

~~Increasing Openness and Connections
throughout the Scientific Workflow~~

The Importance of APIs in RDM

(and other assorted acronyms)



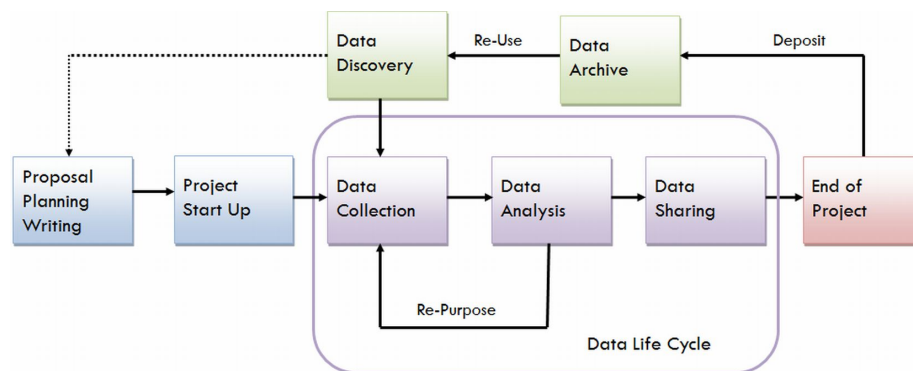
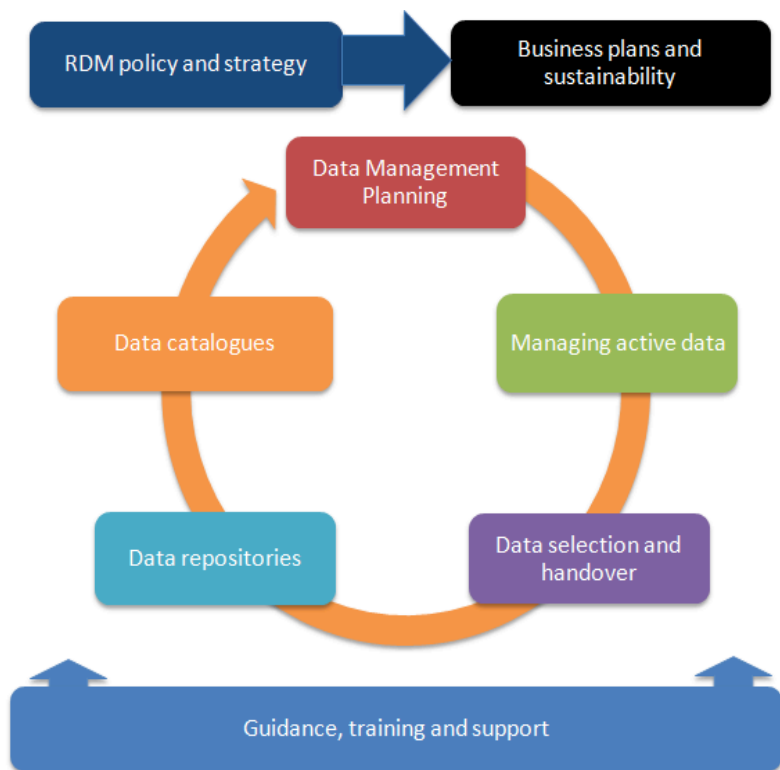
Dan Valen
Product Specialist
dan@figshare.com



Elizabeth Quigley
User Experience Lead
equigley@iq.harvard.edu



Recommended paths for RDM



Van den Eynden, V., Corti, L., Woollard, M. & Bishop, L. (2009). *Managing and Sharing Data: A Best Practice Guide for Researchers*. Retrieved 02/06/2010, from <http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>

Strasser, Carly; Cook, Robert; Michener, William; & Budden, Amber. (2012). *Primer on Data Management: What you always wanted to know*. UC Office of the President: California Digital Library. Retrieved from: <http://escholarship.org/uc/item/7tf5q7n3>

Jones, S., Pryor, G. & Whyte, A. (2013). 'How to Develop Research Data Management Services - a guide for HEIs'. DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: <http://www.dcc.ac.uk/resources/how-guides> - See more at: <http://www.dcc.ac.uk/resources/how-guides/how-develop-rdm-services#sthash.xYof7KTN.dpuf>

101 INNOVATIONS IN SCHOLARLY COMMUNICATION



Jeroen Bosman @jeroenbosman
Utrecht University Library

THE CHANGING RESEARCH WORKFLOW



Bianca Kramer @MsPhelps
Utrecht University Library

Science is in transition. This poster gives an impression of the exploratory phase of a project aiming to chart innovation in scholarly information and communication flows from evolutionary and network perspectives.

We intend to address the questions of what drives innovation and how these innovations change research workflows and may contribute to more open, efficient and good science.

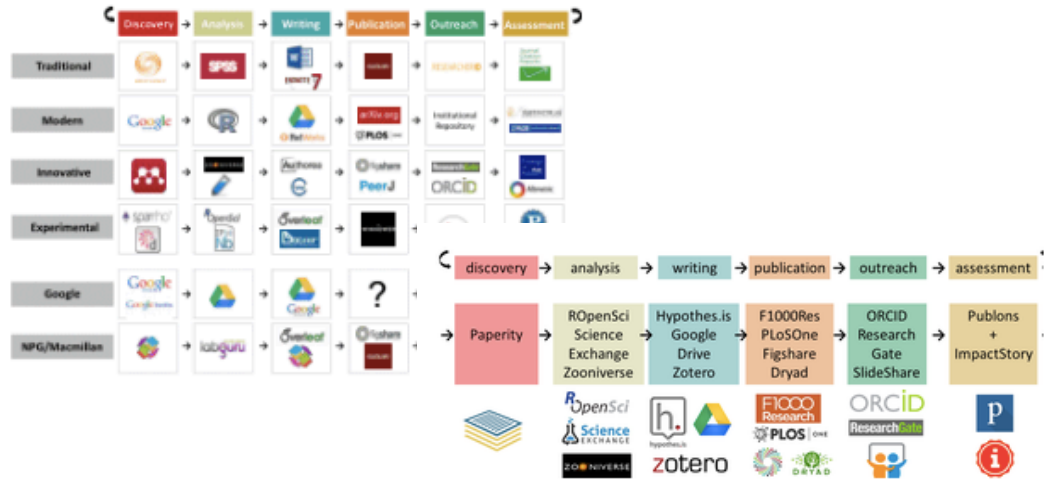
101 Innovative tools and sites in 6 research workflow phases (< 2000 - 2015)



Most important developments in 6 research workflow phases

	Discovery	Analysis	Writing	Publication	Outreach	Assessment
Trends	social discovery tools	dashboards & crowdsourced science	collaborative online writing	Open Access & data publication	scholarly social media	altmetrics (altmetrics)
Expectations	growing importance of data discovery	more online analysis tools	more integration with publication & assessment tools	more use of "publish this, judge later"	use of abstracts for monitoring outreach	more open and post-publication peer review
Uncertainties	support for tailored search and technology	willingness to share in analysis phase	acceptance of collaborative online writing	effect of journal/publisher status	requirements of funders & institutions	who pays for costly evaluation assessment?
Opportunities	discoveries based on aggregated data sets	open literature	researcher tags while writing/editing	reader-side paper formatting	using repositories for institutional visibility	using authors' publication and affiliation IDs
Challenges	real semantic search (taxonomy & metadata)	reproducibility	safety/library of online writing	globalization of online journals/academic institutions	making outreach a primary discussion	quality of measuring tools
Most important long-term development	multidisciplinary + citation-enhanced databases	collaboration + data-driven	online writing platforms	Open Access	more & better connected researcher profiles	importance of social networks + open publication contributions
Potentially most disruptive development	emerging alternative research + contextualized recommendations	open science	collaborative writing + integration with publishing	circumventing traditional publishers	public access to research findings, also for agenda setting	moving away from single quantitative indicators

Typical workflow examples



Open Science

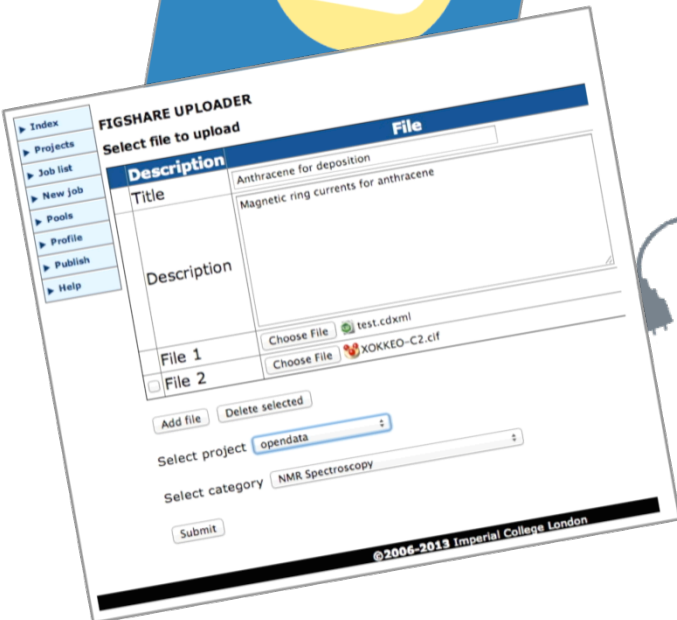
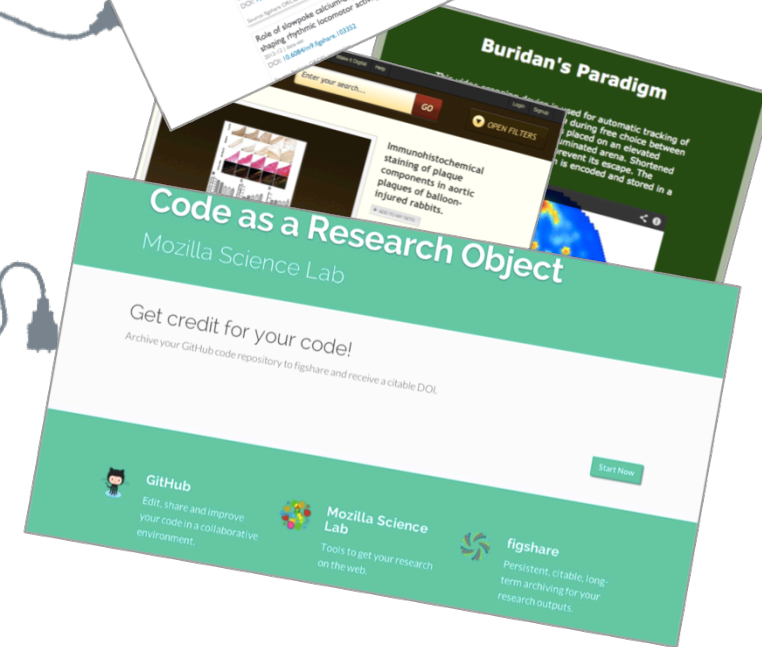
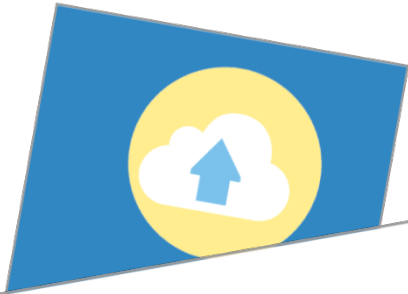
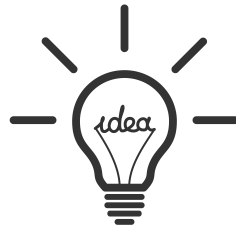
See more: <http://innoscholcomm.silk.co/page/Workflows>
<https://101innovations.wordpress.com/>

Kramer, Bianca; Bosman, Jeroen (2015): 101 Innovations in Scholarly Communication - the Changing Research Workflow. [figshare](https://figshare.com/).
<http://dx.doi.org/10.6084/m9.figshare.1286826>
Retrieved 16:29, Oct 13, 2015 (GMT)

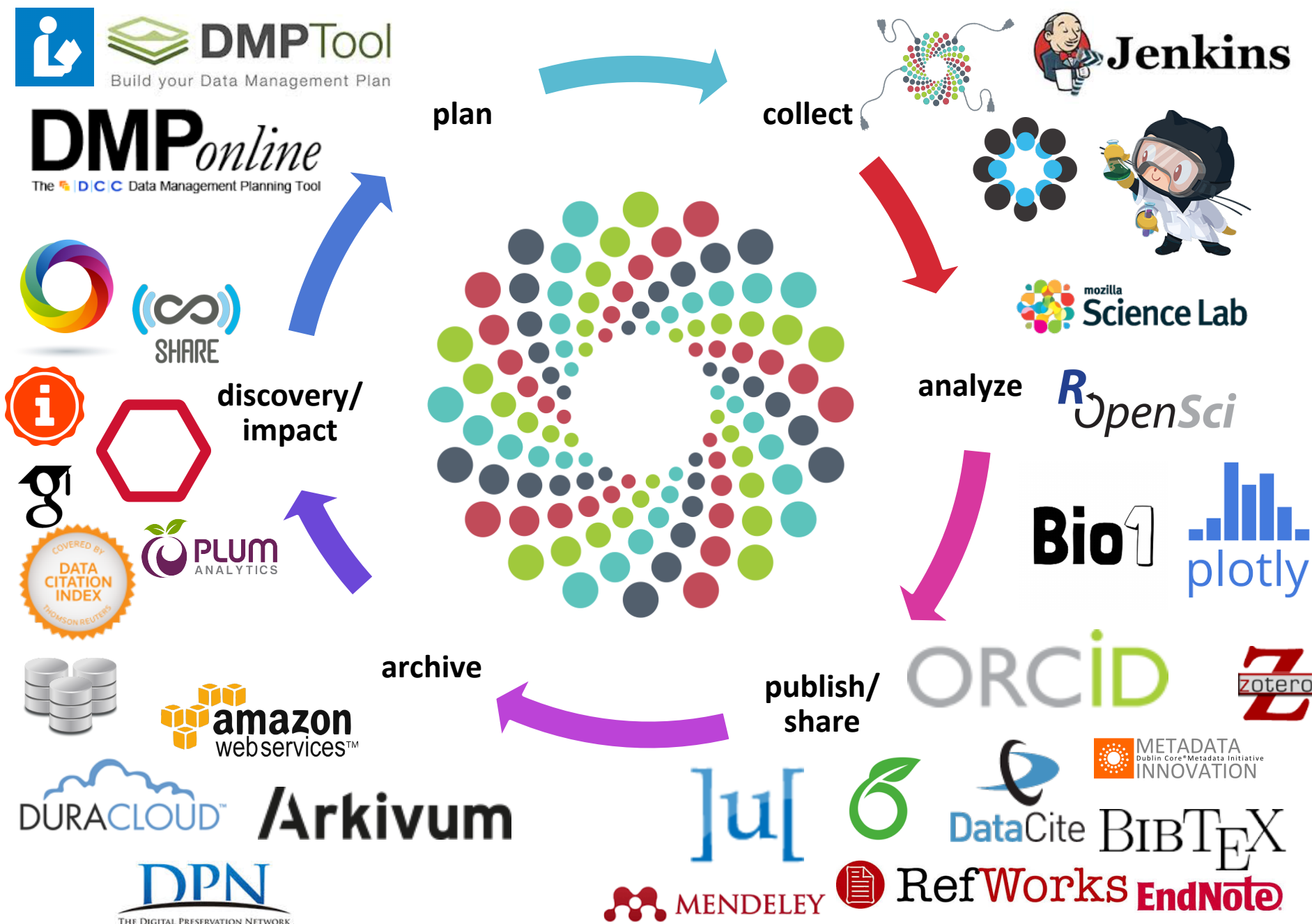
figshare's API

The figshare API allows users to push data to figshare, or pull data off of the platform

This allows you to build applications on top of academic research



figshare integrations in action



An R open science example...



id	group	genotype	treatment	machine	other	date	timeofday	median_spei	distance_tr	turning_angl	meander	activitytime	act_bouts_5	pause_durat	numb_pause	activitytime
1	100_CSJC	CS_JC_2nder_CS_JC	none	A	2ndross	9/20/13	#####	17.6122386	510.940185	7.08158437	0.39205073	143.5	3.14999998	1.5	19	138.9
2	101_CSJC	CS_JC_2nder_CS_JC	none	B	2ndross	9/20/13	#####	13.3093301	231.145092	11.8844266	0.7789247	155.3	3.79999995	0.9000001	31	135.7
3	102_CSJC	CS_JC_2nder_CS_JC	none	C	2ndross	9/20/13	#####	15.9188682	368.327127	9.29123379	0.55608839	159.3	3.79999995	0.9000001	13	159.2
4	103_CSJC	CS_JC_2nder_CS_JC	none	D	2ndross	9/20/13	#####	16.5038787	717.444474	15.8494973	0.45214688	273.600001	11.3000002	1.04999995	14	271
5	104_CSTP	CS_TP_2nder_CS_TP	none	A	2ndross	9/20/13	#####	15.5501774	777.795393	6.55529888	0.40337478	291.6	6.79999995	0.50000012	8	293.8
6	105_CSTP	CS_TP_2nder_CS_TP	none	B	2ndross	9/20/13	#####	16.1131212	455.115117	7.40149444	0.43970117	150.5	2.89999998	3	17	146.2
7	106_CSTP	CS_TP_2nder_CS_TP	none	C	2ndross	9/20/13	#####	17.7771567	794.539736	6.59162962	0.36481201	272.3	16.35	1.20000005	9	270.6
8	107_CSTP	CS_TP_2nder_CS_TP	none	D	2ndross	9/20/13	#####	16.1620151	812.967026	7.38395334	0.44849885	294	39.3999999	1.04999995	4	294
9	108_C8V5	CS_BS_2nder_CS_BS	none	A	2ndross	9/20/13	#####	14.8886688	518.828231	6.35658234	0.38641675	229.199999	7.54999995	6.00000014	25	230.8
10	109_C8V5	CS_BS_2nder_CS_BS	none	B	2ndross	9/20/13	#####	14.2972058	400.366699	5.71643304	0.38361199	175.7	5.95000005	1.64999998	16	171.4
11	110_C8V5	CS_BS_2nder_CS_BS	none	C	2ndross	9/20/13	#####	16.2253154	657.396599	6.62661286	0.38946413	226.6	8.20000005	0.75	18	258.499999
12	111_C8V5	CS_BS_2nder_CS_BS	none	D	2ndross	9/20/13	#####	14.0429198	795.970077	6.88686922	0.45239103	281.5	26.4499999	1.89999984	9	280.7
13	112_CSIC	CS_HS_2nder_CS_HS	none	A	2ndross	9/20/13	#####	15.0008946	638.145575	5.94732506	0.37019237	256.000001	6.4000001	2.09999999	19	249
14	113_CSIC	CS_HS_2nder_CS_HS	none	B	2ndross	9/20/13	#####	17.5043204	703.351364	5.35418829	0.28268952	236.699999	9.20000005	2.60000002	18	234.5
15	114_CSIC	CS_HS_2nder_CS_HS	none	C	2ndross	9/20/13	#####	10.8642468	174.942796	9.3507279	0.78705473	160.2	17.9000001	0.90000001	11	165.2
16	115_CSIC	CS_HS_2nder_CS_HS	none	D	2ndross	9/20/13	#####	15.0400845	333.16162	5.9165218	0.36168682	142.5	5	4.4000001	21	137.4
17	116_CSTZ	CS_TZ_2nder_CS_TZ	none	A	2ndross	9/20/13	#####	16.4458536	559.448644	6.64451914	0.37078207	198	11.6999999	0.74999988	10	197
18	117_CSTZ	CS_TZ_2nder_CS_TZ	none	B	2ndross	9/20/13	#####	12.8965403	60.2270539	11.6131909	0.8283955	35.4000001	1.29999995	4	11	35.5000002
19	118_CSTZ	CS_TZ_2nder_CS_TZ	none	C	2ndross	9/20/13	#####	21.6350837	832.369509	5.74553669	0.24687918	268.4	3.89999998	0.70000005	29	270
20	119_CSTZ	CS_TZ_2nder_CS_TZ	none	D	2ndross	9/20/13	#####	14.7033046	523.296729	7.74348163	0.49171641	231.2	6.20000005	1.15000001	22	226.9
21	120_CSIC	CS_JC_2nder_CS_JC	none	A	2ndross	9/20/13	#####	16.4713165	641.471823	7.42486007	0.42834449	246.7	6.20000005	0.78999999	19	241.899999
22	121_CSIC	CS_JC_2nder_CS_JC	none	B	2ndross	9/20/13	#####	18.0458379	685.768452	7.90810691	0.40873037	257.7	3.64999998	6.00000014	29	247.700001
23	122_CSIC	CS_JC_2nder_CS_JC	none	D	2ndross	9/20/13	#####	15.820831	500.481894	7.42417077	0.45075403	220.9	2.60000002	1.29999995	40	208.1
24	123_C8TP	CS_TP_2nder_CS_TP	none	A	2ndross	9/20/13	#####	15.803955	491.485263	8.4448159	0.52401448	204.9	5.24999998	1.44999999	28	193.3
25	124_C8TP	CS_TP_2nder_CS_TP	none	B	2ndross	9/20/13	#####	16.4713165	641.471823	7.42486007	0.42834449	246.7	6.20000005	0.78999999	19	241.899999
26	125_C8TP	CS_TP_2nder_CS_TP	none	C	2ndross	9/20/13	#####	15.9714384	772.39629	6.96620733	0.42136408	273.3	21.6	1.70000005	7	270.8
27	126_C8V5	CS_BS_2nder_CS_BS	none	A	2ndross	9/20/13	#####	12.6565204	450.47614	7.6914075	0.5550748	255.599999	3.20000005	1.30000007	24	251
28	127_C8V5	CS_BS_2nder_CS_BS	none	B	2ndross	9/20/13	#####	17.3626962	710.185765	6.25425327	0.33375696	241.3	16.7	2.75	10	235.5
29	128_CSTZ	CS_TZ_2nder_CS_TZ	none	A	2ndross	9/20/13	#####	16.5003325	283.521089	7.3730861	0.42159624	101.7	45.5	46.3	3	104.6
30	129_CSTZ	CS_TZ_2nder_CS_TZ	none	B	2ndross	9/20/13	#####	15.14244	550.273449	8.13010235	0.49868875	214.6	40.5	7.29999995	5	215.9
31	130_CSTZ	CS_TZ_2nder_CS_TZ	none	C	2ndross	9/20/13	#####	16.4412787	718.485579	7.29914751	0.40790347	274.6	7	0.79999995	3	276.2
32	131_CSTZ	CS_TZ_2nder_CS_TZ	none	D	2ndross	9/20/13	#####	16.4092038	489.699153	6.71172445	0.38769446	180.9	19.9499999	4.94999993	8	178.9
33	132_CSIC	CS_JC_2nder_CS_JC	none	A	2ndross	9/26/13	#####	9.4284638	2.6851132	26.3236799	2.30794334	1.59999999	1.59999999	148.85	2	2.29999999
34	133_CSIC	CS_JC_2nder_CS_JC	none	B	2ndross	9/26/13	#####	18.7399411	331.782623	7.33729051	0.37669148	121.5	6.4000001	3.29999995	17	115.8
35	134_CSIC	CS_JC_2nder_CS_JC	none	C	2ndross	9/26/13	#####	17.2576718	135.484593	9.96032396	0.36530732	57.09999999	0.90000001	25	50.80000019	
36	135_CSIC	CS_JC_2nder_CS_JC	none	D	2ndross	9/26/13	#####	17.1818269	353.970673	8.388457	0.45681235	146.2	3.79999995	2.78999995	32	137.999999
37	136_CSTP	CS_TP_2nder_CS_TP	none	A	2ndross	9/26/13	#####	13.4731757	424.55242	8.52774683	0.60971847	233.3	8.59999999	0.79999999	21	236.1
38	137_CSTP	CS_TP_2nder_CS_TP	none	B	2ndross	9/26/13	#####	16.9825144	622.835999	9.59538348	0.48577073	222.4	11.8500001	1.89999998	14	217.999999
39	138_CSTP	CS_TP_2nder_CS_TP	none	C	2ndross	9/26/13	#####	14.9235309	392.965722	7.22771816	0.47280038	183.3	4.45000005	2.09999999	27	172.1

This repository Search Explore Git Blog Help thevandalen + + +

jcolomb / CeTrAn

Watch 3 Star 4 Fork 3

CeTrAn is a R script to analyse and publish (on figshare) trajectory data, principally data obtained from Buritrack a tracker designed for fruit flies running in the Buridan assay, see also buridan.sourceforge.net <http://buridan.sourceforge.net>

41 commits 1 branch 1 release 1 contributor

branch: master CeTrAn / +

changes to accommodate new plotting and testing with the new data.

- `.Rproj/user/F4DCDC` readme with license 2 years ago
- `CeTrAn` changes to accommodate new plotting and testing with the new data. 19 days ago
- `docu` add r3.0.0 test in docu 2 years ago
- `example_data` add data 2 years ago
- `gignore` Initial commit 2 years ago
- `CeTrAn.Rproj` add studio project, start decomposing test_figshare 2 years ago
- `README.md` Update README.md 2 years ago
- `first commit` first commit 2 years ago
- `figshare upload code.R` pqa3d+ test figshare 2 years ago



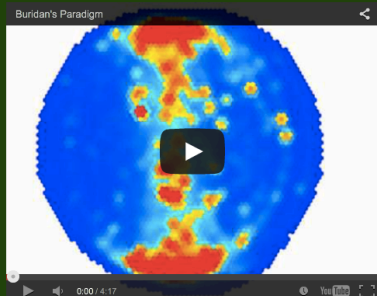
share

analyze

publish

Buridan's Paradigm

This video scanning device is used for automatic tracking of a walking fruit fly fly *Drosophila* during free choice between two visual landmarks. The fly is placed on an elevated platform in the center of the illuminated arena. Shortened wings and a water-filled moat prevent its escape. The position on the circular platform is encoded and stored in a plain text file for later analysis.



Buridan's paradigm was originally devised by Götz K.G. (1980: "Visual guidance in *Drosophila*." Basic Life Sci. 16: 391-407) in the MPI for Biological Cybernetics.

Figshare My data search figshare (titles, tags, authors, etc.) Browse Upload Dan Valen

Sub-strains of *Drosophila* Canton-S differ markedly in their locomotor behavior

jcolomb-Buridan_5CS-4a295e

4CS_data
output_table.csv
PCAresult_2datapointout.csv

289 views 1 shares

Published on 02 May 2014 - 13:02 (GMT)
Filesize is 6.40 MB

Categories
Behavioral Neuroscience

Authors
julien colomb
Björn Brembs

Tags
buridan • open data • ratat
wild-type • drosophila
locomotion • behavior

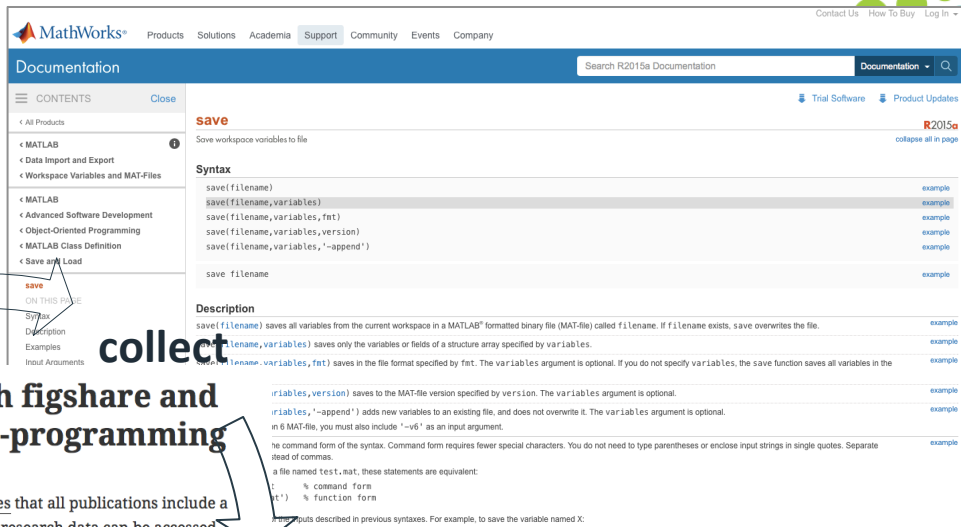
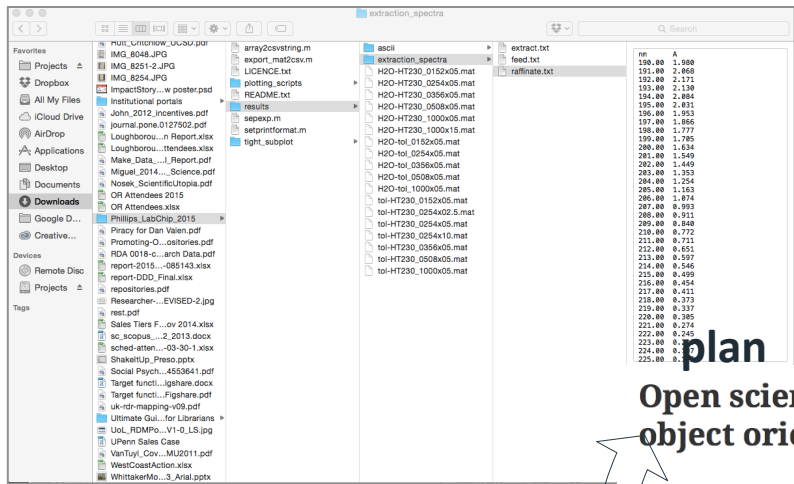
License (what's this?)
CC0

Export
Export to RefWorks
Export to SciSpace
Export to Ref. manager
Export to Mendeley

Description
Raw data for the F1000 Research paper "Sub-strains of *Drosophila* Canton-S differ markedly in their locomotor behavior"



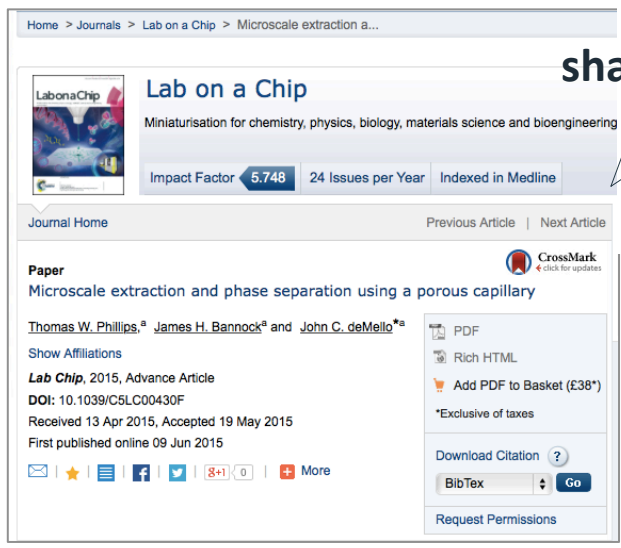
Object oriented programming with figshare...



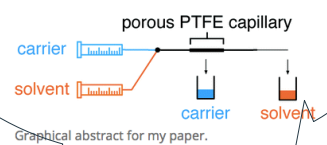
plan collect

Open science with figshare and object orientated-programming

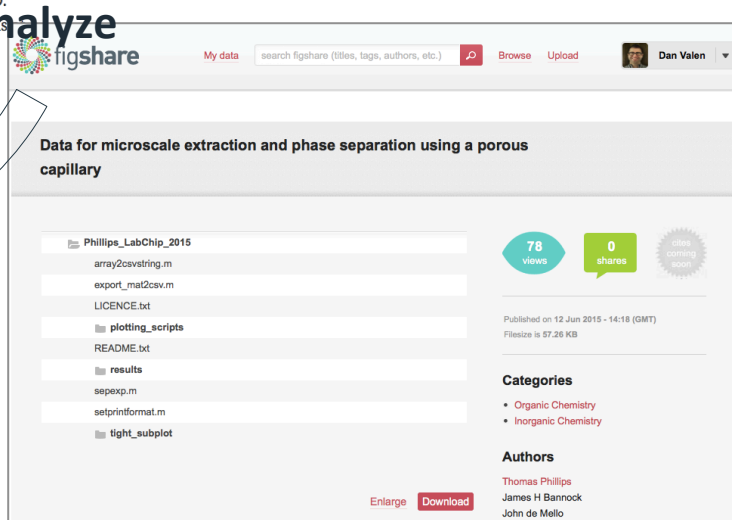
From 1st May 2015, the EPSRC requires that all publications include a statement saying how the underlying research data can be accessed. Technically, you can simply include an email address to contact for the data, but I think that's hardly in the spirit of open science. In this post, I want to describe how I used object-orientated programming (OOP) and figshare to meet this requirement for my latest paper in *Lab on a Chip*. You can download the data and MATLAB code to reproduce the graphs.



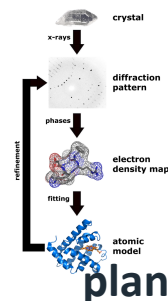
share



publish



A hardware to figshare open science example



plan

collect



share

analyze

publish

FIGSHARE UPLOADER
Select file to upload

Description	File
Title	Anthracene for deposition
Description	Magnetic ring currents for anthracene

File 1 Choose File test.cdxml
File 2 Choose File XOKKEO-C2.cif

Add file Delete selected

Select project opendata
Select category NMR Spectroscopy

Submit

©2006-2013 Imperial College London

ORCID
Connecting Research and Researchers

FOR RESEARCHERS FOR ORGANIZATIONS ABOUT HELP SIGN IN

1,476,618 ORCID iDs and counting. See more

Henry Rzepa
ORCID ID
orcid.org/0000-0002-8635-8390

Biography
Professor of Computational chemistry, Imperial College London

Education (1)
Imperial College London: London, London, United Kingdom
Source: Henry Rzepa Created: 2014-05-30

Employment (1)
Imperial College London: London, London, United Kingdom
Source: Henry Rzepa Created: 2014-05-30

Works (1320)

Conference access statistics and the browsing habits of ECHET96 readers
ECHET96
1996-07-04 | conference-paper
DOI: 10042a3v2p
URL: http://www.ch.ica.ac.uk/ectoc/echet96/papers/122/index.htm
Source: Henry Rzepa Preferred source

Anthracene for deposition
data-set
DOI: 10.6084/m9.figshare.739138
Source: figshare ORCID integration Preferred source

Asymmetric epoxidation: a combined laboratory and molecular modelling experiment
2014-04 | data-set
DOI: 10.6084/m9.figshare.988346
Source: figshare ORCID integration Preferred source

Chemical datuments as scientific enablers
data-set
DOI: 10.6084/m9.figshare.759405

figshare
My data search figshare (files, tags, authors, etc.) Browse Upload Dan Valen

Gaussian 4

input.gjf
checkpoint
em.xml
logfile.log
structure.pdb
wavefnct

Share this:
Cite this:

structure.pdb - Download

.GJF .XML .LOG .WFN

Description
Archive of work produced using the resources of the Imperial College High Performance Computing Service.

Links
http://www.ch.ica.ac.uk/ectoc/echet96/papers/122/index.htm
http://www.ncbi.nlm.nih.gov/pubmed/24761107
http://www.chemeduk.com/search.aspx?cid=694&cid2=694&cid3=694&cid4=694&cid5=694&cid6=694&cid7=694&cid8=694&cid9=694&cid10=694

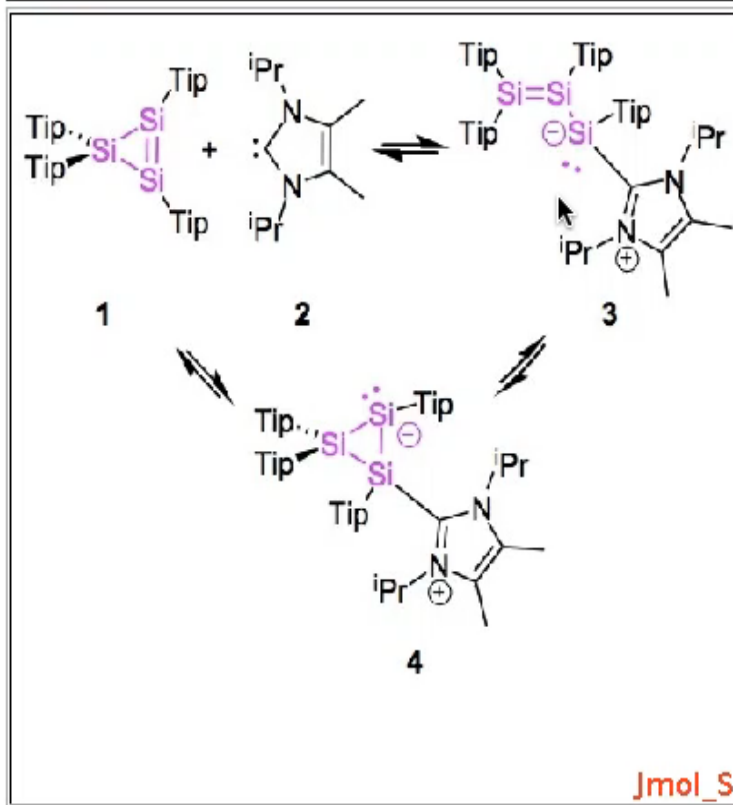
License
CC-BY

Export
Export to RefWorks
Export to EndNote

Interactive data within the browser

Interactivity box. Calculated properties for Si₃ clusters.

ΔG_{298}^a	3D model and v(Si-Si) animation	NMR and UV-Vis Spectra ^e
-3209.006417^b -3208.558059 -3209.746651^d	1 <input type="text"/>	1 <input type="text"/>
-540.412657^b -540.327293^c -540.584999^d	2 <input type="text"/>	
-3749.440365^b -13.3 ^B -3748.901271^c -10.0 ^B -3750.297509^d +21.4 ^B	3 <input type="text"/>	3 <input type="text"/>
-3749.434450^b -9.6 ^B -3750.261410^d +44.1 ^B	4 <input type="text"/>	
-3208.984595^b +13.7 ^B	5 <input type="text"/>	



^aCalculated using Gaussian 09, revision C.01, at the 6-311G(d,p)/6-311G(2df)[Si only] basis set. Free Energies in Hartree (relative to reactant in kcal mol⁻¹). Energies are hyperlinked to a OAI-PMH compliant Digital repository identifier, resolved as

633

views

0

shares

Published on 15 Jul 2013 - 06:44

Filesize is 12.16 KB

This data is part of a publication
Equilibrium between a cyclohexadiene
an isolable base adduct of
silylene



Categories

- Inorganic Chemistry

Authors

David Scheschkewitz
Michael J. Cowley
Volker Huch
Henry S. Rzepa

Tags

- silicon

License (what's this?)

CC-BY

Export

- Export to RefWorks
- Export to BibTeX
- Export to Ref manager

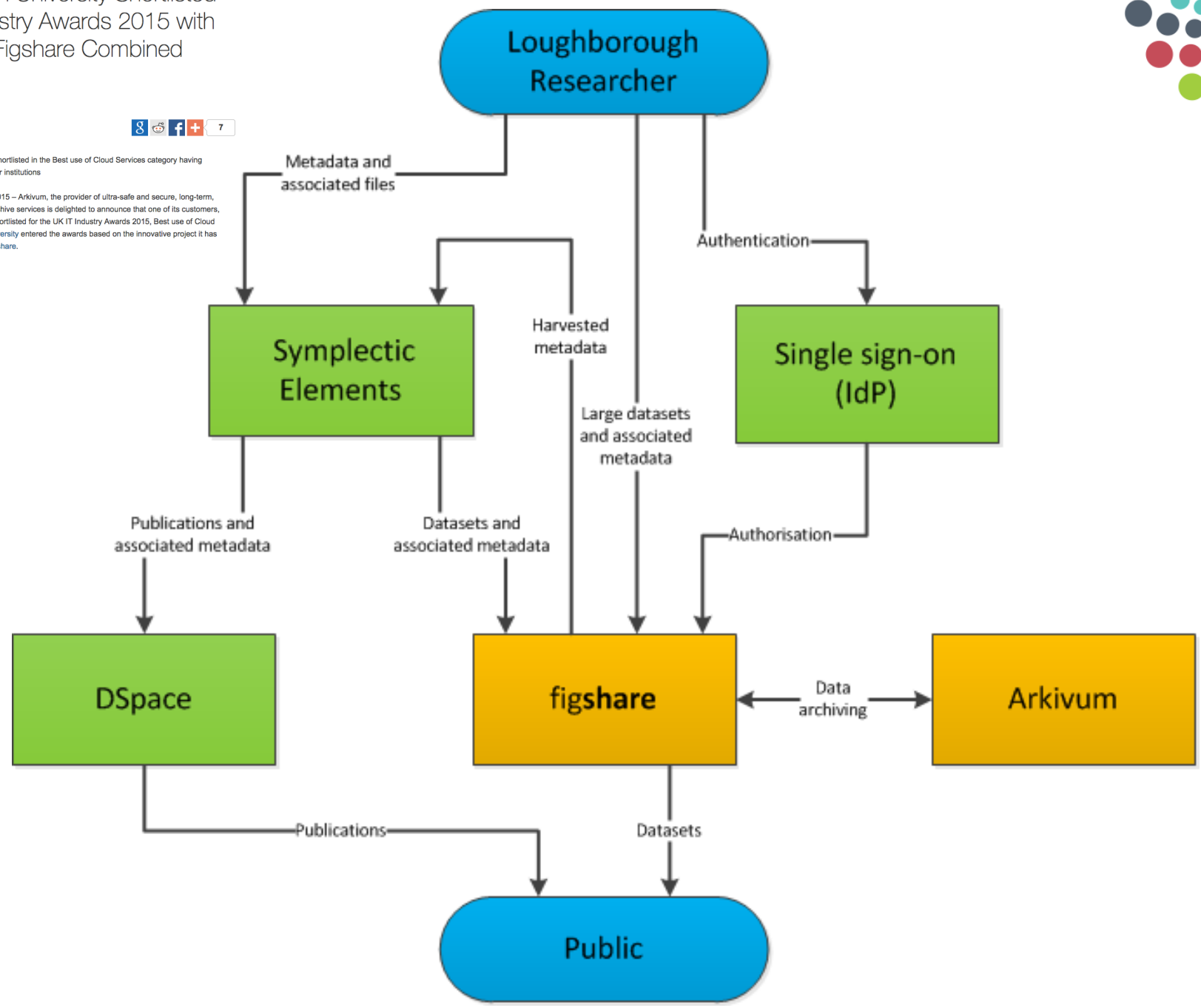
Loughborough University Shortlisted for UK IT Industry Awards 2015 with Arkivum and Figshare Combined Solution



Tuesday 25 August 2015 7

Loughborough University has been shortlisted in the Best use of Cloud Services category having deployed Arkivum100 and Figshare for institutions

Chippenham, Wiltshire, 25th August 2015 – Arkivum, the provider of ultra-safe and secure, long-term, large-scale digital data storage and archive services is delighted to announce that one of its customers, Loughborough University, has been shortlisted for the UK IT Industry Awards 2015, Best use of Cloud Services category. Loughborough University entered the awards based on the innovative project it has undertaken with both Arkivum and Figshare.



Dataverse APIs

- Data Deposit (SWORD)
- Search
- Data Access
 - Basic file access
 - Multiple files download
 - All format download for Tabular files
 - Data Variable metadata access
- Native
 - Exposes most of the UI functionality in the application via a REST-based API

Learn more about the APIs with the [API Guides](#)

Want to test out an API? Use: apitest.dataverse.org

Current Integrations with **Dataverse** APIs



In Progress & Future Integrations with **Dataverse** APIs



@archivematica.®



iRODS



DataTags

The
Dataverse
Project 

The Dataverse Project logo icon, featuring three orange circles of varying sizes connected by thin lines.

Dataverse API Integrations by Users

- [R package](#) by Thomas Leeper
- [Visualization of Datasets on Historical Maps](#) by CLIO INFRA
- [Anti-Slavery Petitions of Massachusetts](#) by Garth Griffin

Interested in building off the APIs or code base?
Check out the Dataverse source code on [Github](#).

Support simulated and derived data

data



UNIVERSITY OF CALIFORNIA



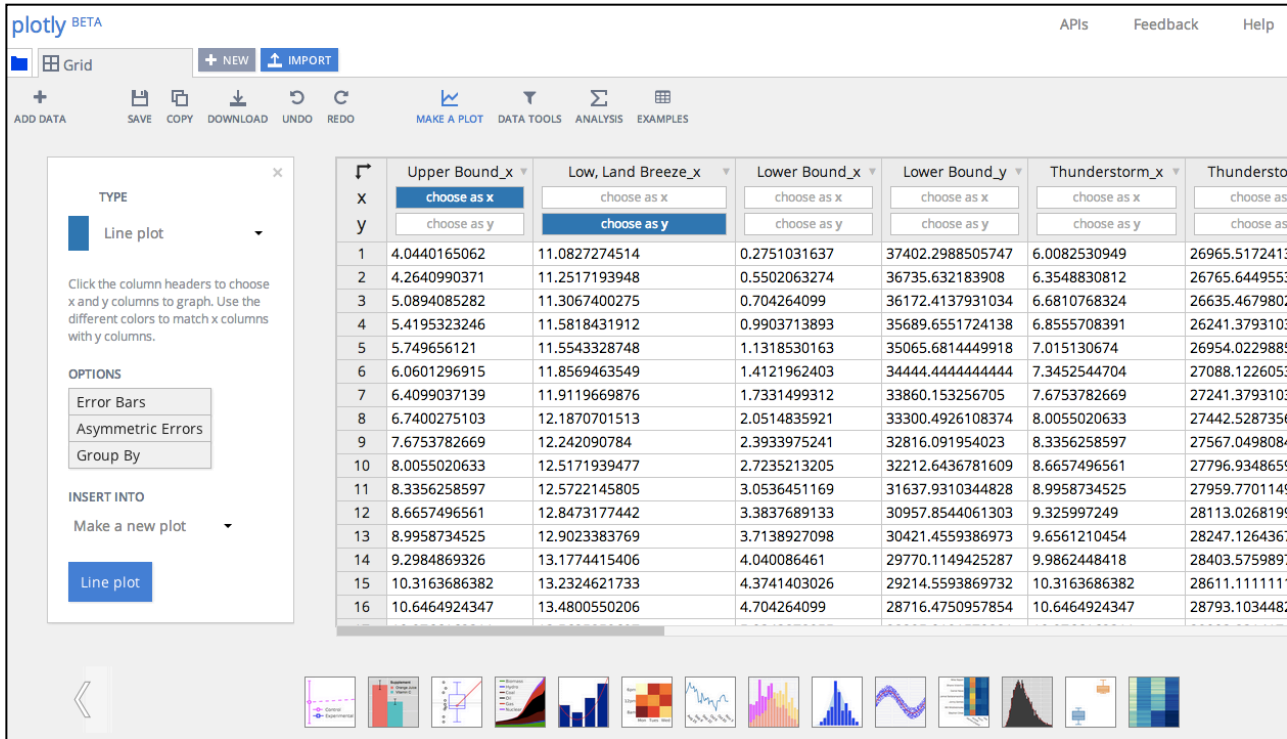
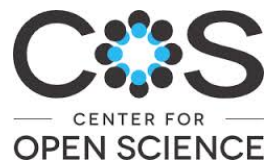
DSpace



F1000 FACULTY of 1000



3TU.Datacentrum



Any questions?

Get in touch!



@figshare



dan@figshare.com



@dataverseorg



equigley@iq.harvard.edu

