Welcome to the Library Carpentry Etherpad!

This pad is synchronized as you type, so that everyone viewing this page sees the same text. This allows you to collaborate seamlessly on documents.

Use of this service is restricted to members of The Carpentries community; this is not for general purpose use (for that, try etherpad.wikimedia.org).

Users are expected to follow our code of conduct: https://docs.carpentries.org/topic_folders/policies/code-of-conduct.html

All content is publicly available under the Creative Commons Attribution License:
https://creativecommons.org/licenses/by/4.0/

Library Carpentry Community Calls:
----
16 May 2019 @ 14:00 UTC
See time and date in your zone:
https://www.timeanddate.com/worldclock/fixedtime.html?
msg=Library+Carpentry+Community+Call&iso=20190516T14&ah=1
We will be using Zoom for the call: https://ucop.zoom.us/my/ce
rdmann

0-20 min
Alice Meadows (https://twitter.com/alicejmeadows), Communications Director, ORCID
Terry Law (https://www.pnnl.gov/science/staff/staff_info.asp?staff_num=8519), User Program Services, Project Manager, Pacific Northwest National Laboratory
Topic: What's New at ORCID? Learn about new features available via ORCID's API 3.0 to connect and share more of your own and community's contributions.

20-50 min
Community Updates
Attending:
    Sheila Rabun, LYRASIS, sheila.rabun@lyrasis.org
    Gabriela Mejias, ORCID, g.mejias@orcid.org
    Melroy Almeida, AAF, melroy.almeida@aaf.edu.au
    kay Steel, Federation University Australia kay.steel@federation.edu.au
    Debbie Martindale, Rhodes University, d.martindale@ru.ac.za
    Marié Roux, Stellenbosch University, mr@sun.ac.za
    Molly O'Brien, UMass Dartmouth, mobrien9@umassd.edu
    Rebecca Hedreen, Southern Connecticut State University, hedreenr1@southernct.edu
    Kirikova Alexandra, Kursk State University, sashakirikova@mail.ru
    Christina Miskey, University of Nevada Las Vegas, christina.miskey@unlv.edu
    Alainna wrigley, UCOP, alainna@hotmail.com
    Frank Berkemeier, University of Muenster, frank.berkemeier@uni-muenster.de
    Sebastian Herwig, University of Muenster, s.herwig@uni-muenster.de
    Dan McCloskey, Massachusetts General Hospital, dmccloskey@mgh.harvard.edu

•   Eleanor Colla, University of New England, ecolla@une.edu.au

- Paloma Marín Arraiza,TU Wien Bibliothek paloma.arraiza@tuwien.ac.at
- Thomas Hilberer, Universität Tübingen, hilberer@uni-tuebingen.de
- Henrik Åkerfelt, Karolinska Institutet, henrik.akerfelt@ki.se
- Birgid Schlindwein, Technial University of Munich, birgid.schlindwein@ub.tum.de
- Shawna Sadler, Sadler Studio. shawna@sadlerstudio.com
- Marc Kaulisch, Leipzig University, marc.kaulisch@zv.uni-leipzig.de
- Eva Babonich, Graz University of Technology, eva.babonich@tugraz.at
- Aubrey Kirkpatrick, aubrey.kirkpatrick@vuw.ac.nz
- Joseph Payares, Momentum Engineering, j_payares@yahoo.com
- Jesse Johnston, Library of Congress, jesjohn@loc.gov
- Gerlinde Maxl, TU Graz Library, gerlinde.maxl@tugraz.at
- Ute Grimmel-Holzwarth, Eberhard Karls Universität Tübingen, UL, openaccess@ub.uni-tuebingen.de
- Axel Braun, Eberhard Karls Universität Tübingen, IT, axel.braun@uni-tuebingen.de
- Markus Wust, Eberhard Karls Universität Tübingen, Information, markus.wust@uni-tuebingen.de
- Helen Baer, Colorado State University, helen.baer@colostate.edu
- Joan Kolarik, Weizmann Institute, joan.kolarik@weizmann.ac.il
- Martin Mehlberg, University of Bremen, martin.mehlberg@suub.uni-bremen.de
- Trevor Burrows, Purdue University, dburrows@purdue.edu
- Christine Wells, UCLA, crwells@ucla.edu
- Marco Winkler, University of Potsdam, mawinkl@uni-potsdam.de
- Philipp Rumpf, University of Bamberg, philipp.rumpf@uni-bamberg.de
- Jens Schwidder, KOBV, schwidder@zib.de
- Bernette Sherman, ORCID, b.sherman@orcid.org
- peter.rempis@uni-tuebingen.de
- Paul Pival, University of Calgary, ppival@ucalgary.ca
- Steven Pryor, University of Missouri, pryors@missouri.edu
- Wendy Robertson, University of Iowa, wendy-robertson@uiowa.edu
- Claudia Piesche, University of Bayreuth, claudia.piesche@uni-bayreuth.de

Participants: 28

Slides:https://docs.google.com/presentation/d/1nK5sGZGtzJLcFjx8CbOFhRiZ4boeJRKDWZKmYI1ZYQE/edit#slide=id.g49c08e27e0_2_0
Survey: https://orcid.org/blog/2019/05/14/listening-our-users-what-we-learned-our-2019-community-survey

Questions:
   Does EMSL use the Data Science GRID for facilities? USe all the numbers RINGGOLD, GRID...
   Each time data is updated, you have to push all data (the full schema) to ORCID?
   Replacement document sent - there is some validation

   How fun! Are any of the other national research labs that host resources and facilities doing the same as PPNL? And couldn't that org without a unique org ID obtain one via GRID, which ORCID supports?
   Yes a number of orgs are doing this - Los Alamos, Oakridge
   Making it easier for individual researchers
   Data quality and user standpoint - track changes with email addresses, etc

   Does their integration help with meeting their OSTI reporting requirements?
   Helps office of science - takes whole year to determine distinct researchers (45K?)

Determine what resources these facilties need

Recommend talking to IT - sit in secure zone - multiple connections to receive authentications

ORCID API and Registry resources: https://members.orcid.org/api
  ORCID Public API

- Public API for ORCID - read publicly available data (any ORCID iD holder can create a Public API client): https://members.orcid.org/api/accessing-public-api
- Anyone can use the public API Swagger to test API functionality (no credentials needed): https://pub.orcid.org/v2.0/ (no client needed)

  ORCID Sandbox Member API

- Any individual or organization can freely request access to ORCID's sandbox testing server to try out the member API: https://orcid.org/content/register-client-application-sandbox
- Workflows: https://members.orcid.org/api
- API 3.0 technical documentation: https://github.com/ORCID/ORCID-Source/blob/master/orcid-model/src/main/resources/record_3.0/README.md

ORCID Annual public data file

- Annual public data releases usually in October (ORCID's birthday) coincides with OA Week: https://orcid.org/content/orcid-public-data-file-use-policy

Notes:
  Today is the official launch date of the new API
  Real power of ORCID is for researchers to share what information they would like to share with various services
  Information about yourself - work, education, etc
  New API adds affiliation types - enable researchers to get credit for qualifications, memberships (association), service (peer reviewer, board)
  Connect research resources (facilities)
  Tokens allow researcher to give permission once, can be shared with other organizations using that information
  Upgrade to this API is much smoother - improve backwards compatability
  New work types such as preprints and annotations and physical objects (only works in 3.0 API)
  Grouping versions of the same output (only works in 3.0 API)
  Improved documentation
  Contact consortial lead or support@orcid.org
  Or join the API Knowledgebase
  Survey results:
    Librarians are the biggest enthusiasts of ORCID (thank you!)
    What info researchers want to share all their outputs and contributions
    Added conference contributions!

PNNL use of ORCID / EMSL (provides scientific resources to the community)
Recognition outside of just acknowledgements
For annual reports, need to be able to link/uniquely identify researchers/facilities

ORCID was identified as solution

Build around automated workflows and integrate into systems (badging, safety property records, etc)

System - User portal, admin, instrument scheduling, data repository

One stop shop for users - tied ORCID into portal under their profile as part of authentication process

Present permissions - yes/no - authenticates w ORCID

Policy indicates that they have a year to publish data - one year embargo

Publicly released data available via faceted search

In order to download data - requires registration and associates w/ ORCID (tracking)

Link shown in ORCID record to infrastructure of facility

Minting DOIs for all awards

Have access to multiple facilities through single review process - so give credit to all organizations - developed partner tables to link w/in ORCID records

Associating resesarch with facilities to better track usage of resources

ORCID schema find and replace - submit larger data push each time with all the new information overwriting in ORCID schema with all the new relationships

Recognition of contributions from joint appointments, other levels of contributions

github/PNNL-EMSL orcid repo where sharing work

Technical specs Java to Tomcat

50-60 min
Library Carpentry Updates

Archived Calls
----
6 August 2018 @ 20:00 UTC
See time and date in your zone: https://www.timeanddate.com/worldclock/fixedtime.html?msg=Library+Carpentry+Community+Call&iso=20180806T20&p1=1440&ah=1
We will be using Zoom for the call: https://carpentries.zoom.us/j/918430249

0-20 min
Presentation can be found at https://github.com/LibraryCarpentry/governance/tree/master/community-calls
Shelley Stall (https://twitter.com/ShelleyStall) Director of Data Programs at American Geophysical Union
Topic: Brief update on Findable, Accessible, Interoperable, and Reusable (FAIR) data & software and how can librarians get involved
See https://www.force11.org/group/fairgroup/fairprinciples
Slide 3: AGU's Position Statement on Data: https://sciencepolicy.agu.org/files/2013/07/AGU-Data-Position-Statement-Final-2015.pdf
Slide 4: Belmont Forum, Data Management Skills Gap Analysis, April 7,2017, http://bfe-inf.org/document/skills-gap-analysis
Slides 5-7: FAIR Guiding Principles, Article in Nature journal *Scientific Data*: Wilkinson, M. D. *et al.*The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data*3:160018 doi: 10.1038/sdata.2016.18 (2016).
Slides 8 - 15: Enabling FAIR Data Project Website with Commitment Statement, http://www.copdess.org/enabling-fair-data-project/
Slides 17-18: Registry of Repositories, re3data.org
  For **tune your repository record** use this guide: http://bit.ly/RepoGuide
  Summarize your update in an email: info@re3data.org

Slide 19: **European Commission, Interim Report on FAIR Data & Action Plan**
Turning FAIR data into reality: Interim report from the European Commission Expert Group on FAIR data, https://doi.org/10.5281/zenodo.1285271
FAIR Data Action Plan: Interim recommendations and actions from the European Commission Expert Group on FAIR data, https://doi.org/10.5281/zenodo.1285289
Slide 20: Data Management Training Clearinghouse, http://dmtclearinghouse.esipfed.org

Repository Finder Tool:
   https://repositoryfinder.datacite.org/ (This is still in development!!!)
   Note: re3data records are being used for tool
   Share your feedback here:
     Why is Core Trust Seal the only option, there are other certifications out there and this may unfairly penalize certain repositories
     re3data data may be missing discipline repositories and the records need to be normalized

How can Shelley engage with institutions?
Suggestions:
   Write story for:
   Chronicle for Higher Ed
   Inside Higher Ed
   Scholarly Kitchen
   LSE Blog
   SHARE

Also suggestions on FAIR webinar ideas?
Suggestions:

Mark: What thoughts do you have regarding the numerous tenure policies where data management and/or curation does not count toward tenure? It is about changing culture there as well.
National Academies Design for Open Science Report
AGU Fall Meeting Session on Credit (Changing Culture)
Systems/tools should inform researchers they are getting credit (e.g. Make Data Count) More info on this at makedatacount.org

Scott: Berkeley Labs does Carpentries workshops, other labs, reached out to them?
Shelley: Doesn't have a labs list to reach out to

Juliane: Harvard Catalyst, how can we help the researchers, difficult to connect with the researchers (different initiatives at Harvard)
Can be really challenging getting institutions onboard

20-50 min
Community Updates
Attending:
Sarah Seymore, University of Oregon
Jonathan Cain, University of Oregon
Mark Laufersweiler / University of Oklahoma
Doug Joubert, NIH Library

Annika Rockenberger, National Library of Norway (maybe - it's pretty late in the day)
Peter Oxley, Weill Cornell Medicine (Samuel J. Wood Library)
Drew Heles, Johns Hopkins University
Daniel Bangert (Research Data Alliance Europe/University of Göttingen)
Chris Erdmann (The Carpentries/CDL)
Scott Peterson, UC Berkeley
Josh Quan, UC Berkeley
Anna Sackman, UC Berkeley
John Chodacki - CDL
Ariel Deardorff - UCSF
Shari Laster - ASU

50-60 min
Library Carpentry Updates

How can I implement a FAIR Data & Software workshop
TIB Hannover was a good example https://librarycarpentry.org/blog/2018/07/24/tib-hannover-fair-report/
(slides and material listed in the blog post)
Trouble is, it is 5 days long and more of a summer school at the moment
Would be great to test it out over a series of days
Waiting for the organizers to share their material via a Library Carpentry lesson which we can collaborate
on as a community
Australian libraries w/ ARDC have express interest in doing something like this
The workshops allow librarians that are not familiar with Carpentries tools but familiar with the content
of the slides to partner with Carpentries instructors and to potentially engage with the Carpentries
commmunity.

Blog post will be out shortly on What is a Library Carpentry workshop?
Answer: we value flexibility, so half day, one day, two day, multi day are fine
Report your workshop to https://software-carpentry.org/workshops/request/ and Chris
Overview of our core lessons:
https://librarycarpentry.github.io/lc-overview/

Curriculum Advisory Committee of 5 members has been selected to provide guidance to community on
lesson ideas
There have been many ideas: DH, data privacy, Python, Wikidata...

----
24 September @ 20:00 UTC
See time and date in your zone: https://www.timeanddate.com/worldclock/fixedtime.html?
msg=Library+Carpentry+Community+Call&iso=20180924T20&ah=1
We will be using Zoom for the call: https://carpentries.zoom.us/j/555458528

0-20 min
Tara Vancil (https://twitter.com/taravancil) In charge of words, product, website at Beaker, a browser for
the peer to peer web
See https://beakerbrowser.com/

Topic: Peer to peer web for librarians
Slides: carpentry.taravancil.com (dat:// or https://)
Install Beaker: https://beakerbrowser.com/install

Attending:

- Chris Erdmann, Library Carpentry Community and Development Director, @libcce
- Wladimir Labeikovsky, Bioinformationist; Univ. Colorado Health Sciences Library, @nucholab
- Eric Olson, Engagement and Partnerships Lead, ORCID, @olsonscholcomm
- Mark Laufersweiler / University of Oklahoma / @laufers
- Joshua Dull, Yale University Library
- Amy Hodge | Stanford University Libraries | amyhodge@stanford.edu
- Lauren Ko, University of North Texas
- Shari Laster, Arizona State University,
- Scott Peterson UC Berkeley
- Konrad Förstner (ZB MED, Germany) @konradfoerstner
- Erika Mias @erikamias

- Chrome has 60% of market share, has influence on how web will work
- Peer to peer protocol instead of https, send out request to whole network, anyone online will respond
- Your laptop can function as server, authoring tool
- If your laptop goes offline, if others have cached your site/file then you will still be able to access (archiving, scraping easier)
- dat is the name of the peer to peer protocol used (in Beaker)
- in Beaker browser, you can see how many people are viewing the website at that time in Browser (can see files)
- Create project folder, website from browser (demo site created)
- No server involved, you can help seed website, dat address (per Tara, like long IP address) can be shared to download files
- Editor is directly in the browser
- Fritter alternative to Twitter and files are completely open to review

? Sounds outstanding for finding "popular" files via peer to peer, but what about obscure content that few would be seeking (like specialized researchers)?
if researcher in lab that has important set of data to publish, has support of university to duplicate content on several devices?, publish from laptop, should be duplicating elsewhere
non profit dat working with labs, stanford is one of them, blog posts on website on what went well and what went wrong, data project dot org
webs of trust are going to be important in peer to peer web, social graph is powerful tool for defining authenticity, demonstrated in Fritter, who to follow uses web of trust
copyright will be a challenge in peer to peer web

datproject.org

Example of researcher using Beaker here at Univ Colorado Med School:
dat://adef21aa8bbac5e93b0c20a97c6f57f93150cf4e7f5eb1eb522eb88e682309bc/docs/Placeholder.html

Universities are in good place, computing resources to support researchers, students, librarians -

universities run seeding services

Cities, Belgium example, explore every citizen having website, city as friendly seeding service for city

Companies and institutions have caps on bandwidth and what does this mean for net neutrality
Remains to be seen

Storage is cheaper, bandwidth is more expensive, as data is more popular, cost is shared between visitors

I'm already having to use streisand/vpn just to use dat:// applications from Uni :\
Universities have locked down peer to peer activities, dat get better at net traversal, universities get more flexible

Q: Is Electron/Javascript an interim building material for dat infrastructure/applications?
Yes
Based on Google Chrome's browser engine

Dat in the Lab links:
   - http://uc3.cdlib.org/2018/09/19/dat-in-the-lab-announcing-the-dat-anacapa-container/
   - https://blog.datproject.org/2018/04/24/data-sharing-at-institutions-and-beyond-with-dat/

20-50 min
Community Updates
Attending:

50-60 min
Library Carpentry Updates
New website and becoming a lesson organization
Call for Maintainers
Consortium approach to Carpentries membership is proving to be popular (NESCLiC is a good example)
Recommend reading:
   https://librarycarpentry.org/blog/2018/08/22/library-carpentry-strategic-value/

----
1 October @ 20:00 UTC
See time and date in your zone: https://www.timeanddate.com/worldclock/fixedtime.html?
msg=Library+Carpentry+Community+Call&iso=20181001T20&ah=1
We will be using Zoom for the call: https://carpentries.zoom.us/j/712284220

0-20 min
Tim Head (https://twitter.com/betatim) Python and Stats Consultant, Building Binder
Topic: Binder, a tool to help researchers make their code immediately reproducible
See https://mybinder.org/

   • slides https://hackmd.io/p/BJDmTh0KQ#/
   • if you want to watch a (much) longer version of this https://www.youtube.com/watch?
     v=MhyzZAbOCGw&list=PL055Epbe6d5b572IRmYAHkUgcq3y6K3Ae&index=34&t=1950s
   • If you know mybinder.org already make a + on the next row, if you don't make a - on the row
     below
        • ++++++

- ------
- https://mybinder.org/v2/gh/betatim/openrefineder/master?urlpath=%2Fopenrefine
- Questions (list your questions below this bullet point). Maybe answering questions is the best use of our time.
    - Do you have a list of Frequently Asked Questions?
        - Yes see https://mybinder.readthedocs.io/en/latest/faq.html
    - How are researchers finding Jupyter/Binder instances? Via GitHub? Binder? (no information about how people are finding, see someone else use it and then use it)
    - Are there any societies/journals using Binder? Libraries? How can we (librarians) help you?
        - Need connections with journals (connect with shelley Stall/AGU)
        - Pangeo hub gets less traffic, pick that hub
        - Universities have resources to spare for researchers / people that use mybinder are people that need 1 or 2 CPUs where it is difficult to get at Uni / example of high schools that use mybinder, they have challenges getting access to compute
        - Libraries should host binder hubs
    - If there are multiple Binder hubs, are they connected to each other, or do you have to pick one?
    - Discuss what you need w/ regard to DOIs? Also archiving?
        - Need machine readable list of files with DOI, with every DOI get differences
    - Is there anyone in your group that wants to come to Jupyter in the Triangle? https://libcce.github.io/TriangleJupyter/
    - How does Binder compair to CodeOcean?
        - Can do everything in CodeOcean except for MatLab, user experience is better, difficult to buy MyBinder.org
    - Yay, OpenRefine in Binder! How did that work to set up?
        - psot link to github repo
        - https://github.com/binder-project
        - Could you give some other examples beyond Jupyter notebooks? Would you use this to set up an R Studio instance for example?
        - https://github.com/binder-examples/
    - What would the security implications be of opening up BinderHubs from our library to the world?
    - What are the differences/similarities between JupyterHub and BinderHub (overlay for JupyterHub)?
    - JupyterLab?
    - If one wished to run a no-install-needed workshop wherein all participants simply launched the needed environment in their browsers, how many participants could one realistically support? What would the bandwidth requirements be?
        - less then 100 users from GitHub repo (more we need to talk to them) best place to ask is our chat http://gitter.im/jupyterhub/binder
        - list all the tools in requirements file/anything from Ubuntu
        - Binder is about 18 months old now

Difficult for everyone in the world to build Docker containers (environments for running code & data), we need to make it easier
Example paper: https://arxiv.org/abs/1803.05268 -> how long would it take you to reproduce a figure in the paper -> includes GitHub repository "tbd-nets"
Seeing more GitHub repositories with "launch binder" badges -> message showing Docker loading

requirements.txt in Python (also environment.yml) or R has DESCRIPTION files for listing dependencies (Julia has a REQUIRE file)

Firewall rules in place to prevent bitcoin miners

3 public binder hubs (gesis, pangeo in addition to main one)

Looking for a way to be sustainable and keep Binder open!

There is a Binder for OpenRefine please share here!

20-50 min

Community Updates

Attending:

- Mark Laufersweiler / Univesirty of Oklahoma / @laufers
- Chris Erdmann / Carpentries & CDL / @libcce
- Ariel Deardorff/ UCSF Library/ @arieldeardorff
- Drew Heles / JHU Libraries
- Tim Head / Project Binder
- Tracy Teal / The Carpentries
- Nina Exner / Virginia Commonwealth University / @Z669_E9
- Scott Peterson / UC Berkeley
- Annelise
- Andrew Starodubtsev
- Peter Oxley
- Eric Olson
-

50-60 min

Library Carpentry Updates

----

5 November @ 21:00 UTC

See time and date in your sone: https://www.timeanddate.com/worldclock/fixedtime.html?msg=Library+Carpentry+Community+Call&iso=20181105T21&ah=1

We will be using Zoom for the call: https://carpentries.zoom.us/j/104441531

0-20 min

Konrad Förstner, Professor for Information Literacy at the Technical University of Cologne and Head of Information services at ZB MED - Information Center for Life Sciences (https://www.zbmed.de/en/)

Topic: Leveraging The Carpentries to grow library services

20-50 min

Community Updates

Attending:

Chris Erdmann/The Carpentries & CDL/@libcc

Ahmed Hemedan/ Bioinformatics PhD student at Luxembourg university/@ahmed7emedan

Greg Wilson

Tracy Teal

Daniel Bangert

Mark Laufersweiler

Paul Pival
...

Notes and recording:
https://github.com/LibraryCarpentry/governance/blob/master/community-calls/2018-11-06-Foerstner-ZB-MED.md

50-60 min
Library Carpentry Updates
LC is now an official Lesson Program
https://carpentries.org/blog/2018/11/welcoming-library-carpentry/

----
3 December 2018 @ 21:00 UTC
See time and date in your sone: https://www.timeanddate.com/worldclock/fixedtime.html?msg=Library+Carpentry+Community+Call&iso=20181203T21&ah=1
We will be using Zoom for the call: https://carpentries.zoom.us/j/132912782

0-20 min
Ed Summers, Software Developer, Maryland Institute for Technology in the Humanities and DocNow
Topic: Talk and tour of the Documenting the Now (https://www.docnow.io/) project and tools to support social media & web archiving efforts
Slides: https://news.docnow.io/documenting-the-now-ethics-white-paper-43477929ea3e
Recording of previous talk (similar to this one): https://inkdroid.org/2018/11/08/ferguson-principles/

Similar talk to talk at internet preservation conference / adapt for Carpentries community
Will demo the tools
Where archiving fits in with The Carpentries
By 2025 could use up to 20% of the world's electricity and additional projections
Article - Tsunami of data could consume one fifth of global electricity 2025
Andrers Andrea - Tsunami of Data
We can't archive everything
Internet archives the web, but how much of it is in it
130 trillian URLs in Internet Archive in 2016 (extremely low percentage)
1-5% of material from university gets archived (a sliver of a sliver)
TimeMapper: Black Lives Matter Twitter Collection 2014 (What would it take to collect the tweets in Ferguson?)
Used Twarc to collect 13M tweets from Ferguson, relayed what they were doing via a blog, continued to archive other events (https://github.com/DocNow/twarc)
4 years later and a Mellon grant to improve Warc, but create some new tools
% twarc search Ferguson > tweets.json (like searching Twitter UI, via API, export in JSON)
Moving window that you have to operate in, limited time to collect tweets, otherwise you have to pay for access
http://docnow.io/catalog - Jekyll site cataloging curated datesets (Padilla collection on Las Vegas mass shooting example) - datasets are not the original datasets they have ids not the raw data, part of Twitter agreement, they do release the ids of the tweets
Hydrate function https://github.com/DocNow/hydrator will retrieve JSON data for the tweets based on the ids collected
Desktop app called Hydrator, you can point at an id dataset, and then it will write out as JSON or as a

CSV

unshrtn tool (https://github.com/DocNow/unshrtn), a microservice where you can give URL and it can provide additional information about the URL

diff engine tool (born out of fake news/presidential election) researchers were interested in how news is shaped over time, point at RSS feed to watch the content and periodically check, will tweet if there are changes, create web front end/tracker, easier to use?

docnow app for researchers studying events in social media, query of twitter and tries to summarize, to help you iterate to find a query that you are happy with -> data that is collected (will hopefully help archivists communicate curated collections?) -> speak to needs of archivists

Slack channel great tool for collaboration in the docnow team, brought together archivists, researchers, technologists, developers...

Ferguson Principles - for archivists, documentation efforts, terms of service congruent with values (ethics), apply traditional practices where relevant,

Mellon grant - sustain current work, continue development of docnow tool, host series of activist workshops that use social media as part of their process/raising awareness, curriculum development (made available as open access after testing)

On the catalog page, many projects are welcome, for instance, Aretha Franklin dataset (if you want to add a dataset, you can edit a yaml file and PR or create an issue ticket but want to make it more accessible

docnow demo - interactive dashboard for browsing through tweet collections to create dataset collections

Hook up docnow functionality with webrecorder

Ethics and webscraping whether a curriculum would be of interest to The Carpentries community?

Nina Exner: Are the Ferguson Principles published, yes, but buried at the moment, docnow.io -> get involved -> media site -> ethics white paper has link to PDF with Principles

For people in their Slack channel, they run an instance that people can poke out, don't broadcast because it is in half functioning state, but plan to publish at some point

Interest in working on lesson development with DocNow: Shari, Yasmeen, Paul

Mark notes IRB and being careful with regard to involvement

We won't certify - legal reasons

Nina Exner:

   I do RCR (the ethics Mark mentioned) training, and sometimes they're looking for added content to enrich the curricula

   They favor case discussions

Mark: Supplimental material?

20-50 min
Community Updates
Attending:
   Paul Pival - University of Calgary
   Anton Angelo - University of Canterbury, New Zealand.
   Chris Erdmann / Carpentries & CDL
   Nina Exner, Virginia commonwealth university
   Laura Wrubel
   Mark Laufersweiler
   Annelise Sklar
   Lauren Ko - University of North Texas
   Shari Laster
   Iris Jiang
   aft23?
Tim Dennis - UCLA

Juliane Schneider - Harvard Catalyst
Scott Peterson, UC Berkeley

50-60 min
Library Carpentry Updates

----

28 January 2019 @ 21:00 UTC
See time and date in your zone:
https://www.timeanddate.com/worldclock/fixedtime.html?
msg=Library+Carpentry+Community+Call&iso=20190128T21&ah=1
We will be using Zoom for the call: https://carpentries.zoom.us/j/830088889

0-20 min
Susanna-Assunta Sansone (https://twitter.com/SusannaASansone), Associate Director, Associate
Professor (Oxford e-Research Centre, Department of Engineering Science), Peter McQuilton, Project
Coordinator (FAIRsharing.org, Oxford e-Research Centre, Department of Engineering Science)
Topic: FAIRsharing (https://fairsharing.org/) is a curated, informative and educational resource on data
and metadata standards, inter-related to databases and data policies; see pre-print "FAIRsharing, a
cohesive community approach to the growth in standards, repositories and policies"
https://doi.org/10.1101/245183

20-50 min
Community Updates
Attending:
    Chris Erdmann, The Carpentries/California Digital Library
    Sherry Lake, University of Virginia
    Mark Laufersweiler, University of Oklahoma
    Lora Leligdon, Dartmouth
    Scott Peterson, UC Berkeley
    Nina Exner, Virginia Com monwealth Univ.
    Tom Morrell, Caltech

Notes:
    Start with exploring FAIRsharing.org
    Guidance to consumers of standards, databases, and data policies
    A number of stakeholders are involved
    Standards split into 4 types: formats, terms, guidelines, identifiers (PLoS Thesaurus
https://doi.org/10.25504/FAIRsharing.2ndf9r vocab as example, INVIVO? as Guidelines example)
    What format does this resource accept - linking to other resources - triangle of relationships
    NIH - What format do I have to publish my data in, where, etc
    Status indicators - R is ready for use! Dev is for something in development (create record to tell people
that they are working on something) U is uncertain (stopped answering their calls, let them know if you
know more about a U resource) and D is for deprecated, moved somewhere else, no longer available
    Curator team but groups can claim page to verify accuracy - linking allow for notifications for curation
teams/owners

GBIF example of a single database record https://doi.org/10.25504/FAIRsharing.zv11j3 - DOI for record, homepage, where it is developed, when it was created, scope, recommended by other services or if it has been recommended by journals and/or published as in this case by PloS, Gigascience et al

Understand the people behind the resource as well

IVOA example of Collection (grouping standards and/or repositories relevant to a community/project) https://fairsharing.org/collection/IVOA - capture visaul map of all standards avaialble - 3rd party widget can be embedded in page

Scientific Data journal example of a Recommendation (grouping standards and/or reposiories that a data policy, by journals of funders, recommend) https://fairsharing.org/recommendation/ScientificData - filtering/sort by option and have visuals to help understand the resource use better

FAIRsharing has its own ontology terms to classify the records and enable searchers and filtering according to domain and discipline

Coverage started in the life science, but expanded since last year, now mainly from Chemistry and Physics, but other domains are growing such as Humanities

FAIRsharing enables findability of resources - uses schema.org for improving discovery (via Google for example)

Allow you to assess resources whether they are FAIR - Findable, Accessible, Interoperable, Reusable

Working with other services and tools as registry, lookup services via API, working with NIH, Go FAIR and others

Working under Force11 and RDA as a Working Group https://rd-alliance.org/group/fairsharing-registry-connecting-data-policies-standards-databases-wg/outcomes/fairsharing - registry and set of recommendations

Libraries can use FAIRsharing to browser for resources, use it to inform the creation of DMPs (data stewardship tools), formally cite resources via DOI (how to cite resources page)

Tell us what is missing (point them to the right people)

Create a FAIRsharing collection (like the one they created for Astronomy), to help users find what they are looking for

Has a team and advisory board

Find out more about FAIRsharing - read preprint

Overview of FAIRsharing, the communities and stakeholders involved: see pre-print "FAIRsharing, a cohesive community approach to the growth in standards, repositories and policies" https://doi.org/10.1101/245183

Questions:

What objects are in the database? Repositories, or individual datasets?

We are not talking about datasets, we are talking about the metadata around the resource.

Do they work with re3data.org?

Been talking with them for many years to collaborate. But have slightliy difference missions. Capture standards and repositories

Standards are hard to surface (IVOA - Astronomy example); communities tends to know what us used in their area but not necessarrly in others

For journal editors, for example, whose work may span over different disciplines, knowing what are the community norms (e.g. the standards used) is useful

How can that information be surfaced, for people to get at, to connect with people involved with the standards.

Working with IUPAC to surface chemistry standards, also working with Agriculture

Surfacing information across domains for multidisciplinary projects

Can they send us a standard email that we can send to resources to add their data (Humanities is an

example), help identify the people, or be the people
    Create an account with FAIRsharing, anything you add will be checked by curators
    contact@fairsharing.org
    API - doing it on a case by case basis as far as granting access via Memorandum of Understanding
    data is under CC-BY SA license https://fairsharing.org/licence
    Data stewardship wizard via Elixir - metadata in FAIRsharing will be powering their tool
    Form that you can fill in asking for access - keep track of how the data is being used -
https://www.fairsharing.org/API
    Academic project - running since 2011, embedded in several funded infrastructure and research
projects - intention to remain open
    Only place where you can find these kind of relationships between standards, repositories and data
policies (and that tracks the evolution and adoptions of these resources), uniquely placed w/ other
intiaitvies Go FAIR, RDA, CODATA, FORCE11
    Interested in expanding educational offering - have FAQ - https://www.fairsharing.org/educational
Help with more informative links/examples
Funding for hackathons
Mozilla sprints and FAIR sprint with FAIRsharing:
https://librarycarpentry.org/Top-10-FAIR/

50-60 min
Library Carpentry Updates
https://twitter.com/libcce/status/1088871022357553153
Hex stickers

Flexibility valued in Library Carpentry
Including workshops in upcoming workshops page
How do we get feedback in from self organized workshops -what are the feedback mechanisms?
Follow guidelines - then fine
Understanding our learners - thinking of our curriculum
Libs limited in time, likely to go to something that is spread out
Call CAC together to hash out roadmap for new curriculum
Start looking at additional core lessons for LC
Easier to split up workshops over course of weeks/months
Library Carpentries for NorCal, host at CDL
Oklahoma Summer Library Carpentry, 3rd day tech day
OER Library Carpentry lesson discussion
Some elements taught at Oklahoma
Author Carpentry, will it become part of The Carpentries? Library Carpentry should talk to Author
Carpentry, see if there is a way to work together.
We need to have a discussion on OER - guest speaker???

----
25 February 2019 @ 21:00 UTC
See time and date in your zone:
https://www.timeanddate.com/worldclock/fixedtime.html?
msg=Library+Carpentry+Community+Call&iso=20190225T21&ah=1
We will be using Zoom for the call: https://carpentries.zoom.us/j/612323803

0-20 min

Maria Gould (https://twitter.com/uc3cdl), UC3 Product Manager (California Digital Library, University of California Curation Center)

Topic: Research Organization Registry Community (ROR) (https://www.ror.org) is a community-led project to develop an open, sustainable, usable, and unique identifier for every research organization in the world.

See time and date in your zone.

Slides: https://docs.google.com/presentation/d/1vWXzdlK0zZlQMDEERb6nItLkZLa7H77j4ASh6B-6g9Y/edit#slide=id.g4e284dda59_0_12

20-50 min

Community Updates

16 Attending

Notes:

    Maria at UC Berkeley before working in scholarly communications

    Now at California Digital Library & ROR

    How many of you work with affiliation data?

- Work with IR, interested in org info, for impact related questions? ROI on research funding?
    - Interested in PID for affiliations, researchers that associate with different orgs, sometimes have 5 profiles? Can be difficult to get research metrics as a result.
    - I support researchers with research data management and I am interested to hear how PIDs could be implemented for institutions (or instruments/labs).
    - I work with our data repository, and am interested in learning about affiliation data to incorporate into this.

    ROR will be CC0, institutions/libraries shouldn't have to pay for access, do this in a community orientated way

    Kick off meeting in Dublin at PIDapalooza, initial MVR/demo

And if the meta data in ROR is not correct, easy to get corrected?

What policies and procedures should we have for metadata curation? ROR is tied to GRID at the moment, next phase is to get input from community and get policies and procedures in place so that community can make changes. Email ROR/Maria to make changes or visit GitHub org/repo for ROR to submit an issue to make a change.

University of Oklahoma is not UO but is OUAnd not sure about the foreign language is listed.

In searching for my institution I see a listing for the institution itself and one for our medical school. Are there plans to create IDs for each college (the business school, law school etc.) or is the medical school as far as you plan to go?

GRID considered medical institutions as separate from parent insitution

If have feedback on OpenRefine reconciliation service, please let them know

Like to see ROR ids implemented in data publication process - Dryad use case

Forming a community advisory group, interested in joining, contact Maria/ROR

I'd love to see an unique ID for the research centers at my institution as they really need to demonstrate impact, but that sounds very much out of scope for ROR.

How does this differ from the ringgold id?

One difference w/ Ringgold is that it's commercial

What kind of response have you had from publishers?

ROR infrastructure & data are intended to be community-owned & open

It would be useful for organizational IDs to have trees, as it were. Like DOI handles. Where centers can have an ID that's a sub-ID of the larger university or organization.

In its current incarnation does the ROR ID allow for full institution name as well as a preferred short form? E.g., University of Pennsylvania vs Penn; University of California, Berkeley vs Cal

Thank you Maria for this informative talk! Will head off to bed now.

Might also be helpful to include labels for the organization name used in the major rankings. E.g., does US News & World Report use the same form as Times Higher Ed rankings for The Ohio State University vs Ohio State University

I like Nina's (and Jake's) idea. Assuming the institution would need to create and sustain the "trees" registration with ROR.

I really like the open, non-commercial aspect of ROR and I'm curious about your funding model. What sustainability plans are there for ROR?

Also form of name used by major grant organizations, such as NIH: The Scripps Research Institute vs Scripps Research vs TSRI vs Scripps Florida

Build in parent-child, relationships (basic relationality)

Interest from not only publishers, but also different perspectives, ORCID will integrate into their metadata, some publishers that are talking to ROR, interested in joining the community advisory group

Trying to be as inclusive as possible

It's possible to add variants like CDL for California Digital Library but short forms for institutions like the Penn example above

So checking if the institution change is appropriate? Or is this what you will be developing?

Clarivate or Scopus integration on the horizon, or publishers (or ORCID) first?

Want to be integrated as broadly as possible, talking to Clarivate, yes

Will the data in the ROR record be presented as authoritative or is the plan to have something similar to "Twitter verified"?

Will the data in the ROR record be presented as authoritative or is the plan to have something similar to "Twitter verified"?Meaning a badge that shows the record has been approved by the organization

Like this idea, who is authorized to make a change on behalf of the org

Should be a place where these research centers exist in ROR, but haven't done this, put in fork to start building out the ids, haven't created policies to wait for community feedback

At a previous institution, we had one person responsible for ORCID that then left and we couldn't figure out who had access/permissionredundancy will be helpful and having multiple contacts verified

50-60 min
Library Carpentry Updates
Top 10 FAIR material
Plug in to Moodle or similar?
Not enough people using, more people wanting to develop, everyone wants to have their own specialty
I'm interested in a "modularized" approach, spread over a term. OpenRefine has been well-recieved, because it's practical, saves researcher time, etc.
What do librarians want to learn vs our patrons
Sprint was outreach and not prescriptive, people felt can play around
Deliberate focus on Things, or they hang out, hope
Interested in instruction (librarians do this, part of libraries offering?)
librarians come to library carpentry to understand their patrons better
RDAP May 16-17 Mozsprint after?
Anton thinking about creating virtualbox, will Slack

----

6 March 2019 @ 21:00 UTC
See time and date in your zone:
https://www.timeanddate.com/worldclock/fixedtime.html?
msg=Library+Carpentry+Community+Call&iso=20190306T21&ah=1
We will be using Zoom for the call: https://carpentries.zoom.us/j/678945076

0-20 min
Carli Spina (http://carlispina.com/about.html), Associate Professor and the Head of Research &
Instructional Services at the Fashion Institute of Technology
Topic: This session introduces Universal Design for Learning and provides advice on how to teach in a
manner that is accessible and inclusive of those with disabilities and who use assistive devices.
Slides: https://docs.google.com/presentation/d/1oKCm7U8mbTvCW1l8-OtIcfo1BGLcEtdt8IJ6-
oGCh3Y/edit?ts=5c7e7511

20-50 min
Community Updates
Attending:
Count: 9
Kari L. Jordan
Chris Erdmann
Nina Exner, VCU
Mark Laufersweiler, OU, @laufers
Anna Oates

Notes:
    Chairlift example - separates people, only possible to use by someone in a wheelchair and not by others
with mobility challenges or even other groups
    "Universal design is the design of products and environments to be usable by all people, to the greatest
extent possible, without the need for adaptation or specialized design."
- Ronald L. Mace
"Universal Design assumes that the range of human ability is ordinary, not special."   - Elaine
Ostroff, 2001
More inclusive approach
"A framework for designing educational environments and products…that are accessible and optimal for
all learners."
      - Prof. David Rose, 2017

   • Offer multiple ways to recruit interest
         • Allow for choice
         • Authentically explain the relevance and value
         • Offer a distraction-free, non-threatening space

Allow for customization for different learning styles
Learners that have to deal with anxieties
Explain the value of learning this exercise - a different level of engagement
Build in the pathways to demonstrate what learners will get out of the training
Exercises at different levels of difficult to help people feel like they achieved something, also giving them
a chance to be included in process
Avoid mastery in exercises "red pen"
Group work

Personalized goals

Scaffolding and coping strategies after the training

If they set goals at the beginning then reflection after

Provide alternatives for people who have different ways of perceiving material

Some people would rather read or listen - multiple options are helpful

Mathematical symbols/CS can be intimidating - jargon busting is useful, language, and multimedia

Example from Shell lesson - lots of code without explanation, looks foreign, include more description

- Provide alternatives for perception
- Provide alternatives for language, symbols, and mathematical expressions
- Provide support for comprehension
    - Offer background info
    - Emphasize key ideas and patterns
    - Guide information processing
    - Explain how to transfer knowledge to other areas and generalize it

Learning objectives for referring back to later

Context, how can I use this in the real world? How can you transfer the knowlege?

Can you allow people to visualize their learning like drawing

Is there any software that will be problematic for assistive devices - text to speech tools

Allow for responses in different modes - exercises

Paper vs electronically

Carpentries offers interactive approaches - different ways of working and conveying what you have learned

Tracking achievements and goals

Sticky notes but other alternatives saying they are falling behind - are there ways to offer anonymous ways to say if they need help - alleviate the social anxiety

Jargon Busting screenshot - explaining complex terminology which is good - options to focus on terms they have had trouble with (great way to recruit interest) - supporting the collaboration piece (notes or not) - group can select who can report back (alleviate stress) - scaffolding learning techniques - higlighting key points

What else can we do?

Questions/Comments:

   How can we develop exercises that have more explanation?

   Exercises that had link to read more on later

   Have follow up links in Etherpad to share later

   Carpentries doesn't exactly have a lesson open, concentrate on tool, links can be distracting during the exercise

   No answer - hint - support materials for more info then see the answer

   If can't answer all the topics listed in jargon busting - scaffolding learning - so apply this after the training - how they might get the answers later

   Adding more diagrams or pictures to lessons - mental rabbit hole - picture needs a caption or description to describe - are they adding for some people?

   Picture is intended to reproduce what the text is saying - different representation of the same information

   Look at extras - reference- discussion - feedback on how lessons have been taught - extra section on all the figures that have been used - another avenue, not everything has to be answered in the lesson - color blind stickies or different shaped stickies

   Staples has different shapes of stickies

Finite amount of time - what can we do regarding the goal piece - people come together initially to talk about their goals to develop them - ask people to set goals before lesson and submit them to you - think about it before they come - instructors can see what their goals are - understanding individual goals in the beginning - what is one thing you look forward to, what is one thing that intimidates you - allow people to reflect

50-60 min
Library Carpentry Updates

----
29 April 2019 @ 20:00 UTC
See time and date in your zone:
https://www.timeanddate.com/worldclock/fixedtime.html?
msg=Library+Carpentry+Community+Call&iso=20190429T20&ah=1
We will be using Zoom for the call: https://ucop.zoom.us/my/cerdmann

0-40 min
Aleksandra Pawlik (https://twitter.com/aleksandrana), Community Advocate, Stencila
Topic: Stencila (https://stenci.la/) is a set of open source tools supporting research reproducibility and interoperability. During the talk Aleksandra will demonstrate how all the components fit together and how to integrate them within the existing ecosystem of tools, such as spreadsheets, R, RStudio, Jupyter Notebook and more.

Attending:
　　Paul Pival, University of Calgary, ppival@ucalgary.ca
　　Ana Van Gulick, Carnegie Mellon University, anavangulick@cmu.edu
　　Ariel Deardorff, UCSF Library, ariel.deardorff@ucsf.edu
　　Jan Dvorak, Charles University Prague, jan.dvorak@ff.cuni.cz
　　Debbie Martindale, Rhodes University, d.martindale@ru.ac.za
　　Marié Roux, Stellenbosch University, mr@sun.ac.za
　　Colleen Cuddy, Stanford University, ccuddy@stanford.edu
　　Erin Carrillo, Virginia Commonwealth University, emcarrillo@vcu.edu
　　Gail Clement, Caltech Library, gperetsm@caltech
　　Aleksandra Pawlik, Stencila, aleksandra@stenci.la
　　Chris Erdmann, CDL, Christopher.Erdmann@ucop.edu
　　Mark Laufersweiler, Univ Oklahoma, laufers at ou.edu
　Anneliese Taylor, UCSF LIbrary, anneliese.taylor@ucsf.edu
Nokome Bentley, Stencila, nokome@stenci.la

At most, 10 people were on the call

Questions:
　　Do you have any Stencila examples to see in the wild - being used by researchers? For instance, in fisheries? And are journals accepting Stencila living documents? +1

　　　• See the examples repo e.g. https://github.com/stencila/examples/tree/master/elife-30274

　　Is there a Stencila Binder we can run in the cloud? https://stenci.la/blog/2018-11-20-stencila-binder/
　　How does Stencila welcome the different user communities at different stages from the coder to the

beginner? YES! :-) Always and forever! We are open and friendly.

Do you get some support (resources) via the eLife collaboration? We have very close collaboration with eLife and we should be soon able to announce some great news.

https://elifesciences.org/labs/ad58f08d/introducing-elife-s-first-computationally-reproducible-article

Google has been a big promoter of JSON schema for search but also for structured data, is there something on the horizon for Google to run interactive/living docs? That's the idea for our Plugin - Coming Soon (TM) :-)

Is there anything in the wild that is using CodeMeta?

Sustainability model? Possibly services around the software but Stencila will be always open source and freely available.

This is different from Code Ocean how?

If you would like to become a beta tester, please register your interest
https://docs.google.com/forms/d/e/1FAIpQLScmwS5YNeTv_zfgnhLjtuxgIwr6hM7J5aA7037zFTbp9gzclA/viewform

Stencila links
https://stenci.la/
Gitter Chat Channel https://gitter.im/stencila/stencila - come and talk to us, ask for help
Twitter   @stencila
Email hello@stenci.la

The slides are available here
https://docs.google.com/presentation/d/1_RFXN8oJ8PLe_CYicDeMYghU5XgEagmKL9FJb_nWgy0/edit#slide=id.g58eb834f99_0_40

Notes:
3 staff based in New Zealand, 1 person in Seattle, 1 person in Montreal (fully remote)

Open source project - expanding and welcoming new ideas

Integrations with Binder several months ago

Living documents - origins from Bentley, background as fisheries scientist, started off in spreadsheets, then became skilled R user, these were living documents for him, that wasn't working for people he was collaborating with, they were working with Word et al which caused some friction (copy and pasting) not helping with reproducibility and not making the process transparent, not interoperable, R markdown not necessarily friendly - goal of WSIWYG - meet inbetween R markdown vs Word

Received 3 yrs ago a Sloan grant and received 2nd grant but need to become sustainable

Looks similar to a Jupyter Notebook - aim is not to replace the notebook rather to become interoperable with it

Use the power of code and retain it in spreadsheets - also support a spreadsheet environment - a lot of researchers in fisheries work used spreadsheets

Use R code in spreadsheet to produce plot - can see the code running in the background - Stencila saves in its own format

Discussions about what their goals are, one being reproducibility and there are a number of tools that are doing this well, they want to compliment the existing ecosystem, history of Stencila tools helping with collboration

Not valuing one user/research group over the other, make it easier for groups to collaborate, to transfer to advanced coding, versioning, etc

They very much see a spectrum of users

Whatever tool you became used to using, you don't want to give up, Stencila wants to be complimentary

One of their goals - enablers of collaboration

Do a lot of collaboration with eLife - support directly in Stencila JATS format (the underlying XML/metadata that your template gets translated into for the publishers to use)

Fitting into ecosystem, Excel, Word, R, Jupyter, Python, Latex, Google Docs, Sheets

Format ignostic - how they will acheive this is via Stencila Schema (JSON Schema)

Each document can be represented by a tree of JSON nodes

Not introducing their own format, reusing existing formats, existing schemas

Using CodeMeta as well

See Stencila github convert/ and schema/ and hub/ (Google docs and sheets will allow you to have interactive code inside these programs)

gitter stencila channel - open for everyone to come and ask for help

demo of Stencila https://hub.stenci.la/beta/?next=/me/ - like GitHub w/ Projects

Can collaborate even on underlying JATS...

You can have different languages in the same documents

University in New Zealand using for data ethics course

The hub is not just a place where you can run your code et al in the cloud, it is also about the integrators

Stencila is open and allows for granularity, citing software, data, etc while Code Ocean is more at the environment/blob level

Contract Stencila to set up local environments - this is maybe similar to Binder?

There is a Stencila Binder where you can run it with examples

Examples: have an examples repository - recent example with eLife

Will be getting some support from eLife

Aware of some repositories that use CodeMeta but uptake hasn't been that great

Using schema to make research docs more disoverable

40-60 min
Library Carpentry Updates
2-day Library Carpentry at Oklahoma over the summer?
Data Fridays - breakouts of the individual lessons
https://osf.io/ymzu8/