



VCU

Virginia Commonwealth University
VCU Scholars Compass

Mid-Atlantic Chapter of the Medical Library
Association (MAC- MLA) Annual Meeting
Posters/Presentations

Conferences and Events

2022

Free, FAIR, and Fabulous: Five Data Tools to Support Open and Reproducible Research at Your Institution

Christine Nieman

University of Maryland, cnieman@hshsl.umaryland.edu

Katie Pierce-Farrier

University of North Texas Health Science Center, katie.pierce-farrier@unthsc.edu

Sean Corning

UMass Chan Medical School, sean.corning@umassmed.edu

Elizabeth Roth

Medical University of South Carolina, rothel@musc.edu

Follow this and additional works at: https://scholarscompass.vcu.edu/mac-mla_present



Part of the [Medicine and Health Sciences Commons](#)

Downloaded from

https://scholarscompass.vcu.edu/mac-mla_present/1

This Poster is brought to you for free and open access by the Conferences and Events at VCU Scholars Compass. It has been accepted for inclusion in Mid-Atlantic Chapter of the Medical Library Association (MAC- MLA) Annual Meeting Posters/Presentations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

Free, FAIR, and Fabulous: Five data tools to support open and reproducible research at your institution



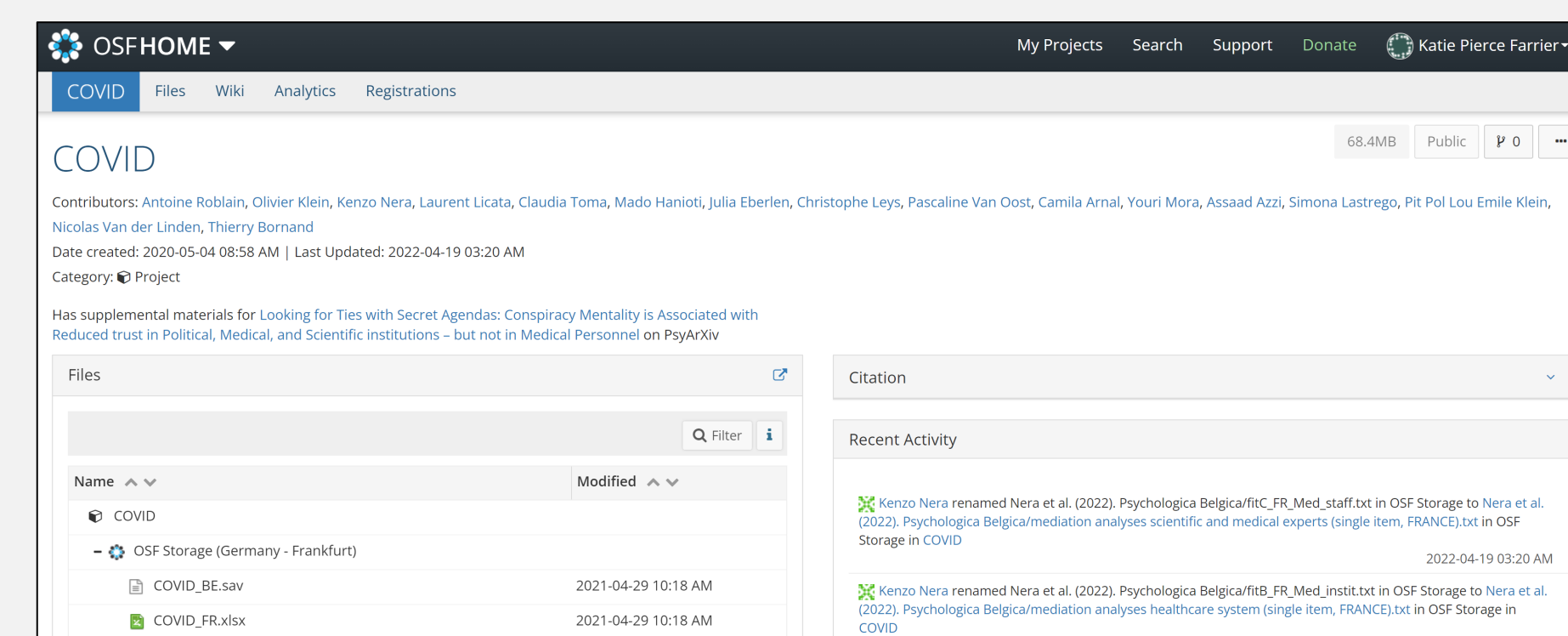
Scan for Product Guides!

Christine Nieman, University of Maryland, Baltimore; Katie Pierce Farrier, University of North Texas Health Science Center; Elizabeth Roth, Medical University of South Carolina; Sean Corning, University of Massachusetts Chan Medical School



Research Stage: Plan, Collect/Capture, Publish, Preserve, Reuse

What it does: The Open Science Framework is a free, open platform for organizing, managing, and sharing research projects that supports and promotes collaboration.



- ❑ Register project plan and assign persistent identifiers (DOI) to project materials
- ❑ Share project materials such as research methods, data, and results with collaborators, and share preprints to gain feedback from peers
- ❑ Store materials and share datasets in accessible, discoverable secure cloud storage

Use it for: Design and plan research projects with collaborators. Manage and share research data, results, and preprints.

Introduction

These tools facilitate open and FAIR data practices across the research data lifecycle. This poster seeks to promote awareness of freely available tools for data management, wrangling, and sharing, that will be of use to librarians, researchers, and support staff in daily work and research projects. With more publishers and funders requiring data management and sharing plans, information research professionals can use freely available tools to help them skillfully manage research data and to facilitate open data practices.

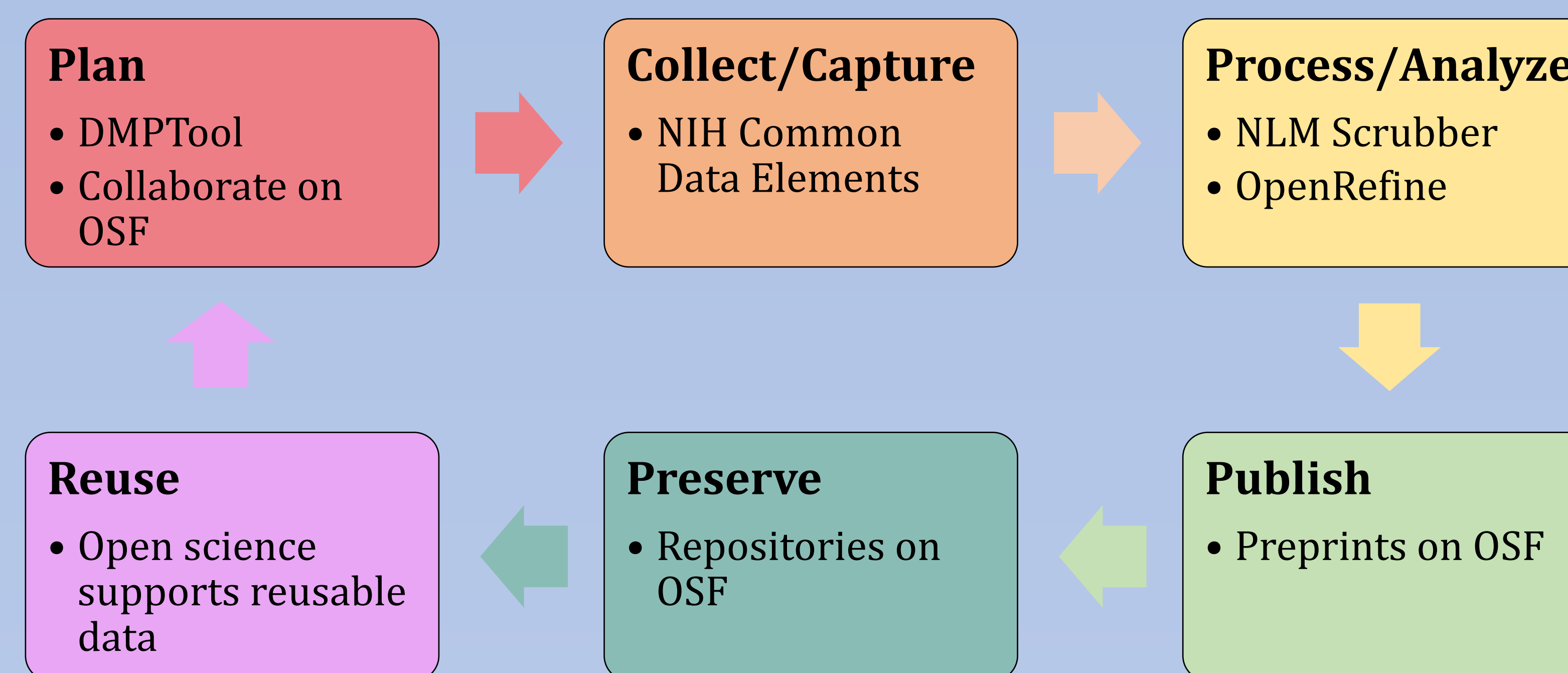
FAIR Principles

Findable: Collaboration tools, preprints, data sharing make data more findable.

Accessible: Free tools, open formats make data more accessible.

Interoperable: Common data elements make data more interoperable.

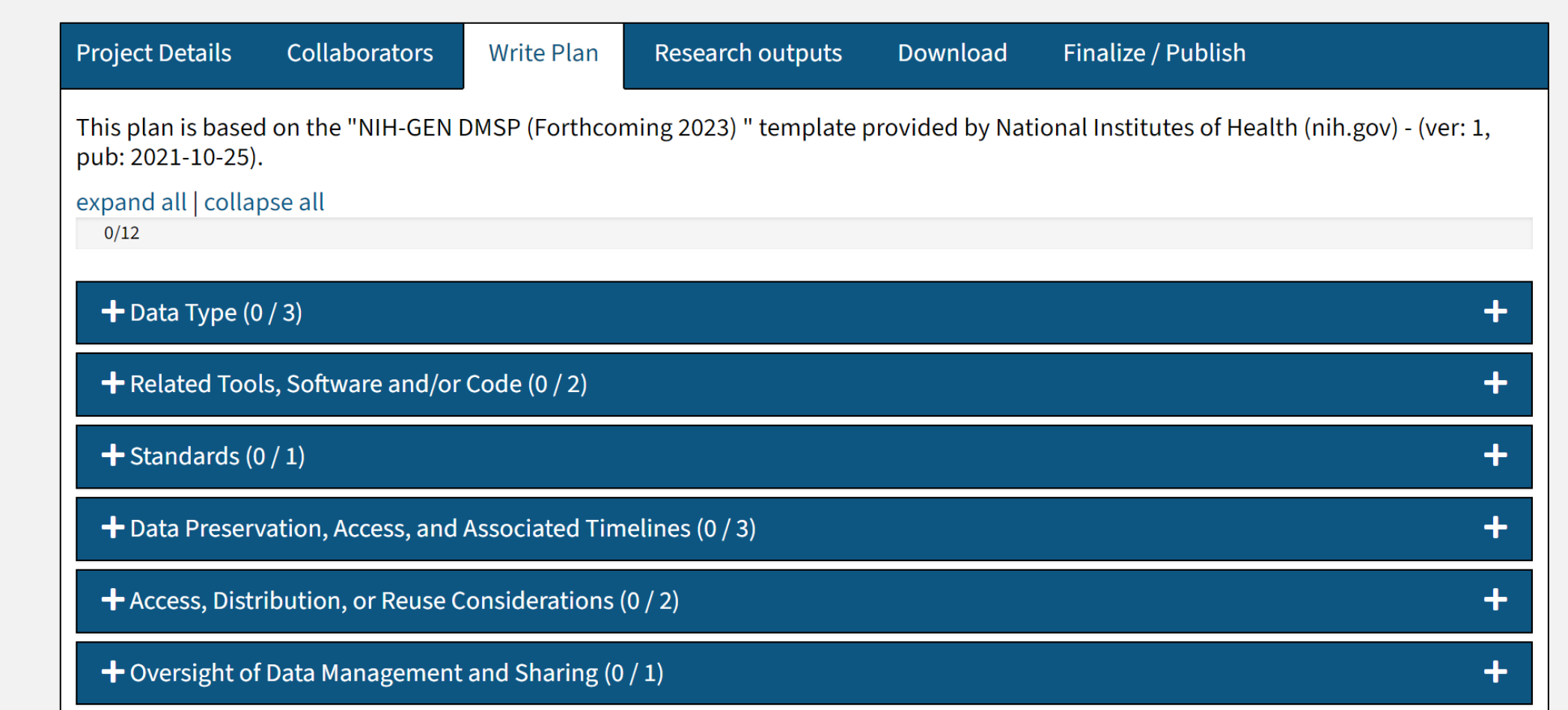
Reusable: Clean, well documented workflows make data reusable and reproducible.



Build your Data Management Plan

Research Stage: Plan

What it does: The DMPTool is a free, open-source, online application that helps researchers create data management plans.



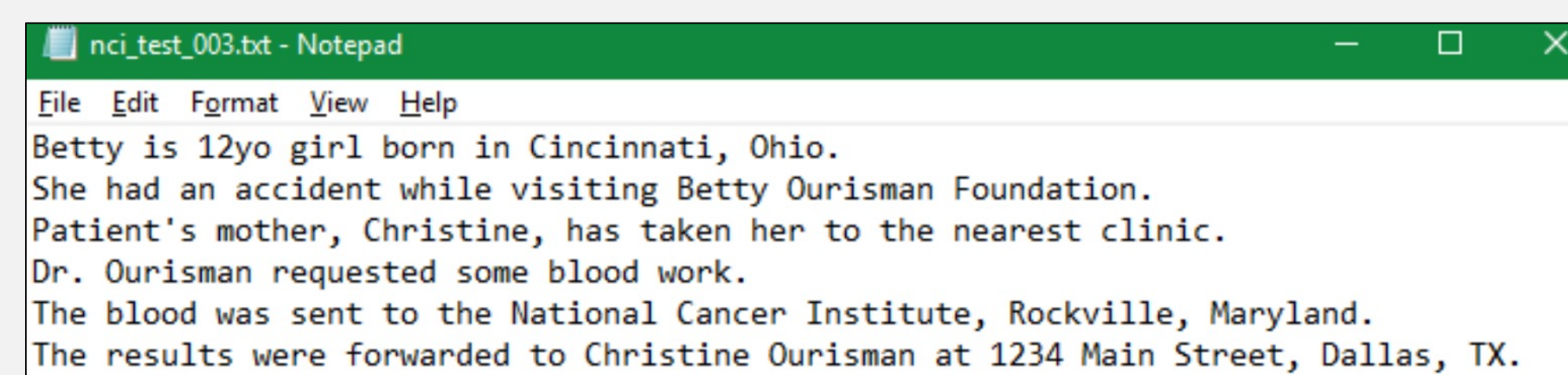
- ❑ Guides researcher through process of creating a data management plan
- ❑ Helps researchers comply with funder and publisher requirements
- ❑ Provides links to funder information, suggested answers, and other resources

Use it for: Create data management plans to comply with funder requirements.

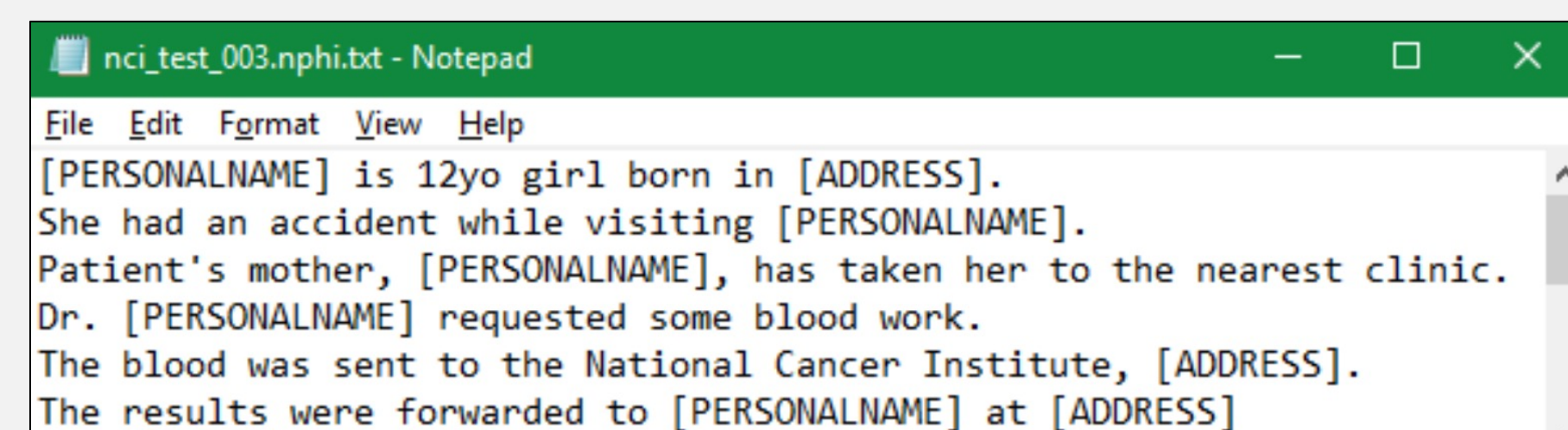


Research Stage: Process/Analyze, Reuse

What it does: The NLM Scrubber is a freely available tool for clinical text de-identification that uses natural language processing to find and remove personally identifiable information (PII).



Before (above) and after deidentification (right)



- ❑ Identify PII and replace with generic labels
- ❑ Allows researchers access to medical data and protects patient privacy by deidentifying electronic medical records (EMRs)
- ❑ Produces HIPAA-compliant datasets for scientific research, publication, and sharing

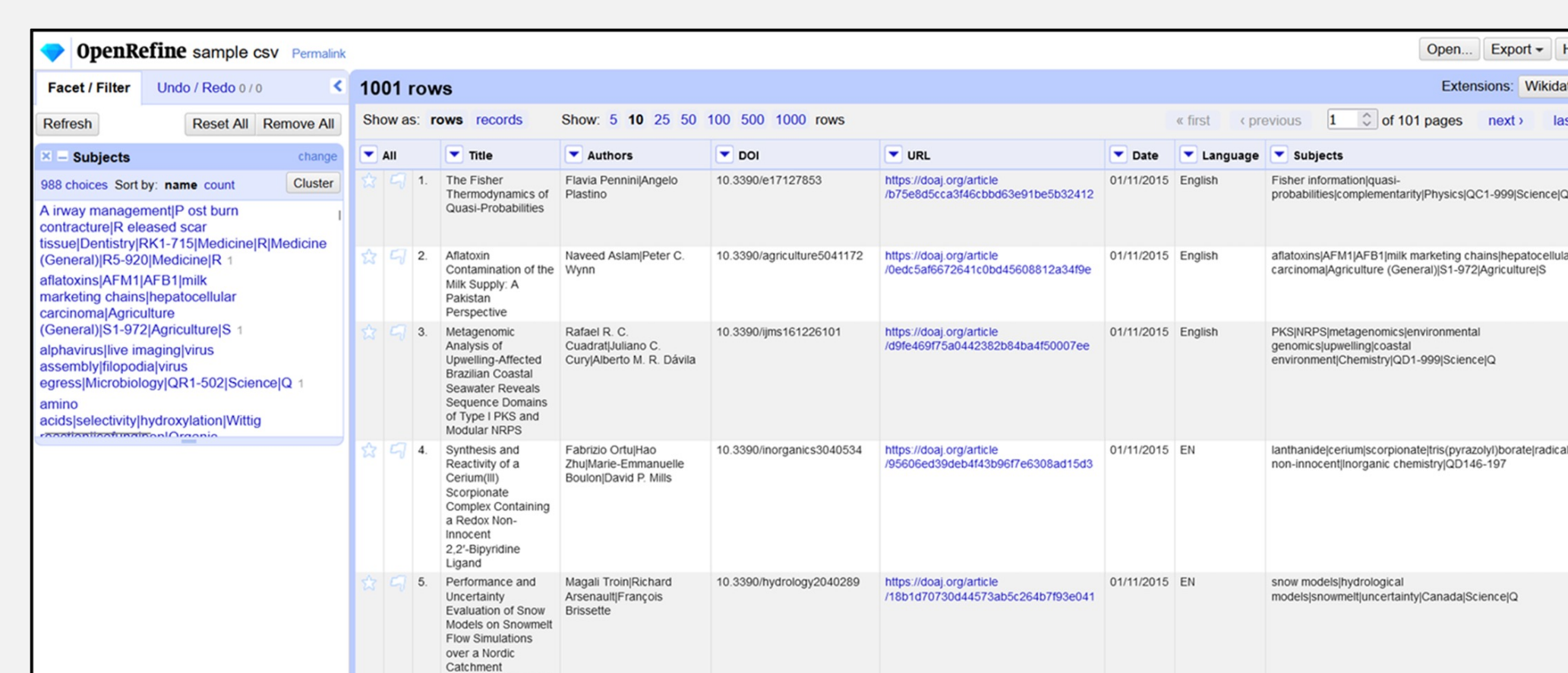
Use it for: Protect patient privacy. Deidentify clinical and medical data.



OpenRefine

Research Stage: Process/Analyze, Reuse

What it does: OpenRefine is a free, open source tool for cleaning, transforming, and formatting data in multiple formats.



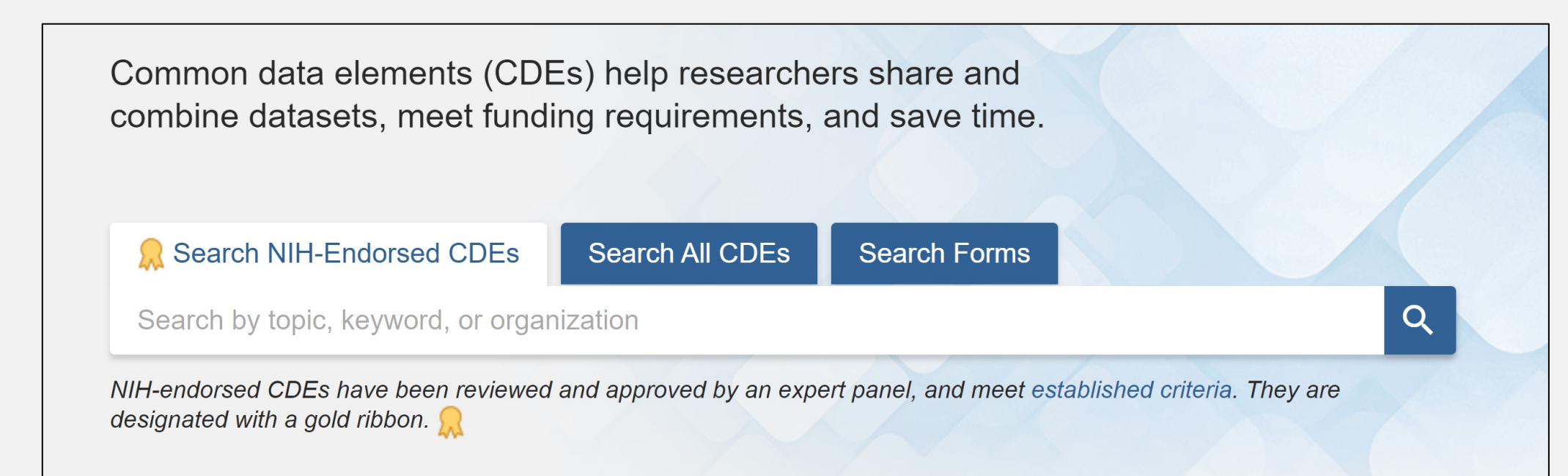
- ❑ Save and export operations for reproducible data manipulation
- ❑ Preserve the original file with version control
- ❑ Quickly make mass edits to fix spelling, formatting or other cell data

Use it for: Clean, process, and document reproducible data transformations.



Research Stage: Plan, Collect/Capture, Reuse

What it does: The NIH CDE Repository is a repository of common data elements (CDE). A CDE is a standardized, precisely defined question, paired with rules for allowable responses. The definitions are human and machine-readable.



- ❑ Access CDEs recommended or required by NIH Institutes and Centers
- ❑ Find standardized forms, questionnaires, and response options to conduct research studies
- ❑ Increase interoperability of research datasets across diverse studies

Use it for: Develop standardized data collection methods. Collect cleaner data that is easier to reuse.