CUL MWG Data Scripting Bootcamp

The Cornell University Library Metadata Working Group is hosting a Data Scripting Bootcamp on February 20th, from 10:00 AM-4:00 PM with hour for lunch. This will be in Mann Library B30A, and we request that participants register here.

Note: this Bootcamp is meant to cover a number of (meta)data scripting topics in a short amount of time, with the focus being on learning through practical use. You won't be an expert in these areas by the end of this bootcamp; you won't even be entirely comfortable with many of these concepts at the end of the bootcamp, and this day won't cover everything.

Instead, this bootcamp is meant to be a way for you to jump in with examples of metadata work you could script in hand. It hopes to give some preliminary preparation for working with metadata tools and scripts that we dive into further detail in other MWG workshops this Spring.

As such, we strongly recommend you register here and provide use with examples of workflows, use cases, or data you'd like to work with using the skills in this workshop. We will try to work your examples into this day!

Proposed Outcomes for Participants:

- Know basics of data structures with particular relevance for library metadata work at Cornell;
- Know basics of working in a Command Line Interface with bash/shell and examples of working with metadata in this environment;
- Know basics of SQL, with examples and use cases for metadata queries and updates in particular;
- Be able to run simple scripts and programs, understanding how this works with relevance for metadata work;
- Know basics of Git and version control, be able to create a repo, push a change, check version of a dataset, etc.

Gratitude & License

This workshop materials reuses open curricula from the Library Carpentry project. We have updated lessons with specific examples for the Cornell University Library Metadata Working Group examples and data needs. We are not endorsed by Software, Data, or Library Carpentry.

We've also added curriculum materials influenced by similar introduction to scripting workshops at Library Technology conferences, so we would like to give a special thanks to Patrick Hochstenbach (Ghent University Library) & Johann Rolschewski (Berlin State Library) for their open and shared work.

As such, this workshop material is licensed under a Creative Commons Attribution 4.0 International License.

Bootcamp Sessions

Note: These times may shift. If you intend to only go to one session, please let us know by registering here. That way, we can alert you to time shifts if /when they occur for the session of interest to you.

Session	Time	Description	Link to Materials
Data Structures & Models	10 - 11 AM	Brief introduction to data structures and models for the metadata worked with in the workshop	TBD
*nix Command Line Interface	11 - 12:30 PM	Overview of the CLI for running simple data scripts. Will introduce some basic bash scripts, as well as how to run existing Python scripts (and begin to debug errors) for analyzing metadata. YOU WON'T LEARN PYTHON, but how to run Python scripts.	TBD
Git for Metadata at CUL	1:30 - 3 PM	Using Git and GitHub for metadata work run previously.	TBD
SQL / MySQL for CUL DATA	3 - 4 PM	Built off the previous, learn the basics of SQL - what it does and how to run basic queries or updates on metadata managed or used through out this workshop.	TBD