

Data Librarianship & Helping Patrons with Data

with Jo and Lindsay



UNCG Data Librarians



Lindsay Gypin

she/her

Data Services Librarian

Liaison to

Educational Leadership and Cultural Foundations

Teacher Education and Higher Education

Educational Research Methodology

lagypin@uncg.edu



Jo Klein

they/he

GIS & Data Visualization Librarian

Liaison to

Geography Environment and Sustainability

Informatics and Analytics

Computer Science

Nanoscience and Nanoengineering (interim)

ejklein@uncg.edu

Data librarians help patrons...

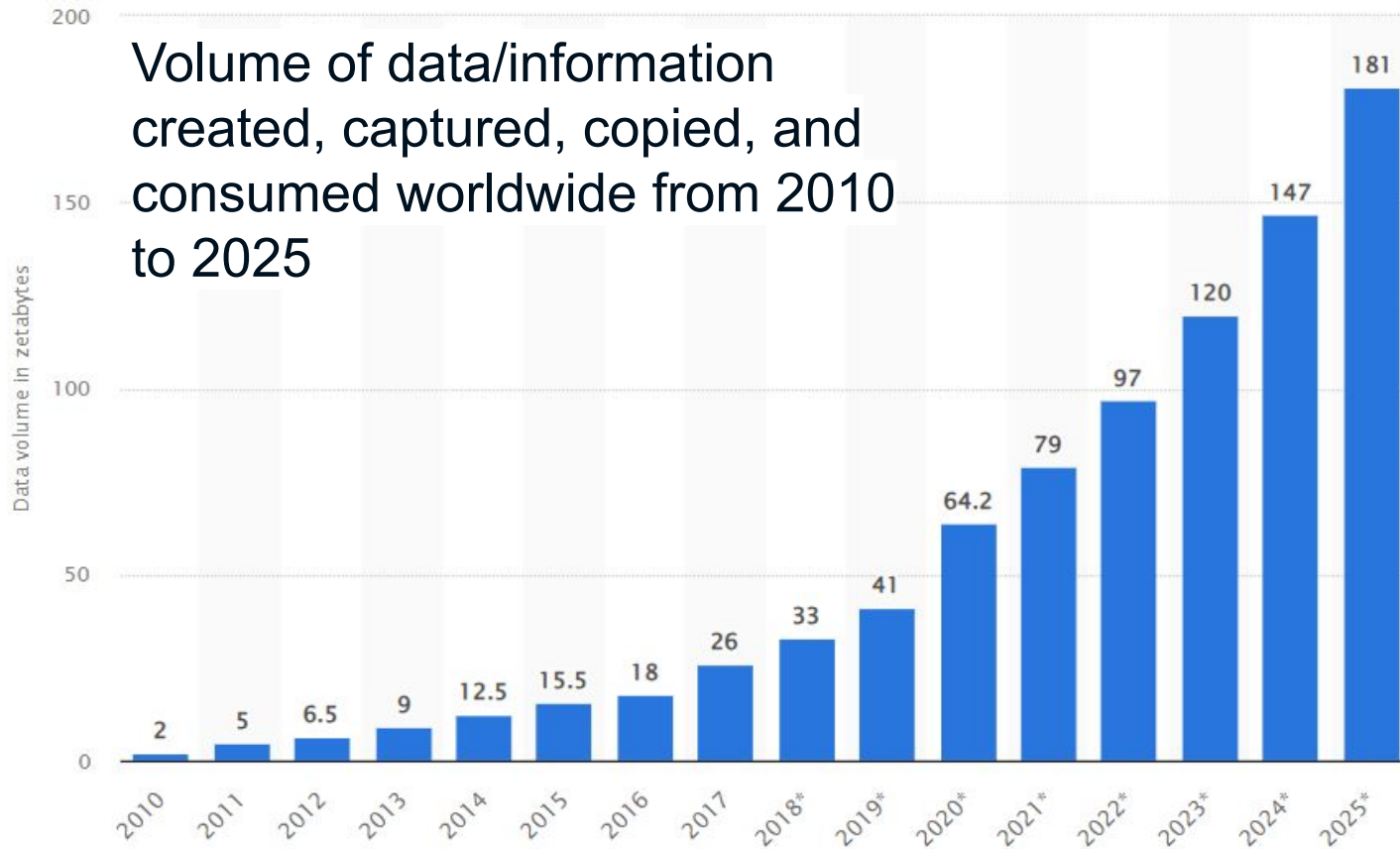
- find and access datasets
- analyze, interpret, and otherwise use those datasets
- organize and manage original research data
- with data ethics and security



Data librarians help patrons...

- find and access datasets
- analyze, interpret, and otherwise use those datasets
- organize and manage original research data
- with data ethics and security





From: <https://www.statista.com/statistics/871513/worldwide-data-created/>

Where does data come from?



Data Lifecycle

Plan - description of data including collection and management

Collect - collecting and digitizing data

Assure - data quality check

Describe - metadata standards

Preserve - submitted to a long-term archive

Discover - data and metadata are located and obtained by interested users

Integrate - combine data from various sources

Analyze - interpret data



Data Lifecycle

Plan - description of data including collection and management

Collect - collecting and digitizing data

Assure - data quality check

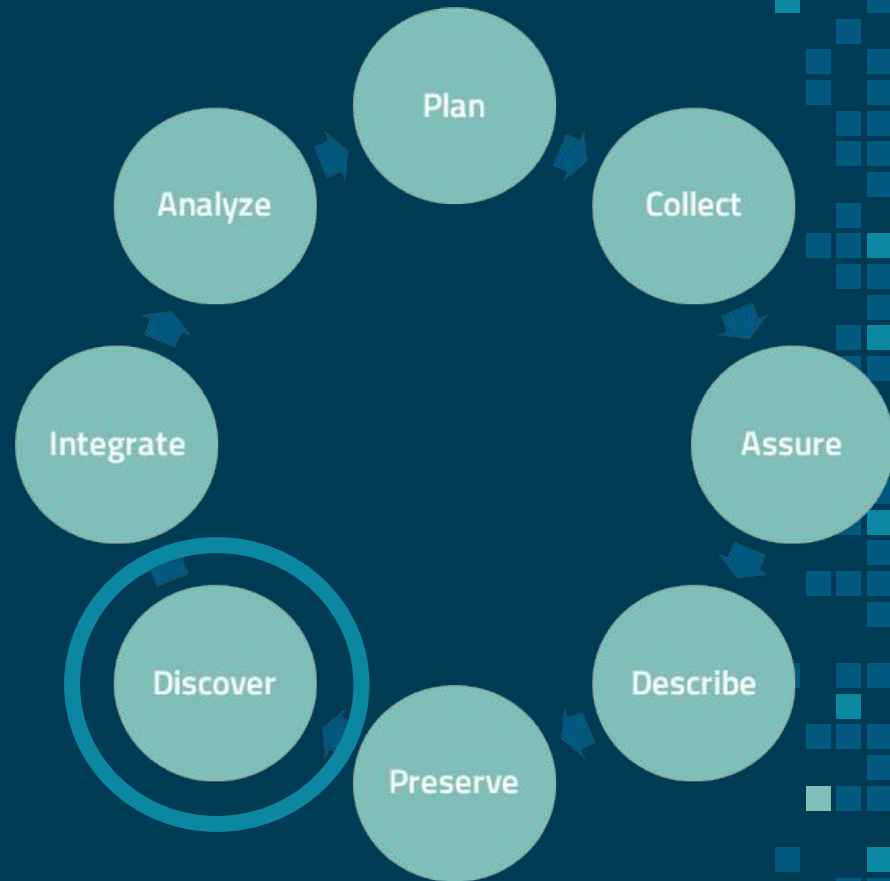
Describe - metadata standards

Preserve - submitted to a long-term archive

Discover - data and metadata are located and obtained by interested users

Integrate - combine data from various sources

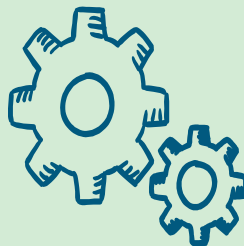
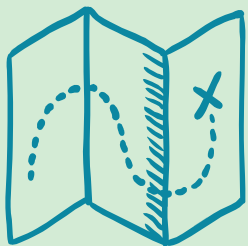
Analyze - interpret data



Secondary Data:

Data collected for a primary project or purpose and then shared for reuse in other, *secondary* projects



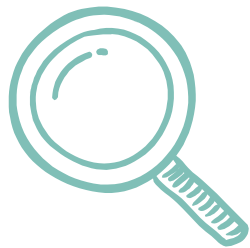


FINDABLE

ACCESSIBLE

INTER-
OPERABLE

REUSABLE



FINDABLE

Where should I
look for data?



Photo by [Cookie the Pom](#) on [Unsplash](#)

How do you normally discover data?

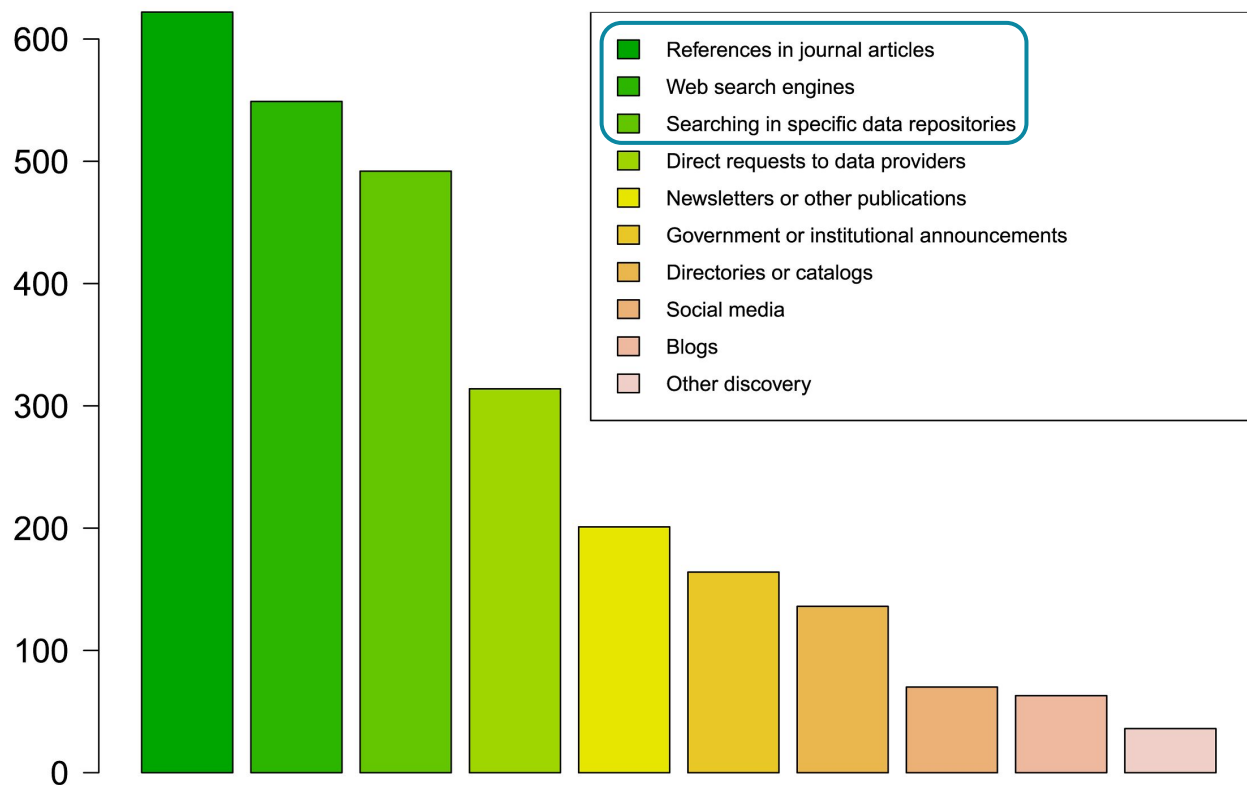
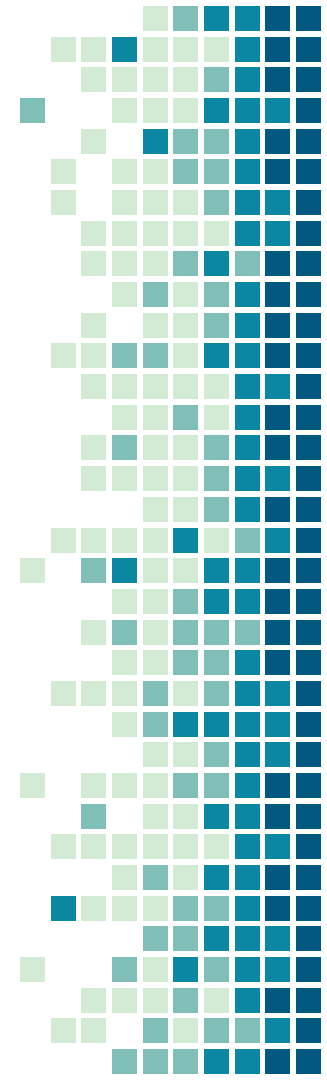


Image: "Fig 9: Discovery of data." How 774 surveyed users and providers of open data in the scope of global environmental change normally discover data, in Schmidt, B., Gemeinholzer, B., & Treloar, A. (2016). Open data in global environmental research: The Belmont Forum's open data survey. *PLOS ONE*, 11(1). <https://doi.org/10.1371/journal.pone.0146695>

Data Repositories

- [Re3data.org](https://re3data.org) - discipline specific data repositories
- [ICPSR](https://icpsr.umich.edu) - social and behavioural science data (and so much more)
- [UNC Dataverse](https://dataverse.unc.edu) - institutional data repository
- [Dryad](https://www.dryad.org) - general data repository (can be institutional)
- [Figshare](https://www.figshare.com) - open and general data repository (can be institutional)

Find more on the [Data Research Guide](#)

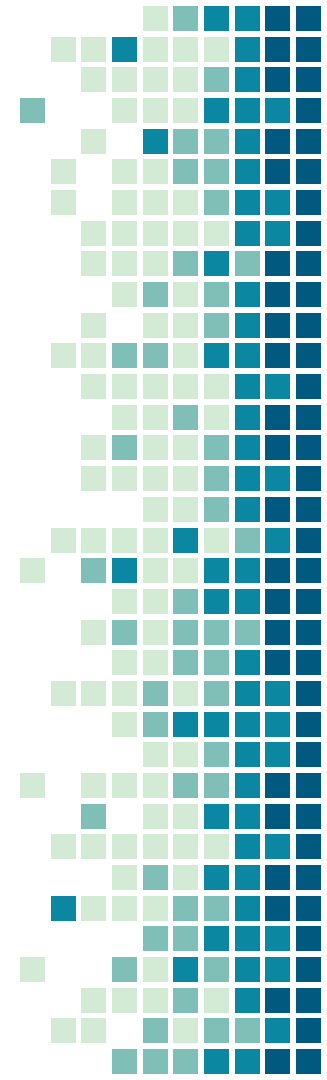


“Web” search engines

Google Dataset Search 

Portals & Browsers

- [Data.gov](https://data.gov)
- [Healthdata.gov](https://healthdata.gov), data.census.gov
- [Library databases](#)
 - PolicyMap, SimplyAnalytics, Social Explorer, Data Planet, and more



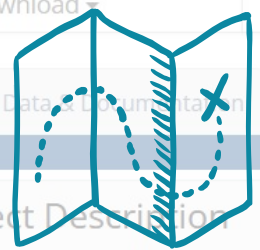
Child Care and Children with Special Needs: Challenges for Low Income Families, Maine, United States, 2002-2005 (ICPSR 27001)

Version Date: [Study](#) | [Share this page](#)

Principal Investigator: [Edmund S. Muskie](#) Institute of Public Affairs; [Lisa A. Morris](#), Edmund S. Muskie Institute of Public Affairs

[https://doi.org/10.31233/osf.io/ztq91.v2](#)

Version V2 (see more versions)



Download

Analyze Online

Access Restricted Data

[At a Glance](#) [Data & Documentation](#) [Variables](#) [Data-related Publications](#) [Export Metadata](#)

Project Description

Summary

ACCESSIBLE

This project is a study of low income families of children with special needs primarily on child care, employment, and balancing work and family. It is a synthesis of existing national and state-level data sets, statewide surveys of parents and child care providers, and a field study to look at these issues at the local level in three selected communities in the state of Maine: Portland, Lewiston/Auburn, and Presque Isle. While the primary focus was on access to child care, this project also looked at the related issues of welfare reform,

4,050
Downloads *
[Usage Report](#)
* past three years

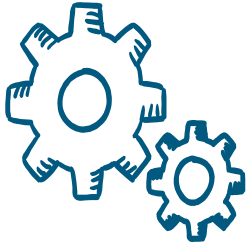
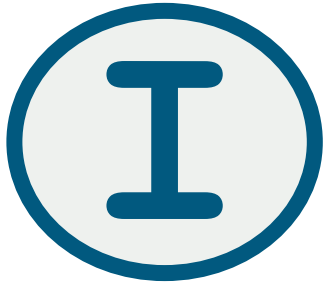
4
[Data-related Publications](#)

Notes

- The public-use data files in this collection are available for access by the general public. Access does not require affiliation with an ICPSR member institution.
- One or more files in this data collection have special restrictions. Restricted data files are not available for direct download.

“ *Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike.*

From *The Open Data Handbook*
by the Open Knowledge
Foundation



INTER-
OPERABLE

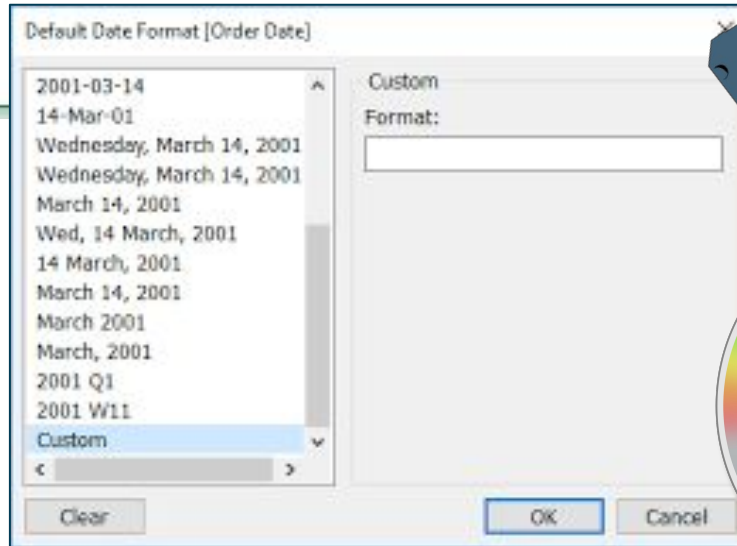
Couldn't Open the Workbook



We're sorry, we couldn't open your workbook. It's possibly corrupt or using a file format that's not supported.

Do you want to try and open this file in Excel?

[Give Feedback](#)





REUSABLE

	A	B	C	D	E	F
1	PROJID	AGE	GENDER	CNTRY	GUILT	DRIVE
2	0002357	3/19/1991	0	06	1	1
	0002357	9/14/1988	0	08	1	
	0002357	1/17/1949	0	12	0	0
	0002357	11/21/1967	1	02	1	3
	0002357	11/20/1997	0	05		2
	0002357	10/12/1951	0	07	3	0
	0002357	1/29/1965	0	10	3	5
	0002357	4/2/1990	1	08	4	4
	0002357	1/4/1989	1			
11	0002357	8/26/1984	0			
12	0002357	4/15/1980	1			
13	0002379	6/20/2001	0			
14	0002379	11/21/1947	1			
	0002379	4/22/1979	0			
	0002379	6/1/1951	0			
	0002379	11/11/1997	1			
	0002379	9/6/1990	1			
	0002379	9/12/1953	0			
	0002379	4/4/1970	1			
21	0002357	7/26/1980	1			
22	0002357	1/3/1999	1			
23	0002357	1/3/1979	0			
24	0002379	5/31/1952	0			

(Spoof data randomly generated by Jo Klein in Google Sheets)

Variable Name	Variable Label	Values: Labels and Codes
PROJID	Project code	7-digit ID
AGE	Age	Date of birth: Mo/day/year
GENDER	Gender	Gender of respondent: 0 = male; 1 = female. Leave blank if missing.
CNTRY	Country born in	Country person is born in: 01 = Argentina; 02 = Bolivia; 03 = Chile; 04 = Colombia; 05 = Cuba; 06 = Ecuador; 07 = El Salvador; 08 = Guatemala; 09 = Honduras; 10 = Mexico; 11 = Nicaragua; 12 = Peru; 13 = Puerto Rico. Leave blank if no data.
GUILT	Feels guilty or sorry	1 = daily or almost daily; 2 = at least once a week, but less than daily; 3 = at least once a month, but less than weekly; 4 = less than once a month; 0 = never. Leave blank if no data.
DRIVE	Drinks and drives	1 = 20 or more days; 2 = 10–19 days; 3 = 6–9 days; 4 = 3–5 days; 5 = 1–2 days; 0 = never. Leave blank if no data.

The data reference interview

- **Who** is the source of the data, i.e. who collected and organized it or published it?
- **What** is your data about, and what data characteristics are you looking for?
- **When** was the data produced, what time period does the data cover, and what else was happening around those times?
- **Where** is the coverage of the data, and where was the data collected and published?
- **Why** was the data collected?
- **How** was the data collected?



Lindsay Gypin

she/her

Data Services Librarian

lagypin@uncg.edu

Schedule a consultation with Lindsay:
uncg.libcal.com/appointments/gypin



Jo Klein

they/he

GIS & Data Visualization Librarian

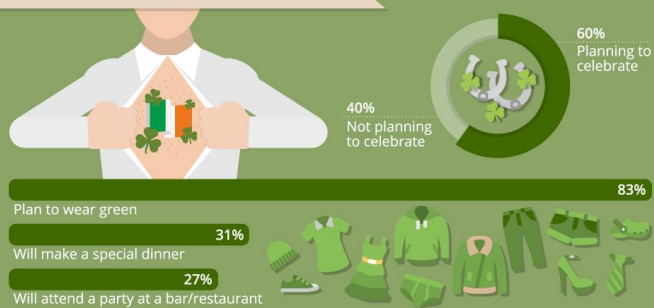
ejklein@uncg.edu

Schedule a consultation with Jo:
go.uncg.edu/joklein-appt

St. Patrick's Day In The U.S.

By the Numbers

How Americans will celebrate their inner Irishness



Some key Guinness facts

Number of countries where Guinness is...



49



150

Seconds needed to pour the perfect pint



119.5 sec.

Calories in a pint



210 cal.

Glasses of Guinness enjoyed worldwide on St. Patrick's Day



13 million

Cases of Irish whiskey consumed in the U.S.

Dye needed to turn the Chicago River green



Key U.S. spending trends

Average planned spending per person

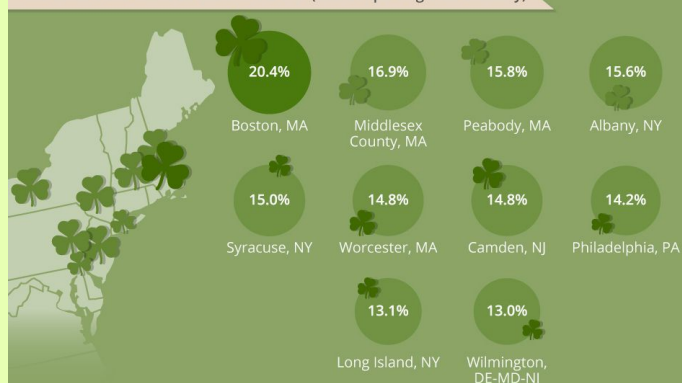
\$37.92 2017

\$39.65 2018

Total planned St. Patrick's Day spending (in billion USD)



The most Irish American cities in the U.S. (share reporting Irish ancestry)



Sources: NRF, IRI, Grocery Headquarters, Marriott Traveler, Food Republic, census.gov, Board Bia

@StatistaCharts

statista