

**DEEP BLUE REPOSITORY
AND RESEARCH DATA SERVICES**

IMPACT REPORT 2022



MICHIGAN PUBLISHING
UNIVERSITY OF MICHIGAN LIBRARY

TABLE OF CONTENTS

INTRODUCTION

Welcome to Deep Blue Repositories	3
-----------------------------------	---

TIMELINE

Deep Blue Docs and Data	4-5
-------------------------	-----

KEY STATS

Deep Blue Docs	6
Deep Blue Data	7

INFRASTRUCTURE

Reinforcing U-M's Research Infrastructure	8-9
Facilitating Research Access	10

CAMPUS CONNECTIONS

Supporting Scholarship across Campus	11-12
Supporting Student Scholarship	13

DATA CURATION

How Data Curation Increases Research Value	14
The Data Curation Network	15

DEIA

Adding Diversity to the Data Mix	16
Spotlight: The Global Feminisms Project	17

ACKNOWLEDGEMENTS

With Gratitude	18
----------------	----

**Thank you to the sponsors
and funders who
have enabled
our work, including:**

LYRISIS
The ALA Office for Diversity,
Literacy and Outreach Services
The Alfred P. Sloan Foundation
The National Institutes of Health
The National Science Foundation

MICHIGAN PUBLISHING

WELCOME TO DEEP BLUE REPOSITORIES

Deep Blue is essential to Michigan Publishing's goal of increased openness, accessibility, and equity through all portions of the scholarly lifecycle.

Since Deep Blue Repository and Research Data Services launched under the University of Michigan Library in 2006, it has grown to be a nucleus for the preservation and promotion of the work of University of Michigan scholars.

Deep Blue Documents: Dissertations and More

Deep Blue Documents began as a simple dissertation deposit venue and has grown to house over 140,000 pieces of scholarship, including articles, chapters, conference presentations, and other works created by U-M community members. Most of the authors who choose to deposit their research with Deep Blue decide to make that content freely viewable and searchable to anyone regardless of institutional affiliation.

Deep Blue Data: Making Open Data Available to All

Deep Blue Data includes close to 500 data sets developed by or used in the support of research activities at the University of Michigan. As a founding member of the multi-institutional Data Curation Network, Deep Blue Data and the University of Michigan have helped to create a network of specialists dedicated to advancing open research and to ensuring data can be reused and understood.

The 2022 Deep Blue Impact Report

This report shares how both Deep Blue Documents and Deep Blue Data have influenced both the University of Michigan as well as the broader scholarly landscape. The stories contained here highlight Deep Blue's role in advancing the goals of academic innovation, equity, and increased faculty discourse with the public.

In its first decade and a half, Deep Blue Repository and Research Data Services has seen over 100 million downloads of its documents and datasets. We hope the impact that Deep Blue has had on the world inspires you just as it inspires us, and that you join us in making Deep Blue's journey to its next 100 million downloads even more successful than its first.

Visit deepblue.lib.umich.edu to learn more.



2006

**DEEP BLUE DOCS
LAUNCHES UNDER
U-M LIBRARY**

2007

PILOT DISSERTATION DEPOSIT
PROJECT WITH U-M RACKHAM
GRADUATE SCHOOL BEGINS

2008

DEEP BLUE LINKS BECOME
AVAILABLE IN PUBMED CENTRAL

2009

**DISSERTATION
DEPOSITS GO
MAINSTREAM IN
DEEP BLUE DOCS**

2009

DEEP BLUE DOCS JOINS
MICHIGAN PUBLISHING

2010

DEEP BLUE DOCS BEGINS HOSTING
THE BENTLEY DIGITAL ARCHIVES

2013

PLANNING BEGINS FOR
RESEARCH DATA SERVICES (RDS)
AND
DEEP BLUE DATA REPOSITORY

2011

**10 million
Downloads**

2016

**DEEP BLUE DATA AND RDS
OFFICIALLY LAUNCH
and
U-M HELPS CREATE
THE
DATA CURATION NETWORK**

2017

DEEP BLUE DATA INSTALLS GLOBUS

2019

U-M LIBRARY CHAIRS WORKING
GROUP ON PUBLIC ACCESS TO
RESEARCH DATA

TODAY

**100 million
Downloads**

2016

**100,000
DEPOSITS**

2018

RDS/DEEP BLUE DATA AND
DEEP BLUE DOCUMENTS MERGE TO
FORM DEEP BLUE REPOSITORIES

2021

14 INSTITUTIONS JOIN
THE DATA CURATION NETWORK
and
U-M LIBRARY RECEIVES 2 GRANTS
TO RESEARCH DATA COMMUNITIES

2022

**500 DATASETS IN
DEEP BLUE DATA**

KEY STATS - DEEP BLUE DOCUMENTS

10 million new web visitors last year
More than 1.2 million pageviews

143,534 works composed of 191,075 files as of Feb 2022

Over **4,200** deposits in 2021, with an average of 5-6 thousand deposits per year over the past 15 years and most users making 1-10 deposits annually

Deposits from all three University of Michigan campuses, including articles, theses, dissertations, multimedia images and recordings, as well as books and chapters, conference papers, learning objects, presentation materials, technical reports, and working papers

Deep Blue Docs currently holds **12.29 (as of Nov 2021) terabytes** of storage

**100 million
downloads**

KEY STATS - DEEP BLUE DATA



over **500**
datasets

87 datasets deposited in 2021

All datasets in the repository reviewed and optimized by Deep Blue staff to ensure they are FAIR—

Findable
Accessible
Interoperable
Reusable

Most popular depositing disciplines:

Science (279 datasets)
Engineering (115 datasets)
Health Sciences (99 datasets)
Humanities (72)
Social Sciences (38)

14,000 new web visitors in the past year
More than 37,000 pageviews

REINFORCING U-M'S RESEARCH INFRASTRUCTURE

Deep Blue Repository and Research Data Services might seem like a relatively small department within Michigan Publishing and the university library, but it provides integral services to the U-M research community. Deep Blue helps ensure U-M scholars are in compliance with the grant and open access requirements so that those sources can continue funding innovative research on campus. The repositories serve as critical support for both the publishing and storage of the university's research outputs—in the form of published articles, data sets, reports and multimedia.

The Deep Blue repositories house hundreds of thousands of files created by Michigan faculty, staff and students. Deep Blue's team not only maintains the databases that enable the storage and search of these materials, but constantly improves upon the service to better meet the needs of researchers. In recent years, they have added DOI links for all Deep Blue Docs and Data deposits, ensuring a permanent and universal pathway to access and cite the research contained in the repositories. They have established a system for embedding README for datasets within the metadata itself to enhance the understanding and sharing of the data and created a streamlined user interface for easy navigation.

Deep Blue's team is available to help U-M researchers successfully deposit their work in the repositories, and the work done by Deep Blue staff ensures that the legacy and research of U-M scholars is preserved both at the university and within their field.

"With Deep Blue, I know the documents will be there and the links will work. So when I cite something from Deep Blue in a bibliography of a printed published book, I am confident the reader will find the material in many years to come."

REINFORCING U-M'S RESEARCH INFRASTRUCTURE

What the U-M Research Community Values about Deep Blue:

“The size of my data was big and it was not easy to share it with others, [with] Deep Blue I can share it with anyone.”

“I like that there are persistent identifiers and that others, including those outside U-M, can get to materials we produce.”

“I found the curators at Deep Blue to be responsive and helpful.”

“The data upload process is simple and user-friendly.”

“The download statistics are awesome each month—giving me a picture of what people might be reading.”

Deep Blue Data plays a unique role in the early stages of research by helping to fulfill grant funding requirements that ask for a data management plan. Between March 2020 and February 2021, 85 grant applications and awards filed at the university listed Deep Blue Data in their data management plans. A recent survey found that more than half Deep Blue Data depositors included or planned to include the repository in future grant requests as well.

Deep Blue's staff play a critical role in helping researchers understand and navigate these grant requirements by proactively reaching out to the Principal Investigators of the projects listing Deep Blue Data in their DMP to introduce them to the library's research data services and collect additional useful information to better assist the researchers in the final publication of their datasets.

Several 2021 surveys highlight the level of support Deep Blue has given to depositors. 92% of respondents said that Deep Blue Docs completely or mostly met their needs, and 98% of respondents would recommend Deep Blue Data services to their colleagues. As one member of the research community explained,

“Deep Blue has been a transformative experience for me as a scholar.”

The support that the Deep Blue Repositories provide will only become more important as data and document needs grow in the future.


FACILITATING RESEARCH ACCESS

One of Deep Blue's core functions is ensuring access to research. In a recent survey of Deep Blue depositors, 90% of respondents said that they decided to use the repository because they wanted to provide greater access to the material(s) so that other people could find and read them.

Making these deposits discoverable and usable involves more than simply hosting the unaltered materials on a website. Downloading large data sets can be difficult if not impossible over the internet. Deep Blue was a leader among institutional repositories in trying to resolve this issue and was among the first to adopt the use of Globus, a platform that divides files into smaller packages for transfer. Globus also has a resumption feature that means if a transfer connection is interrupted the system will attempt to reconnect and continue the download for up to three days, which allows people to access much larger files than they could in the past even if they are located somewhere with unstable internet access.

Partnering with Globus has made Deep Blue Data unique among institutional data repositories because it has allowed Deep Blue to increase the size of the data that it can accept and disseminate. In October of 2017, Deep Blue Data housed just under 2.94 TB of data, but by December 2021, the repository had grown to house nearly 25 TB of data, including seven works that are over 1 TB.

Deep Blue Data is one of the few institutional data repositories in the country that can provide support for files as large as the low terabytes and remains a leader in this space.



"I love the idea that my deposits will be permanent and that they can be accessed by anyone in the world."

SUPPORTING SCHOLARSHIP ACROSS CAMPUS

The repositories and staff of Deep Blue are more than just infrastructure.

They uniquely enable research and publishing projects in various departments at the University of Michigan. These case studies represent just a few of the ways DBRDS partners with faculty, staff, and students to produce meaningful research and scholarly publications.

Department of English Language and Literature

In 2019, the staff of Deep Blue began talking with the English Department about how to publish their MFA theses in a secure and accessible place. The 2020 graduating MFA class was the first to have their theses published in Deep Blue Docs. The repository's growing Dissertations and Theses collection now brings together more than 36,000 works by U-M masters and Ph.D students to share with the world.

Department of Family Medicine

On March 20, 2020, U-M faculty members at the Department of Family Medicine and the editorial board of the journal *Annals of Family Medicine* realized that they needed a place to publish rapid-peer-reviewed work and practitioner essays on the novel coronavirus. As a result, Deep Blue launched the "COVID-19: *Annals of Family Medicine*" collection, and at the time of this writing, the collection provides access to over 100 publications, and continues to grow as Deep Blue Docs provides reliable and up-to-date information on the pandemic and best practices for treating those who suffer from COVID-19.

School of Education

In 2010, the NYT magazine wrote an article about Professor Deborah Ball, Dean of U-M's School of Education from 2005-2016, and her teaching techniques. To show, rather than just tell, what happened when a student had the idea that the number "6" could be both odd and even, the article linked to a video of her early days as a third grade teacher, hosted in Deep Blue. For a few months after that, this became one of the most-watched videos hosted by the repository.

Kelsey Museum of Archaeology

Aphrodisias is one of the most important archaeological sites of the Greek and Roman periods in Turkey. The Aphrodisias Regional Survey began in 2005 to investigate the interaction between human habitation and the natural environment from prehistory to the present day, and the results of the Aphrodisias Regional Survey published in 2012 by Christopher Ratté and Peter De Staebler is one of the few collections in Deep Blue dedicated to a single project. The Deep Blue collection provides electronic open access to: (1) the entire catalogue of archaeological points of interest recorded by the survey; the complete photographic archive; high-resolution scans of all the hand-drawn architectural drawings; high-resolution versions of all the digital maps and plans; (2) the entire catalogue of all the contexts targeted for systematic collection of surface finds; and (3) the entire intensive (transect) survey database, including all the data recorded for each subtransect of the survey.

Department of Radiology

In 2020, Deep Blue Data staff worked with Yuni Dewaraja, a professor in the Radiology department, and her team to host data sets they had developed to run their "Dosimetry Challenge." The challenge promoted Dosimetry as a field of research. Dewaraja and her team wrote up their "Dosimetry Challenge" in the Journal of Nuclear Medicine and even included the Director of Deep Blue Repository and Research Data Services, Jake Carlson, as an author.

Michigan Institute for Data Science

Since 2020, Deep Blue has collaborated with the Michigan Institute for Data Science (MIDAS) on the annual Reproducibility Challenge. Reproducibility of research results has been a significant challenge across scientific fields. The Reproducibility Challenge assesses and highlights examples of best practices for reproducible work across diverse fields. In addition to celebrating exceptional workflows, the projects showcased as part of the Challenge can act as models for reproducibility that others can follow.

Museum of Zoology

Museum of Zoology: anniversary, simultaneous print volume publication w/ OA pdf stored on Deep Blue Docs. Here is something Susan wrote up for a conference in Dublin, Ireland Feb 2020 <http://www.ijdc.net/article/view/700/575>

SUPPORTING STUDENT SCHOLARSHIP

Students choose to attend the University of Michigan because it produces some of the best public research in the country - and because it allows students to be part of that research.

By making research enduring, available, and usable, Deep Blue supports scholars and helps create richer networks of knowledge.

Deep Blue Data and Document repositories create the opportunity for student publication. Through the Deep Blue repository, students are able to publish dissertations and theses like “Detroit-Upper Peninsula Energy Burden Survey: a research project report and initial analysis of results” by graduate student Kate Hutchens in the School for Environment and Sustainability. This report highlights the work that was done in partnership with a Detroit-based non-profit organization to conduct a survey of residents of the City of Detroit and six counties in Michigan's Upper Peninsula.

Dissertations and theses are the most used (and requested) type of material in Deep Blue. Deep Blue Documents materials are also indexed by Google Scholar, which increases the accessibility of students' Honors and Master's theses. This helps students establish a research brand and builds a base for their future careers.

This student and research is not just available to those at U-M, but is available to anyone, anywhere, including the general public and non-academic community. Globally, people want to engage with the research done at the University of Michigan, and Deep Blue facilitates that access.

According to Google Analytics, in 2021, nearly two-thirds of the visitors to the Deep Blue platform came from outside of the United States and represented over 140 countries. These visitors explored data on medical imaging, ecological modeling, clinical trials, quantifying news bias and more. By collaborating with Deep Blue U-M students spread their work, and the University of Michigan name, around the world.

HOW DATA CURATION INCREASES RESEARCH VALUE

"I tend to assume that my data are self-explanatory, but the curation review provides the point of view of someone who is not an expert on the topic, which is extremely valuable."

Deep Blue acts as more than just a static holding pen for datasets. The staff spend hours assessing, organizing and augmenting the data through a process called data curation. In its broad form, the phrase data curation involves all of the actions taken by those overseeing and stewarding the data to ensure that the data are findable, accessible, interoperable and reusable by others (often referred to as the FAIR principles), and to enhance the long-term value of the data through preservation.

At Michigan, data curation is a collaboration between the scholars depositing the data and the data specialists at Deep Blue who know how to translate that information into meaningful, enduring data packages. Curation work includes multiple activities such as developing documentation and contextualization on the provenance of the datasets, creating metadata associated with the data to make it more discoverable, reviewing data files and code for quality control, organization and completeness, and minting a persistent identifier for the data so that it can be shared and cited.

As a leader in the field of data curation, the staff of Deep Blue collaborate with other institutional repositories via the Data Curation Network to develop a better understanding of curation activities that researchers value. This collaboration culminated in the article "How Important is Data Curation? Gaps and Opportunities for Academic Libraries," which included a study of 91 participants at six different academic institutions. The study indicated that though researchers were active in data curation for their projects, they would like even more support from their home institutions. The investments in the highly valued work of data curation done by libraries is part of the reason scholars value research intensive universities like Michigan so highly.

As one Deep Blue Depositor said,

"When submitting our own datasets, we have a limited view into the repository as a whole, which restricts our own ability to standardize and/or provide information that makes our data easier to find and use by the community. Curators, who see the whole repository as a unified entity, provide that critical perspective."

THE DATA CURATION NETWORK

The University of Michigan has the highest research volume of U.S. public research universities. This research produces a wide range of data sets in a variety of types and formats. Deep Blue Data hosts the span of this data from field observations in notebooks to computer based simulations.

The sheer diversity of the types and formats of the data submitted to Deep Blue Data can make it a challenge to curate. No single data curator can possess the knowledge and expertise needed to work deeply with all kinds of data.



To address this challenge, the data librarians at the University of Michigan came together with librarians from five peer institutions to develop the Data Curation Network (DCN) with support from a grant by the Alfred P. Sloan foundation. Since its launch in 2018, the DCN has expanded from six member institutions to fifteen, including a general data repository (Dryad) and a specialized repository (The Michael J. Fox Foundation for Parkinson's Research).

The DCN has also played an important role in advancing data curation in libraries and in research communities. With support from the Institute of Museum and Library Services (IMLS), the DCN designed and delivered a day and a half long workshop to teach librarians and other information professionals how to curate research data.

In 2021, the DCN was awarded two grants from the National Science Foundation (NSF). The first grant is centered on understanding the challenges and needs of "data communities", groups of researchers from different institutions and disciplines who come together over a need to access data sets of mutual interest. The second is to develop a better understanding of the current data sharing practices of researchers and to identify the costs of these practices.

More information about the DCN can be found on its website:
<https://datacurationnetwork.org/>

ADDING DIVERSITY TO THE DATA MIX

The University of Michigan Library has continued to push the boundaries of what it means to be inclusive. By allowing open access to data, who has access to content begins to change the way we conduct research and build a more diverse data infrastructure. “The way data functions is shifting,” said Rachel Woodbrook, Data Curation Librarian at Michigan Publishing and lead PI. “There’s a lot of change in the way that data is viewed, is cited, is valued, and who is involved in that data.”

Repositories are shifting the narrative of who contributes to and consumes data, by supporting initiatives and projects that allow more voices to be heard. Sharing data outside of projects is a relatively new expectation for researchers. “Additionally, there is a tragic legacy of extracting data from communities without care or concern for the consequences of having done so or the rights of people to have access to information about themselves that could benefit them,” said Rachel.

Through a recent grant from LYRASIS, a consortium that focuses on supporting libraries, Deep Blue has joined forces with the National Center for Institutional Diversity (NCID) to bring about the Open Data Toolkit Based on Diversity, Equity, Inclusion and Access (DEI&A) Principles, a project that provides guidance on best practices in collecting, managing, utilizing, sharing, and curating research data for the public good. This allows researchers to understand the historical and contemporary issues related to identity, difference, culture, representation, power, oppression and inequality as they occur and affect individuals, groups and communities, in order to conduct more ethical research on behalf of the communities involved.

In July of 2021, the first version of the toolkit was made available through the library, compiling resources from a global perspective and from a diverse set of backgrounds, allowing researchers to practice DEI&A values throughout their work. The resource includes books, websites, and articles that can help researchers answer questions and tackle different scenarios when it comes to providing groups with the most empathy, care, and access to the research project. "It's incumbent on all of us to partner and advocate for clear guidelines and get support for researchers when making those decisions," said Rachel. "These are not always easy conversations, but we can help researchers who come to us to assess their data, how to distribute it, and who has access to it."

Open Data DEIA Toolkit 1.0

 Reset

 Share


[Link to full spreadsheet](#)

Linked Title ^	Creator(s)	Description	Why we chose it
Anti-Racism Data Reuse Guide	ICPSR	The purpose of this document is to advance racially equitable data reuse practices for researchers to ensure that analysis does not harm the population(s) being studied or contribute to further marginalization of racialized communities.	This resource focuses on secondary data reuse including evaluating data and analysis methods, and is recently-created.
Data Ethics Canvas	The Open Data Institute	The Data Ethics Canvas is a tool for anyone who collects, shares or uses data. It helps identify and manage ethical issues – at the start of a project that uses data, and throughout. It encourages you to ask important questions about projects that use data, and reflect on their responses.	This tool can be used at multiple points throughout research without feeling redundant, and provides a clear outline through which to approach various parts of the research process.
Equitable Open Data Report	Detroit Digital Justice Coalition and Detroit Community Technology Project	This report has on-the-ground advice for researchers working with data from at-risk/marginalized communities, based on interviews with Detroit residents considering the potential benefits and harms of various data sets currently available on the City's Open Data Portal. The report assesses what actions can be taken by the City to maximize benefits and minimize harms, and investigates open data policies and tools in place in other cities that model their vision for data justice.	Easy to follow advice and general issues to think about; this resource is focused on open government data but includes direct perspectives from community members thinking about their data being used and shared.

Screenshot from the Open Data DEIA Toolkit 1.0, August 23, 2022,
<https://um-deia-data-toolkit.github.io/home/Toolkit.html>

SPOTLIGHT: GLOBAL FEMINISMS PROJECT

Since its inception, Deep Blue's mission has been making the research of University of Michigan faculty, staff, and students available and accessible to the world. "We know through the statistics we gather on downloads that we reach a global audience, and that the vast majority of use comes from outside of the U-M community," said Jim Ottaviani, a librarian with Deep Blue Docs. Recently surpassing 100 million downloads, the continued appetite for U-M's research, scholarship, and creativity is supported by the Michigan community, as well as the world at large.

Deep Blue has continued to support the mission to connect with global communities through its collaborative campus efforts. With its support, the Institute for Research on Women and Gender (IRWG), worked to archive, share, and promote feminist scholarship from an international perspective. The Global Feminisms Project examines the history of feminist activism, women's movements, and academic women's studies in China, India, Poland, and the United States. In collaboration with universities across the globe, and scholars right here at home, the project examines the histories of feminism in local contexts by collecting videotaped oral histories of women's movement activists and women's studies scholars in each country. The project's drive to share feminist knowledge with the world, from individuals, movements, and research from a global perspective, is exactly what drew Deep Blue in.

Professor Abigail Stewart, project co-director, requested the help of Deep Blue in the early stages of the project through the IRWG's library liaison. The project would need a host, one that was capable of a worldwide reach. Deep Blue continues to support the project, assisting in hosting materials and providing video support to the project. "We continue to grow this collection of oral histories by adding new videos and enhanced transcripts regularly," said Jim. As the site grows, the collaboration between Deep Blue and IRWG allows the Global Feminisms Project to address issues that reach across disciplines, bringing together understandings of activism, historical contexts, and aspects of identity formation.

"Supporting researchers wherever and however they work remains a core value of our libraries. For Deep Blue that means making the work of U-M researchers freely available to as many people as possible, and to preserve it so that future researchers can build on what we know today to improve life for everyone."

- Jim Ottaviani

WITH GRATITUDE

**Thank you to the staff of Deep Blue Repositories
and Research Data Services for their impact on our
work and collaboration in this report:**

Jake Carlson
Jim Ottaviani
Martha Stuit
Rachel Woodbrook
Susan Borda

**A special thanks to the members of the Michigan
Publishing Sales, Marketing, and Outreach team
who made this report a reality:**

Danielle Coty-Fattal
Briana Johnson
Jamie Jones
Natasha Robertson
Theresa Schmid
Kristen Twardowski



deepblue.lib.umich.edu
publishing.umich.edu

© 2022 Regents of the University of Michigan

Regents of the University of Michigan

Jordan B. Acker; Michael J. Behm; Mark J. Bernstein; Paul W. Brown;
Sarah Hubbard; Denise Ilitch; Ron Weiser; Katherine E. White; Mark S.
Schlissel (ex officio)

The University of Michigan, as an equal opportunity/affirmative action employer, complies with all applicable federal and state laws regarding nondiscrimination and affirmative action. The University of Michigan is committed to a policy of equal opportunity for all persons and does not discriminate on the basis of race, color, national origin, age, marital status, sex, sexual orientation, gender identity, gender expression, disability, religion, height, weight, or veteran status in employment, educational programs and activities, and admissions. Inquiries or complaints may be addressed to the Senior Director for Institutional Equity, and Title IX/Section 504/ADA Coordinator, Office for Institutional Equity, 2072 Administrative Services Building, Ann Arbor, Michigan 48109-1432, 734-763-0235, TTY 734-647-1388, institutional.equity@umich.edu. For other University of Michigan information call 734-764-1817.

2022