

CUR Remote Undergraduate Research Resources

As undergraduate research opportunities move to remote settings, your institution's library and your librarians can be great resources to seek out. Librarians understand how information is created and organized. In addition, our librarians are the ones making decisions about database access and ultimately, what information is available to you and your students to conduct research.

Below are just a sampling of ways and resources to create meaningful and engaging undergraduate research experiences in a virtual setting. Our work is divided into several sections and we hope you can use this document as a source of inspiration as you design your undergraduate research projects. This is a living document, so feel free to provide feedback to the authors for additions.

This document was created by a team from Penn State University Libraries. Please reach out if you have any questions.

- Hailley Fargo | Student Engagement Coordinator | hmf14@psu.edu
- Jen Jarson | Head Librarian, Penn State Lehigh Valley | jmj12@psu.edu
- Christina Riehman-Murphy | Reference & Instruction Librarian, Penn State Abington | cer20@psu.edu

TABLE OF CONTENTS

[Librarians as Collaborators](#)

[Online Guides](#)

[Workshops](#)

[Case Studies](#)

[Data Sets](#)

[Library-subscription Databases](#)

[Government, NGO, and IGO Databases](#)

[United States Government](#)

[International Governments](#)

[Intergovernmental Organizations](#)

[Non-governmental Organizations](#)

[Research Centers and Professional Organizations](#)

[Discovering More Data Sets](#)

[Search Engines and Repositories](#)

[Library Guides](#)

[Institutions and Professional Organizations](#)

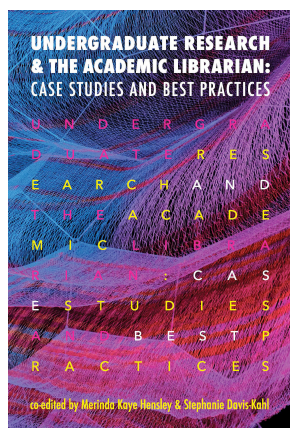
[Transcription, Oral History, & Other Research Projects](#)

[Transcription Projects](#)

[Oral History](#)

[Other research projects](#)

Librarians as Collaborators



If you're curious about how librarians and faculty can collaborate to create undergraduate research experiences, [*Undergraduate Research & the Academic Librarian: Case Studies and Best Practices*](#), edited by Merinda Kaye Hensley and Stephanie David-Kahl, is a great place to start. These chapters showcase the advantages of collaborating in order to create independent and class-based undergraduate research projects. This book is also available [online](#).

Online Guides

Many librarians create guides to help their users access selected, disciplinary resources, or gain training on certain research tools. For example, the University of Illinois has a selection of [“Savvy Researcher” guides](#) that can be a good starting spot. Additionally, the University of Minnesota has [a landing page for research resources](#), connecting their community to the resources they need.

Some libraries will also create guides for data sets. A few examples are below, but also refer to the [Data Sets](#) section of this guide.

- [Carnegie Mellon University Library Business & Economics Data Sets](#)
- [UC Berkeley Open Data Sets](#)
- [Webster University Library Medicine & Biological Science Statistics](#)
- [Sacramento State University Library Datasets and Statistics](#)
- [Vanderbilt University Library Social Sciences Data Sets](#)

Workshops

Finally, in these remote situations, many libraries are hosting virtual research workshops. These workshops are well-suited for faculty, graduate students, and undergraduates, helping to provide foundational skills. For some examples of this, see Penn State Libraries' [press release on our Remote Research Seminar Series](#). Many librarians can offer workshops that support the beginning stages of research for either classes or research groups. For example, the Penn State Libraries [has foundational skills and joining the scholarly](#)

[conversation workshops](#) that work well with students beginning undergraduate research experiences.

Case Studies

In their 2019 article [Undergraduate Research for Online Students](#), Faulconer and Gruss made the case for bringing undergraduate research to online students, noting that while there are perhaps different barriers to entry for this population, that the known numerous and significant benefits of participating in undergraduate research make it well worth exploring how to engage remote and distance students in this high impact practice. This area is ripe for exploration as there is very little known about how or even if higher education institutions are providing these opportunities. Now with the sudden and wide-scale shift to remote learning, it's imperative that institutions, faculty researchers, and librarians share the work they're doing to begin to address this need.

While there are few case studies looking at Undergraduate Research Experiences (URE) for distance undergraduate students, there are nevertheless many examples of undergraduates doing research with digital collections, with large public datasets, and with social media data that can be used as models for implementing similar UREs in remote settings.

Digital humanities projects are well-suited for distance and residential learning. At Penn State Abington, the literary informatics librarian, who is located at a different campus, teaches undergraduate research students [corpus analysis](#) using Voyant in order to analyze the recipes they are transcribing through the Folger Shakespeare's [Digital Transcription Portal](#). Elsewhere, a librarian and archivist collaborated to coteach a [distance liberal arts course](#) to undergraduate students through the Council of Public Liberal Arts colleges (COPLAC) to engage them with the curation and creation of digital scholarly projects. They chose broad themes and allowed students to explore local stories using primary sources but they also used this course to explore best practices for managing digital projects in a distance environment.

Members of CUR might be familiar with how [geoscience instructors](#) leveraged the Paleobiology Database (PDBD) to engage students and institutions, which might not otherwise have the time or equipment, with undergraduate research. [Project EDDIE](#) is a cross-institutional pedagogical collaborative which develops flexible models that guide undergraduate students through the exploration of large publicly available data sets that focus on environmental science and education as part of a course-based undergraduate research experience (CURE). Bethune-Cookman University's Center for Undergraduate Research runs an [interdisciplinary summer research program](#) where all of the students

worked with the same large criminal justice dataset from the Florida Department of Juvenile Justice but worked in teams to examine discipline-specific facets of the data. [Communication faculty](#) at Tarleton State University used a first-year seminar course to engage students with Twitter data around activism and multiculturalism, a project that could be customized for many social science fields. Institutions looking to develop strong support programs should look to the University of Maryland-College Park's First-Year Innovation & Research Experience (FIRE) which focuses on [big-data analysis and visualization](#) as research methods "in a manner that does not require facilities, equipment, consumables, protocols, or an established research plan."

What is common among many of these case studies is that they are designed to be interdisciplinary which makes them scalable and customizable to research projects of all scopes and sizes. The faculty running these projects noted that students gained crucial research skills, particularly in the highly marketable and necessary areas of visualization, statistics and analysis. Some noted that using publicly available data and digital artifacts eliminated the barrier of needing IRB approval and the subsequent training for students that requires. Many of the teams who worked with large datasets mentioned the importance of humanizing and contextualizing the numbers, which can be overwhelming, especially for undergraduates. Doing so requires collaboration with humanities faculty and projects which will go a long way towards making distance undergraduate research experiences accessible and rewarding for all involved. And engaging students with public sources of information and novel methods and technologies gives them the necessary data literacy and inquiry skills needed for lifelong learning.

Data Sets

The databases and websites featured here offer a snapshot of the search tools and sources available to help you discover data sets. Some of the tools and sources are multi-disciplinary while others are subject-specific. Consult with librarians at your institution for assistance locating additional data sets specific to your needs.

Library-subscription Databases

Your library may subscribe to databases that provide data sets relevant for your discipline and research area. The following selected list highlights some options offered by many libraries. Consult with your librarians or check your library's website for access to these and other subscription databases.

- [Data Citation Index](#) - "Search directly across millions of records from hundreds of evaluated data repositories in the sciences, social sciences, and humanities."
- [Data Planet Statistical Datasets](#) - "The Data Planet repository contains more than 157 billion data points from more than 80 source organizations." Subject categories include agriculture, banking and finance, criminal justice, education, energy, government, health, housing, business and commerce, trade, labor, military, natural resources, population, prices and cost of living, transportation, and more.
- [ICPSR](#) - "An international consortium of more than 750 academic institutions and research organizations, Inter-university Consortium for Political and Social Research (ICPSR) provides leadership and training in data access, curation, and methods of analysis for the social science research community. ICPSR maintains a data archive of more than 250,000 files of research in the social and behavioral sciences."
- [iPoll Databank](#) - "The iPOLL database is a unique, rich data source for US public opinion questions and answers. The system allows for users to sift through over 700,000 questions archived from national public opinion surveys since 1935."
- [Mergent Online](#) - "Business and financial data"
- [SAGE Research Methods Datasets](#) - "SAGE Research Methods Datasets provide that opportunity for users looking to master quantitative or qualitative data analysis. Each title provides sample data that has been optimized for demonstrating a method, accompanied by a how-to guide that explains the method at hand and walks users through the analysis using the data provided. Quantitative examples also include software guides for statistical analysis using SPSS, R, Stata, or Python."

- [SAGE Stats](#) - “Features statistical data series created from more than 400,000 government and non-government datasets, covering popular topics of research interest for U.S. states, counties, cities, metropolitan areas, and ZIP codes.”
- [Wharton Research Data Services \(WRDS\)](#) - “WRDS provides researchers with one location to access over 350 terabytes of data across multiple disciplines including Accounting, Banking, Economics, ESG, Finance, Healthcare, Insurance, Marketing, and Statistics.”

Government, NGO, and IGO Databases

Governments, nongovernmental organizations (NGOs), and intergovernmental organizations (IGOs) collect a wide variety of data and make some of this data available to the public. The following selected list is intended to highlight the range and types of datasets available. To locate others, check government, NGO, and IGO websites or consult with your librarians.

- United States Government
 - [Bureau of Justice Statistics](#) - “The United States' primary source for criminal justice statistics. BJS mission: To collect, analyze, publish, and disseminate information on crime, criminal offenders, victims of crime, and the operation of justice systems at all levels of government.”
 - [Bureau of Labor Statistics](#) - “The Bureau of Labor Statistics measures labor market activity, working conditions, price changes, and productivity in the U.S. economy to support public and private decision making.”
 - [Centers for Disease Control and Prevention Data](#) -Subjects include: health, injury and violence, disease, vaccination, tobacco use, disability, and more.
 - City level open data projects like [Open Data Philly](#) and [NYC Open Data](#)
 - [Climate Data Online](#) (National Oceanic and Atmospheric Administration) - “Climate Data Online (CDO) provides free access to NCDC's archive of global historical weather and climate data in addition to station history information.”
 - [Data.gov](#) - “Under the terms of the 2013 Federal Open Data Policy, newly-generated government data is required to be made available in open, machine-readable formats, while continuing to ensure privacy and security.”
 - [Environmental Dataset Gateway](#) and [Envirofacts](#) (Environmental Protection Agency) - “EPA's official open data catalog” and “Your one-stop source for environmental information.”
 - [FRED: Federal Reserve Economic Data](#) (Federal Reserve Bank of St. Louis) - Economic data

- [HealthData.gov](#) (U.S. Department of Health and Human Services) - “On this site, you can find data on a wide range of topics, including environmental health, medical devices, Medicare & Medicaid, social services, community health, mental health, and substance abuse. The data is collected and supplied from agencies from the U.S. Department of Health and Human Services as well as state partners. This includes the Centers for Medicare and Medicaid Services, Centers for Disease Control and Prevention, Food and Drug Administration, and the Agency for Health Care Research and Quality, among others.”
- NASA - [EarthData](#) and [Space Science Data Coordinated Archive](#) and [Astrophysics Data System](#)
- [National Center for Biotechnology Information](#) (U.S. National Library of Medicine) - “NCBI has been charged with creating automated systems for storing and analyzing knowledge about molecular biology, biochemistry, and genetics; facilitating the use of such databases and software by the research and medical community; coordinating efforts to gather biotechnology information both nationally and internationally; and performing research into advanced methods of computer-based information processing for analyzing the structure and function of biologically important molecules.”
- [National Library of Medicine Data Discovery](#) - “Access to datasets from selected NLM resources.”
- [National Longitudinal Surveys](#) (Bureau of Labor Statistics) - “A set of surveys designed to gather information at multiple points in time on the labor market activities and other significant life events of several groups of men and women.”
- [State Data Center Program](#) - “The SDC lead agencies organize a statewide network of coordinating and affiliate agencies. This network works in partnership with the Census Bureau through the Data Users Branch (DUB) and the Census Bureau's Regional Offices for data dissemination at the local level.”
- [U.S. Census](#) - “The Census Bureau is the leading source of quality data about the nation's people and economy.”
- [U.S. Department of Energy Data Explorer](#) - “Search tool for finding DOE-funded, publicly available, scientific data submitted by data centers, repositories, and other organizations funded by the Department.... Includes data Project, data Collection, and individual Dataset records.”
- [U.S. Energy Information Administration Open Data](#) - “Data for electricity, petroleum, and natural gas.”

- [U.S. Geological Survey](#) - USGS provides “science about the natural hazards that threaten lives and livelihoods, the water, energy, minerals, and other natural resources we rely on, the health of our ecosystems and environment, and the impacts of climate and land-use change. “
- International Governments
 - [Central Bank and Monetary Authority Websites](#) (Bank for International Settlements) - A directory by country. Visit each bank’s website to check on data availability.
 - [National Statistical Offices](#) (United Nations) - “The United Nations Statistics Division ... has developed a central repository of country profiles of statistical systems.” Visit each country’s site to check on data availability.
- Intergovernmental Organizations
 - [Organisation for Economic Co-operation and Development Data](#) (OECD) - Works on “establishing evidence-based international standards and finding solutions to a range of social, economic and environmental challenges.”
 - [UNdata](#) (United Nations) - “The numerous databases or tables collectively known as "datamarts" contain over 60 million data points and cover a wide range of statistical themes including agriculture, crime, communication, development assistance, education, energy, environment, finance, gender, health, labour market, manufacturing, national accounts, population and migration, science and technology, tourism, transport and trade.”
 - [UNESCO Institute of Statistics](#) - “Internationally-comparable data on education, science, culture and communication. As the official statistical agency of UNESCO, the UIS produces a wide range of indicators in UNESCO’s fields of action by working with national statistical offices, line ministries and other statistical organizations.”
 - [World Bank Open Data](#) - “Global development data”
- Non-governmental Organizations
 - [Worldwide NGO Directory](#) (World Association of Non-Governmental Organizations)

Research Centers and Professional Organizations

Research centers and institutes engaged in data collection sometimes make their data available to the public. The following selected list provides just a snapshot of these options. To locate others, check the websites of research centers with which you’re already familiar or consult with your librarians.

- [American National Election Studies](#) - "High quality data from its own surveys on voting, public opinion, and political participation."
- [Association of Religion Data Archives](#) - Religion data
- [ASTI](#) (International Food Policy Research Institute) - "Data on agricultural research systems across the developing world"
- [Archaeology Data Service](#) (University of York) - "Accredited digital repository for heritage data generated by UK-based fieldwork and research."
- [Center for Applied Internet Data Analysis](#) - "Collection, curation, and sharing of data for scientific analysis of Internet traffic, topology, routing, performance, and security-related events."
- [Crystallography Open Database](#) - "Open-access collection of crystal structures of organic, inorganic, metal-organics compounds and minerals, excluding biopolymers."
- [Cornell Institute for Social and Economic Research](#) - "Extensive data collection with emphasis on labor, political, and social behavior, demography, economics, and health. The collection includes federal and state census files, administrative records, public opinion surveys, economic and social data from national and international organizations, along with studies compiled by individual researchers."
- [General Social Survey](#) (National Data Program for the Social Sciences) - "The GSS contains a standard core of demographic, behavioral, and attitudinal questions, plus topics of special interest. Among the topics covered are civil liberties, crime and violence, intergroup tolerance, morality, national spending priorities, psychological well-being, social mobility, and stress and traumatic events."
- [Geological Society of America Data Repository](#) - Geoscience research
- [National Archive of Data on Arts and Culture](#) (ICPSR) - "A repository that facilitates research on arts and culture by acquiring data, particularly those funded by federal agencies and other organizations, and sharing those data with researchers, policymakers, people in the arts and culture field, and the general public."
- [PEW Research Center](#) - "We study U.S. politics and policy; journalism and media; internet, science and technology; religion and public life; Hispanic trends; global attitudes and trends; and U.S. social and demographic trends."
- [Stanford Large Network Dataset Collection](#) (Stanford University) - Datasets of social and information networks.
- [UK Data Service](#) - "Collection includes major UK government-sponsored surveys, cross-national surveys, longitudinal studies, UK census data, international aggregate, business data, and qualitative data."

- [ZINC15](#) - “Free database of commercially-available compounds for virtual screening. ZINC contains over 230 million purchasable compounds in ready-to-dock, 3D formats.”

Discovering More Data Sets

- Search Engines and Repositories

Explore search engines and repositories that are specific to data sets such as:

- [DataHub](#) - “Thousands of datasets from financial market data and population growth to cryptocurrency prices.”
- [Dryad Digital Repository](#) - “A curated resource that makes research data discoverable, freely reusable, and citable. Dryad provides a general-purpose home for a wide diversity of data types. Dryad originated from an initiative among a group of leading journals and scientific societies ... Our mission is to provide the infrastructure for, and promote the re-use of, data underlying the scholarly literature.”
- [Google Dataset Search](#) (Google)
- [Google Public Data Search](#) (Google)
- [Harvard Dataverse](#) (Harvard University)
- [Qualitative Data Repository](#) - “Dedicated archive for storing and sharing digital data (and accompanying documentation) generated or collected through qualitative and multi-method research in the social sciences.”
- [Registry of Open Data on AWS](#) - “Datasets made publicly available through AWS services.”
- [Registry of Research Data Repositories](#) (re3data.org)

- Library Guides

Look for guides or compilations from your institution’s library that may point you to subscription-based or open databases and data sets such as:

- [Biomedical and Life Sciences Data](#) (Brown University Library)
- Data Resources in the Health Sciences: [Clinical Data](#) and [Scientific Data](#) (University of Washington Health Sciences Library)
- [Data Sets for Social Sciences: Social Data](#) (Vanderbilt University Libraries)
- [Data Sets for Quantitative Research: Public Use Datasets](#) (University of Missouri Libraries)

- [Datasets: Where to Find Them](#) - Note: Covers a wide variety of disciplines (George A. Smathers Libraries, University of Florida)
- Institutions and Professional Organizations

Check websites of professional organizations that may point you to subscription-based or open databases and data sets such as:

- [Behavioral and Social Science Data Sets and Repositories](#) (American Psychological Association)
- [Duke Clinical and Translational Science Institute](#) (Duke University)
- [Scientific Data: Recommended Data Repositories](#) (Nature)

Transcription, Oral History, & Other Research Projects

These are research projects already underway and could become an independent project an undergraduate takes on. For all of these projects, it's important to think through ways to scaffold this research. Scaffolding includes information on why these projects matter, how researchers use this data to inform future research projects, and other instruction on how to do this work.

The projects linked below are works in progress, where students can help contribute to a growing set of transcribed, primary source documents. There are also databases with fully transcribed primary sources, that could be leveraged into an undergraduate research project. If you'd like to learn more about those databases, please consult with your local librarian to see what access your institution has.

Transcription Projects

- [Newberry Transcribe](#)
- [Homer Multitext Project](#). These researchers are particularly interested in UGR
- [DIY History](#) from the University of Iowa
- [The Ryan White Letters Transcription Project](#)
- [Vast Early America Resources](#) from Omohundro Institute
- [By the People](#) from the Library of Congress
- [EMROC](#) which is part of the larger Early Modern Manuscripts Online Project [EMMO](#)
- [EMPOP](#) which is part of the larger Early Modern Manuscripts Online Project

Oral History

- [COVID-19 Oral History Project](#) from IUPUI

Other research projects

- [Nickels and Dimes Research Opportunities](#), from Northern Illinois University
- [The Newberry Library's Resources](#) that you can engage with at a distance