

2021-01-14

Introduction to OpenRefine

Carruthers, Matthew

<https://hdl.handle.net/2027.42/165329>

<http://creativecommons.org/publicdomain/zero/1.0/>

Downloaded from Deep Blue, University of Michigan's institutional repository

CLEANING AND PROCESSING DATA WITH OPENREFINE



2020-02-18

Matt Carruthers

mcarruth@umich.edu

CAVEAT

There is a lot of stuff we won't be able to cover.



WHAT IS OPENREFINE?

- “A free, open source, power tool for working with messy data.”
- Better for data processing than programs like Excel
- A little more accessible than languages like R or Python



WHAT CAN OPENREFINE DO?

- OpenRefine can help with:
 - Originally designed for spreadsheets / tabular data
 - But the list of formats is expanding (and now includes MARC!)
 - Resolving inconsistencies in formatting, content, or terminology
 - Getting an overview of large data sets
 - Splitting or joining data
 - Reconciling local data with other data sets
 - Enhancing local data with data from other sources



OPENREFINE: THE BASICS

- <http://openrefine.org/download.html>
- Can run on any platform (Mac, Windows, Linux)
- Program runs locally on your machine, but the interface is through a web browser
 - Recommended: Chrome, Firefox, Safari
 - Don't use Internet Explorer or Edge. OpenRefine doesn't like Microsoft browsers.





DEMO

GETTING HELP WITH OPENREFINE

- Github wiki (can be difficult to navigate)
 - <https://github.com/OpenRefine/OpenRefine/wiki>
 - <https://github.com/OpenRefine/OpenRefine/wiki/Documentation-For-Users>
 - <https://github.com/OpenRefine/OpenRefine/wiki/Recipes>
 - <https://github.com/OpenRefine/OpenRefine/wiki/External-Resources>
 - <https://github.com/OpenRefine/OpenRefine/wiki/GREL-Functions>
- <https://datacarpentry.org/openrefine-socialsci/>
- <https://programminghistorian.org/en/lessons/cleaning-data-with-openrefine>
- <https://programminghistorian.org/en/lessons/fetch-and-parse-data-with-openrefine>
- <http://kb.refinepro.com/>
- <http://groups.google.com/d/forum/openrefine>



QUESTIONS?

mcarruth@umich.edu

