

Interdisciplinary Tools to Support Your Research Across the Data Lifecycle

Nickoal Eichmann-Kalwara (she/her)

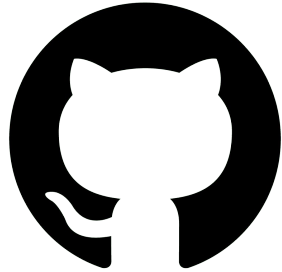
Associate Professor, Digital Scholarship Librarian

Center for Research Data and Digital Scholarship, CU Boulder

nickoal.eichmann@colorado.edu

bit.ly/nickoal-appt

This week, you're learning about...

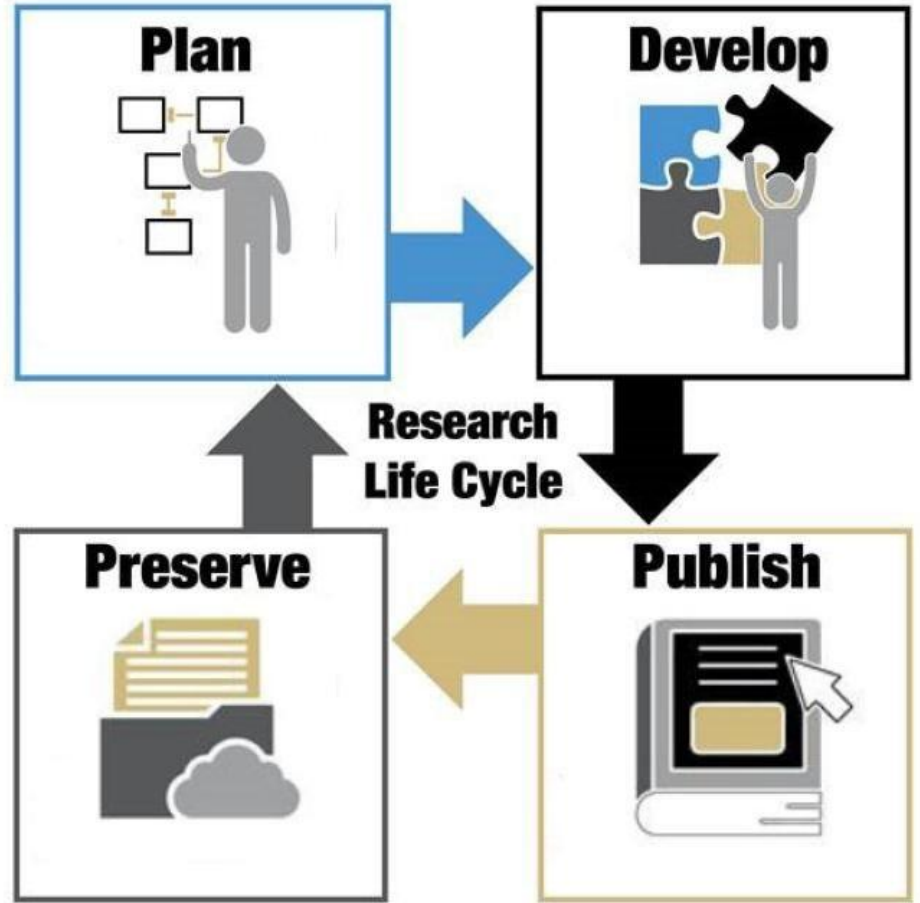


GitHub



python™

The Research / Data Lifecycle



zotero

Citation management software - Free

<https://www.zotero.org/>

My Library

- Book Reviews
- Colonial Medicine**
- Dissertation
- Science and Empire
- Teaching
 - Mapping
 - Open Access
 - Text Mining
 - Visualization
- My Publications
- Duplicate Items
- Trash

Group Libraries

- Grant Proposal
- Research Lab
- Topic Modeling

To Read 19th century Acclimatization Age Appetite Blood Cemetery Children Climate Colonies Competition Creoles Crossing Degeneration Diet Digestion Disease Doctors Drugs Electric Eels Empiricism Expertise Food France Geography Global Guyane Hair Indies Indigenous medicine Intemperance Language Lemonade Medicine Mortality Piment Poison Practice Professionalism Regeneration Secrets

Title	Creator	Year
▶ Guerre, maladie, empire. Les services de santé militaires en ...	Zaugg	2016
▶ Officiers de santé et soignantes créoles face à la fièvre jaune	Nobi	2016
▶ The Emergence of Tropical Medicine in France	Osborne	2014
▶ Colonial Disease, Translation, and Enlightenment: Franco-Briti...	Charters	2014
▶ Trading in Drugs through Philadelphia in the Eighteenth Centu...	Wilson	2013
▶ The Medicines Trade in the Portuguese Atlantic World: Acquisi...	Walker	2013
▶ Leprosy and Slavery in Suriname: Godfried Schilling and the Fr...	Snelders	2013
▶ Medical Experimentation and Race in the Eighteenth-century ...	Schiebinger	2013
▶ The Circulation of Bodily Knowledge in the Seventeenth-centu...	Gómez	2013
▶ Circulation of Medicine in the Early Modern Atlantic World	Cook and Walker	2013
▶ Synthesis of scholarship on "medicines" to restore focus o...		
▶ Full Text PDF		
▶ Colonial Medical Encounters in the Nineteenth Century: The Fr...	Thoral	2012
▶ Networks in Tropical Medicine: Internationalism, Colonialism, a...	Neill	2012
▶ Early Clinical Features of Dengue Virus Infection in Nicaraguan...	Biswas et al.	2012
▶ Medicine in an age of commerce and empire: Britain and its tr...	Harrison	2010
▶ Finding the "Ideal Diet": Nutrition, Culture, and Dietary Practi...	Neill	2009
▶ Battles of the Self: War and Subjectivity in Early Modern France	Pichichero	2008
▶ The Experiments of Ramón M. Termeyer SJ on the Electric Eel ...	de Asúa	2008
▶ Psychiatry and Empire	Mahone and Vaughan	2007
▶ Medicine and the Market in England and Its Colonies, C.1450-...	Jenner and Wallis	2007
▶ Matters of exchange: commerce, medicine, and science in the...	Cook	2007
▶ A Horrible Tragedy in the French Atlantic	Rothschild	2006
▶ "Neither of meate nor drinke, but what the Doctor alloweth": ...	Chakrabarti	2006
▶ Transnationalism in the colonies: Cooperation, rivalry, and rac...	Neill	2005
▶ Variolation, Vaccination and Popular Resistance in Early Coloni...	Brimnes	2004
▶ "Syphilis, Opiomania, and Pederasty": Colonial Constructions ...	Proschan	2003
▶ Choosing Scientific Patrimony: Sir Ronald Ross, Alphonse Lav...	Guillemin	2002
▶ Madness and Colonization: Psychiatry in the British and Frenc...	Keller	2001
▶ The Colonial Machine: French Science and Colonization in the ...	McClellan and Rego...	2000
▶ From medical astrology to medical astronomy: sol-lunar and pl...	Harrison	2000
▶ Disease and Empire: The Health of European Troops in the Co...	Bynum	2000
▶ Climate & Constitutions: Health, Race, Environment and Bri...	Harrison	1999

Info Notes Tags Related

Item Type Journal Article

Title Circulation of Medicine in the Early Modern Atlantic World

▼ Author Cook, Harold J.

▼ Author Walker, Timothy D.

Abstract The search for powerful drugs has caused people and commodities to move around the globe for many centuries, as it still does...

Publication Social History of Medicine

Volume 26

Issue 3

Pages 337-351

Date 2013/08/01 y m d

Series Atlantic World

Series Title

Series Text

Journal Abbr Soc Hist Med

Language en

DOI 10.1093/shm/hkt013

ISSN 0951-631X

Short Title

URL https://academic.oup.com/shm/article/26/3...

Accessed 1/24/2018, 10:17:12 AM

Archive

Loc. in Archive

Library Catalog

Call Number

Rights

Extra

Date Added 1/24/2018, 10:17:12 AM

Modified 1/24/2018, 11:50:15 AM



OpenRefine

Data cleaning, reconciliation - Free, open source

<https://openrefine.org/>

Cluster & edit column "Place"

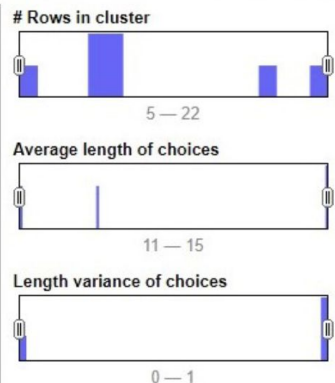
This feature helps you find groups of different cell values that might be alternative representations of the same thing. For example, the two strings "New York" and "new york" are very likely to refer to the same concept and just have capitalization differences, and "Gödel" and "Godel" probably refer to the same person. Find out more...

Method

Keying Function

7 clusters found

Cluster size	Row Count	Values in cluster	Merge?	New cell value
2	10	<ul style="list-style-type: none">PRINCESS FIELD (6 rows)PRINCESS FIELD (4 rows)	<input type="checkbox"/>	<input type="text" value="PRINCESS FIELD"/>
2	22	<ul style="list-style-type: none">MOUNT ZION (18 rows)MOUNT ZION (4 rows)	<input type="checkbox"/>	<input type="text" value="MOUNT ZION"/>
2	10	<ul style="list-style-type: none">SWABYS HOPE (5 rows)SWABYS HOPE (5 rows)	<input type="checkbox"/>	<input type="text" value="SWABYS HOPE"/>
2	9	<ul style="list-style-type: none">BROGUE HILL (5 rows)BROGUE HILL (4 rows)	<input type="checkbox"/>	<input type="text" value="BROGUE HILL"/>
2	9	<ul style="list-style-type: none">BALLARDS RIVER (7 rows)BALLARDS RIVER (2 rows)	<input type="checkbox"/>	<input type="text" value="BALLARDS RIVER"/>
2	5	<ul style="list-style-type: none">SPRING MOUNTAIN (3 rows)MOUNTAIN SPRING (2 rows)	<input type="checkbox"/>	<input type="text" value="SPRING MOUNTAIN"/>
2	19	<ul style="list-style-type: none">GRAVEL HILL (17 rows)GRAVEL HILL (2 rows)	<input type="checkbox"/>	<input type="text" value="GRAVEL HILL"/>



Select all Deselect all

Export clusters

Merge selected & re-cluster

Merge selected & Close

Close

RAWGraphs

Data Visualization - Free, open source
rawgraphs.io

RAWGraphs 2.0 beta

1. Load your data

- Paste your data
- Upload your data
- Try our data samples
- SPARQL query
- From URL
- Open your project

Copy and paste your data from other applications or websites. You can use tabular (TSV, CSV, DSV) or JS

TAGUETTE



FEEDING YOUR QUALITATIVE NEEDS

Qualitative Data Analysis (QDA) - free

<https://www.taguette.org/>

Project info | Documents | Highlights

Backlight

Recent studies demonstrated that the reproducibility of previously published computational experiments is inadequate. Many of these published computational experiments are not reproducible, because they never recorded or preserved their computational environment. This environment consists of artifacts such as packages installed in the language, libraries installed on the host system, file names, and directory hierarchy. Researchers have created reproducibility tools to help mitigate this problem, but they do nothing for the experiments that already exist in online repositories. This situation is not improving, as researchers continue to publish results every year without using reproducibility tools, likely due to benign neglect as it is common to believe publishing the code and data is sufficient for reproducibility. To clarify the gap between what existing reproducibility tools are capable of and this issue with published experiments, we define a framework to distinguish between actions taken by a researcher to facilitate reproducibility in the presence of a computational environment and actions taken by a researcher to enable reproduction of an experiment when that environment has been lost. The difference between these approaches in reproducibility lies in the availability of a computational environment. Researchers that provide access to the original computational environment perform proactive reproducibility, while those who do not enable only retroactive reproducibility. We present Reproducibility as a Service (RaaS), which is, to our knowledge, the first reproducibility tool explicitly designed to facilitate retroactive reproducibility. We demonstrate how RaaS can fix many of the common errors found in R scripts on Harvard's Dataverse and preserve the recreated computational environment.

iii

Lay Summary

One of the pillars of modern science is computation. Research software allows scientists to quickly and accurately analyze large amounts of data. When scientists publish their results, it is critical that their peers can then repeat their computations.

Unfortunately, researchers have discovered they cannot reproduce the findings of openly published analyses. This failure is not necessarily due to an incorrect experiment or flawed science. Instead, the software that scientists use to analyze their data fails if the original authors do not take careful steps ahead of time to preserve it. Researchers created new technologies to help facilitate this preservation, but they are not (yet) commonly used. We created a new tool, Reproducibility as a Service, that helps facilitate reproducibility for publicly available analyses that lack this preservation.



dedoose

Great Research Made Easy

Qualitative Data Analysis (QDA) - proprietary (\$)

<https://www.dedoose.com/>

The screenshot displays the dedoose web application interface. The main content area is divided into several sections:

- Project Overview:** Shows project details like 'Description', 'Page', and 'Tag'.
- Data Table:** A table with columns for 'File', 'Added', 'Status', and 'File'. It lists several files with their respective dates and statuses.
- Analysis Tools:** Includes a 'Scores by Description' section with a bar chart showing the distribution of scores for different languages (Bilingual, English, Spanish).
- Codebook:** A section for defining and managing codes.

A circular callout highlights a list of categories in the left sidebar:

- Similarity to Parent Child Reading
- Parent Education History
- Pre-Writing Activities
- Reading by Mother
 - Reading Routine
 - Reading Routine
 - Morning Reading Routine
 - Bedtime Reading Routine
 - Reading Frequency
 - Reading Duration

Mapping



- QGIS: Free and Open Source
- <https://qgis.org/>



ArcGIS

- ArcGIS: Proprietary / institutional accounts
- CSU: <https://gis.colostate.edu/resources/esri-software/>
- CUB: <https://libguides.colorado.edu/gis/access>

ScrollyTeller

Generate scroll-driven stories with Google Docs



ArcGIS StoryMaps

Scrollytelling

- Free and Open Source
 - <https://github.com/Davidson-Library/scrolly>
 - Based on Google Doc template
 - **Example:** [Growing with the City](#)
-
- Proprietary / institutional accounts
 - <https://storymaps.arcgis.com>
 - **Examples:**
 - [Biodiversity of the Colorado State University Lands](#)
 - [Chicanx Movement & Los Seis de Boulder](#)



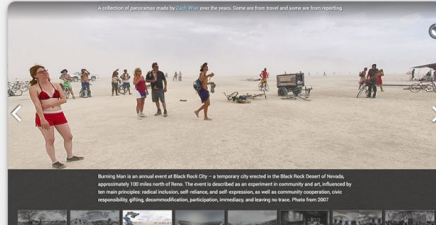
Knight Lab Tools

Digital storytelling, digital humanities - free
<https://knightlab.northwestern.edu/>



Juxtapose^{JS}

Easy-to-make frame comparisons.



Scene^{VR}

Easy-to-make VR stories.

strange thumping rhythm on the bass string. ▶ “If I
I get killed, please don’t bury my soul.” There’s a blu
l,” with 16-bar, four-line stanzas, that begins by repe
ies, ▶ “My mother told me just before she died,” AA
aning the words, each time with achingly subtle
notes blue enough to flirt with tonal chaos. Generatio
rough “Motherless Child,” field melodies and work sc

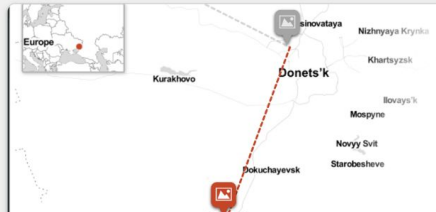
Soundcite^{JS}

Seamless inline audio.



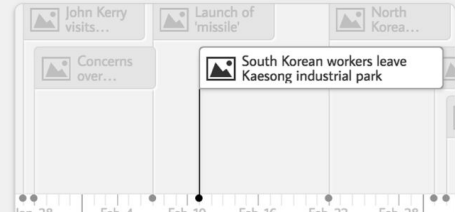
Storyline^{JS}

Tell the story behind the numbers.



StoryMap^{JS}

Maps that tell stories.



Timeline^{JS}

Easy-to-make, beautiful timelines.



CollectionBuilder

Digital exhibits and archives - free, open source

<https://collectionbuilder.github.io/>



CollectionBuilder Examples

Digital Collection and Exhibit Websites Built with CollectionBuilder



Oral History (as) Data

Oral History archives - free, open source

<https://oralhistoryasdata.github.io/>



Oral History (as) Data

Analyze and publish coded oral history and qualitative interviews

Learn more

Browse Transcripts

Browse by Subject



Interdisciplinary Tools to Support Your Research Across the Data Lifecycle

Nickoal Eichmann-Kalwara (she/her)
Associate Professor, Digital Scholarship Librarian
Center for Research Data and Digital Scholarship, CU Boulder

nickoal.eichmann@colorado.edu

bit.ly/nickoal-appt

THANKS!