

# Securely store your supplementary material on Figshare

While making it discoverable and citable

Presented by: Petr Čermák  
@petrscience



# MOTIVATION: WHY to publish data?

90% of the world's total  
data was created in last  
2 years [1]

2.5 quintillion bytes of  
data every day [1]

## Why to publish data

Transparent research

Reuse of already measured data (TOF datasets...)

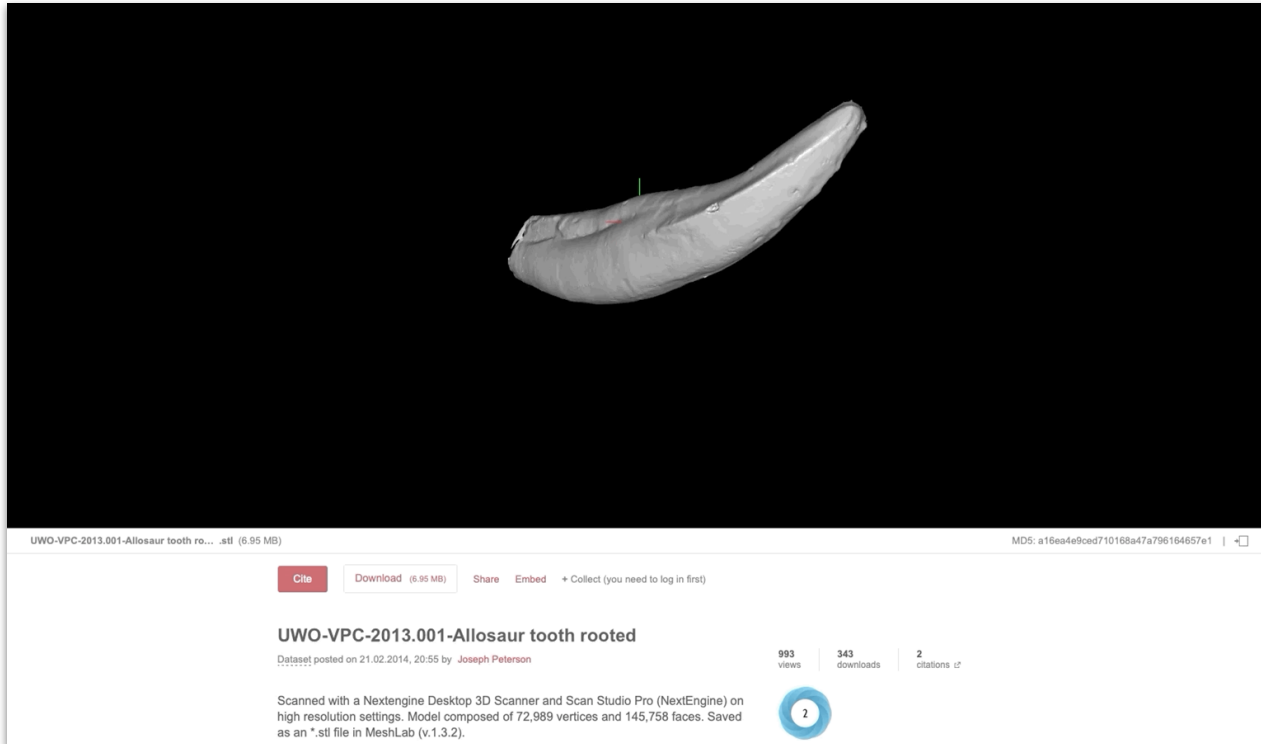
AI search and analysis of datasets (in future)

## How to publish data

Automatically by Institution (ILL, ISIS)

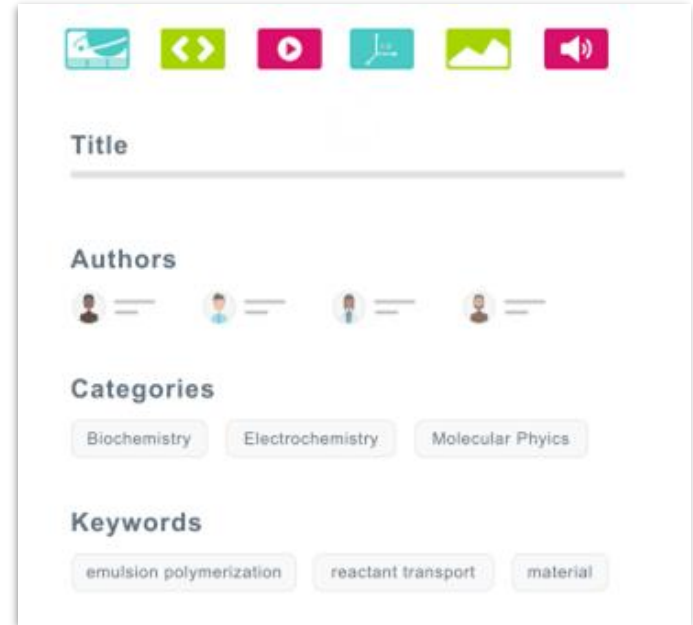
Manually: for example FigShare

Figshare is a cloud-based platform for securely storing your research data while making it discoverable and citable.



# Why Figshare?

- ✓ Store your outputs in any file format
- ✓ Securely collaborate with researchers from other institutions
- ✓ Get a citable DOI for public outputs
- ✓ Demonstrate impact with Altmetrics, usage metrics, & citation counts
- ✓ Help others discover your outputs on Google Scholar
- ✓ 20 GB private space  
Unlimited public space (5GB per file)



The image shows a screenshot of the Figshare metadata form. At the top, there is a row of six icons representing different file formats: a blue document icon, a green code icon, a red play button icon, a teal line graph icon, a green landscape icon, and a red speaker icon. Below the icons, the form has several sections:

- Title:** A text input field with a horizontal line below it.
- Authors:** A section with four user avatars, each followed by an equals sign (=), indicating a list of authors.
- Categories:** A section with three rounded rectangular buttons labeled "Biochemistry", "Electrochemistry", and "Molecular Physics".
- Keywords:** A section with three rounded rectangular buttons labeled "emulsion polymerization", "reactant transport", and "material".

# As open as possible, as closed as necessary

Your research outputs should be as open as possible, but control over access is important.

File(s) not publicly available

Reason: Data is hosted at PI's institution (Surrey). Link provided in description.

Share Cite + Collect (you need to log in first)

### Object-Based audio drama scenes, stored as BWF files with ADM headers

01.06.2016, 12:01 by James Stephen Woodcock, Chris Pike, Phil Coleman, Frank Melchior, Andreas Franck, Adrian Hilton

This dataset contains 3 object-based audio drama scenes, commissioned as part of the S3A project. The metadata required to render the scenes is stored in the BWF header as an XML chunk using the Audio Definition Model. A suite of C++ libraries for creating and editing ADM data and reading from and writing to BWF files is available from BBC R&D <http://www.bbc.co.uk/rd/publications/audio-definition-model-software>.


To access the data, please follow this link [http://cvssp.org/data/s3a/public/radiodrama\\_register.php](http://cvssp.org/data/s3a/public/radiodrama_register.php).

For further information, please email James Woodcock ([j.s.woodcock@salford.ac.uk](mailto:j.s.woodcock@salford.ac.uk)).



FUNDING  
EP/L000539/1 (EPSRC)

491 views | 0 downloads | 0 citations

4

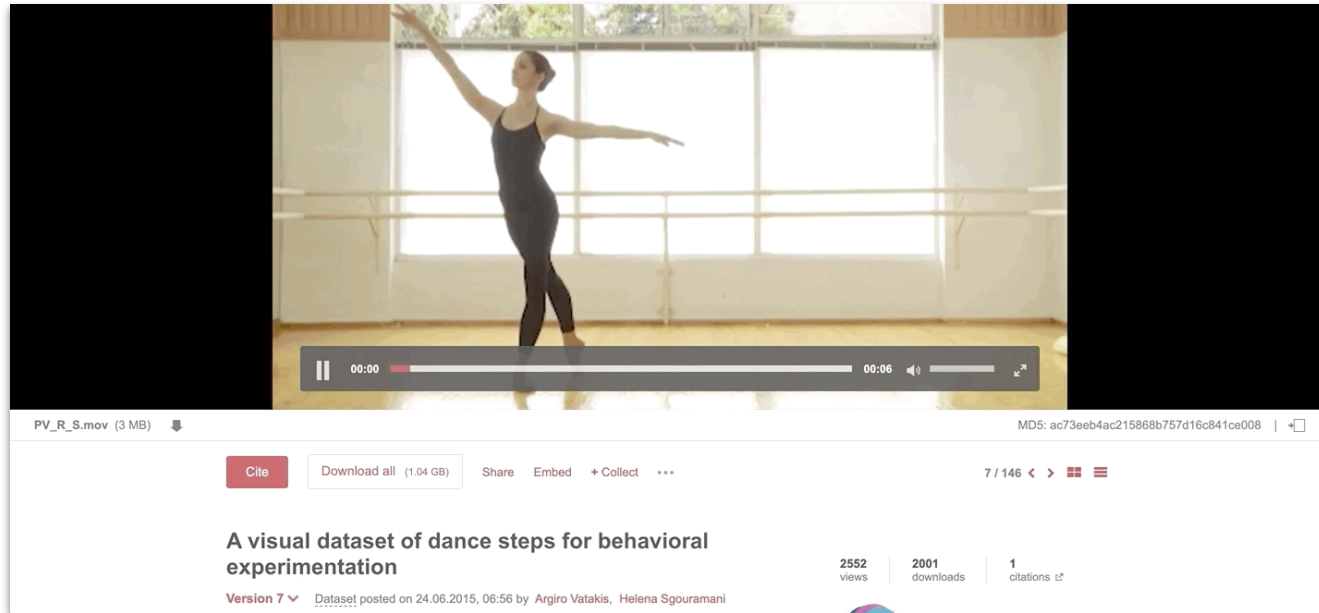


CATEGORIES  
• Acoustics and Noise Control (excl. Architectural Acoustics)

-  Apply embargo
-  Make file(s) confidential
-  Generate private link
-  Reserve Digital Object Identifier

# What can you upload to Figshare?

We accept any file format and aim to preview files in the browser.



PV\_R\_S.mov (3 MB) MD5: ac73eeb4ac215868b757d16c841ce008

Cite Download all (1.04 GB) Share Embed + Collect ... 7 / 146 < > ☰ ☷

**A visual dataset of dance steps for behavioral experimentation**

Version 7 Dataset posted on 24.06.2015, 06:56 by Argiro Vatakis, Helena Sgouramani

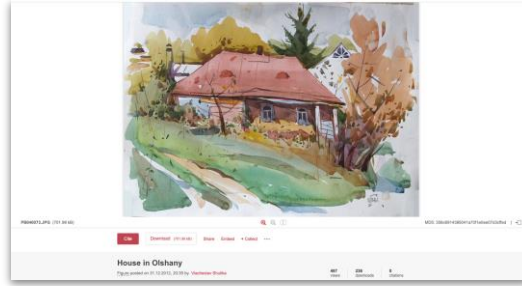
2552 views | 2001 downloads | 1 citations

Vatakis, Argiro; Sgouramani, Helena (2015): A visual dataset of dance steps for behavioral experimentation. figshare. Dataset.  
<https://doi.org/10.6084/m9.figshare.1453169.v7>

# Audio



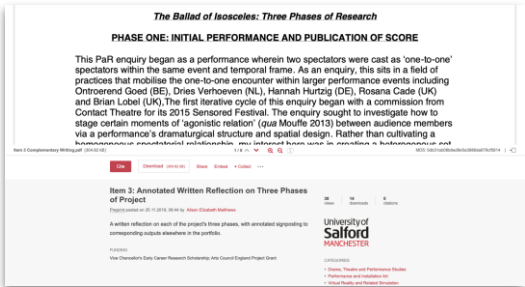
# Images



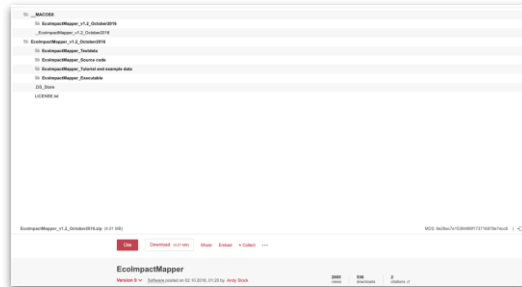
# 3D files



# Papers/docs



# Code



# Posters/Talks



# Projects

→ Upload items to the project

→ Add collaborators or viewers. Invite institutional colleagues, figshare.com users, or invite people with no Figshare account to join your project

→ Keep it private or make it public

The screenshot displays a list of three items within a project feed, each with a timestamp, a user profile picture, a title, and a set of action icons.

Timestamp	User Profile	Title	Category	Icons
24.4.2018 11:58		FOSTER Open Science Bootcamp Notes	JOURNAL CONTRIBUTION	Comment, Settings
24.4.2018 11:54		Highlighted Open Science Training Handbook	JOURNAL CONTRIBUTION	Green dot, Comment, Settings
24.4.2018 11:39		Welcome, everyone! I'm going to start uploading some materials into this project. If you have anything relevant to upload, please feel free! Any questions, let me know.		Edit, Comment, Settings



# Collections

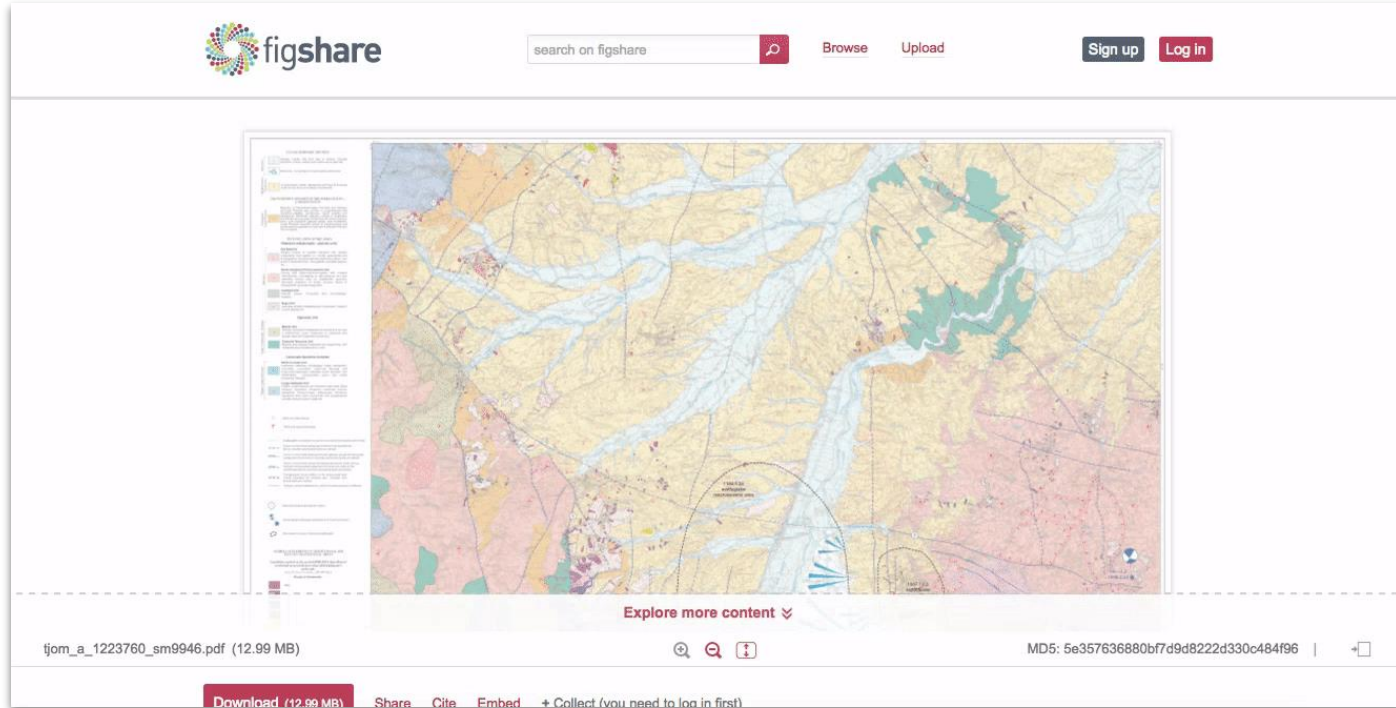
- ✓ Pull from your data or public data
- ✓ Use to showcase your data
- ✓ Get a citable DOI for the collection
- ✓ Supports versions
- ✓ Tracks Altmetrics, usage metrics, and citations

The screenshot shows a Figshare collection page. At the top, the title is 'Online appendix: the happiness of software developers' with 3236 views and 9 citations. Below the title, there is a '+ Follow' button, 'Version 9', and the publication date 'Published on 04 Jul 2017 - 11:56 by Daniel Graziotin'. The description reads: 'Datasets, supplemental material, replication kit for a series of studies on the happiness of software developers.' There is a 'CITE THIS COLLECTION' section with a 'DataCite' dropdown menu. The citation text is: 'Graziotin, Daniel; Fagerholm, Fabian; Wang, Xiaofeng; Abrahamsson, Pekka (2017): Online appendix: the happiness of software developers. figshare. Collection. <https://doi.org/10.6084/m9.figshare.c.3355707.v9>'. Below this, there is a 'SHARE' section with buttons for 'Share 0', 'Tweet', 'Share', and 'email'. On the right side, there is a circular badge with the number '35'. Below that, the 'AUTHORS' section lists 'Daniel Graziotin', 'Fabian Fagerholm', 'Xiaofeng Wang', and 'Pekka Abrahamsson'. The 'CATEGORIES' section lists 'Computer Software', 'Software Engineering', and 'Industrial and Organisational Psychology'. The 'KEYWORD(S)' section lists 'psychology', 'software engineering', 'affect', 'emotions', 'mood', 'happiness', 'software developer', 'quantitative data', 'qualitative data', 'survey', 'questionnaire', 'large-scale study', and 'GitHub'.

Graziotin, Daniel; Fagerholm, Fabian; Wang, Xiaofeng; Abrahamsson, Pekka (2017): Online appendix: the happiness of software developers. figshare. Collection.

<https://doi.org/10.6084/m9.figshare.c.3355707.v9>

# Altmetric



The screenshot displays the Figshare interface for a document. At the top, the Figshare logo is on the left, followed by a search bar containing "search on figshare" and a magnifying glass icon. To the right of the search bar are "Browse" and "Upload" links. Further right are "Sign up" and "Log in" buttons. The main content area features a large map with a legend on the left side. The map shows a river network and various colored regions. Below the map, there is a red button labeled "Explore more content" with a downward arrow. At the bottom of the viewer, the file name "tjom\_a\_1223760\_sm9946.pdf (12.99 MB)" is displayed on the left, and the MD5 hash "MD5: 5e357636880b7d9d8222d330c484f96" is on the right. Below the file name are icons for search, zoom, and download. At the very bottom, there are buttons for "Download (12.99 MB)", "Share", "Cite", "Embed", and "+ Collect (you need to log in first)".

# Figshare API

The image displays a Figshare API interface. The top navigation bar includes the Figshare logo, a search bar, and user information for Megan Harde... Below this, there are two rows of dataset thumbnails. Each thumbnail contains a plot titled 'Altitude Offset' and a file name with its size. The first row shows files for 2014 (146.31 kB, 145.54 kB) and a text file (0.31 kB). The second row shows files for 2014 (188.37 kB, 149.9 kB) and a text file (0.31 kB). A 'Download all (1.01 GB)' button is visible. The main content area shows a list of datasets under 'SE Texas - 2015', with a detailed view of the 'Launch: Wednesday, 04 March 2015' dataset. This view includes a satellite map showing a trajectory, a metadata table, and a list of project details.

Year	File 1 Size	File 2 Size	Text File Size
2014	146.31 kB	145.54 kB	0.31 kB
2014	188.37 kB	149.9 kB	0.31 kB

Station	Principal Investigator	Co-Investigator	Latitude (deg)	Longitude (deg)	Elevation (m)	Launch Date	Launch Time (UT)	Ozonesonde Type Number
University of Houston Houston Texas USA	Gary A. Morris (St. Edward's University)	Barry Lefer (University of Houston Houston TX USA)	29.72	95.34	19	20150304	23:02:18	EnSci/DMT 2227878

- STATION : University of Houston Houston Texas USA
- Station Principal Investigator : Gary A. Morris (St. Edward's University)
- Station Co-Investigator : Barry Lefer (University of Houston Houston TX USA)
- Latitude (deg) : 29.72
- Longitude (deg) : 95.34
- Elevation (m) : 19
- Launch Date : 20150304
- Launch Time (UT) : 23:02:18
- Ozonesonde Type Number : EnSci/DMT 2227878

- Data File (.dat)
- Trajectory Map (.kml)

# GitHub integration

The screenshot displays the GitHub interface for the 'Online Labour Index' repository. The top navigation bar shows 66 commits, 1 branch, 1 release, and 2 contributors. Below this, the 'Branch: master' dropdown and 'New pull request' button are visible. A list of recent commits by 'martinjhndley' is shown, including 'Merge pull request #7 from okassi/master' and several updates to 'employer\_dashboard', 'worker\_dashboard', and 'README.md'. The main content area shows a grid of files, including 'codebook.pdf', 'worker\_countryda...', 'OLdata\_2018-11-27.txt', 'Countries\_contine...', 'bcountrydata\_201...', 'OLI 2016-11-03.pdf', 'PlotBuyerCountry...', and 'Plot28ma.R'. At the bottom, the repository title 'Online Labour Index: Measuring the Online Gig Economy for Policy and Research' is displayed, along with statistics: 13944 views, 77860 downloads, and 0 citations. The repository description states it is the data underlying the Online Labour Index and provides the URL 'http://labour.oii.ox.ac.uk online-labour-index/' for details. Categories include 'Business Information Systems' and 'Economic Models and Forecasting'.

# Only data is not enough

- Well written electronic logbook
- Correct metadata
- Reproducible data analysis!

## Clicking Origin/Excel

- Fast to do
- Hard to understand later

Book1				
	A(X)	B(Y)	C(Y)	D(Y)
Long Name				
Units				
Comments				
F(x)=				
1	0.1	-0.68486	-1.08629	-1.34996
2	0.2	-0.72482	-1.13174	-1.37348
3	0.3	-0.73473	-1.14712	-1.38268
4	0.4	-0.73968	-1.15597	-1.38798
5	0.5	-0.74304	-1.16109	-1.39064
6	0.6	-0.74498	-1.16392	-1.39294
7	0.7	-0.74622	-1.16605	-1.39417
8	0.8	-0.74764	-1.16728	-1.3947
9	0.9	-0.74781	-1.16817	-1.39541

## Scripting Python/Matlab

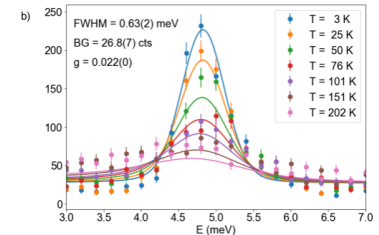
- More work to do properly
- Reproducible
- You can publish!

```
In [14]: ylist = []
for r in P2:
    r.plot(params=False, multi=True)
    # save fits to ascii
    ylist.append(r.yy)

ylist = [r.xx] + ylist

# save raw data
# save fits to cfs-by/comp/ing-fits.txt, array(ylist).T)

text(3.1, 285, "BG = %.1f(%.0f) cts" % (r.parameterict["bg_bkgd"].value, r.parameterict["bg_bkgd"].error*1e1))
text(3.1, 238, "FWHM = %.2f(%.0f) meV" % (r.parameterict["conv_width"].value, r.parameterict["conv_width"].error*1e2))
text(3.1, 189, "g = %.3f(%.0f)" % (r.parameterict["g_g"].value, r.parameterict["g_g"].error*1e3))
text(3.4, 285, "D")
xlim(3, 7)
title("")
xlabel("E (meV)")
ylabel("cts (comp. u.)")
legend()
savefig("cfs-by/comp/ing-pdf")
show()
```



It will be legen....Wait for it, we will publish something during HANDS-ON....

Questions?

...dary!