What the U.S. Digital Analytics Program (DAP) Data Can & Cannot Tell Us

Investigating U.S. Government Website Usage Over Time

ANN MARSHALL, PURDUE UNIVERSITY FORT WAYNE
OLDOOZ DIANAT, REVENUE NSW

May 30th at IASSIST 2019 in Sydney
Overview & Objectives

• Brief Intro to Digital Analytics Program (DAP) website
• Irony of DAP Data
• My project: creating an (initial) DAP Archive
• My data journey with the project
• Data visualizations
• Conclusions
Digital Analytics Program (DAP) Public Website

334,200
people on government websites now

There were 3.63 billion visits over the past 90 days.
Background: About the Digital Analytics Program (DAP)

- [https://analytics.usa.gov/](https://analytics.usa.gov/) went live in 2015
- 60+ agencies and over 5,000 sites
- a project of the General Services Administration (GSA)
- DAP staff provide trainings to agencies
- Help agencies do successful outreach
Does DAP Include All Government Websites?

Digital Analytics Program (DAP)
Last updated May 14, 2019

The Digital Analytics Program (DAP) is a central, shared web analytics service for federal agencies of the United States government, managed by the General Services Administration. Participating federal agencies report web analytics to a central account, and have access to all reported data. Data from the program is also shared publicly at analytics.usa.gov.

Participation in the Digital Analytics Program was last measured on May 14, 2019.

By Domain

<table>
<thead>
<tr>
<th>Domain</th>
<th>Agency</th>
<th>Participates in DAP?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020census.gov</td>
<td>Department of Commerce</td>
<td>Yes</td>
</tr>
<tr>
<td>9-11commission.gov</td>
<td>National Archives and Records Administration</td>
<td>No</td>
</tr>
</tbody>
</table>

https://pulse.cio.gov/analytics/domains/
Irony of DAP Data: Changes Every Five Minutes

236,388

people on government websites now

There were 3.69 billion visits over the past 90 days.

https://analytics.usa.gov/
Irony of DAP Data: Download Options

- Download option: "All pages people are visiting"

- Prior 30 days is the largest archive download option

- Monthly data doesn't list daily data

https://analytics.usa.gov/data/
The Irony of DAP Data: A Snapshot vs. An Archive

“It’s really interesting to see the zeitgeist of what’s happening on government sites ... the snapshot is great, but we’re going to see a lot of value from looking at this data over time and asking, ‘what are the trends?’”

Gwynne Kostin
Digital Services Innovation Center, G.S.A.
https://bits.blogs.nytimes.com/2015/03/20/government-website-analytics/
My project: Create an Initial Archive:
Data Downloaded for 571 Days (at 1 p.m. Eastern)

"All pages people are visiting"
My Data Journey with this Project

• Python – automatically download data daily
• Missing data
• R Studio: combine the files, "data wrangling," reshape the data
• Microsoft Access
• Timesearcher – initial visualization
• Datawrapper – interactive visualization
# Formatting Date Column Before Analysis

1. list.files(path = ".")
2. dbformat <- read_csv("iassist2.csv")
3. library(resize)
4. library(Stringr)
5. dbformatdate_col <- str_sub(dbformat$fileName, 19, 26)
6. dbformatdate_col <- str_replace_all(dbformatdate_col, "", "/")
7. dbformatdate_col <- as.date(dbformatdate_col, "%m/%d/%y")
8. class(dbformatdate_col)
9. dbformatdate_col <- str_sub(dbformat$fileName, 19, 26)
10. dbformatdate_new <- as.date(dbformatdate_col) - 1
11. dbformatdayofwk <- weekdays(dbformatdate_col)
12. write.csv(dbformat, file = "iassist2_formatted.csv")

R (Rstudio)
TimeSearcher

http://www.cs.umd.edu/hcil/timesearcher/
Looking at the Data: Why Data Over Time Matters

What occasion in 2017 was the highest number of website hits on a single day?
NASA’s eclipse livestream drew more than 1M viewers at peak, Digital Analytics Program says

https://www.fedscoop.com/solar-eclipse-livestream-numbers-nasa/
Website visits at 1:00 p.m.

August 15-30, 2017 Visits: NASA | Eclipse Live

Eclipse Live
Aug 21
184,354
### "Eclipse" Pages with Sum of 2017 Visitors

<table>
<thead>
<tr>
<th>page_title</th>
<th>page</th>
<th>sum of active visitors</th>
</tr>
</thead>
<tbody>
<tr>
<td>NASA</td>
<td>nasa.gov/connect/chat/lunar_eclipse.html</td>
<td>72</td>
</tr>
<tr>
<td>Total Solar Eclipse 2017</td>
<td>eclipse2017.nasa.gov/</td>
<td>5887</td>
</tr>
<tr>
<td>Total Solar Eclipse 2017</td>
<td>eclipse2017.nasa.gov/?linkid=39562814</td>
<td>70</td>
</tr>
<tr>
<td>Total Solar Eclipse 2017</td>
<td>eclipse2017.nasa.gov/?mbid=synd_msntravel</td>
<td>42</td>
</tr>
<tr>
<td>2017 Eclipse - Natural Phenomena (U.S.)</td>
<td>nps.gov/subjects/naturalphenomena/2017eclipse.htm</td>
<td>351</td>
</tr>
<tr>
<td>2017 Solar Eclipse</td>
<td>weather.gov/pqr/eclipse</td>
<td>573</td>
</tr>
<tr>
<td>2017 Solar Eclipse</td>
<td>weather.gov/lnk/eclipse_2017</td>
<td>149</td>
</tr>
<tr>
<td>2017 Solar Eclipse</td>
<td>weather.gov/boi/eclipse</td>
<td>81</td>
</tr>
<tr>
<td>2017 Solar Eclipse</td>
<td>weather.gov/cys/eclipse</td>
<td>81</td>
</tr>
<tr>
<td>2017 Solar Eclipse</td>
<td>weather.gov/epz/2017solareclipse</td>
<td>34</td>
</tr>
<tr>
<td>2017 Solar Eclipse</td>
<td>weather.gov/slc/eclipse</td>
<td>32</td>
</tr>
<tr>
<td>2017 Solar Eclipse</td>
<td>w2.weather.gov/pqr/eclipse</td>
<td>10</td>
</tr>
<tr>
<td>2017 Solar Eclipse - Blue Ridge Parkway</td>
<td>nps.gov/blri/planyourvisit/2017-solar-eclipse.htm</td>
<td>26</td>
</tr>
<tr>
<td>2017 Solar Eclipse - Fort Donelson Nat</td>
<td>nps.gov/fodn/planyourvisit/2017-solar-eclipse.htm</td>
<td>11</td>
</tr>
<tr>
<td>2017 Solar Eclipse - Great Smoky Moun</td>
<td>nps.gov/grsm/planyourvisit/2017-solar-eclipse.htm</td>
<td>510</td>
</tr>
<tr>
<td>2017 Solar Eclipse - NWS Riverton, Wy</td>
<td>wy.weather.gov/rw/eclipse</td>
<td>182</td>
</tr>
<tr>
<td>2017 Solar Eclipse - San Juan Island</td>
<td>nps.gov/sajh/planyourvisit/2017-solar-eclipse.htm</td>
<td>16</td>
</tr>
<tr>
<td>2017 Solar Eclipse - Yellowstone Nation</td>
<td>nps.gov/yell/planyourvisit/eclipse.htm</td>
<td>50</td>
</tr>
<tr>
<td>2017 Solar Eclipse Highlights</td>
<td>nasa.gov/feature/2017-solar-eclipse-highlights/</td>
<td>132</td>
</tr>
<tr>
<td>2017 Solar Eclipse Live</td>
<td>nasa.gov/eclipselive/</td>
<td>865507</td>
</tr>
</tbody>
</table>

Record: | < 536 of 536 | ▼ ▲ | ▼ ▲ No Filter | Search |
Visualizations of 2017 Data
Available at: https://guides.library.pfw.edu/dapdot
Visualizations of 2017 Data (missing data)

Available at: https://guides.library.pfw.edu/dapdot

Missing data = 86 days
Downloaded data = 279 days (at 1 p.m.)
Visualizations of 2017 Data: Data Wrapper

Available at: [https://guides.library.pfw.edu/dapdot](https://guides.library.pfw.edu/dapdot)
What Happened on this Day?
9,977 visits on August 30, 2017 at 1 p.m.

Disaster Assistance Website

https://www.disasterassistance.gov/

- National Flood Insurance Program (NFIP): [How to file a claim](https://www.disasterassistance.gov/)
  - Learn more about NFIP.

- COMMUNITY LEADERS
  - Search for resources that may be able to help your community.
  - Search now

Enter your full address to see if your area has been declared for Individual Assistance:

Example Address: 123 Main Street, Washington, DC 55555

Note: Your address and personal data will not be stored.

Find Assistance  Apply Online  Check Status
What Happened on this Day?
1,347 visits on January 25, 2017 at 1 p.m.

https://petitions.whitehouse.gov/

Sign a Petition

Add your name to these petitions and help them reach their goal.

View Petitions With Updates

- **Foreign Policy & Government & Regulatory Reform**: Immediately release Donald Trump's full tax returns, with all information needed to verify emoluments clause compliance. 331,028 SIGNED, 100,000 GOAL

- **Government & Regulatory Reform**: Divest or put in a blind trust all of the President's business and financial assets. 100,078 SIGNED, 100,000 GOAL
What Happened on this Day?
4,014 visits on November 1, 2017 at 1 p.m.

https://www.defense.gov/

IMMEDIATE RELEASE

Nov. 1, 2017
News Release
Press Operations
Release No: NR-379-17

Department of Defense Releases Budget Figure for 2017 Military Intelligence Program (MIP)

Officials Detail DoD Support During Unprecedented Hurricane Season...
https://dod.defense.gov › News ▼

Nov 1, 2017 - Officials Detail DoD Support During Unprecedented Hurricane Season. Defense Department officials told the Senate Homeland Security and Governmental Affairs Committee yesterday how DoD has supported federal, state and local officials in responding to the unprecedented series of hurricanes this year.
Data Visualization: Top Website Visits

2017 Top Website Visits

Max Visitors = highest number of visitors on a single day.
Sum Visitors = total number of visits for all days.

<table>
<thead>
<tr>
<th>Website</th>
<th>Max Visitors</th>
<th>Sum Visitors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welcome</td>
<td>USPS</td>
<td>555,065</td>
</tr>
<tr>
<td>Where’s My Refund? - It’s Quick, Easy and Secure</td>
<td>352,799</td>
<td></td>
</tr>
<tr>
<td>USPS Tracking</td>
<td>351,932</td>
<td></td>
</tr>
<tr>
<td>National Weather Service</td>
<td>351,314</td>
<td></td>
</tr>
<tr>
<td>myUSCIS - Case Status</td>
<td>316,816</td>
<td></td>
</tr>
<tr>
<td>Home</td>
<td>DisasterAssistance.gov</td>
<td>Access to Disaster Help and Resources</td>
</tr>
<tr>
<td>StudentLoans.gov</td>
<td>209,122</td>
<td></td>
</tr>
<tr>
<td>Latest Earthquakes</td>
<td>201,552</td>
<td></td>
</tr>
<tr>
<td>Eclipse Live Stream</td>
<td>Total Solar Eclipse 2017</td>
<td>184,160</td>
</tr>
<tr>
<td>Internal Revenue Service</td>
<td>181,803</td>
<td></td>
</tr>
<tr>
<td>my Social Security Update Message</td>
<td>180,328</td>
<td></td>
</tr>
<tr>
<td>StudentLoans.gov</td>
<td>178,559</td>
<td></td>
</tr>
</tbody>
</table>
Conclusion: Potential of DAP Data Over Time

- High variability in DAP Data Over Time
- Identify important dates for further study
- Important Data for Social Scientists and Historians
- Suggests the need for better government website archiving
Conclusion:
Challenges of DAP Data Over Time

- Only 64% of Government Websites Participate
- "Messy Data" – URLs/pages are ephemeral
- Informal Public Archive Not Sufficient
  - Missing Data
  - Daily totals a better measure
- What archive of DAP (or related) data should we have?
Questions & Discussion

thank you

https://guides.library.pfw.edu/dapdot

ANN MARSHALL, PURDUE UNIVERSITY FORT WAYNE
OLDOOZ DIANAT, REVENUE NSW