

Midwest Big Data Hackathon

October 6-7, 2018

Iowa City IA, USA

bigdata.uiowa.edu



Datasets + Tools

- [Twitter REST API's](#) - Provide programmatic access to read and write Twitter data.
- [Yelp](#) - 2.7M reviews and 649K tips by Yelp's users and user businesses.
- [BestBuy Open Product Data](#) - Programmatic access to BestBuy's product and store data.
- [The New York Times API's](#) - Programmatic access to The New York Times' content. It includes articles, books, events etc.
- [The Guardian API's](#) - Programmatic access to The Guardian's content.
- [GDELT Database](#) - Most comprehensive database of human society.
- [MIT Human Dynamics Lab Datasets](#) - Several mobile data sets that contain dynamics of several communities of about 100 people each.
- [Million Song Dataset](#) - Collection of audio features and metadata for a million contemporary popular music tracks
- [Wikipedia Dumps](#) - Wikipedia database dumps in any language.

- [Google Books Ngram Dataset](#) - The dataset that used in Google Books Ngram Viewer.
- [International Cancer Genome Consortium](#) - Dataset that includes information about cancer patients and their genomes.
- [GeoNames Data](#) - Geographical database covers all countries and contains over eleven million place names.
- [NOAA Ocean CO₂ Dataset](#) - More than 30 years worth of surface ocean CO₂ observations.
- [NOAA Climate Datasets](#) - Datasets that are provided by NOAA on climate.
- [The Movie DB API](#) - Allows programmatic access to TMDb that includes information about movies, tv shows, actors and actresses.
- [US Patent Citation Network](#) - Consists of all the utility patents granted between January 1, 1963 and December 30, 1999. totaling 3,923,922 patents
- [HAZUS](#) - Consists of hazard data, boundary map data, and a proxy for the general building stock (GBS) in the continental United States, Hawaii and the U.S. held Territories.
- [ACT](#) - Test events dataset

Project Ideas

TBD