Tumblr Image Bot
a friendly social-media robot

Bill Levay
LIS 664 | Fall 2014

Why?

- The ARChive of Contemporary Music website features many image galleries depicting items from the collection

- Since launching the site in May 2014, web traffic to the galleries has been low

- The ARC's social media posts have relatively low reach and low engagement (e.g., avg interaction per tweet = 1)

Page path level 1	Pageviews ↓	Unique Pageviews	Avg. Time on Page	Bounce Rate	% Exit
	12,387 % of Total: 100.00% (12,387)	**5,123** % of Total: 100.00% (5,123)	**00:00:36** Site Avg: 00:00:36 (0.00%)	**0.65%** Site Avg: 0.65% (0.00%)	**22.41%** Site Avg: 22.41% (0.00%)
1. /	**3,344** (27.00%)	1,350 (26.35%)	00:00:31	0.69%	24.34%
2. /blog/	**2,009** (16.22%)	871 (17.00%)	00:00:37	0.50%	33.90%
3. /catalogs/	**1,638** (13.22%)	661 (12.90%)	00:00:39	0.00%	18.99%
4. /about/	**1,478** (11.93%)	608 (11.87%)	00:00:44	0.92%	16.10%
5. /galleries/	**1,223** (9.87%)	521 (10.17%)	00:00:29	0.76%	15.62%
6. /support/	**578** (4.67%)	242 (4.72%)	00:00:54	1.15%	22.15%
7. /services/	**456** (3.68%)	180 (3.51%)	00:00:41	0.00%	14.04%
8. /news/	**295** (2.38%)	121 (2.36%)	00:00:33	3.57%	14.24%
9. /category/	**220** (1.78%)	95 (1.85%)	00:00:21	0.00%	6.36%
10. /tag/	**195** (1.57%)	81 (1.58%)	00:00:18	0.00%	24.62%

Google Analytics snapshot | Sept. 1-30, 2014

Page path level 2 ⍰	Pageviews ⍰ ↓	Unique Pageviews ⍰	Avg. Time on Page ⍰	Bounce Rate ⍰	% Exit ⍰
	1,223 % of Total: 9.87% (12,387)	**521** % of Total: 10.17% (5,123)	**00:00:29** Site Avg: 00:00:36 (-18.79%)	**0.76%** Site Avg: 0.65% (16.84%)	**15.62%** Site Avg: 22.41% (-30.31%)
1. □ /	**253** (20.69%)	89 (17.08%)	00:00:23	0.00%	6.32%
2. □ /blues/	**154** (12.59%)	65 (12.48%)	00:00:47	0.00%	19.48%
3. □ /cha-cha-cha/	**140** (11.45%)	65 (12.48%)	00:00:07	0.00%	16.43%
4. □ /almost-beatles/	**116** (9.48%)	51 (9.79%)	00:00:40	0.00%	18.10%
5. □ /45-centers-inserts/	**100** (8.18%)	42 (8.06%)	00:00:23	0.00%	7.00%
6. □ /adopt/	**90** (7.36%)	40 (7.68%)	00:00:29	0.00%	13.33%
7. □ /la-punk-posters/	**88** (7.20%)	43 (8.25%)	00:00:57	3.85%	30.68%
8. □ /the-first-12-sun-lps/	**80** (6.54%)	34 (6.53%)	00:00:13	0.00%	12.50%
9. □ /space/	**74** (6.05%)	33 (6.33%)	00:00:17	0.00%	24.32%
10. □ /bellydance/	**64** (5.23%)	30 (5.76%)	00:00:20	0.00%	29.69%
11. □ /pop-music-pulp/	**36** (2.94%)	15 (2.88%)	00:01:07	0.00%	13.89%
12. □ /adopt	**26** (2.13%)	13 (2.50%)	00:00:46	0.00%	11.54%
13. □ /arc-graphics/	**2** (0.16%)	1 (0.19%)	00:02:18	0.00%	0.00%

Google Analytics snapshot | Sept. 1-30, 2014

A Potential Solution

By repurposing fun & quirky digital content in the context of social media, perhaps we can better engage followers, attract new users, and drive new traffic to the site, potentially attracting new donors.
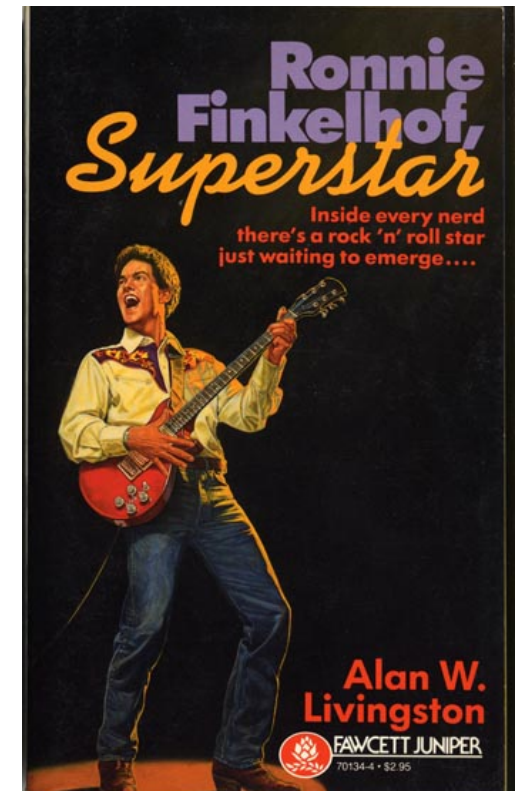
A Python Script, in Two Parts

- Web Scraper: Using Beautiful Soup, scrape each gallery (generated, in this case, by the popular WordPress plugin Next-Gen Gallery) and write image data to a JSON file

```html
<!-- Thumbnails -->
        <div id="ngg-image-0" class="ngg-gallery-thumbnail-box" >
            <div class="ngg-gallery-thumbnail">
<a href="http://arcmusic.org/wp-content/gallery/books/platter.jpg"
   title="The Golden Platter / by Joe Green"
   data-src="http://arcmusic.org/wp-content/gallery/books/platter.jpg"
   data-thumbnail="http://arcmusic.org/wp-content/gallery/books/thumbs/thumbs_platter.jpg"
   data-image-id="256"
   data-title="The Golden Platter / by Joe Green"
   data-description="The Golden Platter / by Joe Green"
   class="ngg-fancybox" rel="475b5aa03bac617c80d1f1ec5036f6fc">
 <img
     title="The Golden Platter / by Joe Green"
     alt="The Golden Platter / by Joe Green"
     src="http://arcmusic.org/wp-content/gallery/books/thumbs/thumbs_platter.jpg"
     width="220"
     height="369"
     style="max-width:none;"
    />
 </a>
</div>
                                    </div>
```

A Python Script, in Two Parts

- Tumblr Bot: Using Pytumblr, a Python Tumblr API client, build and send *x* photo posts to Tumblr along with appropriate caption text and tags
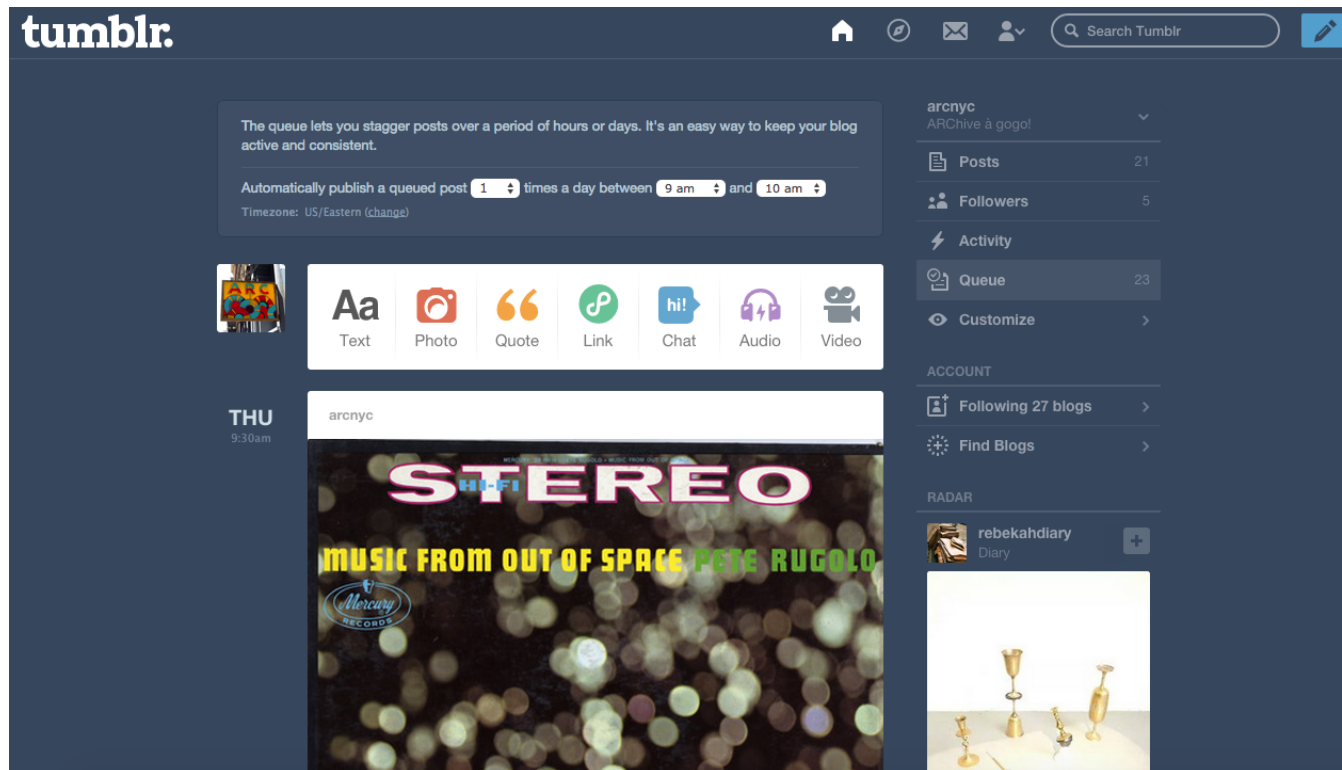
Why Tumblr?

- It's free

- Many themes support a photo-gallery style layout

- Built-in queue with post scheduler

- Can serve as a social media hub – Tumblr posts can easily beget Twitter and FB photo posts

Challenges

- Getting Pytumblr to install successfully was difficult

- After writing the scripts specifically with the ARChive of Contemporary Music in mind, I went back and moved all ARC-specific data into a separate settings file

- This leaves the Tumblr-post code clean and generic – in theory, someone else could use this code as their own Tumblr bot, repurposing their own images

- This process was much more difficult for the web scraper script

Demo Time

Future Development

- Continue "abstracting" the code

- Finish the README.md documentation

- Continue analysis on ARChive web traffic and social media engagement

- Systematically change variables in social media posts (time of day, caption text, tags/hashtags) and observe effect

- Set up web scraping "profiles" for other popular image-gallery generators, e.g., CONTENTdm, Omeka, Flickr, Tumblr itself?