BTAA Geospatial Data Project
Analytics Working Group - Final Report
November 2018

Executive Summary

The Analytics Working Group was initially charged with determining which analytics would be most useful to collect and visualize, examining baseline usage analytics and determining what would count as "success" for the Geoportal (in terms of analytics), as well as generating ideas for promotion in order to increase use of the site. The goal would be to generate recommendations to hand off to committees (Interface, Collection Development, and Communications).

The group brainstormed ideas for what analytics to collect and connected those analytics to specific project goals, strategic directions, and needs from specific committees. Once a list was generated, the group met with the project Application Developer and Tableau experts at UMN to determine what was possible and how difficult specific analytics would be to collect and visualize.

Finally, the group created a weighted priority matrix of analytics in order to make final recommendations.

Analytics

The analytics identified as high priority can be group to answer questions in these categories:
- What is in the collection?
- What is the geographic coverage of the collection?
- Where are the users geographically located?
- By user location, what geographies are of interest?
- What collection materials are of most interest to users?
- What collection materials are users downloading?

What is in the collection?

Priority analytics here include being able to create charts showing a filtered output of data records by metadata facets – specifically Genre, Institution, Collection, Time Period, Subject, and Place. The ability to select from multiple categories and multiple values within categories would also be critical.

What is the geographic coverage of the collection?

This would be achieved by a map showing the bounding boxes of data records in the collection. In addition, being able to filter data records by metadata facets – specifically Genre, Institution, Collection, Time Period, Subject, and Place – would also be critical.

Where are the users geographically located?

Priority analytics here would be the ability to create charts showing the geographical location of users by city or state. If possible, a map output where the geographic region (city or state) indicates the number of users. Additionally, being able to filter data records by metadata facets – specifically Genre, Institution, Collection, Time Period, Subject, and Place – would also be helpful.

By user location, what geographies are of interest?

This analytics would ideally give project team members the ability to select a geographic place and then see a chart showing the location (city/state) of the users who visited those data record pages.

What collection materials are of most interest to users?

Similar to the "What is in the collection?" category, but using Page View data rather than the entire collection. Priority analytics here include being able to create charts showing a filtered

output of data records by metadata facets – specifically Genre, Institution, Collection, Time Period, Subject, and Place. The ability to select from multiple categories and multiple values within categories would also be critical.

What collection materials are users downloading?

The ability to see what is being downloaded is the most important analytic here. In addition, the ability to create charts showing a filtered output of data records by metadata facets – specifically Genre, Institution, Collection, Time Period, Subject, and Place. The ability to select from multiple categories and multiple values within categories would also be critical.

What other analytics information should we be collecting?

In Google Analytics, the project currently collects Event information for Downloads and Spatial Search. The group recommends adding More Details Links, Metadata button, Web Services button, and the Facet "more »" link to the Event data we currently collect.

Priorities

Recognizing that not all of the analytics will be available at first, and in order to give the Application Developer and Tableau Expert a sense of priorities, this is the priority order, as determined by the Weighted Priority Matrix:

1. Where are the users geographically located?
2. What collection materials are of most interest to users?
3. What is in the collection?
4. What is the geographic coverage of the collection?
5. What collection materials are users downloading?
6. By user location, what geographies are of interest?

In addition, the group recommends adding the analytics Event information listed above.

Connections to Goals & Committees

While the analytics may be utilized by any project member or committee, the group identified some specific connections between the analytics categories and existing goals and committees:

- Where are the users geographically located?
  - Communications Committee
  - Collection Development Committee
  - Sustainability Goal

- What collection materials are of most interest to users?
  - Communications Committee
  - Collection Development Committee
  - Sustainability Goal
- What is in the collection?
  - Collection Development Committee
  - Sustainability Goal
- What is the geographic coverage of the collection?
  - Communications Committee
  - Collection Development Committee
- What collection materials are users downloading?
  - Communications Committee
  - Collection Development Committee
- By user location, what geographies are of interest?
  - Communications Committee
  - Collection Development Committee
- What other analytics information should we be collecting?
  - Interface Committee
  - Metadata Committee

Future Work

In addition to the analytics mentioned above, the group saw value in collecting some form of analytics around session duration and "bounce rates on home page vs search results page vs item pages." These fall under the category of "Can users find what they need?" and could be achieved through analytics analysis or by the use of some other tool, such as Crazy Egg (for session duration) or a user survey.

In the time allotted, the group was not able to determine what would count as "success" for the Geoportal (in terms of analytics) or generate ideas for promotion in order to increase the use of the site – the group recommends that the Steering Committee charge another short-term group or pass this work on to the relevant committee(s).

Documents

Analytics Weighted Prioritization Matrix
Meeting Notes
Analytics & Goals
Analytics Brainstorming