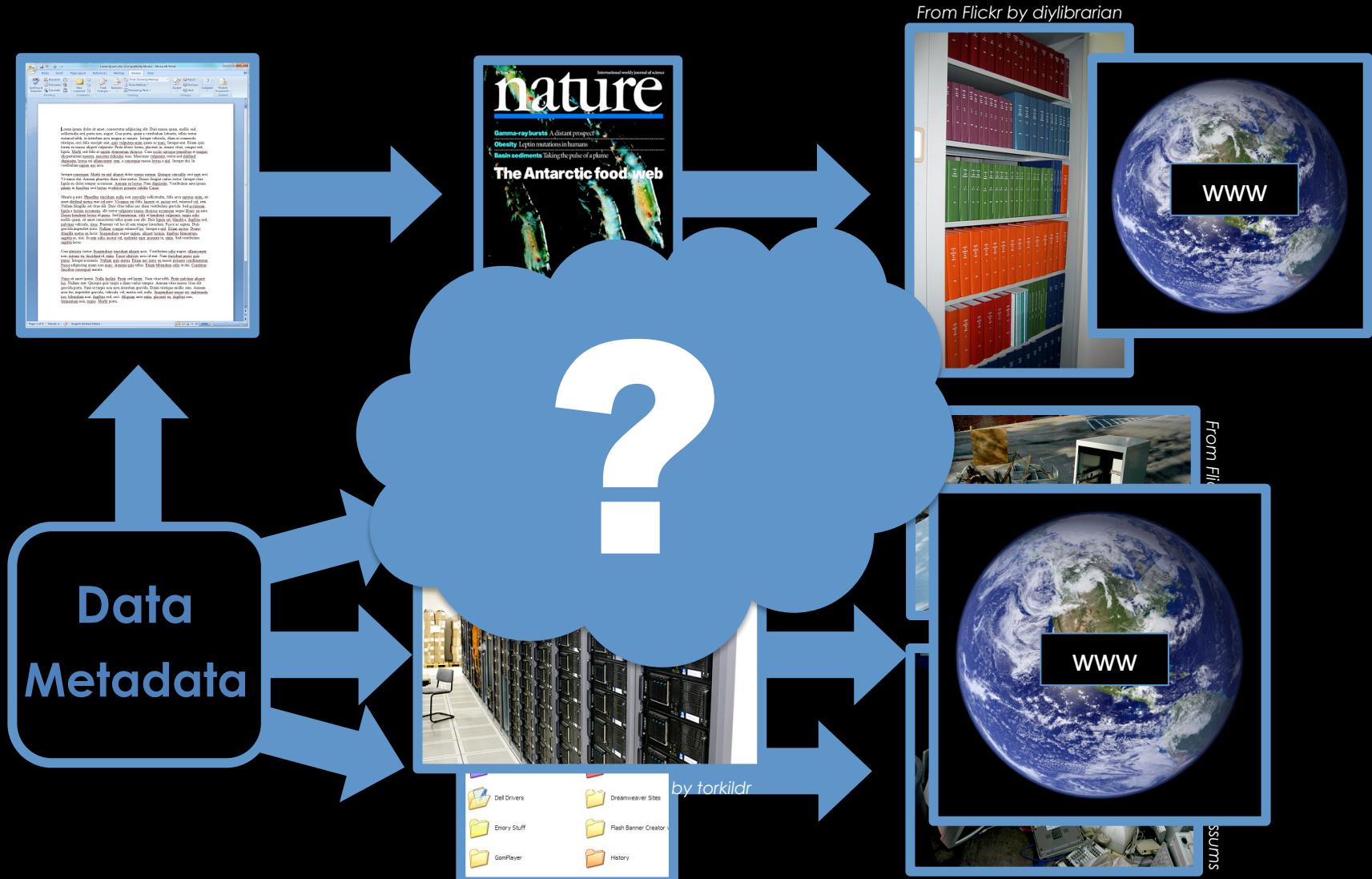




Data Curation for Excel

Carly Strasser



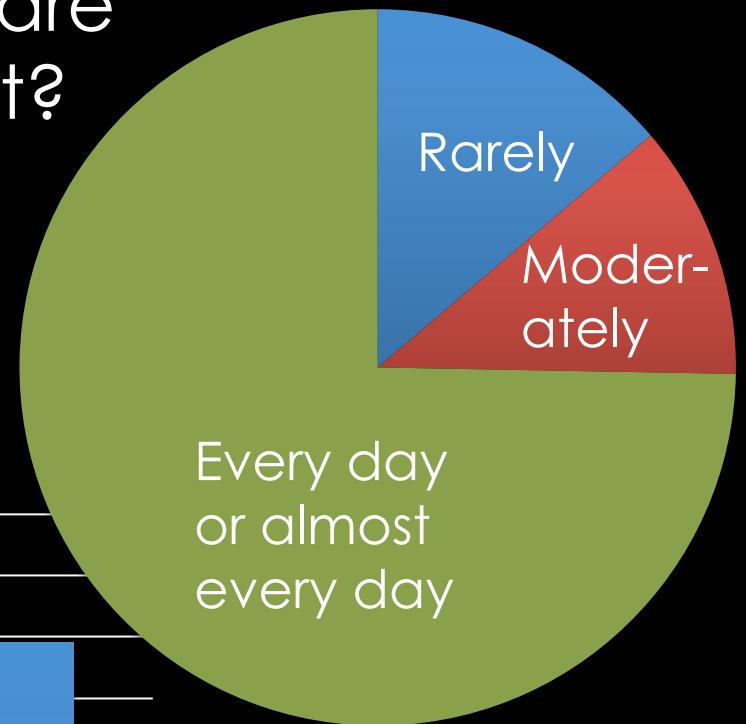


Intercept
researchers where
they already work

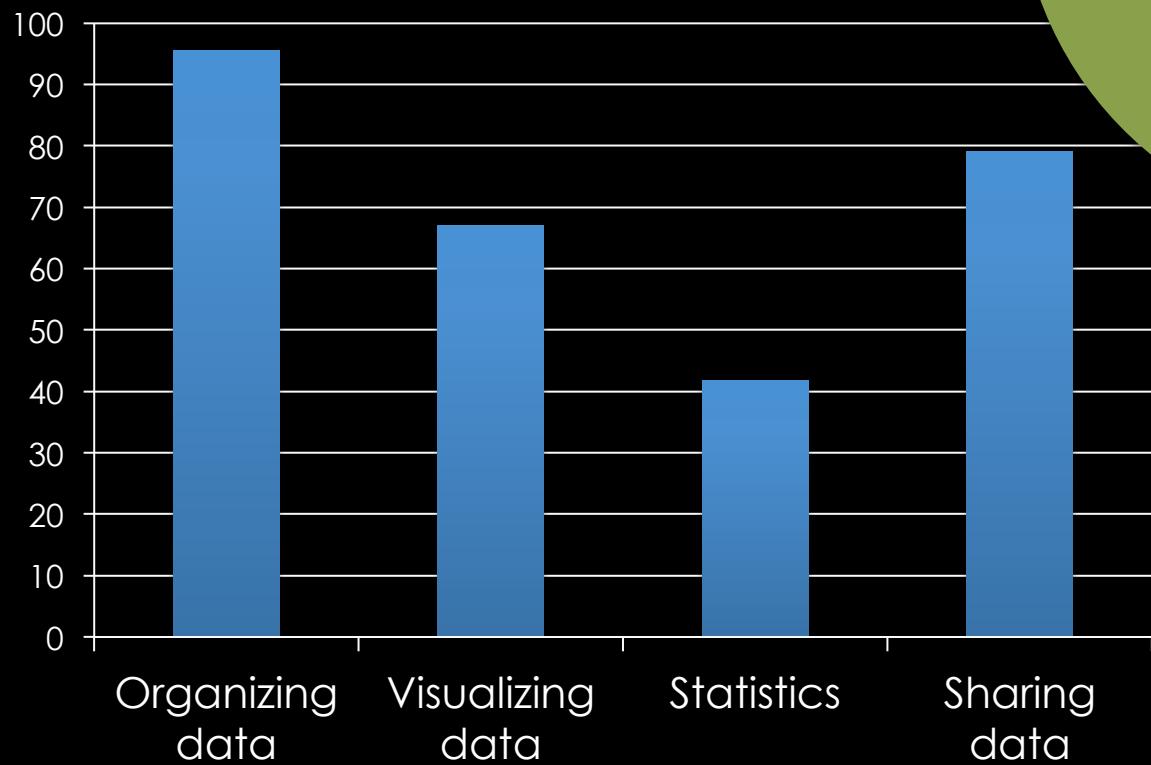




How often are they using it?



What are they using it for?



GORDON AND BETTY
MOORE
FOUNDATION

Microsoft®
Research



University of California
CDL
California Digital Library

Data**ONE**



Data
management
& organization

Facilitate
Archiving
Sharing
Publishing

Data Reuse &
Reproducibility

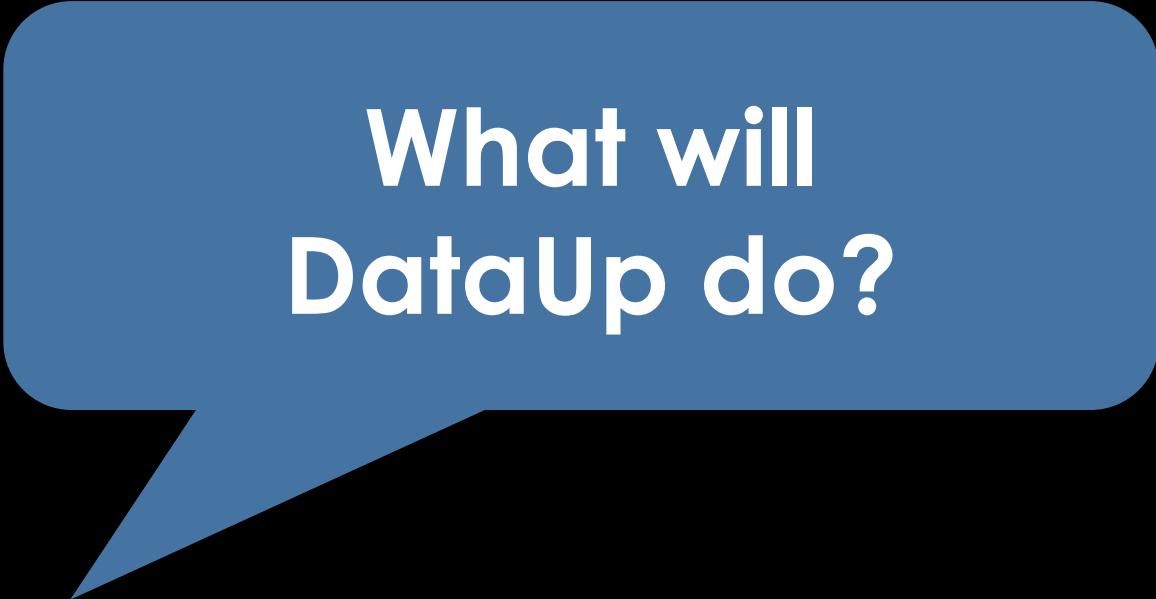


Open
Source Tool

Add-in & Web
Application

Earth,
environmental,
ecological
researchers



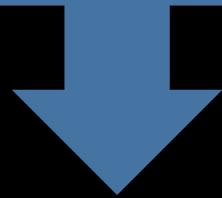


**What will
DataUp do?**



**What do
scientists need?**

Requirements



Features

Best practices check

Generate metadata

Generate citation

Post data to repository

Best Practices Check

copes_log4.xlsx - Microsoft Excel

Home Insert Page Layout Formulas Data Review View DataUp

Post to Repository Check Data Description Column Description Describe Data Generate Citation

T19 f_x 3100

	A	L	M	N	O	P	Q
1		S8T15_R3		S12T15_R1		S12T15	
2	23-Oct		300		380		200+
3	25-Oct		450		900		200+
12	27-Oct		500		810	765	450+
13	29-Oct		400		1450	1100	350
14	31-Oct		350	500	200	1150	550
15	2-Nov	1000	200	100	800	840	850
16	4-Nov	2425	250	1600	550	2080	250
17	6-Nov	1300	165	800	760	1640	300
18	8-Nov	1450	300	1600	600	1440	180
19	10-Nov	1050	250	2120	500	2200	160
20	12-Nov	1320	300	1400	950	1500	50
21	14-Nov	1240	300	680	400	1640	50
22	16-Nov	720	630	840	450	760	180
23	18-Nov	1000	420	880	700	600	80
24	20-Nov	0	400	440	200	480	<50
25	22-Nov	90	350	510	250	180	200
26	24-Nov	30	350	480	100	60	200
27	26-Nov	90	180	390	100	240	<100
28	28-Nov	60	<100	330	<100	0	<100
29	30-Nov	30	<50	120	<50	30	<50
30	2-Dec	0	<50	90	<25	0	<25
31	4-Dec	0	<25	0	<10	0	0
32	6-Dec		0	0	0		
33	8-Dec						
34	10-Dec						
35	12-Dec						
36	14-Dec						
37	16-Dec						

Error Details

Remove Selected

Fixable Errors Non fixable Errors

Oct_7to15(6) Oct17to26(4) Oct27toNov16(6) scope_data_No...(6) Notes (6) red_days

Embedded comments Remove Advice

Embedded charts, tables, pictures Remove Advice

Color coded text or cell shading Remove Advice

Non-contiguous data Remove Advice

Blank Cells Remove Advice

Special Characters Remove Advice

Generate Metadata

A	B	C	D	E	F	G	H	I	J	K	L
1 Name	Value										
2 Creator:First name	Carly										
3 Creator:Last name	Strasser										
4 Creator:Address											
5 Creator:City											
6 Creator: State/province											
7 Creator:Postal code											
8 Creator:Country											
9 Creator:Email	cah.1208@yahoo.ca										
10 Creator:Phone											
11 Creator:Organization											
12 Title of dataset	Title here										
13 Today's date											
14 Abstract	abstract										
15 Repository name and contact information											
16 url for data											
17 Data Contact Person: First name											
18 Data Contact Person:Last name											
19 Data Contact Person:Address											
20 Data Contact Person:City											
21 Data Contact Person:State/province											
22 Data Contact Person:Postal code											
23 Data Contact Person:Country											
24 Data Contact Person: Phone											
25 Data Contact Person:Email											
26 Data Contact Person:Organization											
27 keyword(s)	keyword										
28 Keyword thesaurus used											
29 Geographic coverage:Description											
30 Geographic coverage:West bounding coordinate											
31 Geographic coverage:East bounding coordinate											
32 Geographic coverage:North bounding coordinate											
33 Geographic coverage:South bounding coordinate											
34 Temporal coverage:Description											
35 Temporal coverage:Beginning date											
36 Temporal coverage:Ending date											
37 Project title											
38 Project description											
39 Funding											
40 Intellectual rights											
41 Data table name											
42 Data table description											
43 Identifier											
44 Citation											
45											
46											
47											

CREATE METADATA

It generates and applies metadata based on a pre-defined schema. Some of the metadata values are automatically populated by the system – the remaining must be supplied by you. Some of the metadata fields are required, designated by red asterisk, and some are optional. Please make sure you provide at least all the required metadata values.

Data descriptions **Column descriptions**

CLEAR ALL **RESET**

Creator:First name *	Carly
Creator:Last name *	Strasser
Creator:Address	
Creator:City	
Creator: State/province	
Creator:Postal code	
Creator:Country	
Creator:Email *	cah.1208@yahoo.ca
Creator:Phone	
Creator:Organization	
Title of dataset *	Title here

SAVE **CANCEL**

scope_data_Nov17toDec8 / Notes / red_days / pink_days / scope_data_Dec9to... / phytoplankton&counting / Metadata

Ready 100%

Create a Data Citation

CREATE CITATION

Steps:

1. Retrieve an identifier from the repository for your dataset (Read this important information about your unique identifier [here](#))
2. Check the rest of the information below for accuracy
3. Make any necessary changes or edits to the information
4. Click 'Generate Citation' to create your data citation

Unique identifier

Publication year

Dataset title

Dataset version

Publisher (Data owner)

Citation

CREATE CITATION

Steps:

1. Retrieve an identifier from the repository for your dataset (Read this important information about your unique identifier [here](#))
2. Check the rest of the information below for accuracy
3. Make any necessary changes or edits to the information
4. Click 'Generate Citation' to create your data citation

Unique identifier

Publication year

Dataset title

Dataset version

Publisher (Data owner)

Citation



UC3EZID
long-term identifiers made easy

Upload to a Repository

FILE POST
Post as XLSX

issues descriptions citation

You are now ready to post your curated document to the repository of your choice. You can select the repository from the list, and provide your credentials for the repository.

Repository Name*

Repository Type

I accept [User Agreement](#)



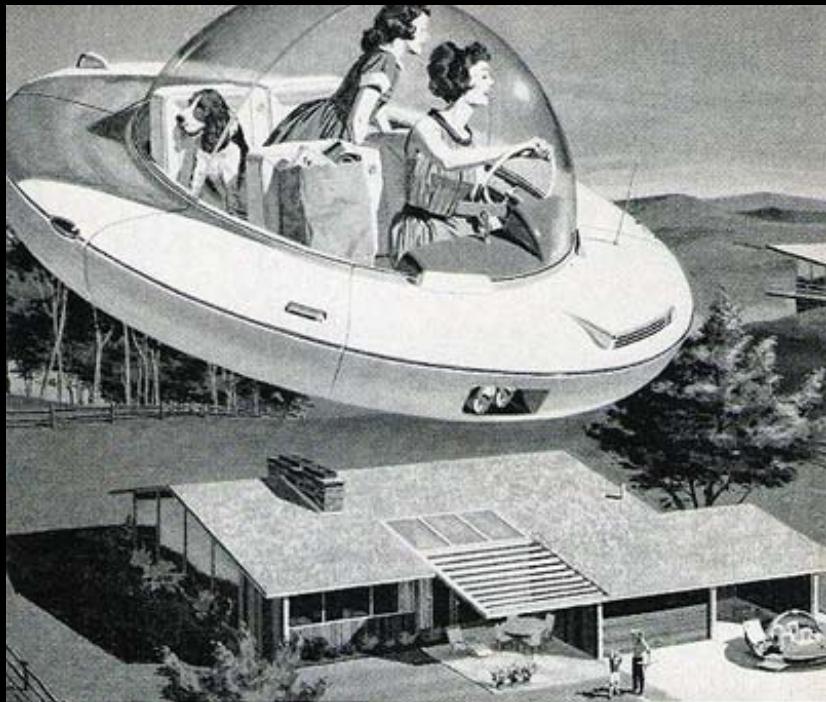
Merritt Data The logo consists of the word "Data" in a large, dark teal sans-serif font, with a stylized globe icon integrated into the letter "a". To the right of "Data" is the word "ONE" in a smaller, dark teal sans-serif font, with a stylized globe icon integrated into the letter "O".

Data Repository

for

Anyone
Anywhere

From animationresources.org



Build
community

Add
repositories

Add metadata
schema



UC3
UC Curation Center



University of California
CDL
California Digital Library

DataONE

dataup.cdlib.org

@DataUpCDL

facebook.com/DataUpCDL

carlystrasser.net

carly.strasser@ucop.edu

[@carlystrasser](https://twitter.com/carlystrasser)