": { "Romania": { "Romania": 1, "Iasi": 1 }, "United States": { "United States": { "Unknown": 1 "Unknown": 1</pre>	
20 0	1



Retrieve repository accounts

OR23 Workshop / Get accounts	🖺 Save 🗸 👓 🥖 🗐
GET v https://api.figsh.com/v2/account/institution/accounts?limit=1&offset=0	Send ~
Params • Authorization • Headers (6) Body Pre-request Script Tests Settings	Cookies
💿 none 🜑 form-data 🜑 x-www-form-urlencoded 🜑 raw 🜑 binary 🜑 GraphQL	
This request does not have a body	
Body Cookies Headers (14) Test Results	Size: 934 B 🖺 Save as Example 👓
Pretty Raw Preview Visualize JSON V	r_ Q
<pre>1 [2 { 3 "id": 1997142, 4 "first_name": "figshare admin", 5 "last_name": "global", 6 "email": "globaladmin@figsh.com.bk", 7 "active": 1, 8 "institution_id": 3682, 9 "institution_user_id": "", 10 "quota": 10737418240, 11 "used_quota": 159246054, 12 "user_id": 2750658 13 } 14]</pre>	I



Create account

OR23 Workshop / Create account	🖺 Save 🗸 👓
POST · https://api.figsh.com/v2/account/institution/accounts	Send
Params Authorization • Headers (8) Body • Pre-request Script Tests Settings	Cookies
🔵 none 🜑 form-data 🜑 x-www-form-urlencoded 💿 raw 🜑 binary 🜑 GraphQL JSON 🗸	Beautify
<pre>1 { 2</pre>	
Body Cookies Headers (10) Test Results	Size: 473 B 🖺 Save as Example 👓
Pretty Raw Preview Visualize JSON ~	r Q
1 { 2 "account_id": 2139246 3 }	I



PROJECT TIME

Postman

Use Postman to test out the endpoints you need for your project.

If your project requires chaining API calls, decide what language you will use. There are resources for <u>Python</u> <u>here</u>.

Generate code using Postman

- Postman can generate source code in the language of your choice for the performed requests
- You can easily prototype using a *notebook*, for example <u>https://colab.research.google.com/</u>



Generate code using Postman

GET Retrieve item metadata	+ 000						New Environm	nent	~	
OR23 Workshop / Retrieve	e item metadata	🖺 Save 🗸	000	۶ 📮	.11	Code snippet				\times
GET ~ https://	/api.figsh.com/v2/articles/8418909		Sen	d V	Ē	Python - Request	s v	ক্ষ		ſ
Params Auth Headers (5) Body Pre-req. Tests Settings			Cookies		1 import r 2 3 url = "h	requests https://api.fi	ash com/v2/a	rticles	s /
Туре					:@:	8418 4		g511.com/ v2/ u	, creter	
Inherit auth from p \vee					i	5 payload 6 headers 7				
The authorization header will be automatically generated when you send	This request is using No Auth fr	om collection OR23	Workshop.	I		head	e = requests.r lers=headers,	-		
Response				~		9 10 print (re 11	esponse.text)			



Generate code using Postman

C	0		Jntitled0.ipynb ☆ Edit View Insert Runtime Tools Help <u>All changes saved</u>
≣	-	+ Cod	e + Text
Q	V Os	D	import requests
[x }			<pre>url = "https://api.figsh.com/v2/articles/8418909"</pre>
			<pre>payload = {} headers = {}</pre>
			response = requests.request("GET", url, headers=headers, data=payload)
			<pre>print(response.text)</pre>
			{"files": [{"id": 830411360, "name": "aes.JPG", "size": 172621, "is_link



Tea Break 30 min

Building an application

Requirements

- Retrieve the items from https://global.figsh.com
- Retrieve the number of views for each item
- Display the data on a HTML page



Tools

- Server:
 - Python: <u>https://www.python.org/</u>
 - Flask: <u>https://flask.palletsprojects.com/en/2.3.x/</u>
- Web interface:
 - React: <u>https://react.dev/</u>
- Development environment:
 - Visual Studio Code: <u>https://code.visualstudio.com/</u>
 - Python in Visual Studio Code:

https://code.visualstudio.com/docs/python/python-tutorial

- Code generation:
 - ChatGPT: <u>https://chat.openai.com/</u>

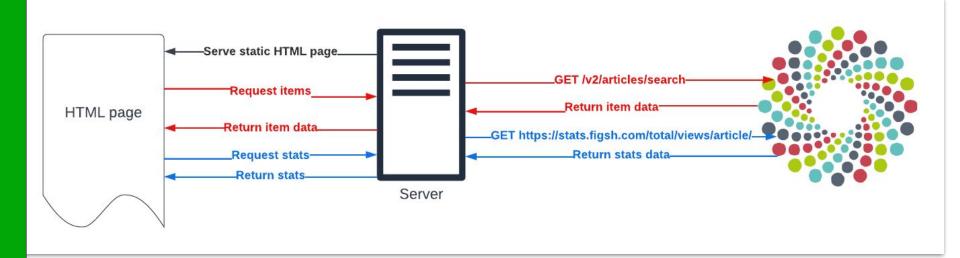


CORS

- Abbreviation for cross-origin resource sharing
- Mechanism which indicates from what origins (<u>http://localhost:5000</u> in our case) a browser is allowed to make HTTP API requests
- Security feature which prevents malicious sites stealing user information from legitimate sites
- By default Figshare's API does not allow cross-origin requests; solutions:
 - Make API requests via a (proxy) server
 - Contact <u>support@figshare.com</u>



System diagram





Further documentation

- Python Flask: <u>https://flask.palletsprojects.com/en/2.3.x/</u>
- Python requests: <u>https://requests.readthedocs.io/en/latest/</u>
- React: <u>https://react.dev/</u>
- Fetch: <u>https://developer.mozilla.org/en-US/docs/Web/API/Fetch_API</u>
- Promise:

https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/Promise



And that's everything

- What we've covered:
 - Basics of HTTP APIs
 - Overview of Figshare's APIs
 - How to use Postman to perform HTTP requests
 - How to prototype applications that make use of an API
 - How to build a basic application that displays data from Figshare's API



And that's everything

- What we'll make available to you:
 - This presentation: <u>https://amckennafoster.github.io/figshare-api-workshop.github.io/assets/Figshare-API-</u> <u>Workshop-Open-Repositories-2023.pdf</u>
 - The Postman collection with the performed requests: <u>https://www.postman.com/tudor/workspace/public-workspace/collection/329625-1a755e</u> <u>1d-811a-48b8-8471-860e3f09f0ec</u>
 - The ChatGPT chats:
 - Workshop:

https://chat.openai.com/share/28e38008-350f-420f-b53a-9710d9987817 (or PDF)

• Complete:

https://chat.openai.com/share/fea4383b-3b32-4608-9699-69aca739aa6f (or PDF)

• The workshop website:

https://amckennafoster.github.io/figshare-api-workshop.github.io/index.html



Q&A & PROJECT TIME

Use this time to ask questions and work on your project.

Thank you!

And don't hesitate to contact us during OR 2023 and beyond!