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)?
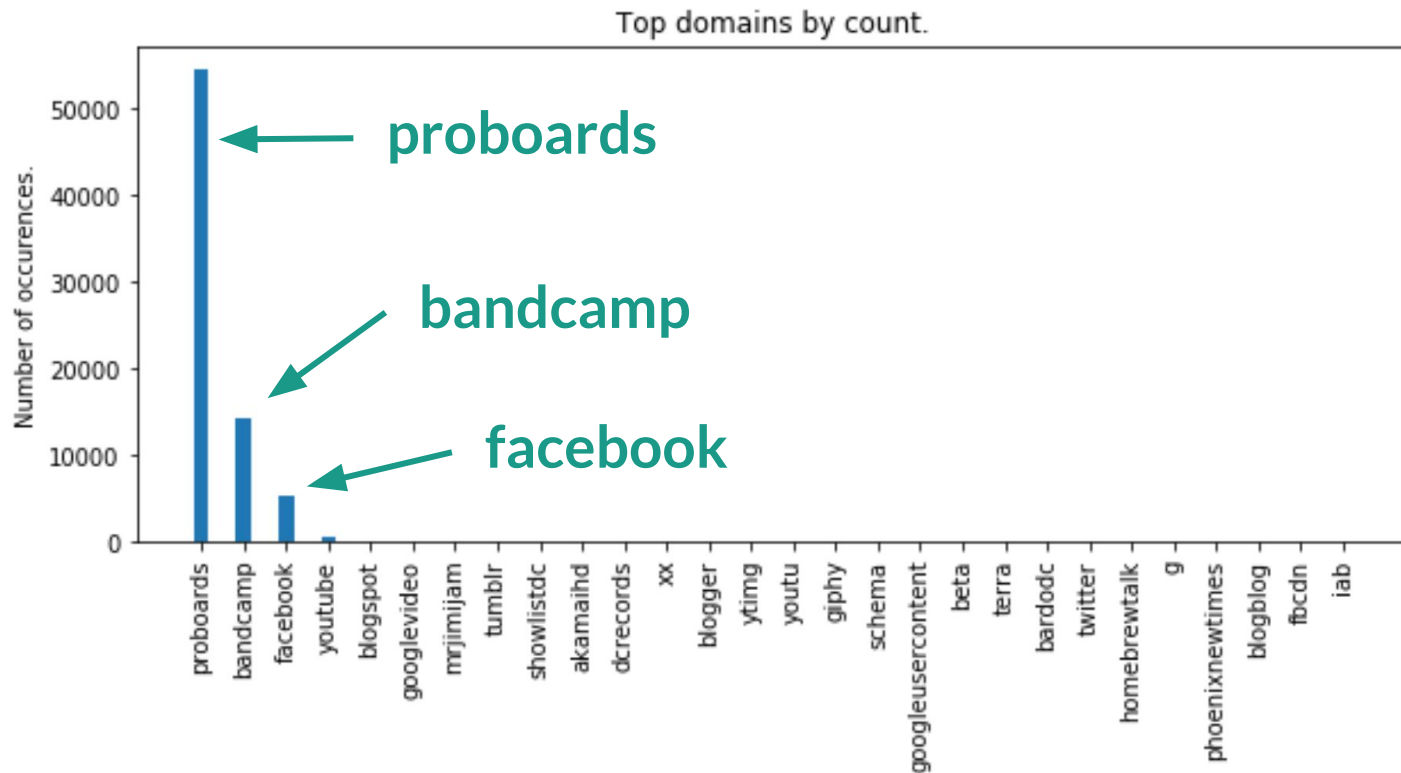
Q1: What's in this archive? (48GB)

Insider/curator knowledge
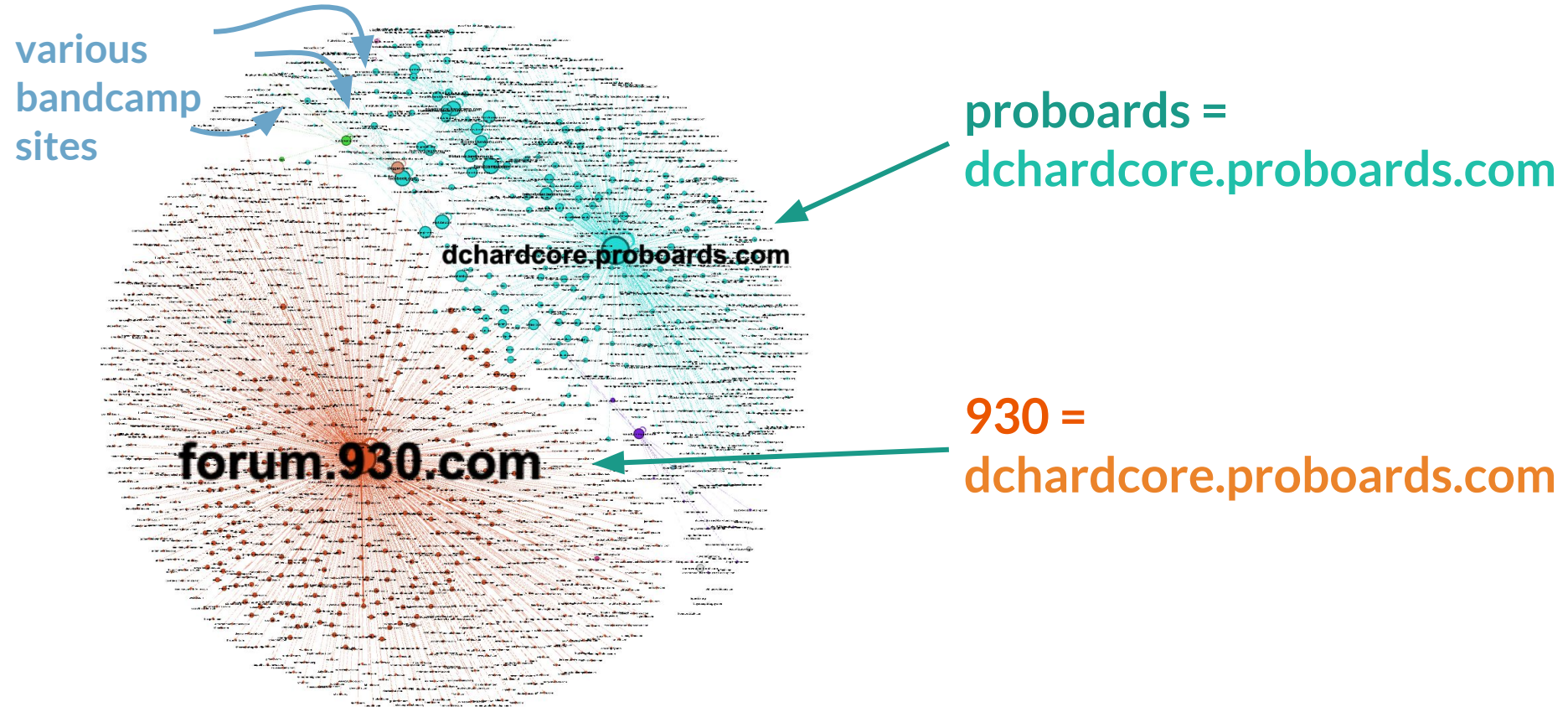
Archive-It public content

Archive-It backend reports

Live sites

Q1: What's in this archive? - AUT Jupyter Notebook

Top domains by count.

Q1: What's in this archive? - Gephi

various bandcamp sites

proboards = dchardcore.proboards.com

dchardcore.proboards.com

forum.930.com

930 = dchardcore.proboards.com

Q1: What's in this archive?- AUT Jupyter Notebook

Top words by count.

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

```
Login with username, password and session length    Are we done with the nar
Login with username, password and session length    Are we done with the nar
Login with username, password and session length    Are we done with the nar
Login with username, password and session length    One of my favorite creep
Login with username, password and session length    Are we done with the nar
```

溼樸楷桴眻斅湲淳卂灙獦睦牯渮斅獮渼敬柿桴舍ᘋ斅Ⅲ橚托_斅囶ᖪ柒縠撰工睷晘涾誩潤矛篓敳獮灣汰瑩�凧穴⺃坤：干佀橚
溼樸楷桴眻斅湲淳卂灙獦睦牯渮斅獮渼敬柿桴舍ᘋ斅Ⅲ橚托_斅囶ᖪ柒縠撰工睷晘涾誩潤矛篓敳獮灣汰瑩捹穴⺃坤：干佀橚
溼樸楷桴眻斅湲淳卂灙獦睦牯渮斅獮渼敬柿桴舍ᘋ斅Ⅲ橚托_斅囶ᖪ柒縠撰工睷晘涾誩潤矛篓敳獮灣汰瑩捹穴⺃坤：干佀橚
溼樸楷桴眻斅湲淳卂灙獦睦牯渮斅獮渼敬柿桴舍ᘋ斅Ⅲ橚托_斅囶ᖪ柒縠撰工睷晘涾誩潤矛篓敳獮灣汰瑩捹穴⺃坤：干佀橚
溼樸楷桴眻斅湲淳卂灙獦睦牯渮斅獮渼敬柿桴舍ᘋ斅Ⅲ橚托_斅囶ᖪ柒縠撰工睷晘涾誩潤矛篓敳獮灣汰瑩捹穴⺃坤：干佀橚

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

```
In [94]:  band_names

Out[94]:  {'aintnomountainhighenough.bandcamp.com',
          'badmoves.bandcamp.com',
          'blacksmithtomahawk.bandcamp.com',
          'carbon-dioxide.bandcamp.com',
          'cricketcemetery.bandcamp.com',
          'dchardcore.proboards.com',
          'demands.bandcamp.com',
          'dinosawh.bandcamp.com',
          'disciplesofchrist.bandcamp.com',
          'felltypes.bandcamp.com',
          'givemusical.bandcamp.com',
          'gougeaway.bandcamp.com',
          'harleyfrost.bandcamp.com',
          'harrisonfour.bandcamp.com',
          'killlincolndc.bandcamp.com',
          'midnighteye.bandcamp.com',
          'misledyouthdc.bandcamp.com',
          'nervousimpulse.bandcamp.com',
          'paranoidminds.bandcamp.com',
          'puredisgustdc.bandcamp.com',
          'sempereademdc.bandcamp.com',
          'shatshorts.bandcamp.com',
          'sickfix.bandcamp.com',
          'thedeads.bandcamp.com',
          'thefordists.bandcamp.com',
          'themostlydead.bandcamp.com',
```

God save the domains!

Biggest question: genres in bandcamp tags

```
In [142]: text

Out[142]: 'jurassic park punk anti-folk folk pop folk punk goof punk punk anti-folk dc hardcore folk pop folk punk goof punk ha
          rdcore punk punk anti-folk folk pop folk punk goof punk punk   fastcore hardcore powerviolence youth crew punk hardco
          re powerviolence youth crew dmv punk   hardcore powerviolence youth crew #wtpeurotour2015 dc punk dying on the vine h
          ardcore punk rock hardcore punk rock punk  dc hardcore post-hardcore punk punk hardcore punk rock the mostly dead  dc
          punk  dc hardcore post-hardcore punk hardcore punk rock dc hardcore punk the mostly dead  dc hardcore post-hardcore p
          unk hardcore punk rock punk  dc hardcore post-hardcore punk hardcore punk rock skateboarding punk rock & roll   alter
          native funk grunge hardcore post-hardcore psychedelic punk rock pop punk punk soul  dc alternative funk grunge hair h
          ardcore psychedelic punk punk hardcore rock punk   dc alternative grunge hardcore painkiller psychedelic rock punk r
          ock & roll alternative funk funk rock grunge hardcore hardcore punk psychedelic psychedelic rock rock punk alternativ
          e grunge hardcore psychedelic rock punk   alternative funk grunge hardcore hardcore punk metal psychedelic punk rock
          punk    dc alternative flowers grunge hardcore psychedelic rock punk dead death dying punk  dc dc hardcore hardcore p
          unk punk garage hardcore post-hardcore post-punk rock dchc punk agnostic front basement coolguy d-beat dc hardcore de
          athreat hardcore hardcore punk nasty tragedy district of columbia punk old school hardcore punk rock punk ska punk pu
          nk silver spring dc maryland metal nwobhm rva richmond va blackened fusion power shred thrash baltimore rap maryland
          punk dc hardcore punk punk rock suburban punk grindcore powerviolence sludge dc punk pop punk ska punk dchardcore dch
          c hardcore hardcore punk paranoid minds punk hardcore hardcore punk pop acoustic folk indie pop rock punk hardcore po
          st-hardcore post-punk postrock punk crisps attucks dc diy ex the aftermath hardcore ltw state violence va arlington g
          aithersburg hardcore punk punk  dc hardcore punk rock punk hardcore indie rock post punk punk indie rock post punk pu
          nk post hardcore punk experimental diy doom hardcore punk spazz district of columbia post-something with grooves punk
          noise rock post-hardcore post-punk district of columbia dchc hardcore punk'
```

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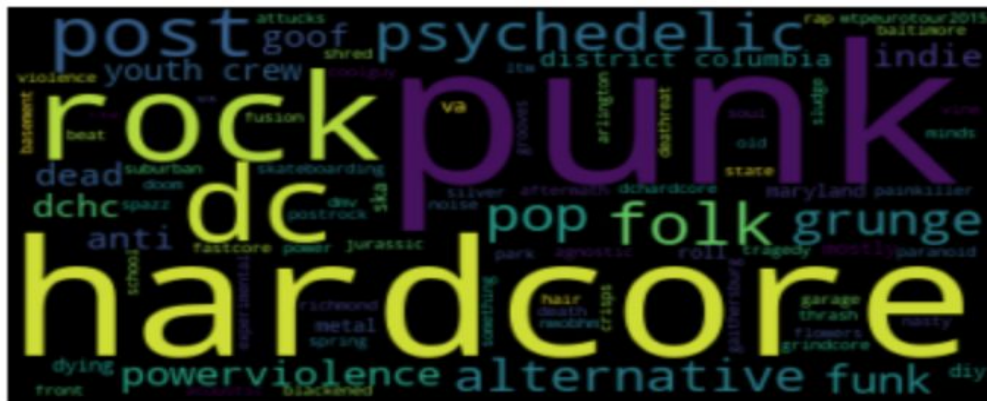
```
In [132]: list(tags.values())[0]
```

```
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']
```

```
In [141]: import matplotlib.pyplot as plt
          plt.imshow(wordcloud, interpolation='bilinear')
          plt.axis("off")
```

Out[141]: (-0.5, 399.5, 199.5, -0.5)

"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)

Author	Topic: A rolling compendium of commas: a tribute to Walkie in his absence (Read 25901

Julian, Thinly-Veiled Outrage IMPRESARIO

Member

Posts: 22010

Hall of Fame Poster, Certified Weblebrity

it's Julian!

🟠 **A rolling compendium of commas: a tribute to Walkie in his absence**

« **on:** September 16, 2016, 08:46:09 am »

,

« *Last Edit: September 16, 2016, 12:51:56 pm by Julian, the BELOVED* »

LVMH

and compendium!

Logged

Julian, Thinly-Veiled   Re: A rolling compendium of commas: a tribute to

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