

# Endangered Data Week



How University Libraries Can Get Involved

*Joseph Koivisto*  
*University of Maryland, College Park*

*jkoivist@umd.edu*  
*@Joseph\_Koivisto*

**Where we are**

8 November 2016

# Where we are

- **02 January 2017:** HR 05 - Remove requirement for long-term data analysis for ACA costs<sup>1</sup>
- **24 January 2017:** EPA cuts climate change page from website<sup>2</sup>
- **03 February 2017:** USDA blacks out animal welfare data<sup>3</sup>
- **08 March 2017:** HR 1430 -- Honest and Open New EPA Science Treatment Act<sup>4</sup>

<sup>1</sup> <https://docs.house.gov/billsthisweek/20170102/BILLS-115hres5-PIH-FINAL.pdf>

<sup>2</sup> <https://www.reuters.com/article/us-usa-trump-epa-climatechange-idUSKBN15906G>

<sup>3</sup> <http://www.sciencemag.org/news/2017/02/usda-blacks-out-animal-welfare-information>

<sup>4</sup> <https://www.congress.gov/bill/115th-congress/house-bill/1430>

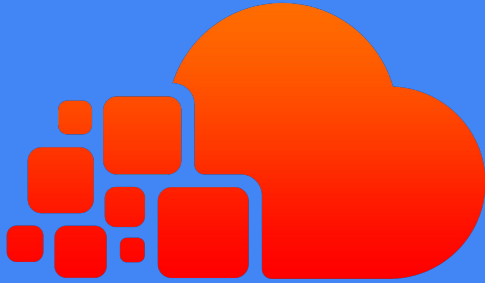
# Where we are



## Coalition Calls for Urgent Investigation of Destruction of Records on Family Separation Crisis

---

**Congress must demand full transparency and accountability to protect children from disappearing**



Endangered  
Data Week

**Enter EDW...**

# Endangered Data Week

Endangered Data Week is a new, collaborative effort, coordinated across campuses, nonprofits, libraries, citizen science initiatives, and cultural heritage institutions, to shed light on public datasets that are in danger of being ...

**Hold on a  
second...**

**What do we  
mean by  
'endangered'?**





- **Deleted**
- **Repressed**
- **Mishandled**
- **Lost**

- **Deleted**
- **Repressed**
- **Mishandled**
- **Lost**
- ***Never collected***

# Endangered Data Week

Promote care for endangered collections by:

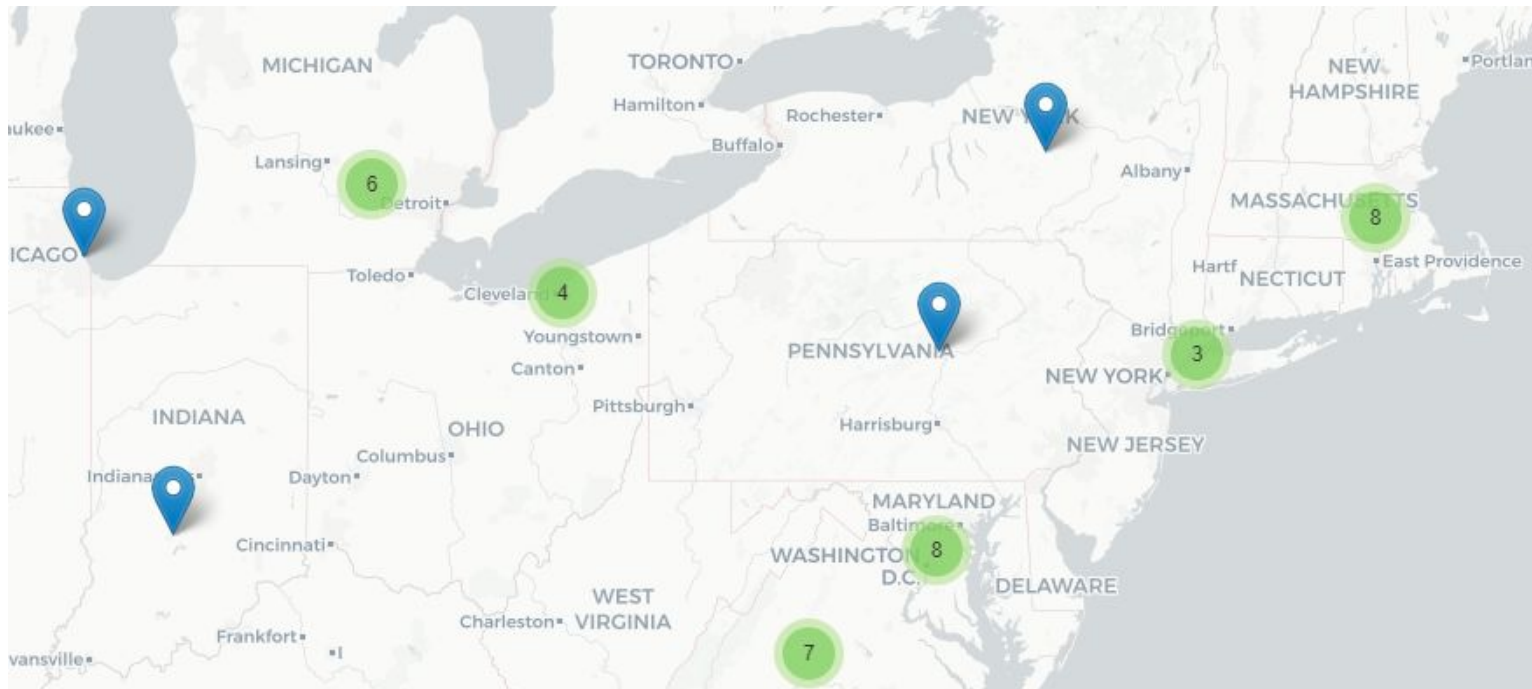
1. publicizing the availability of datasets
2. increasing critical engagement with them, including through visualization and analysis
3. encouraging political activism for open data policies
4. fostering of data skills through workshops on curation, documentation and discovery, improved access, and preservation

**Nota Bene**

**These  
problems  
aren't new**

**What happens  
during  
Endangered Data  
Week?**

# EDW: What happens?



<https://endangereddataweek.org/map/?sorts%5Bdate%5D=1>



# EDW: What happens?

## 1. Publicizing & analyzing public datasets

- a. Organizing and Using Government Data
- b. Data Sharing in Public Health Emergencies
- c. Exploring Chicago's Open Data
- d. Introduction to the City of Charlottesville's Open Data Portal
- e. Open Baltimore: The Impact of Local Data
- f. Census Data: Access, Importance, and the Future
- g. Citizen Science: Health Hackathon

# EDW: What happens?

## 2. Contributing to preservation efforts

- a. Data Refuge<sup>1</sup>
- b. Environmental Data Governance Initiative<sup>2</sup>
- c. Documenting The Now<sup>3</sup>
- d. Internet Archive<sup>4</sup>

<sup>1</sup> <https://www.datarefuge.org/dataset>

<sup>2</sup> <https://enviroidatagov.org/>

<sup>3</sup> <https://www.docnow.io/>

<sup>4</sup> <https://archive.org/web/>

# EDW: What happens?

3. Encourage political activism for open data policies
  - a. Environmental Activism Under the Trump Administration<sup>1</sup>
  - b. Open Data Advocacy & Letter Writing<sup>2</sup>

<sup>1</sup> <https://github.com/endangereddataweek/resources/tree/master/talk-environmental-data-activism-under-trump>

<sup>2</sup> <https://github.com/endangereddataweek/resources/blob/master/resource-advocacy.md>

# EDW: What happens?

## 4. Fostering of data skills through workshops

- |                 |                |
|-----------------|----------------|
| a. Unix shell   | e. Git         |
| b. Tableau      | f. Open Refine |
| c. R            | g. DataLumos   |
| d. Web Scraping | h. Dataverse   |

# EDW: What happens?

## 5. Socializing!

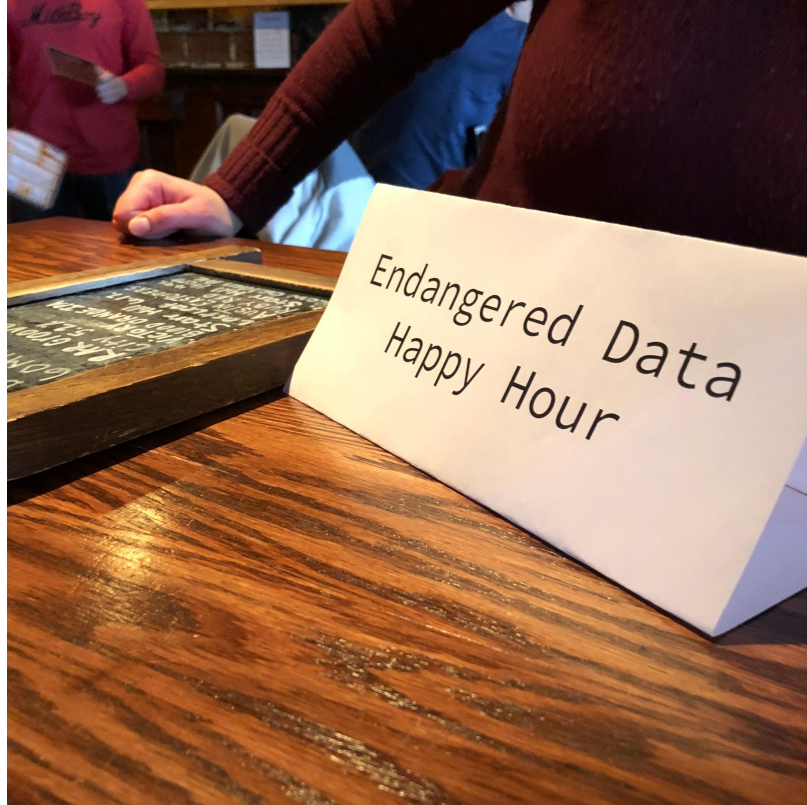


Photo credit: Kate Dohe

**What can you do  
for Endangered  
Data Week  
2019?**

# Think small, think local...

- Avoid becoming overly ambitious
- Think about applicable skills or local data
  - What will grab the attention of those around you
- Seek out collaborations within your institutional and local community
  - What are interdisciplinary connections you can make?
  - Who has expertise that can make your event more well-rounded?
- Be realistic about what you can accomplish with what you've got

# A note on thinking local...

- State, municipal, and local data at *even greater risk* of being endangered
- Infrastructure, laws, mandates, oversight less likely to be well-wrought, mature
- Perhaps a lower bar for EDW engagement?
  - Closer geographically
  - Smaller scale
  - Local relevance



# What you can do

1. Publicizing & analyzing public datasets
  - a. Your data and you!
  - b. Data hackathon: What's in there? What isn't?
2. Contributing to preservation efforts
  - a. Data rescue events (<http://www.ppehlab.org/datarescueworkflow/>)
  - b. Ad hoc preservation

# What you can do

3. Encourage political activism for open data policies
  - a. Teach-ins
  - b. Letter writing
  - c. Advocacy meetings with administrators and legislators
4. Fostering of data skills through workshops
  - a. Good data stewardship starts at home!

# What you can do — Practical steps edition

1. Get started early (no, really, get started early)
2. Doing workshops? Figure out tech now!
3. Review EDW resources page for event ideas & tutorials (<https://github.com/endangereddataweek/resources>)
4. Add events to EDW web page (<https://endangereddataweek.org/>)\*

\* Also a good resource for past events: <https://endangereddataweek.org/archives/>

**Endangered Data:**  
**Less of an event, more of**  
**a lifestyle**

**What are you  
going to do?**

