



Anna Krystalli

Nationality: Greek | r.rse.eu@gmail.com | annakrystalli@googlemail.com |
<https://www.r-rse.eu/> | <https://twitter.com/annakrystalli> |
<https://github.com/annakrystalli> | Dani Xrysonisos, 84100, Syros, Greece

WORK EXPERIENCE

20 MAY 2022 – CURRENT – Ermoupoli, Syros, Greece
RESEARCH SOFTWARE ENGINEERING CONSULTANT – R-RSE SERVICES

Have just returned to my homeland of Greece and am in the process of setting up a Research Software Engineering Consultancy specialising in the statistical programming language R.

31 MAY 2017 – 8 MAY 2022 – Sheffield, United Kingdom
RESEARCH SOFTWARE ENGINEER – UNIVERSITY OF SHEFFIELD

Was part of a team of RSEs working with researchers to build more robust analysis pipelines and software, promote best practice in research programming and digital resource management and facilitate the shift to a more robust, open, transparent and collaborative research culture.

SELECTED SOFTWARE PROJECTS (with links to publicly available materials where possible):

SeabirddietDB: The goal of seabirddietDB project was to provide access and tools to interact with a database of seabird diets collected around the British Isles. The project involved processing, cleaning, validating and packaging client supplied Seabird Diet dataset and associated metadata into an R package available on GitHub. Functionality to filter and plot data as well as an accompanying metadata webpage was also developed.

- Dataset R package: <https://github.com/annakrystalli/seabirddietDB>
- Metadata webpage: <https://annakrystalli.me/seabirddietDB/>

SedMaps: Shiny app developed to provide visualisation, extraction, spatial selection and summarising of raster maps of the north-west European Shelf sedimentary environment. The app allows users to extract gridded data and metadata as full layers, use in-built administrative maritime boundary polygons to extract data or alternatively, use toolbar to draw custom extraction polygons.

- Shiny app source code: <https://github.com/annakrystalli/sedMaps>
- Deployed Shiny app: https://sheffield-university.shinyapps.io/sedmap_shiny/

EMODnetWFS: Developed R package to allow interrogation of and access to EMODnet geographic vector data in R through EMODnet Web Feature Services.

- Package repository: <https://github.com/EMODnet/EMODnetWFS>
- Package documentation: <https://emodnet.github.io/EMODnetWFS/>

CAMP SITE package and Shiny app (on going): The project required formalisation of research code developed by client for a paper and involved refactoring and generalising simulation code into a user friendly and installable R package as well as refactoring a prototype shiny app to explore simulation results.

- Package source code: <https://github.com/annakrystalli/CAMP SITE>
- Shiny App source code: <https://github.com/annakrystalli/campsite-shiny>
- Deployed Shiny app: <https://sheffield-university.shinyapps.io/macro-evo-dynamics/>

Research Software Engineering consultant for AMRC OWAST group. Remit of the project was to help the Hole Generation group improve their experimental, data processing and result reporting to clients pipeline and included:

- Planning and designing next generation results database and ingress application.
- Containerisation & refactoring of prototype Shiny app using Golem framework, designed to interact with experimental results database to make application more robust, efficient & portable.

31 MAY 2017 – 8 MAY 2022 – Sheffield, United Kingdom
RESEARCH SOFTWARE ENGINEER – UNIVERSITY OF SHEFFIELD

Another important aspect of my role as a Research Software Engineer was capacity building through advocacy, providing support through code clinics and training, including development of original training materials.

SELECTED TRAINING MATERIALS

- **Reproducible Research Data and Project Management in R:** four day course for NERC Funded ACCE Doctoral partnership 1st year PhD students on research data and project management through R and Rstudio. The course is a mixture of lectures and practical sessions introducing students to best practice and includes chapters on data management and metadata, project management and structure, data munging, iteration,

functions including documenting and testing them, defensive programming, packaging R code, literate programming through Rmarkdown, data visualisation with ggplot2, version control and more. All materials are hosted openly online as a bookdown site with links to the source code: <https://annakrystalli.me/rrresearchACCE20/>

- **Git & GitHub Through GitKraken Client - From Zero To Hero!** Introduction to version control, basic and advanced team collaboration through GitKraken Client. Taught this course monthly on a total of 10 occasions. <https://srse-git-github-zero2hero.netlify.app/>
- **Introduction to GIS in R** Introduction to GIS basics and working with vector and raster data in R. <https://annakrystalli.me/intro-r-gis/>
- **Introduction to Deep Learning in R.** This course was developed in python & R as collaboration with other members of the RSE team. I translated the practical colab notebooks originally in python to R and taught the R versions of the course on 3 occasions. <https://rses-dl-course.github.io/>
- **Dataspice Tutorial:** An Introduction to using R package Dataspice to collect metadata for EMODnet Biology Data Products <https://annakrystalli.me/emodnet-dataspice-tutorial/>

See below for a list of selected talks and other advocacy and community building work

NOV 2014 – JUN 2017 – Sheffield, United Kingdom

RESEARCH DATA SCIENTIST / RESEARCH SOFTWARE ENGINEER – FREELANCE

Collaboration with researchers at Universities of Sheffield, Bath, UK; UNAM, Mexico and Seville, Spain.

Provided analytical and code development support in R, for an international team of Principal Investigators on two related projects:

1. project investigating the evolution of sex roles across the global avian assemblage.
2. project exploring the drivers of the structure of bird trait correlation networks.

Evolution of Sex Roles in Birds project

Goal was to compile a large dataset (1,307,253 data points across 202 variables) on bird reproductive, morphological, life history and ecological traits across as many bird species as possible. Project involved extracting, collating large disparate trait and environmental data for which I developed a series of tools for managing, validating and exploring data and analyses. Code, data and analyses were the primary outputs and the majority were made public. More specifically, scripts were developed to:

- clean, validate and compile trait datasets generated by a number of teams from literature reviews.
- Use BirdLife vector species range data to extract range ecological characteristics from environmental raster data (BioClim). Script parallelised and developed to be run on HPC cluster.
- explore distribution and bi-variate of traits through a shiny app (as the number of variables made Rmarkdown reports too long to review).

Bird Trait Correlation Networks

The next projects was to use the dataset described above and build a phylogenetically corrected network of trait correlations and entailed scripts to perform the phylogenetic correction using taxonomic metadata, build the correlation networks using rnetcarto and visualise them interactively using igraph

Professional, scientific and technical activities

<https://github.com/annakrystalli/rmacroRDM>

<https://github.com/annakrystalli/IUCNextractR>

https://github.com/annakrystalli/bird_trait_networks

<https://rpubs.com/annakrystalli/122371>

<https://figshare.shef.ac.uk/articles/software/>

[Interactive App Enhanced sensitivity to group differences with decision modelling/8109161](https://annakrystalli.shinyapps.io/bird_app/)

https://annakrystalli.shinyapps.io/bird_app/

AUG 2011 – MAY 2017 – Sheffield, United Kingdom

GRADUATE TEACHING INSTRUCTOR / ASSISTANT – UNIVERSITY OF SHEFFIELD

Provided a variety of teaching support including instructing, developing teaching materials and assisting throughout my time as a PhD student and later as a freelancer.

- Codeveloped a one day workshop on tidy data, metadata and open science in R for first year ACCE PhD students from the University of Sheffield, York and Liverpool. Delivered at the University of York (01/03/2015)
- Assisted on a week long postgraduate introductory course to statistical analysis in R for the Dept. of Animal & Plant Science for three consecutive years (2012-14)
- Member of the teaching staff team on the APS 3rd year undergraduate Marine Ecology field course to Anglesey Wales for three consecutive years (2012-14)
- Assisted with a wide variety of undergraduate practical teaching sessions.

SEP 2006 – OCT 2011 – Castle Donington, United Kingdom
BRAND COORDINATOR / DATA OFFICER – ULTRA SPORT EUROPE LTD

Ultra Sport is a distributor of key brands of extreme sports equipment in surfing, snowboarding, wakeboarding, windsurfing and kitesurfing. Professional relationship maintained throughout studies on a freelance basis. Duties included:

- Processing sales and stock purchases, tracking shipments and monitoring stock availability.
- Role was expanded to include providing technical support and advice for marketing, data handling, costing of sales, systems and general IT. Was appointed with populating and maintaining the online shop and dealer locator database on development.
- Providing dedicated customer support to dealers as well as members of the public.
- Organised and carried out events and demos, liaised with media, produced marketing material whilst working to tight deadlines.

Wholesale and retail trade; repair of motor vehicles and motorcycles | info@ultrasporteu.com |

<https://www.ultrasporteu.com/> |

Ultra Sport Europe Ltd, Carnival Way, DE74 2HP, Castle Donington, United Kingdom

MAR 2010 – OCT 2010 – Bristol, United Kingdom
SOCIAL AND MARKET RESEARCH OFFICER – CELLO MRUK

Cello MRUK is a leading global insight and strategic marketing group. Duties included cleaning, coding and analysis of data. In my short stay at the Bristol office of the company I helped improve standard methods by training colleagues in more advanced aspects of Excel functionality.

JAN 2010 – MAY 2010 – Bristol, United Kingdom
TEMPORARY PROJECT OFFICER – SMITH & WILLIAMSON

Smith & Williamson is one of the leading independently owned investment management, financial advisory and accounting groups in the UK. I was employed to realise a large, complex mail out of customised information packs to their entire client base, requiring the integration of Word and Excel and expert use of mail merge functionalities.

AUG 2009 – NOV 2009 – Plymouth, United Kingdom
TEMPORARY PROJECT OFFICER – MARINE BIOLOGY & ECOLOGY RESEARCH CENTRE, UNIVERSITY OF PLYMOUTH

Participated in High Definition video sampling cruises of the Lyme Bay seabed to monitor recovery of the benthos following closure of the Bay to scallop dredging. Performed analysis of resulting video footage and stills (identification and quantification of benthic macrofauna).

31 AUG 2004 – 31 AUG 2006 – Shardlow,, United Kingdom
QUALITY ASSURANCE AUDITOR – SAFEPHARM LABORATORIES LTD (NOW HARLAN LABORATORIES)

Safepharma was a contract research facility involved in pre-clinical chemical safety testing. The primary function of the QA unit was to ensure the work of the company complied with Good Laboratory Practice defined by European and international legislation and comprised:

- Inspection of ongoing studies and facilities to ensure compliance with internal standard methods and operating procedures.
- Auditing of raw data to ensure data was correctly generated, recorded and processed. Auditing of final reports to ensure they were accurate and faithful accounts of the data.
- Reporting of all findings to management and engaging with staff at all levels to resolve them.

MAY 2017 – MAY 2022 – Sheffield, United Kingdom
RESEARCH SOFTWARE ENGINEER – UNIVERSITY OF SHEFFIELD

Was part of a team of RSEs working with researchers to build more robust analysis pipelines and software, promote best practice in research programming and digital resource management and facilitate the shift to more open, transparent and collaborative research culture.

SELECTED SOFTWARE PROJECTS:

- **SeabirdDietDB:** The goal of seabirdDietDB project was to provide access and tools to interact with a database of seabird diets collected around the British Isles. The project involved processing, cleaning, validating and

packaging client supplied Seabird Diet dataset and associated metadata into an R package available on GitHub. Functionality to filter and plot data as well as an accompanying metadata webpage was also developed.

- *Dataset R package*: <https://github.com/annakrystalli/seabirddietDB>
- *Metadata webpage*: <https://annakrystalli.me/seabirddietDB/>
- **SedMaps**: Shiny app developed to provide visualisation, extraction, spatial selection and summarising of raster maps of the north-west European Shelf sedimentary environment. The app allows users to extract gridded data and metadata as full layers, use in-built administrative maritime boundary polygons to extract data or alternatively, use toolbar to draw custom extraction polygons.
- *Shiny app source code*: <https://github.com/annakrystalli/sedMaps>
- *Deployed Shiny app*: https://sheffield-university.shinyapps.io/sedmap_shiny/
- **EMODnetWFS**: The goal of EMODnetWFS is to allow interrogation of and access to EMODnet geographic vector data in R through the EMODnet Web Feature Services
- **CAMP SITE package and Shiny app (on going)**: The project required formalisation of research code developed by client for a paper on the "Opposing effects of competition and selection on macroevolutionary dynamics" (under review). Work involved refactoring and generalising simulation code for running Competition And Multiple-Peak Selection Integrated Trait Evolution (CAMP SITE) model into a user friendly and installable R package. The package then became a dependency of a shiny app, refactored from a client prototype, to visualise replicates of simulation results under different parameter values. The app includes ability to download plots of simulation results and is styled using package bslib. Both outputs were developed to be associated with the paper publication.
- *Package source code*: <https://github.com/annakrystalli/CAMP SITE>
- *Shiny App source code*: <https://github.com/annakrystalli/campsite-shiny>
- *Deployed Shiny app*: <https://sheffield-university.shinyapps.io/macro-evo-dynamics/>
- **Research Software Engineering consultant for AMRC OWAST group**. Remit of the project was to help group improve their experimental, data processing and result reporting to clients pipeline and included:
 - Planning and designing next generation results database and ingress application.
 - Containerisation & refactoring of prototype Shiny app using Golem framework, designed to interact with experimental results database to make application more robust, efficient & portable.

● EDUCATION AND TRAINING

SEP 2010 – MAY 2015 – Department of Animal & Plant Sciences, Sheffield, United Kingdom

PHD IN MARINE MACROECOLOGY – University of Sheffield

Investigated the macroecological temporal dynamics of regional population structure of copepod species at varying temporal and spatial resolutions using generalised linear and mixed modelling techniques. Additionally developed an environmentally informed method of copepod data interpolation incorporating satellite images and using state of the art machine learning techniques. As part of the degree, I also completed a week long training course in Advanced Statistics in R delivered within the Department of Animal & Plant Science.

Field(s) of study

- Natural sciences, mathematics and statistics

Thesis: Abundance-occupancy relationships in North Sea copepods: ventures into time and space

EQF level 8 | <https://theses.whiterose.ac.uk/13831/>

AUG 2007 – JUN 2009 – Drake Circus, Plymouth, United Kingdom

BSC (HONOURS) IN MARINE BIOLOGY & OCEANOGRAPHY – University of Plymouth

Achieved a 1st Class Honours Degree. Awarded Babbage Scholarship of £3000 for performance in 2nd year.

Field(s) of study

- Natural sciences, mathematics and statistics : *Biological and related sciences not elsewhere classified*

Thesis: Long-term trends in abundance and phenology of North Sea Continuous Plankton Recorder (CPR) fish larvae in relation to environmental and trophic drivers

1 | EQF level 6 | Bachelors Degree in Science with Honours

7 DEC 2020 – 8 DEC 2020

REPRODUCIBLE COMPUTATIONAL ENVIRONMENTS USING CONTAINERS: INTRODUCTION TO DOCKER & SINGULARITY – Edinburgh Parallel Computing Centre (EPCC), University of Edinburgh

<https://epccd.github.io/2020-12-08-Containers-Online/index.html>

23 JAN 2020 – 24 JAN 2020

DEEP LEARNING IN R – Workshop at Rstudio::conf2020

4 NOV 2018 – 18 FEB 2019

SOFTWARE CARPENTRY INSTRUCTOR TRAINING – The Carpentries

<https://carpentries.org/become-instructor/>

JUN 2013

SPATIAL STATISTICS IN R – British Ecological Society Macroecology special interest group workshop

JUN 2012

MODEL SELECTION – Workshop at International Statistical Ecology Conference 2012

MAR 2012 – JUL 2012

MACHINE LEARNING – Stanford University - Coursera

<https://www.coursera.org/learn/machine-learning>

OCT 2005

OBSERVATION AND RECORDING FOR AUDITORS – British Association of Research Quality Assurance

JAN 2005

THE AUDITING COURSE – British Association of Research Quality Assurance

● **DIGITAL SKILLS**

My Digital Skills

R | Rstudio | Keras | Git | Ansible | Django | Python | Docker | R Shiny | R package development | Github Actions | SQL | Deep Learning | Machine Learning | Advanced RMarkdown | Scientific Writing Reproducible Research LaTeX Jupyter Rmd | Agile (Scrum & Kanban) | version control | Web: HTML, CSS

● **LANGUAGE SKILLS**

Mother tongue(s): **GREEK** | **ENGLISH**

Other language(s):

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken production	Spoken interaction	
ITALIAN	A2	A2	A2	A2	A1

Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user

● HONOURS AND AWARDS

10 APR 2019

Software Sustainability Fellow 2019 – Software Sustainability Institute

Awarded £3,000 to pursue development of the ReproHack project, one day reproducibility hackathons aiming to help author and participating researchers learn about reproducibility through hands on reviewing and attempting to reproduce papers from associated published code and data.

<https://software.ac.uk/blog/2019-04-11-announcing-2019-software-sustainability-institute-fellows> <https://www.reprohack.org/>

17 SEP 2019

Best Poster at 4th Conference of Research Software Engineering (2019) – Society of Research Software Engineering

Poster Abstract:

The R research community has been extremely fortunate for the existence of rOpenSci. (<https://ropensci.org/>). rOpenSci curate an impressive collection of community contributed packages to support open reproducible research and data access in R. The peer review system used has not only improved the quality of contributed packages, but the practice of reviewing itself has elevated the skills of the whole community involved, by engaging us with best practice. The key to this success is a formalised review process based on detailed recommendations and guidelines (https://ropensci.github.io/dev_guide/). To help reviewers further, rOpenSci have also focused on tooling and automation. One such tool is pkgreviewr (<https://github.com/ropenscilabs/pkgreviewr>), an R package that helps automate some of the steps and guide reviewers through the review process. In this poster I'll present an overview of the rOpenSci review process and the use of pkgreviewr to support it. I'll also tell the story of how the package came to be and how such contributions can lead to deeper involvement with the community.

Prize was sponsored by Microsoft and involved visits to the Microsoft Cambridge Offices for a series of mentorship sessions.

<https://rseconuk2019.sched.com/event/SMZv/p17-pkgreviewr-supporting-ropensci-package-reviews-through-guidance-automation-and-templating> <https://www.microsoft.com/en-us/research/blog/inclusive-environments-empower-big-dreams/>

25 JUN 2017

Network 50 – Mozilla

Named one of 50 People Who Made the Internet a Better Place in 2016.

<https://foundation.mozilla.org/nl/blog/network-50/> <https://storyengine.io/anna-krystalli/>

JUL 2008

Babbage Scholarship – University of Plymouth

Awarded Babbage Scholarship of £3000 for performance in 2nd year.

● PROJECTS

23 NOV 2016 – CURRENT

ReproHack Project

<https://www.reprohack.org/>
https://github.com/reprohack/reprohack_site

ReproHacks represent sandbox environments for practicing Reproducible Research. During a ReproHack, participants attempt to reproduce published research of their choice from a list of proposed papers using publicly available associated code and data. Participants get to work with other people's material in a low pressure environment and record their experiences on a number of key aspects, including reproducibility, transparency and reusability of materials. Authors receive invaluable feedback on the reproducibility and reusability of their work.

ReproHacks started out as part of [OpenCon satellite events](#) (Berlin 2016, London 2017) and the project was developed further as part of my 2019 Software Sustainability Institute Fellowship. The initial scope of the fellowship involved running a few events to raise the profile and promote the activity as valuable practical training in reproducibility. Since then, a core project team has formed and we've received additional funding from the [N8 CIR](#) and [EPSRC](#) to run additional events (such as the [N8 Northern Tour Series](#)), develop the format to accommodate Computationally Intensive Research (such as the HPC ReproHack at the University of Warwick) and develop an online [ReproHack Hub](#) to

administer all ReproHack activities, making it easy for anyone to run their own ReproHack event. As well as the project lead, I was also the main developer of the Hub in Django / Python which is available on GitHub.

● CONFERENCES AND SEMINARS

8 FEB 2022 – Online, First AstraZeneca (AZ) R Conference 2022

Panel discussion: Reproducibility in R

Panelists: Anna Krystalli (University of Sheffield RSE), Diane Beldame (Thinkr), Abhijit Dasgupta (AZ, Machine Learning & AI) and Damian Zajac (AZ, Centralized Monitoring).

24 SEP 2021 – Online, MetaScience 2021 Conference

Panel: Engaging the community in research reproduction

In this panel we discussion how to promote community engagement with reproducing research.

The panelists were: Anna Krystalli is an organizer of reproducibility hackathons ReproHack that engage various people who try to reproduce published software; Stephen Eglon is a co-creator of CODECHECK, a project that checks codes during peer review by voluntary checkers; Sergey Frolov is an experimental physicist who scrutinizes published results, and whose activities recently led to the retraction of a Nature paper; Ethienne Roesch is editor of the replication-specialized journal ReScience C and editor-in-chief of ReScience X. This panel will be moderated by Cassio Amorim, creator of the website SciGen.Report.

A [recording](#) of the panel discussion is available on the Centre for Open Science YouTube channel.

<https://metascience2021.org/events/engaging-the-community-in-research-reproduction/> https://www.youtube.com/embed/nUWFYSF_I9k

8 JUN 2021 – Online, Sheffield RSE Lunchbytes on Clinical Research Software

REDCap: Features, setup and maintenance

Event Description

Short talks on how University of Sheffield researchers and RSEs, and collaborators worldwide, are using [REDCap](#) and other platforms to handle clinical research data. My talk focus on general features and described the setup and maintenance of a REDCap service. [Slides](#) and [recording](#) of the talk are available <https://bit.ly/redcap-lunchbytes> https://docs.google.com/presentation/d/e/2PACX-1vSAWAtWiWC-qeGkK_YPNd7ldG71bP0wWm_chAm-9th9CeNUfatcachw0U_vRI4CSSgMQRWgc1fESM_R/pub?start=true&loop=true&delayms=3000&slide=id.p

11 MAY 2021 – Sheffield RSE LunchBytes

Putting the R into Reproducible Research

Whistle-stop tour of the breadth of available tools, demonstrating the ways R and the Rstudio integrated development environment can be used to underpin more open reproducible research and facilitate best practice.

[Slides](#) and [recording](#) are available.

<http://bit.ly/r-in-repro-research-dc-srse> https://digitalmedia.sheffield.ac.uk/media/Lunch+bytes+6A+Putting+the+R+into+Reproducible+Research/1_kwofatj6

15 DEC 2020 – Online, University College London Tech Socials Series

Putting the R into Reproducible Research: Director's Cut

Whistle-stop tour of the breadth of available tools, demonstrating the ways R and the Rstudio integrated development environment can be used to underpin more open reproducible research and facilitate best practice.

<https://annakrystalli.me/talks/r-in-repro-research-dc-ucl.html#1> <https://www.eventbrite.co.uk/e/kq-codes-technical-social-wed-16th-december-2020-anna-krystalli-tickets-131359733499>

9 JUL 2020 – Online, useR 2020 Conference

useR 2020 Keynote: Computational Reproducibility: from theory to practice

Drawing on experiences as a Research Software Engineer, an editor and reviewer of R packages for rOpenSci and a founding member of the ReproHack reproducibility hackathon team, I take a look at the progress we've made so far, try to identify what we're still missing and offer some thoughts on how to move forward.

[Slides](#) and [recording](#) of keynote are available

<https://www.youtube.com/embed/KHMW8fV2NXo> <https://annakrystalli.me/talks/user2020.html#1> <https://user2020.r-project.org/program/keynotes/>

London, British Ecological Society Quantitative Ecology Special Interest Group Hackathon

Adventures in Open Code & Data: Why I love Hackathons!

Invited talk to kick off the BES Quantitative Ecology Hackathon, a one-day hackathon bringing researchers together in small teams to work with socio-ecological data. I spoke about why I love hackathons, in particular as an opportunity to learn more about open source and open science.

https://annakrystalli.me/talks/bes_hackathon.html <https://www.software.ac.uk/blog/2019-11-19-quantitative-ecology-hackathon>

8 JUL 2019 – British Ecological Society Quantitative & Movement Ecology SIG join meeting

What can RSEs do for U?

Invited talk at the BES Quantitative & Movement Ecology 2019 Annual Meeting to speak about what Research Software Engineers do and how they can support academics with their computational research.

https://annakrystalli.me/talks/bes_sig19.html

20 MAY 2018 – 21 MAY 2018 – Microsoft Reactor, Seattle, Washington, USA

rOpenSci Unconf 2018

The rOpenSci unconference is an annual mostly invite-only event which brings together scientists, developers, and open data enthusiasts from academia, industry, government, and non-profits. Over a few days participants hack on projects related to a wide variety of topics, from open data, data visualization, data publication to open science using R. The agenda is mostly decided during the conference by proposing projects in GitHub issues. I proposed to develop [tools in R and guidance on the most minimal dataset metadata standards, files and formats](#) researchers should record. A team interested in research metadata assembled around the topic and the worked on package **dataspice** which helps researchers create lightweight schema.org descriptions of your datasets and is available on CRAN.

<https://unconf18.ropensci.org/> <https://github.com/ropensci/unconf18/issues/72> <https://github.com/ropensci/dataspice>

3 AUG 2017 – 5 AUG 2017 – GooglePlex, Mountain View California, USA

Science Foo (SciFoo) 2017

Science Foo Camp (scifoo) is an annual, invite only, interdisciplinary scientific [unconference](#) organized by [O'Reilly Media](#), [Digital Science](#), [Alphabet Inc.](#) The meeting has no fixed agenda; the invited scientists, technologists and policy makers set the conference program during the conference itself. I was invited in 2017 and proposed and chaired a well attended session entitled: *"Reproducible Research: Everyone wants it but who is going to teach us to do it?!"*

https://en.wikipedia.org/wiki/Science_Foo_Camp <https://www.digital-science.com/blog/2017/08/collective-noun-science-scifoo-2017/>

OpenCon Berlin ReproHack

OpenCon London ReproHack

17 MAR 2019 – University of Sheffield

The Turing Way: Build a BinderHub

Was part of the Turing Way project team who organised and facilitated a Build a BinderHub Workshop. During the workshop we demonstrated how to build a BinderHub on Microsoft Azure cloud computing resources. We helped participants get started with building a BinderHub on their institution's computing platform and discussed the challenges of maintaining a BinderHub.

The Turing Way is an open developed handbook to support students, their supervisors, funders and journal editors in ensuring that reproducible data science is "too easy not to do".

<https://rse.shef.ac.uk/training/workshop/2019-03-18-build-a-binderhub/> <https://github.com/alan-turing-institute/the-turing-way/tree/main/workshops/build-a-binderhub> <https://the-turing-way.netlify.app/welcome>

● NETWORKS AND MEMBERSHIPS

21 JUN 2018 – CURRENT

rOpenSci Software Review Editor

rOpenSci aims to lower barriers to working with local and remote scientific data sources through carefully vetted, staff- and community-contributed R software tools. We help develop R packages for the sciences via community driven learning, review and maintenance of contributed software in the R ecosystem. At the core of rOpenSci activities lies community review of contributed packages which is administered through GitHub. I have been an editor for rOpenSci since May 2018.

<https://github.com/ropensci/software-review> <https://ropensci.org/> https://ropensci.org/blog/2018/06/22/new_editors/

APR 2016 – MAR 2020

Sheffield R Users Group

Sheffield, UK

The Sheffield R Users hosted monthly talks on all topics related to R. I gave a number of talks and demos and was also co-organiser of the group from 2016 to 2020.

<https://www.meetup.com/SheffieldR-Sheffield-R-Users-Group>

JAN 2016 – SEP 2018

Mozilla Science Lab

The Mozilla Science Lab was an initiative of the Mozilla Foundation from 2013 - 2018 and explored how the power of open source could change the way science was done on the web. The program operated under the belief that a community of peers that work, learn and build together could make research more collaborative, efficient and thrive on the open web.

The program of participatory learning and mentorship in Open Leadership was organised in cohorts and I was part of the inaugural cohort which kicked off with a Working Open Workshop in Berlin 2016. The workshop aimed to prepare up and coming leaders of open science to make the most effective open project possible and build active communities of contributors around them.

After participating as a mentee in the inaugural cohort, I went on to mentor, contribute to the training or as an expert for the next 6 cohorts. I also hosted a Globe Sprint site at the University of Sheffield in 2017 & 2018, facilitated spaces and sessions at the annual MozFest in London in 2016, 2017 and 2019, including a session on [Git & GitHub through Rstudio](#) in 2016. Overall engagement with the community opened my eyes of to the power and potential of open sourcing science.

<https://wiki.mozilla.org/ScienceLab> <http://mozillascience.github.io/working-open-workshop/> https://github.com/annakrystalli/Mozfest_github-rstudio

7 FEB 2021 – 26 MAY 2021

Open Life Sciences

The Open Life Science (OLS) program helps **individuals and stakeholders in research** to become **Open Science ambassadors** and is in many ways an evolution of the sun-setted Mozilla Science Lab initiative. I participated as a mentor in OLS Cohort 3.

<https://openlifesci.org/ols-3>

● CREATIVE WORKS

MAY 2020 – CURRENT

EMODnetWFS R package

Package developed as part of EMODnet Biology Phases III & IV to enable interrogation, filtering and access of EMODnet Web Feature Service data through R.

<https://github.com/EMODnet/EMODnetWFS/> <https://emodnet.github.io/EMODnetWFS/>

DEC 2017

pkgreviewr R package

The goal of pkgreviewr is to facilitate rOpenSci reviewers in their package reviews. It creates a review project containing populated templates of all the files one needs to complete an rOpenSci review. It also clones the source code of the package under review to a convenient location, allowing easy checking and testing.

<https://github.com/ropensci-org/pkgreviewr> <https://docs.ropensci.org/pkgreviewr/>

APR 2018

dataspice R package

The goal of dataspice is to make it easier for researchers to create basic, lightweight, and concise metadata files for their datasets by editing the kind of files they're probably most familiar with: CSVs. These metadata files can then be used to:

- Make useful information available during analysis.
- Create a helpful dataset README webpage for your data.
- Produce more complex metadata formats for richer description of your datasets and to aid dataset discovery.

Metadata fields are based on [Schema.org/Dataset](https://schema.org/Dataset) and other [metadata standards](#) and represent a lowest common denominator which means converting between formats should be relatively straightforward.

This project was a team effort developed as part of rOpenSci unconf 2018 and is now available on CRAN. I've also developed a number of training materials using it, to teach data producers to develop accompanying minimal metadata in R.

<https://github.com/ropensci/dataspice> <https://docs.ropensci.org/dataspice/> <https://annakrystalli.me/emodnet-dataspice-tutorial/>

● **VOLUNTEERING**

6 APR 2016 – 20 APR 2016

Benthic Sampling Research Volunteer

English Channel & Irish Sea

Participated in a 2-week benthic sampling cruise carried out by Bangor University aboard R.V. Prince Madog as part of the Marine Ecosystems Research Programme in the English Channel and Irish Sea. Work involved water column CTD sampling, plankton sampling, dredging and trawling for meio- & macro-fauna, enumeration, identification and measuring biomass of sampled organisms.

OCT 2003 – MAR 2004

Scientific Volunteer - Archipelagos Aigaiou

Ikaria, Greece

- Wrote articles and helped maintain and update the organisation's website.
- Carried out biological surveys of the littoral zone on Ikaria and other islands of the Eastern Aegean visited during research cruises.
- Gained a sound understanding of environmental problems affecting rural island communities as well as major social, political and economic issues affecting their solution.

30 JUN 2001 – 30 AUG 2001

Scientific Volunteer - Operation Wallacea

Hoga Island, Sulawesi, Indonesia

- Participated in data collection using SCUBA to investigate the effects of mangrove loss and the resulting sedimentation on coral reefs. Performed data manipulation and presentation.
- Engaged with local communities in an attempt to promote the sustainable use of reef resources and the establishment of a no fishing area.
- Acquired a good understanding of reef ecology, good reef organism ID skills, a deeper understanding of the interactions between poverty and environmental degradation and enhanced diving skills.