# Useful Links and References

### Uncategorized (perhaps general)

- HHS open data metadata feed https://www.healthdata.gov/data.json
- National Committee on Vital and Health Statistics Recommendations to Maximize the Value of HealthData.gov
- JASON Report "Artificial Intelligence for Health and Health Care" December, 2017 calls out need for open data for training
- JASON Report "A Robust Health Data Infrastructure" April, 2014 calls out need for health APIs
- Global Open Data Index
- Data.gov open data app submission form interesting how they used google docs for citizen engagement
- Project Open Data Dashboard metrics at the department level and HHS' report
- CIO.gov Project Open Data
- PLOS Criteria for Recommended Data Repositories
- FAIR data principles Findable, Accessible, Interoperable, Re-usable
- M-13-13, Open Data Policy Managing Information as an Asset
- Supplemental Guidance on the Implementation of M-13-13 "Open Data Policy Managing Information as an Asset" We maintain an Enterprise Data Inventory per the guidance, but it is currently only populated with data sets hosted by Data.CDC.gov or chronicdata.cdc.gov (through Socrata federation)
- CDC/ATSDR Policy on Releasing and Sharing Data
- Texas Open Data Playbook
- · Seattle Open Data Playbook
- · Seattle Open Data Risk Assessment
- GSA Department Source Code Compliance Dashboard
- DKAN Advice on "Tidy Data"
- Library of Congress Recommended Formats Statement

Visualizations found using CDC data (there are tons, this is just a partial list collected from ongoing social media interactions)

• http://flowingdata.com/2016/06/20/who-still-smokes/ \!TODO compare data.gov's JSON feed with healthdata.gov's JSON feed with Data.CDC.gov's json feed

### Repositories and Repository Resources

- National Electronic Injury Surveillance System All Injury Program, 2013 (ICPSR 36693) @ umich.edu
- Nature.com Recommended Data Repositories
- UPitt Health Services Library System Data Management Home
- Data Seal of Approval 16 Requirements (under Dutch Data Archiving and Networked Services) (Now run by CoreTrustSeal)

## Data Engineering

• The AI Hierarchy of Needs - description of data science/AI functions using the metaphor of Maslow's hierarchy of needs.

### CDC collections of data

- www.cdc.gov/data
- ftp://ftp.cdc.gov/pub/Data/
- ftp://ftp.cdc.gov/pub/Health\_Statistics/NCHS/
- Data.CDC.gov (includes chronicdata.cdc.gov through Socrata federation)
- WONDER.cdc.gov

### Terms of Service

GitHub Amended Terms of Service for Gov

Examples of non-CDC sites scraping and hosting CDC data

Abortion Surveillance System 2009-2013 on DataLumos.org rehosts CSV files and MMWR pdfs (without 2014 file) from CDC Reproductive Health web site