

Exploring the Benefits for Users of Linked Open Data for Digitized Special Collections

Google Analytics Data Summary

Jacob Jett

1 Introduction

In addition to one-on-one user interactions and a planned focus group, additional assessment methods, i.e., site traffic data gathered from Google Analytics and test queries using Google's search engine, were used to produce supplementary benchmark data. The sections below summarize the facts observed from these two data collections.

2 Summary of Google Analytics for Motley and Portraits of Actors Collections

The University Library at the University of Illinois at Urbana-Champaign maintains an extensive super-collection of digital special collections. These digital special collections are accessed and maintained through an instance of OCLC's CONTENTdm collection management system. There are 24 collections altogether containing 20,169 digital items. The Motley Collection of Theatre and Costume Design and the Portraits of Actors collection are the two largest collections out of this super-collection and together account for 7,579 of the digital items¹ (or about 37.6% of the total number items in the super-collection).

Google Analytics was used to monitor Web traffic to and usage of the collections' website maintained by the University Library for a three month period starting in September 2016 and ending in December 2016. During this time the combined total activity across all of the digital collections is as follows:

- Total Pages Accessed: 9,826
- Total Pageviews: 26,981
- Total Unique Pageviews: 20,326
- Total Referrals: 7,454
- Average Time on Page: 55.75 seconds
- Top 5 referrers: images.library.uiuc.edu (526), reddit.com (379), google organic search result A (195), google organic search result B (146), google organic search result C (146)

Activity for the Motley and Portraits of Actors collections in particular during this time period was:

- Total Pages Accessed: 1,294 (Motley – 977, PoA – 317)
- Total Pageviews: 2,473 (Motley – 1,935, PoA – 538)
- Total Unique Pageviews: 1,944 (Motley – 1,497, PoA – 447)
- Total Referrals: 489
- Average Time on Page: 35.39 seconds (Motley – 31.79 seconds, PoA – 46.49 seconds)
- Top 5 referrers: images.library.uiuc.edu (47), images.library.illinois.edu (24), oclc.org (18), google organic search results (15), publish.illinois.edu (11)

From this data we can see that the Motley and Portraits of Actors collections account for:

- 13.2% of all pages accessed,
- 9.2% of total pageviews,

¹ The Motley Collection consists of 4,085 digital items and the Portraits of Actors collection consists of 3,494 digital items.

- 9.6% of total unique pageviews,
- 6.5% of all referrals, and
- that users spent 36.5% less time on their pages than other collections' pages.

When compared directly to one another, the Google Analytics data indicates that:

- Motley Collection pages are accessed three times more often than Portraits of Actors pages,
- Motley Collection pages are viewed three and half times more frequently than Portraits of Actors pages,
- The Motley Collection has three and a third times more unique pageviews than the Portraits of Actors collection, but
- Users spend 46.2% longer on Portraits of Actors collection pages than they do on Motley Collection pages.

In summary, despite accounting for over a third of the University Library's digital special collections, both the Motley Collection and the Portraits of Actors collection account for a relatively small slice of overall usage of those special collections. It is also clear that while users interact with digital items from the Motley Collection more frequently than those of the Portraits of Actors collection, they tend to have longer interactions with digital items from the Portraits of Actors collection. In all this baseline dataset provides a benchmark against which we will be able to observe any measurable variances in Web traffic attributable to the deployment of Linked Open Data resources for the Motley and Portraits of Actors collections during the next phase of this study.

3 Summary of Google Test Queries

From the Google Analytics data, we know that users do access resources using popular search engines like Google, Bing, and Yahoo. Based on the pages these referrals land on it is highly likely that almost all of these result from "known-item searches" and are not serendipitous discoveries. To establish a baseline of how well collection and item pages rank in various Google search results pages, we conducted a series of seven searches using both Google and Google Images. The results are listed in the following tables.

Table 1: Google Search Result Ranking

| Query String | Page Rank |
|--|--|
| "Othello costume design" | Appears on page 2 of results |
| "motley costume designs" | Appears at top of page 1 of results |
| "hamlet costumes" | Results for library on page 3 but no pertinent results in first 10 pages |
| "soldier and his wife" | No pertinent results in first 10 pages |
| "soldier and his wife costumes" | No pertinent results in first 10 pages |
| "soldier and his wife costume designs" | Appears at top of page 1 of results |
| "grocer costume designs" | Appears at top of page 1 of results |

Table 2: Google Image Search Result Ranking

| Query String | Page Rank |
|--|---|
| "Othello costume design" | No pertinent results in first 5 screen-length scrolls |
| "motley costume designs" | No pertinent results in first 5 screen-length scrolls |
| "hamlet costumes" | No pertinent results in first 5 screen-length scrolls |
| "soldier and his wife" | No pertinent results in first 5 screen-length scrolls |
| "soldier and his wife costumes" | No pertinent results in first 5 screen-length scrolls |
| "soldier and his wife costume designs" | No pertinent results in first 5 screen-length scrolls |
| "grocer costume designs" | No pertinent results in first 5 screen-length scrolls |