Punkavists
(aka DC Punk web archive group #2)

Thanks DCPL!!!!

Chase Dooley, Roger Gillis, Rachel Trent, Jennifer Weintraub
Initial questions

What subgenres of punk are represented in the collection? (ie: Queercore, Garage Punk, Skate Punk?)

What can we learn about concerns of the DC scene?

Connections between bands/members/venues (potentially visualize them)?
Q1: What's in this archive? (48GB)

Insider/curator knowledge

Archive-It public content

Archive-It backend reports

Live sites
Q1: What’s in this archive? - AUT Jupyter Notebook

Top domains by count.

- proboards
- bandcamp
- facebook
Q1: What's in this archive? - Gephi

- various bandcamp sites
- proboards = dchardcore.proboards.com
- 930 = dchardcore.proboards.com
- forum.930.com
Q1: What’s in this archive?- AUT Jupyter Notebook

Lots of forum-y words: Post, Title, 2016, like... We wanted to see the rest of this list!
Q1: What's in this archive? - AUT Jupyter Notebook
Rough outlines

Two discussion boards: proboards and forum.930.com

30+ bands’ bandcamp sites

All captured in 2018
Jupyter Notebooks (Thanks, Chase!)

Tag/Genre Analysis of Washington DC Punk Bandcamp sites
God save the domains!
Biggest question: genres in bandcamp tags
Filtering out tags

```
In [125]: tags = dict()
   for k,v in band_tags.items():
       tags[k] = list(set([re.split('about', tag.replace('washington', '').replace('d.c.', '').replace('tags', '').strip() for tag in v]))

In [132]: list(tags.values())[0]
```

```
Out[132]: ['jurassic park punk anti-folk folk pop folk punk goof punk',
                'punk anti-folk dc hardcore folk pop folk punk goof punk hardcore punk',
                'punk anti-folk folk pop folk punk goof punk']
```
"I love a good wordcloud!"
(paraphrasing Ian Milligan)
Top 13 words from 930 Forum (minus stop words)

post title like julian just semi retired wunderkind
whole logged numbers rolling guide integers compendium

See the full list: go.gwu.edu/punk
Some interesting words from 930 Forums

Julian (#9!

Beer (#24!)

Compendium (#20)

Excontradiction (#355) (username)
it's Julian! and compendium!
To Get 930 text

We regenerated full text derivatives just with 930 domain without the boilerplate material from the WARCs. This enabled a cleaner dataset. We got a word frequency list using Python, which we then sorted.
What we’d look at if we had another day (or 10)

Proboards forum (& compare to 930.com)
Topic modeling on forums
Dig into sub-forums & punk sub-genres
Timeline of shows