
Plan Overview

A Data Management Plan created using DMPTool

Title: Analysis of the main focuses of netizens from Weibo topic list and hot search list

Creator: Linzi Huo

Affiliation: Non Partner Institution

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: 47600

Template: NSF-EAR: Earth Sciences

Project abstract:

4.3

Last modified: 12-13-2019

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Analysis of the main focuses of netizens from Weibo topic list and hot search list

Types of data

The Division of Earth Sciences requires that full data sets, derived data products (e.g. model results, output, and workflows), software, and physical collections must be made publicly accessible within two (2) years of final collection.

① 15220091.csv
② 15220091_12.csv
③ 15220091_12.csv

The data type is (.csv)

Use the Octopus Collector to capture the rankings, readings, and tags of Hot Search and Topic Lists on Sina Weibo pages 12 times, and save them in .csv format; For subsequent data cleaning, the format is (.csv)

Data and metadata standards

Standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies).

- ① 15220091.rtf
- ② 15220091_12.csv
- ③ 15220091_12.csv
- ④ 15220091_12.csv

In the secondary folder Weibo topic:

- ① Data topic visualization (.rtf): The content includes the number of times on various topics, the line chart of each category in each category, the average ranking, the average number of readings, and the total reading pictures;
- ② Weibo topic 12 times list (.csv): The content is 12 times Weibo1 content.
- ③ Weibo topic statistics (.csv): 12 times on various topics
- ④ 12-time summary analysis of Weibo topics (.csv): category, number of times, total number of readings, average ranking and average number of readings for various topics of 12 times

In the secondary folder Weibo hot search:

- ① Data search and search (.rtf): The content includes pictures according to the number of hot searches, readings, discussions, and average rankings;
- ② Sorting of hot search data (.csv): 12 categories, times, total readings, average ranking and average readings of various topics.

The i-th data collection of the three-level folder (i = 1,2,3, ..., 12):

The topic general list, Beijing topic local list, hot search list 1, and hot search list 2 are the corresponding categories, names, readings, rankings and other specific content.

Policies for access and sharing

It is the responsibility of researchers and organizations to make results, data, derived data products, and collections available to the research community in a timely manner and at a reasonable cost. In the interest of full and open access, data should be provided at the lowest possible cost to researchers and educators. This cost should, as a first principle, be no more than the marginal cost of filling a specific user request. Data may be made available for secondary use through submission to a national data center, publication in a widely available scientific journal, book or website, through the institutional archives that are standard for a particular discipline (e.g. IRIS for seismological data, UNAVCO for GP data), or through other EAR-specified repositories. Data inventories should be published or entered into a public database periodically and when there is a significant change in type, location or frequency of such observations. Principal Investigators working in coordinated programs may establish (in consultation with other funding agencies and NSF) more stringent data submission procedures.

15220091@bjtu.edu.cn

The data is fully public and can be used by any institution or individual.

The data can be obtained by sending: the subject of the email: hot search topic data on Weibo, and the email address is 15220091@bjtu.edu.cn

Policies and provisions for re-use, re-distribution

Policies and provisions for re-use, re-distribution, and the production of derivatives.

15220091

Anyone and organization can reuse it and redistribute relevant data.

Plans for archiving and preservation of access

