Walking the WALK

Facilitating Interdisciplinary Web Archive Collaboration

UNIVERSITY OF WATERLOO

Nick Ruest (@ruebot)
Ian Milligan (@ianmilligan1)

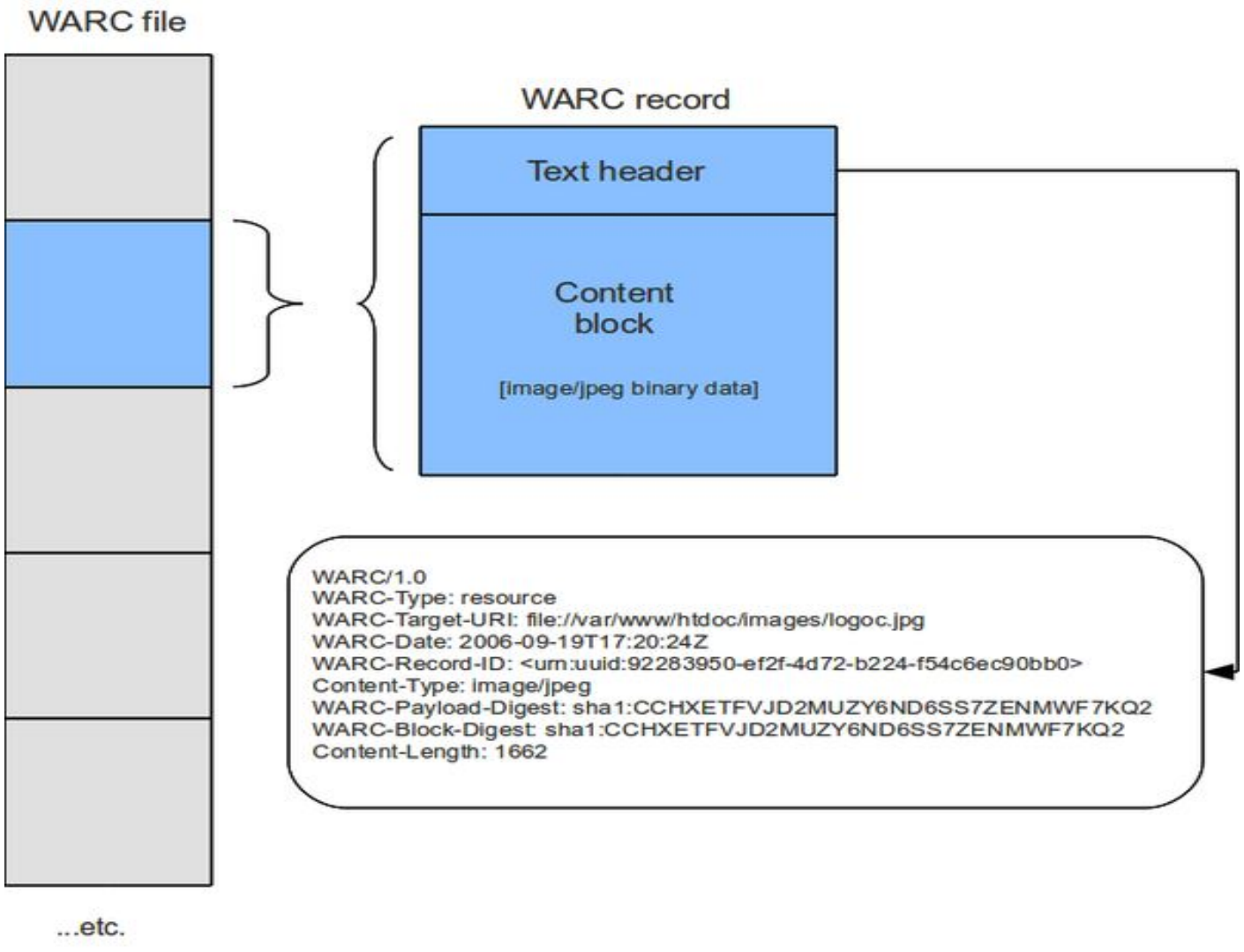
YORK UNIVERSITÉ UNIVERSITY


Why should we even care about web archives?

First, more data than ever before is being preserved...

Second, it'll be saved and delivered to us in very different ways

WARC (ISO 28500:2009)


WARC file
WARC record
Text header
Content block
[image/jpeg binary data]
WARC/1.0
WARC-Type: resource
WARC-Target-URI: file://var/www/htdoc/images/logoc.jpg
WARC-Date: 2006-09-19T17:20:24Z
WARC-Record-ID: <urn:uuid:92283950-ef2f-4d72-b224-f54c6ec90bb0>
Content-Type: image/jpeg
WARC-Payload-Digest: sha1:CCHXETFVJD2MUZY6ND6SS7ZENMWF7KQ2
WARC-Block-Digest: sha1:CCHXETFVJD2MUZY6ND6SS7ZENMWF7KQ2
Content-Length: 1662
...etc.

OccupyWallStreet

The revolution continues worldwide!

Farmers Join Occupy Wall Street, Calling for Food Justice

Posted 5 hours ago on Dec. 2, 2011, 6:21 p.m. EST by OccupyWallSt

As Wall Street's corrupt influence on the economy has grown, the corporate ownership of our food system has hurt the health and livelihood's of some of our most vulnerable communities. This *Sunday, December 4th* food justice activists and occupiers will be traveling from as far as Colorado, Iowa, Maine and Upstate New York to join together for the Occupy Wall Street FARMERS' MARCH.Through a day of dialogue, musical performances, and a march, farmers and their urban allies working for food justice in their communities will form alliances to fight and expose corporate control of the food supply.

Events throughout the day will call and inspire participants to fight against the corporate manipulation of the agriculture system. An industry that is responsible for using chemical toxins tied to soaring obesity rates, heart disease and diabetes and limiting access to affordable, wholesome food to the country's poorest citizens.

Read More...

30 Comments

Occupy Wall Street Goes Home

Posted 1 day ago on Dec. 1, 2011, 3:04 p.m. EST by OccupyWallSt

On December 6th Occupy Wall Street will join in solidarity with a Brooklyn community to re-occupy a foreclosed home. The day of action marks a national kick-off for a new frontier for the

General Inquiries:
general@occupywallst.org
Press Inquiries:
press@occupywallst.org
Press Phone: +1 (347) 292-1444
Help & Directions: +1 (516) 708-4777
Watch: The world we're building
Read: This call to action
Liberty Square Eviction Defense:
Text "@occupyalert" to 23559 to receive alerts in the event of imminent emergency.

Occupy Wall Street is leaderless resistance movement with people of many colors, genders and political persuasions. The one thing we all have in common is that We Are The 99% that will no longer tolerate the greed and corruption of the 1%. We are using the revolutionary Arab Spring tactic to achieve our ends and encourage the use of nonviolence to maximize the safety of all participants.

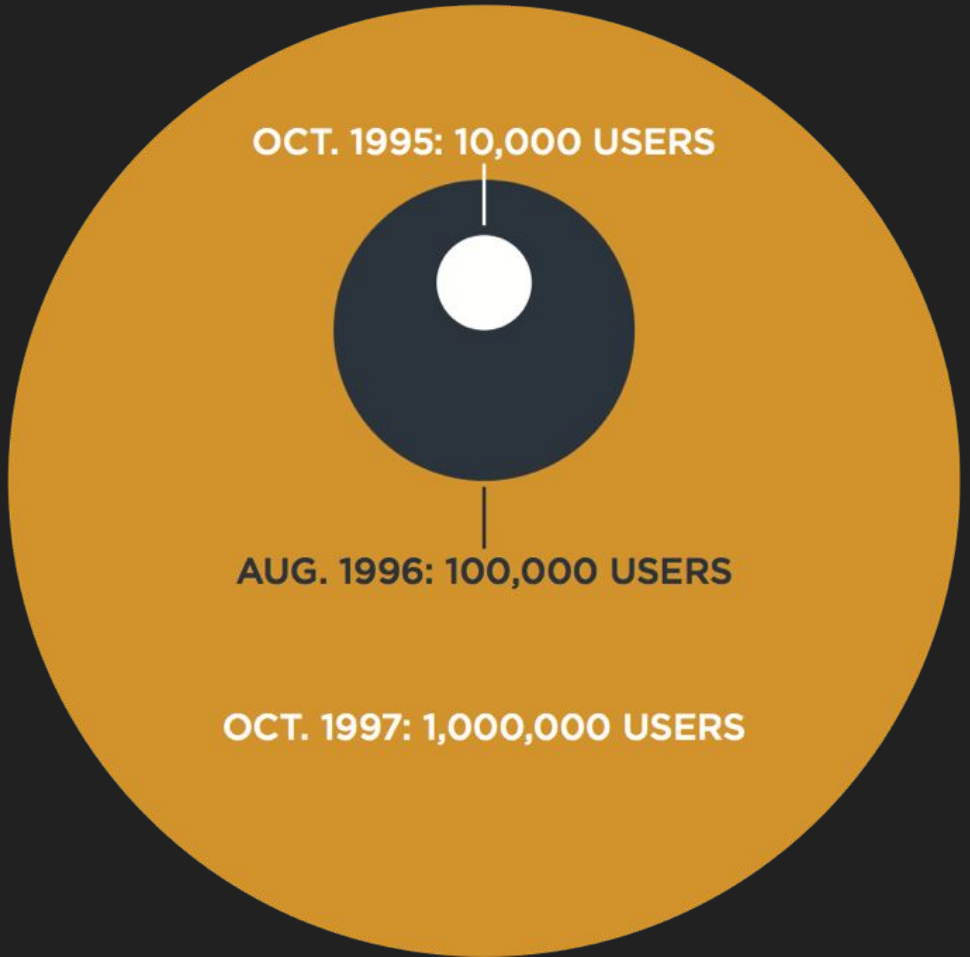
This #ows movement empowers real people to create real change from the bottom up. We want to see a general assembly in every backyard, on every street corner because we don't need Wall Street and we don't need politicians to build a better society.

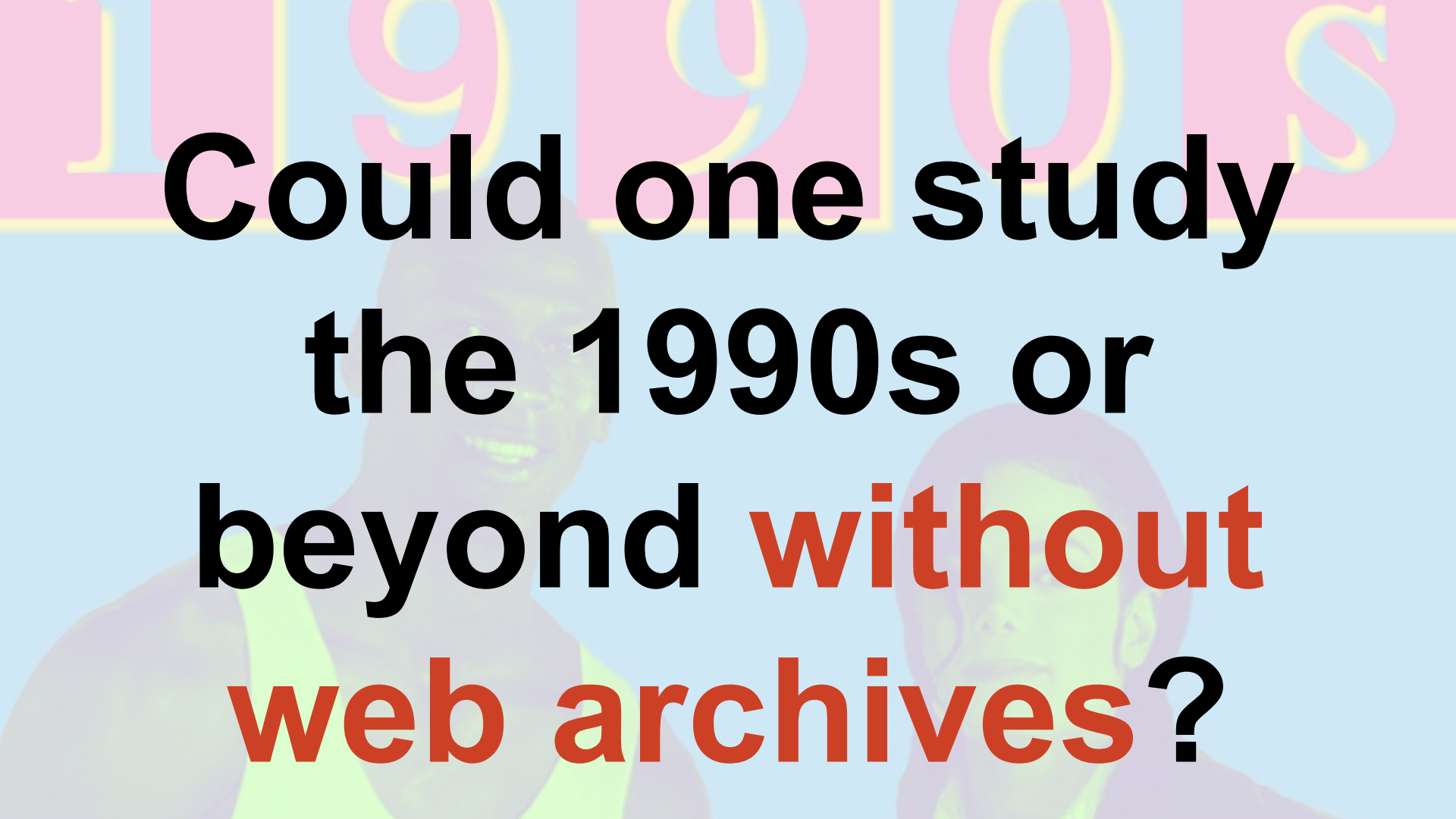
the only solution is WorldRevolution

Click here for NYCGA committee meeting times.


~~Scarcity~~
Abundance

GEOCITIES USERS:




1990s
Could one study the 1990s or beyond without web archives?

And the 1990s are history (as painful as it is to say..)

But right now you have to use the Wayback Machine - requiring you know the URL!


University of Waterloo
Ontario, Canada
prospective students
students
Friday to Thursday, January 21, 2010
search

Microsoft Visual Basic™ 5.0 Site Builder network
CONTROL CREATION EDITION

Visual Basic 5.0 Control Creation Edition Free!

GEOCITIES YOUR HOME ON THE WEB

- ENTER HERE
- INFORMATION
- NEIGHBORHOODS
- WHAT'S NEW
- WHAT'S COOL
- WHAT IS GEOCITIES?

Our communities are home to the most popular collection of FREE HOME PAGES & E-MAIL on the web. Please join or visit one of our 29 neighborhoods today.

* Free Home Pages & Free Member Email

Advertiser Information

Audio DIALweb Update

GeoCities Daily Audio Update -- *Sponsored by IBM VoiceType Simply Speaking*

HOT GEOSTORE

Geotoon of the week

Happy Holidays from all of us at Geocities!

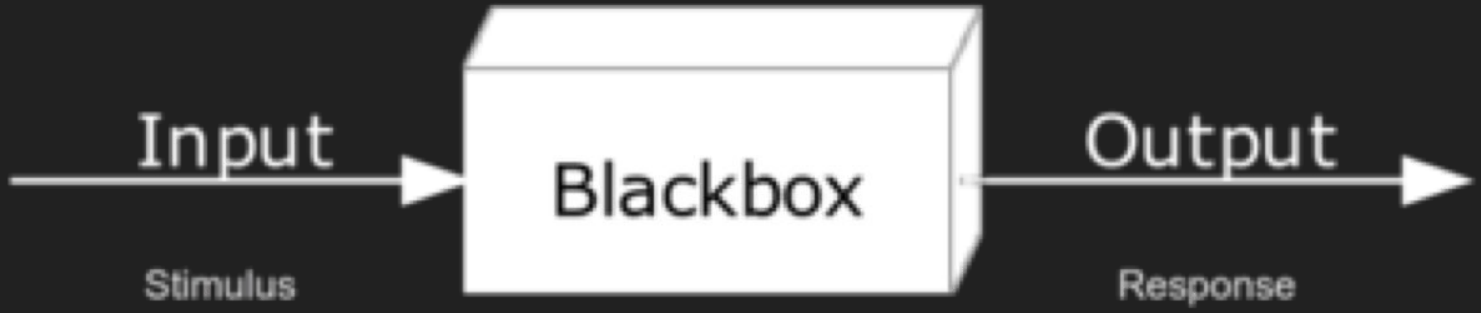
A message from our CEO

Today's Cool Homestead
WallStreet1456
Beginning investors and speculators won't want to miss the Working Class Investor Newsletter.

GeoCities News of the Day - 12/25/96

GEOCITIES LIVE CHRISTMAS TREE! ON CAMERA!

Building a home page for the holidays?
Submit your letters to Santa, favorite holiday recipes and other holiday cheer to our special NorthPole neighborhood. And share your holiday spirit with GeoCitizens around the world im our virtual holiday tree!

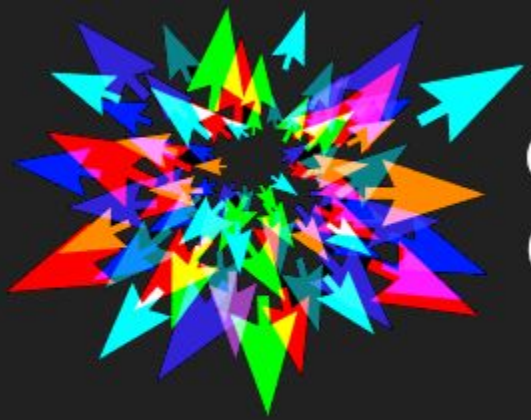

Input
Blackbox
Output
Stimulus
Response

We need interdisciplinary collaboration to tackle this problem!

Social Sciences and Humanities Research Council of Canada

Conseil de recherches en sciences humaines du Canada

Canada



compute canada | calcul canada

Team(s)

We form like Voltron

Ian Milligan

History Faculty Member

Jimmy Lin

Computer Science Faculty Member

Jeremy Wiebe

History PhD Candidate

Alice Zhou

Computer Science Undergraduate

Youngbin Kim

Computer Science Undergraduate

Ryan Deschamps

PhD Candidate, Public Policy

Nick Ruest

Digital Assets Librarian

Collaboration

#Slack & GitHub -> Mentoring

Platforms

CLI tools

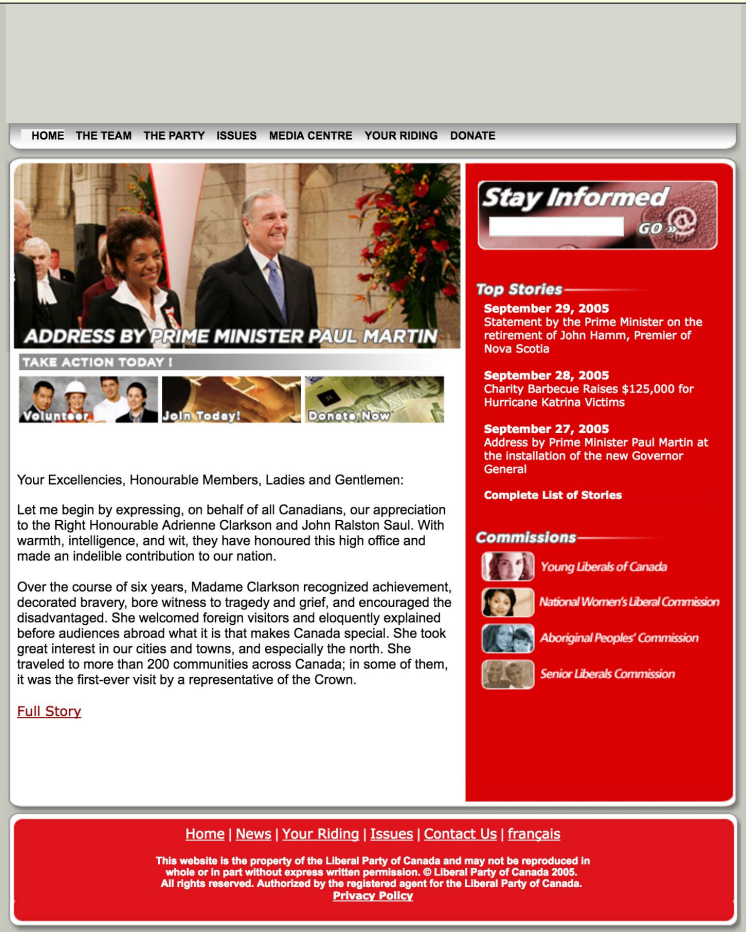
awk, sed, grep, parallel, sort, uniq, wc, jq

Shine

https://github.com/ukwa/shine/

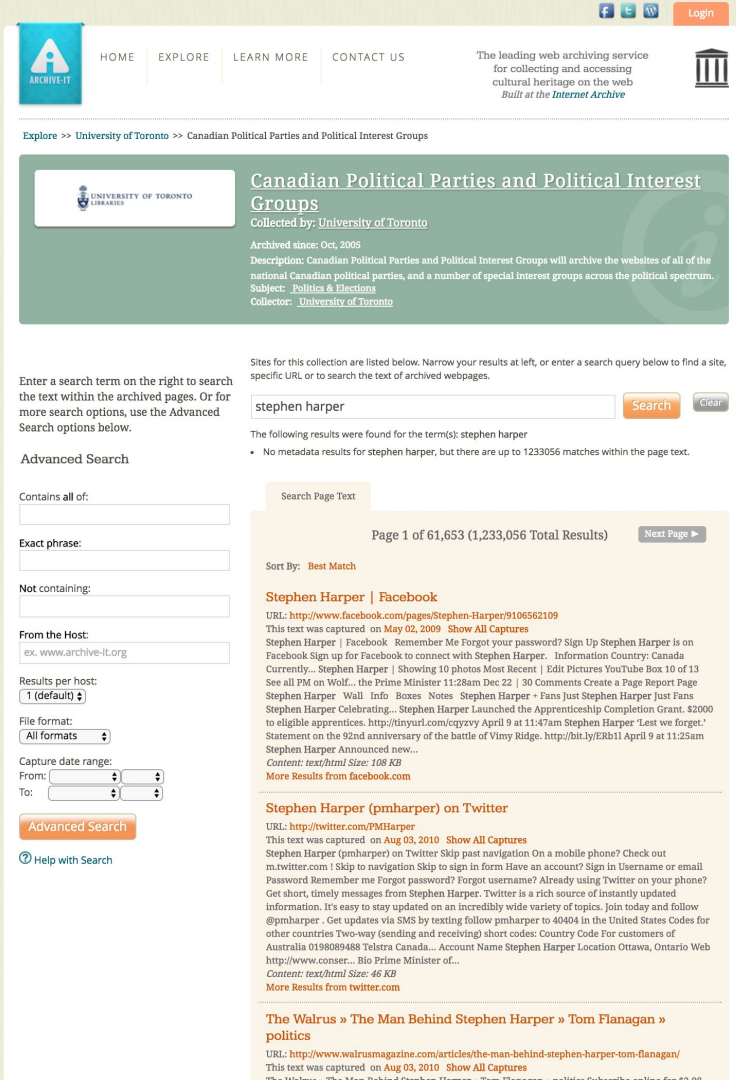
Canadian Political Parties & Political Interest Group Collection (ARCHIVE-IT/Toronto)

- 50 Websites
  - All major political parties
  - Many minor political parties
  - Political interest groups
- Collected quarterly between 2005 and present



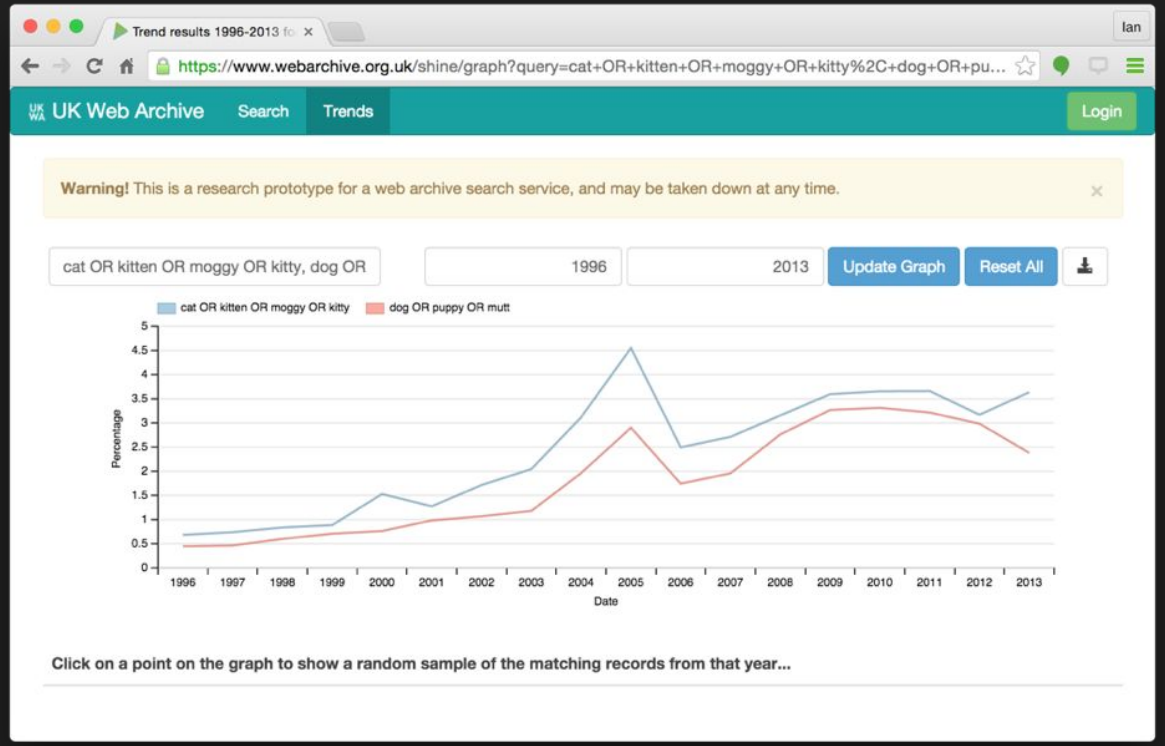
The Current Interface..

- **Very limited** - simple search engine, some advanced options; no facets
- **Great collection**.. But nobody uses them.



14 Million Solr docs!

Shine


Trend results 1996-2013
Ian
https://www.webarchive.org.uk/shine/graph?query=cat+OR+kitten+OR+moggy+OR+kitty%2C+dog+OR+pu...
UK Web Archive
Search
Trends
Login
Warning! This is a research prototype for a web archive search service, and may be taken down at any time.
cat OR kitten OR moggy OR kitty, dog OR
1996
2013
Update Graph
Reset All
cat OR kitten OR moggy OR kitty
dog OR puppy OR mutt
Percentage
Date
Click on a point on the graph to show a random sample of the matching records from that year...

webarchives.ca

Twitter

twarc

twarc-report

#elxn42

Issue 32, 2016-04-25 ISSN 1940-5758

An Open-Source Strategy for Documenting Events: The Case Study of the 42nd Canadian Federal Election on Twitter

This article examines the tools, approaches, collaboration, and findings of the Web Archives for Historical Research Group around the capture and analysis of about 4 million tweets during the 2015 Canadian Federal Election. We hope that national libraries and other heritage institutions will find our model useful as they consider how to capture, preserve, and analyze ongoing events using Twitter.

While Twitter is not a representative sample of broader society – Pew research shows in their study of US users that it skews young, college-educated, and affluent (above $50,000 household income) – Twitter still represents an exponential increase in the amount of information generated, retained, and preserved from 'everyday' people. Therefore, when historians study the 2015 federal election, Twitter will be a prime source.

On August 3, 2015, the team initiated both a Search API and Stream API collection with twarc, a tool developed by Ed Summers, using the hashtag #elxn42. The hashtag referred to the election being Canada's 42nd general federal election (hence 'election 42' or elxn42). Data collection ceased on November 5, 2015, the day after Justin Trudeau was sworn in as the 42nd Prime Minister of Canada. We collected for a total of 102 days, 13 hours and 50 minutes.

To analyze the data set, we took advantage of a number of command line tools, utilities that are available within twarc, twarc-report, and jq. In accordance with the Twitter Developer Agreement & Policy, and after ethical deliberations discussed below, we made the tweet IDs and other derivative data available in a data repository. This allows other people to use our dataset, cite our dataset, and enhance their own research projects by drawing on #elxn42 tweets.

Our analytics included:

- *breaking tweet text down by day to track change over time;*
- *client analysis, allowing us to see how the scale of mobile devices affected medium interactions;*
- *URL analysis, comparing both to Archive-It collections and the Wayback Availability API to add to our understanding of crawl completeness;*
- *and image analysis, using an archive of extracted images.*

Our article introduces our collecting work, ethical considerations, the analysis we have done, and provides a framework for other collecting institutions to do similar work with our off-the-shelf open-source tools. We conclude by ruminating about connecting Twitter archiving with a broader web archiving strategy.

by Nick Ruest and Ian Milligan

Introduction

During the 2015 Canadian federal elections, we captured 3,918,932 tweets written using the #elxn42 hashtag: thoughts on the nature and stature of political candidates or parties, live running commentary during leader debates, exhortations to vote, and witty ripostes or jokes to liven up the long campaign. Political scientists, journalists, and other researchers can use these tweets as evidence of sentiment amongst a

1,203,867 #elxn42 images

Dataset is available here. | Original 32G png available here. | Tiled with deepzoom.py.

Nick Ruest, Web Archives for Historical Research

#ELXN42 TWEETS (42ND CANADIAN FEDERAL ELECTION)

hdl:10864/11311
Version: 2 – Released: Tue Jan 26 17:34:45 EST 2016

< View Previous Study Listing

Use the check boxes next to the file name to download multiple files. Data files will be downloaded in their default format. You can also download all the files in a category by checking the box next to the category name. You will be prompted to save a single archive file. Study files that have restricted access will not be downloaded.

Select all files | Download Selected Files

Total Number of Files: 5 Total Downloads: 13

elxn42-tweet-ids.txt Plain Text - 71 MB - 8 downloads MD5 Checksum: 98b204a8fc0dbae70e5480c5d4a40a50	Download	line-oriented #elxn42 tweet ids
elxn42-tweets-images.txt Plain Text - 54 MB - 1 download MD5 Checksum: 79376146858835c268cec312c4f7d945	Download	line-oriented #elxn42 image urls tweeted
elxn42-tweets-urls.txt Plain Text - 154 MB - 1 download MD5 Checksum: 78c6372996c69f24ba7ea78a0c2a65a8	Download	line-oriented #elxn42 urls tweeted
elxn42-tweets-users.txt Plain Text - 43 MB - 2 downloads MD5 Checksum: 4821eeb35d9b6b78a4d1f3723c5645ef	Download	line-oriented #elxn42 twitter users
elxn42-tweet-tags.txt Plain Text - 1016 KB - 1 download MD5 Checksum: 59b8161616ddd122ff844dd82a8eb4d9	Download	#elxn42 hashtags with counts

#elxn42 web crawl

Description

Consists of a web crawl of unique URLs tweeted with the #elxn42 hashtag. #elxn42 collection took place from August 3, 2015 - November 5, 2015. Unique URLs were extracted from the dataset, and harvested with Heritrix on January 29, 2016 - February 8, 2016.

Download

- warc: #elxn42 web crawl.gz
- cdx: #elxn42 web crawl.cdx
- wat: #elxn42 web crawl.wat.gz
- Seed list: #elxn42 web crawl.txt
- Heritrix configuration: crawler-beans.cxml

In collections

- #elxn42

Details

Title:	#elxn42 web crawl
Creator(s):	Nick Ruest
Note:	Tweet ids: http://hdl.handle.net/10864/11311
Identifier (local):	WEB-20160208134917869-00013-3991~rho.library.yorku.ca~9191-ELXN42
Identifier (md5):	63d707352a6fb45a62889c448154610f
Type:	Website
Subject(s):	#elxn42 Canadian federal election, 2015 Canadian politics Federal politics Canada
Date captured:	2016-01-29
Size:	13GB
File size:	12972947583
PUID:	x-fmt/266
Funding	This research was supported by a research grant -- 435-2015-0011 -- issued by Social Sciences and Humanities Research Council.
Rights:	Use of this resource is governed by the terms and conditions of the Creative Commons "Attribution" License (http://creativecommons.org/licenses/by/2.0/)

#panamapapers

#NDP2016

#MakeDonaldDrumpAgain

#YMMfire

Is this *a* future of collection development?

Institutional vs Research Data

An At-Scale Case Study: Geocities (4.1TB)

Our communities are home to the most popular collection of FREE HOME PAGES & E-MAIL on the web. Please join or visit one of our 29 neighborhoods today.

- NEIGHBORHOODS
- WHAT'S NEW
- WHAT'S COOL
- WHAT IS GEOCITIES?

* Free Home Pages & Free Member Email Advertiser Information

Today's Cool Homestead GeoCities Daily Audio Update

HotSprings 1837

So you hit the snooze bar ten times every morning. You might be lazy. But then again, you might have a sleep disorder. Find out here.

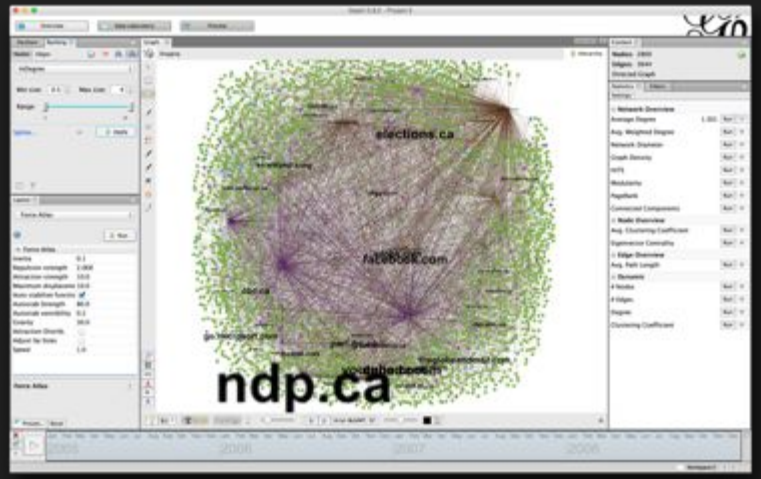
GeoCities News of the Day - 10/22/96

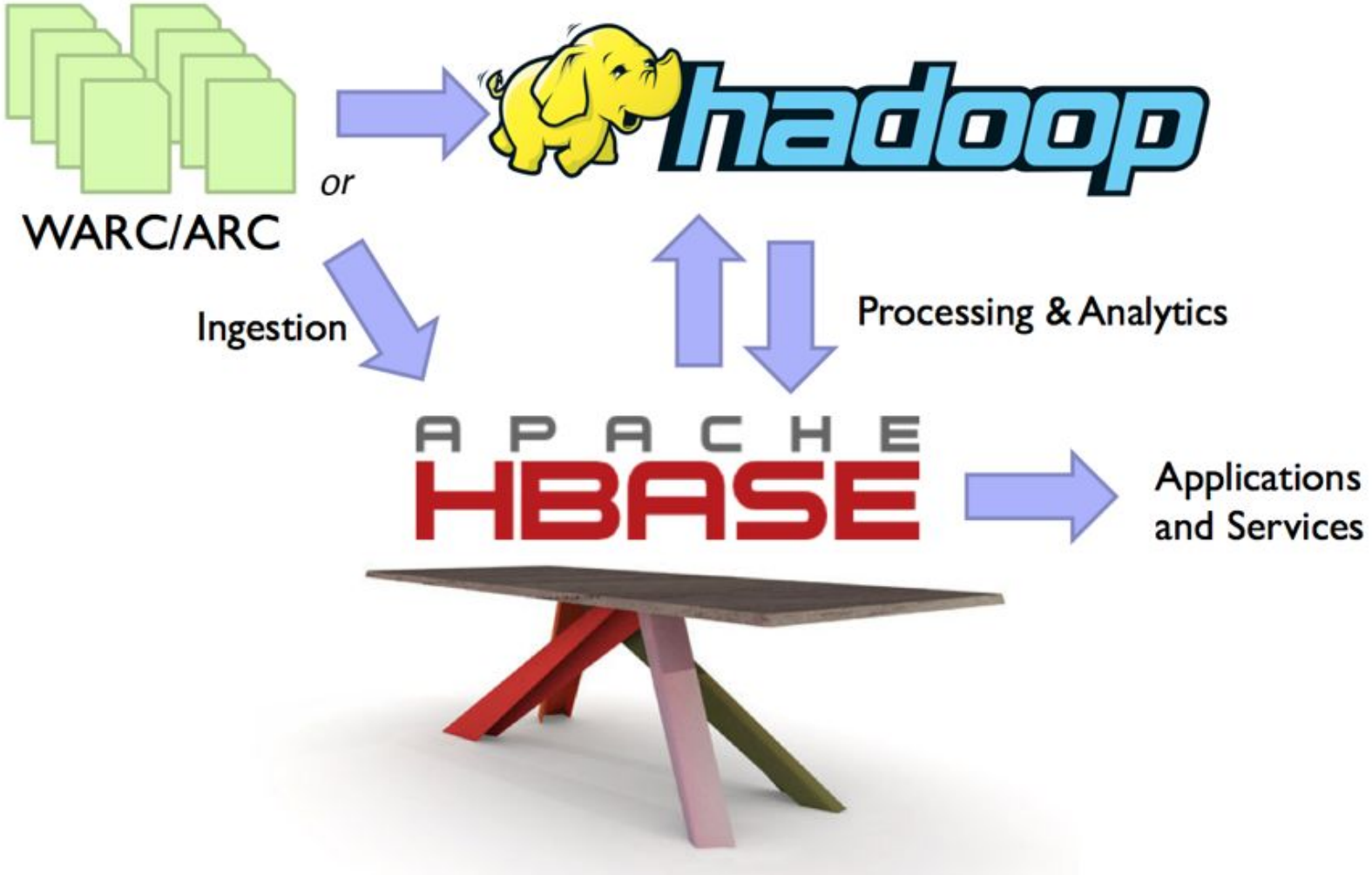
Warcbase

Warcbase

- An open-source platform for managing web archives

- Two main components
  - A flexible data store: your own Wayback Machine
  - Scriptable analytics and data processing

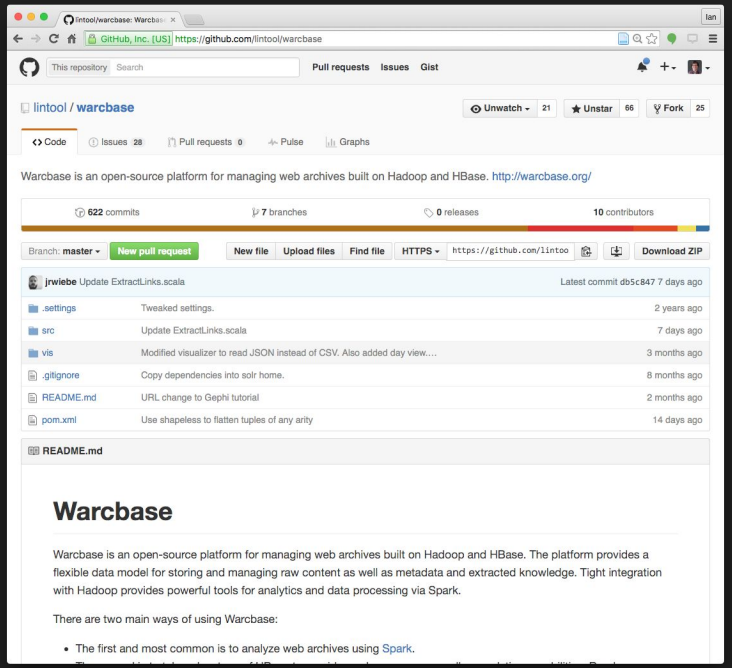



WARC/ARC
or
hadoop
Ingestion
Processing & Analytics
APACHE
HBASE
Applications
and Services

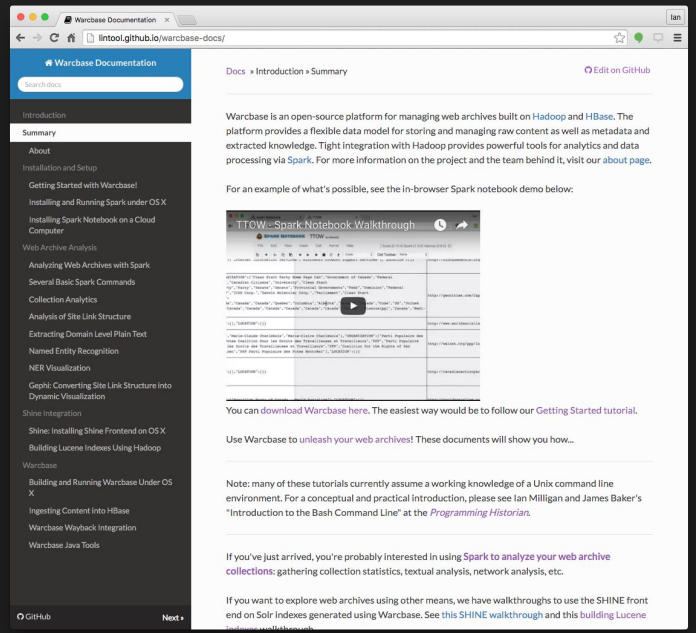
Warcbase

- Scalable
  - From Raspberry Pi to Desktop Computer to Server to Cluster, **all with same scripts and commands**
- Potentially very powerful
  - **Trantor**: 1.2PB of disk, 25 compute nodes (each w/ 128GB memory, 2×6-core Intel Xeon E5 v3 = 3.2TB memory and 300 current-generation Intel cores)
- In active development, led by **Jimmy Lin**, collaborator at the University of Waterloo

You can Warcbase Too! (...and Twarcbase soon!)



warcbase.org



docs.warcbase.org

Let’s do a quick walkthrough of how we’ve used it on one collection: GeoCities

1. i2millig@rho: /mnt/vol1/data_sets/geocities/warcs (ssh)

bash | bash | i2millig@rho: /mnt/vol1/data...

```
GEOCITIES-20091029114236-00191-ia400110.us.archive.org.warc.gz
GEOCITIES-20091029115416-00171-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029123034-00172-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029130439-00173-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029134536-00174-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029140344-00192-ia400110.us.archive.org.warc.gz
GEOCITIES-20091029141553-00193-ia400110.us.archive.org.warc.gz
GEOCITIES-20091029141726-00175-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029144445-00176-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029152151-00177-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029160824-00178-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029164941-00179-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029165037-00194-ia400110.us.archive.org.warc.gz
GEOCITIES-20091029170431-00195-ia400110.us.archive.org.warc.gz
GEOCITIES-20091029171605-00180-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029174154-00181-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029180818-00182-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029182725-00183-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029185858-00184-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029193728-00185-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029194541-00196-ia400110.us.archive.org.warc.gz
GEOCITIES-20091029195911-00197-ia400110.us.archive.org.warc.gz
GEOCITIES-20091029202041-00186-crawling08.us.archive.org.warc.gz
GEOCITIES-20091029221340-00198-ia400110.us.archive.org.warc.gz
GEOCITIES-20091029222459-00199-ia400110.us.archive.org.warc.gz
GEOCITIES-20091030021147-00197-ia400103.us.archive.org.warc.gz
GEOCITIES-20091030021444-00198-ia400103.us.archive.org.warc.gz
GEOCITIES-20091030022413-00171-ia400104.us.archive.org.warc.gz
i2millig@rho:/mnt/vol1/data_sets/geocities/warcs$ du -h
4.1T    .
i2millig@rho:/mnt/vol1/data_sets/geocities/warcs$
```

```
ianmilligan1@Ians-MacBook-Pro:~$ rho
i2millig@rho.library.yorku.ca's password:
Welcome to Ubuntu 14.04.2 LTS (GNU/Linux 3.13.0-32-generic x86_64)

 * Documentation:  https://help.ubuntu.com/

  System information as of Mon Mar  7 13:43:20 EST 2016

  System load:  0.99               Users logged in:       1
  Usage of /:   34.7% of 744.67GB  IP address for em1:     130.63.180.18
  Memory usage: 16%                IP address for em2:     10.0.0.18
  Swap usage:   6%                 IP address for docker0: 172.17.0.1
  Processes:    359

  Graph this data and manage this system at:
    https://landscape.canonical.com/

242 packages can be updated.
130 updates are security updates.

Last login: Mon Mar  7 13:43:21 2016 from 38.123.136.254
i2millig@rho:~$ ./spark-1.5.1/bin/spark-shell --jars ~/warcbase/target/warcbase-0.1.0-SNAPSHOT-fatjar.jar
WARN  NativeCodeLoader - Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Welcome to
      ____              __
     / __/__  ___ _____/ /__
    _\ \/ _ \/ _ `/ __/  '_/
   /___/ .__/\_,_/_/ /_/\_\   version 1.5.1
      /_/

Using Scala version 2.10.4 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_45)
Type in expressions to have them evaluated.
Type :help for more information.
WARN  MetricsSystem - Using default name DAGScheduler for source because spark.app.id is not set.
Spark context available as sc.
SQL context available as sqlContext.

scala> :paste
// Entering paste mode (ctrl-D to finish)

import org.warcbase.spark.matchbox._
import org.warcbase.spark.rdd.RecordRDD._

val r =
RecordLoader.loadWarc("/mnt/vol1/data_sets/geocities/warcs/GEOCITIES-20090808133634-04399-crawling08.us.archive.org.warc.gz", sc)
.keepValidPages()
.map(r => ExtractTopLevelDomain(r.getUrl))
.countItems()
.take(10)

// Exiting paste mode, now interpreting.

INFO  WacWarcInputFormat - Loading file:/mnt/vol1/data_sets/geocities/warcs/GEOCITIES-20090808133634-04399-crawling08.us.archive.org.warc.g
z
import org.warcbase.spark.matchbox._
import org.warcbase.spark.rdd.RecordRDD._
r: Array[(String, Int)] = Array((geocities.com,3748), (www.geocities.com,240), (www.myfilehut.com,12), (asiarooms.com,7), (us.geocities.com
,6), (www.theginge.com,3), (www.angelfire.com,3), (images.quizilla.com,3), (pub28.bravenet.com,3), (ss.webring.yahoo.com,2))

scala>
```

LIVE DEMO

Extracting all URLs

```
1  import org.warcbase.spark.matchbox._
2  import org.warcbase.spark.rdd.RecordRDD._
3
4  val r = RecordLoader.loadWarc("/mnt/vol1/data_sets/geocities/
       warcs", sc)
5  .keepValidPages()
6  .map(r => r.getUrl)
7  .saveAsTextFile("/mnt/vol1/derivative_data/geocities/url-list")
```

Results = 186,761,346 URLs, 9.9GB text file

Extracting a Link Graph

```
1  import org.warcbase.spark.matchbox.{ExtractTopLevelDomain,
       ExtractLinks, RecordLoader}
2  import org.warcbase.spark.rdd.RecordRDD._
3
4  RecordLoader.loadArc("/mnt/vol1/data_sets/geocities/warcs/*", sc)
5  .keepValidPages()
6  .map(r => (r.getCrawldate, ExtractLinks(r.getUrl, r.
       getContentString)))
7  .flatMap(r => r._2.map(f => (r._1, ExtractTopLevelDomain(f._1).
       replaceAll("^\\s*www\\.", ""), ExtractTopLevelDomain(f._2).
       replaceAll("^\\s*www\\.", ""))))
8  .filter(r => r._2 != "" && r._3 != "")
9  .countItems()
10 .filter(r => r._2 > 5)
11 .saveAsTextFile("/mnt/vol1/data_sets/geocities/geocities.
       sitelinks")
```

Results

```
1 ((20090903,http://geocities.com/saganaki2000/ADSLGR/adslgr.htm,
  http://www.adslgr.com),15337)
2 ((20091026,http://geocities.com/saganaki2000/ADSLGR/adslgr.htm,
  http://www.adslgr.com),15337)
3 ((20091027,http://geocities.com/spankbank69hard/,http://pg.photos
  .yahoo.com/ph/spankbank69hard/my_photos/),9807)
4 ((20090903,http://geocities.com/spankbank69hard/index.html,http:/
  /pg.photos.yahoo.com/ph/spankbank69hard/my_photos/),9807)
5 ((20091027,http://geocities.com/CollegePark/Locker/8187/,http://
  www.comercialuruapan.com),8056)
6 ((20090903,http://geocities.com/CollegePark/Locker/8187/,http://
  www.comercialuruapan.com),8056)
```

Creating Entities

403GB of link graph data.

- http://www.geocities.com/EnchantedForest/Grove/1234/index.html
- http://www.geocities.com/EnchantedForest/Grove/1234/pets/cats.html
- http://www.geocities.com/EnchantedForest/Grove/1234/pets/dogs.html
- http://www.geocities.com/EnchantedForest/Grove/1234/pets/rabbits.html

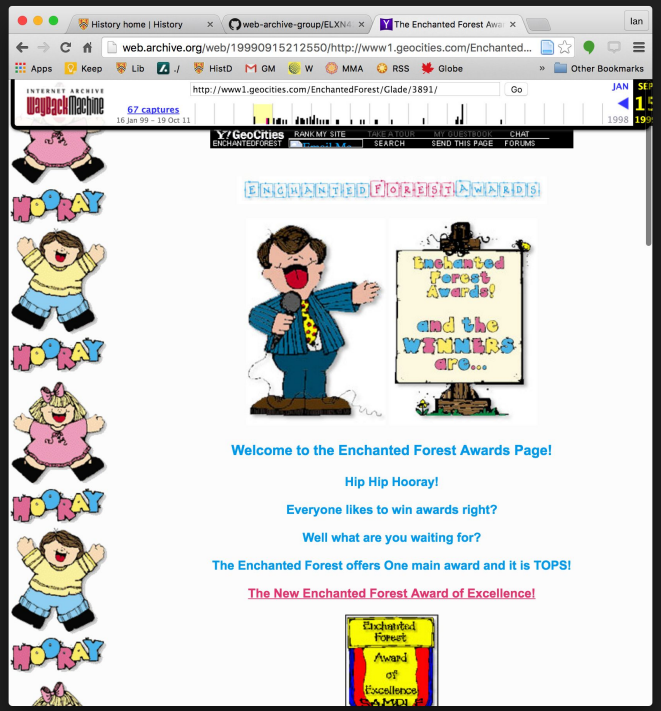
Link Structure

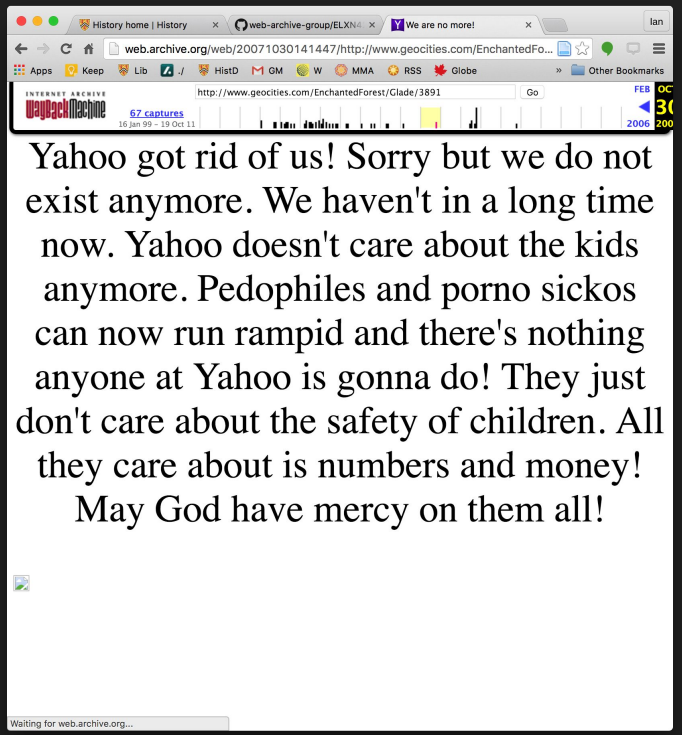
```
1 Source,Target,Weight
2 http://www.geocities.com/EnchantedForest/Meadow/1134,http://www.
  geocities.com/EnchantedForest/1004,83
3 http://www.geocities.com/EnchantedForest/Meadow/1134,http://www.
  geocities.com/EnchantedForest/1004,83
4 http://www.geocities.com/Area51/Stargate/1357,http://www.
  geocities.com/Area51/EnchantedForest/4213,33
5 http://www.geocities.com/Area51/Stargate/1357,http://www.
  geocities.com/Area51/EnchantedForest/4213,33
6 http://www.geocities.com/Eureka/1309,http://www.geocities.com/
  EnchantedForest/Tower/7555,27
7 http://www.geocities.com/Eureka/1309,http://www.geocities.com/
  EnchantedForest/Tower/7555,27
```


http://geocities.com/WestHollywood/9036
http://geocities.com/WestHollywood/2753
http://geocities.com/WestHollywood/Heights/4250
http://geocities.com/WestHollywood/1459
http://geocities.com/WestHollywood/1401
cities.com/WestHollywood/Castro/6902
http://geocities.com/WestHollywood/Village/9414
http://geocities.com/WestHollywood/Park/8021


http://www.geocities.com/EnchantedForest/146
http://www.geocities.com/EnchantedForest/912
http://www.geocities.com/Athens/Acropolis/1465
http://www.geocities.com/EnchantedForest/1001
http://www.geocities.com/EnchantedForest/Meadow/1200
http://www.geocities.com/EnchantedForest/Dell/8854
http://www.geocities.com/Heartland/Hills/7
http://www.geocities.com/EnchantedForest/Glade/3839
http://www.geocities.com/EnchantedForest/1004
http://www.geocities.com/EnchantedForest/Tower/5564
http://www.geocities.com/EnchantedForest/Dell/2948
http://www.geocities.com/EnchantedForest/Tower/9605

EnchantedForest/Glade/3891





WALK Project

Web Archives for Longitudinal Knowledge

WALK

ARCHIVE-IT (Internet Archive subscription service)

25 Canadian partners

130 Publicly-accessible collections

~ 35TB of ARCs/WARCs files

Could we do for Canadian Archive-It Partners what we did for GeoCities and webarchives.ca?

Preparing our dataset!

west cloud

czs-791

nruest

Project

Compute

Overview

Instances

Volumes

Images

Access & Security

Network

Orchestration

Identity

Overview

Limit Summary

Instances
Used 2 of 10

VCPUs
Used 18 of 20

RAM
Used 75GB of 75GB

Floating IPs
Allocated 2 of 2

Security Groups
Used 1 of 10

Volumes
Used 1 of 10

Volume Storage
Used 15.8TB of 15.8TB

Usage Summary

Select a period of time to query its usage:

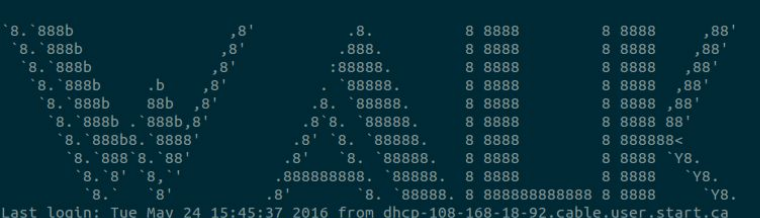
From: 2016-06-01 **To:** 2016-06-01 Submit The date should be in YYYY-mm-dd format.

Active Instances: 2 **Active RAM:** 75GB **This Period's VCPU-Hours:** 437.23 **This Period's GB-Hours:** 11246.43 **This Period's RAM-Hours:** 1865498.75

Usage

Download CSV Summary

Instance Name	VCPUs	Disk	RAM	Time since created
compute-canada-will-accidentally-delete-this	16	412GB	60GB	3 weeks, 1 day




```
Last login: Tue May 24 15:45:37 2016 from dhcp-108-168-18-92.cable.user.start.ca
ubuntu@compute-canada-will-accidentally-delete-this:~$ cd /data/
ubuntu@compute-canada-will-accidentally-delete-this:/data$ ls
alberta_education_curriculum/           circumpolar/  energy_environment/              humanities_computing/              ottawa_shooting_october_2014/    web_archive_general/
alberta_education_curriculum_francais/  cpp/          government_information/          idle_no_more/                      panamapapers/
alberta_floods_2013/                    derby.log     hcf_alberta_online_encyclopedia/ lfrancophonie_de_louest_canadien/  prarie_provinces/
alberta_oil_sands/                      derivatives/  health_sciences_grey_literature/ lost+found/                        transfer.sh
canadian_business_grey_literature/      elxn42/       heritage_community_foundation/   metastore_db/                      university_of_alberta_websites/
ubuntu@compute-canada-will-accidentally-delete-this:/data$
```

```
ARCHIVEIT-1830-NONE-FWPGCP-20111003001238-00494-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB197257-20160219213119199-00000.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003001721-00495-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB197257-20160220144613720-00001.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003002325-00496-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB198318-20160226213020103-00000.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003003144-00497-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB198318-20160227160336248-00001.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003003618-00498-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB199793-20160304231348864-00000.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003004119-00499-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB199793-20160305163910579-00001.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003004447-00500-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB201053-20160311213726199-00000.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003004819-00501-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB201053-20160313041251921-00001.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003005238-00502-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB202130-20160318213027647-00000.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003005548-00503-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB202130-20160319124103352-00001.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003005918-00504-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB203433-20160325213127585-00000.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003010250-00505-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB203433-20160326191120782-00001.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003010637-00506-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB205120-20160401213029030-00000.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003011028-00507-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB205120-20160402165245282-00001.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003011338-00508-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB206270-20160408213022986-00000.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003011722-00509-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB206270-20160409165654512-00001.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003012049-00510-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB207545-20160415214529240-00000.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003012406-00511-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB207545-20160416193001255-00001.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003012710-00512-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB208829-20160422213027236-00000.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003013107-00513-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JOB208829-20160423161249370-00001.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003013415-00514-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-JZPWIF-20120731183011-00000-crawling109.us.archive.org-6683.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003013654-00515-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-KNGFRC-20120619182856-00000-crawling208.us.archive.org-6683.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003014418-00516-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-LFALIS-20120925182725-00000-wbgrp-crawl066.us.archive.org-6682.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003014800-00517-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-LGFTHU-20120605183238-00000-crawling209.us.archive.org-6683.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003015239-00518-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-LODVSM-20120124182717-00000-crawling207.us.archive.org-6681.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003015732-00519-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-MAEEZS-20121002182855-00000-crawling200.us.archive.org-6682.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003020340-00520-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-NPVPJO-20120814182923-00000-crawling203.us.archive.org-6683.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003021438-00521-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-NWUYDC-20120417182849-00000-crawling207.us.archive.org-6682.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003022025-00522-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-OAURLH-20120626182906-00000-crawling212.us.archive.org-6681.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003022757-00523-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-OGHBZD-20120313182936-00000-crawling200.us.archive.org-6683.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003023656-00524-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-OKCKSE-20120703183152-00000-crawling211.us.archive.org-6682.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003024415-00525-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-OMAKFV-20120207182728-00000-crawling203.us.archive.org-6680.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003025319-00526-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-OZWRCH-20130101182808-00000-wbgrp-crawl064.us.archive.org-6681.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003030502-00527-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-PEEQVA-20120320182733-00000-crawling201.us.archive.org-6680.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003031754-00528-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-PNQTNO-20121023181750-00000-wbgrp-crawl066.us.archive.org-6682.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003032523-00529-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-QVRPOA-20121113180904-00000-wbgrp-crawl066.us.archive.org-6680.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003033612-00530-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-RHYTNI-20120724183026-00000-crawling202.us.archive.org-6680.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003034619-00531-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-RJOMJP-20121030182751-00000-crawling212.us.archive.org-6680.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003035531-00532-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-SEPZRL-20120403183145-00000-crawling115.us.archive.org-6683.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003041120-00533-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-SMFGXK-20121009182602-00000-wbgrp-crawl058.us.archive.org-6680.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003042157-00534-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-TBKDVK-20120717183101-00000-crawling204.us.archive.org-6682.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003043149-00535-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-TFPWZS-20120828182859-00000-crawling208.us.archive.org-6681.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003044035-00536-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-TPHTGI-20121030181552-00000-crawling200.us.archive.org-6682.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003045011-00537-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-UCKALS-20130108182813-00000-wbgrp-crawl067.us.archive.org-6683.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003045906-00538-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-VRQQPC-20121106181325-00000-crawling105.us.archive.org-6681.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003050927-00539-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-VTRSQZ-20120911183053-00000-crawling206.us.archive.org-6680.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003051751-00540-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-WAOLMY-20120904182937-00000-crawling114.us.archive.org-6681.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003052903-00541-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-WEJEDU-20120515182839-00000-crawling203.us.archive.org-6683.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003053957-00542-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-WHIZKW-20121218183055-00000-crawling113.us.archive.org-6683.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003055413-00543-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-XFGLFF-20120410183059-00000-crawling109.us.archive.org-6680.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003061251-00544-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-XLUVAV-20111220182844-00000-crawling206.us.archive.org-6682.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003062740-00545-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-XSQRMV-20120508182825-00000-crawling203.us.archive.org-6682.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003064716-00546-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-XUHQKQ-20120501182907-00000-crawling212.us.archive.org-6683.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003070345-00547-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-YEWFRP-20120110182738-00000-crawling114.us.archive.org-6680.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003074116-00548-crawling209.us.archive.org-6681.warc.gz          ARCHIVEIT-1830-WEEKLY-ZQSING-20120131182717-00000-crawling203.us.archive.org-6680.warc.gz
ARCHIVEIT-1830-NONE-FWPGCP-20111003081156-00549-crawling209.us.archive.org-6681.warc.gz
ubuntu@compute-canada-will-accidentally-delete-this:/data/university_of_alberta_websites$
```

Public Access

Shine

Voyant-Tools

```
[nruest@rho:nruest]$ curl "http://localhost:8080/solr/select?q=rob+ford&start=0&rows=2&fl=url%2Ccrawl_date%2Ccontent%2Ctitle%2Ccrawl_year%2Cdomain&wt=json&indent=true"
{
  "responseHeader":{
    "status":0,
    "QTime":0,
    "params":{
      "q":"rob ford",
      "indent":"true",
      "fl":"url,crawl_date,content,title,crawl_year,domain",
      "start":"0",
      "rows":"2",
      "wt":"json"}},
  "response":{"numFound":4616,"start":0,"docs":[
      {
        "url":"http://canadians.org/sites/default/files/styles/large/public/wp-content/uploads/steven-and-rob-ford-240x180_0.jpg?itok=ePOBxrbr",
        "domain":"canadians.org",
        "crawl_date":"2013-08-08T20:45:38Z",
        "crawl_year":"2013",
        "content":[""]},
      {
        "url":"http://www.youtube.com/watch?v=c0sEV98S5cI",
        "domain":"youtube.com",
        "crawl_date":"2014-11-05T22:50:05Z",
        "crawl_year":"2014",
        "content":["Ford bedbug - YouTubeUpload Sign inSearchLoading...This video is unavailable.Watch QueueTV QueueWatch QueueTV QueueRemove allDisconnectLoading...Watch QueueTV Queue__count__/__total__Find out why CloseFord bedbugAnimation: Mike ConstableSubscribeSubscribedUnsubscribe 15Subscription preferencesLoading...Loading...Working...Add toWant to watch this again later?Sign in to add this video to a playlist.Sign inShareMoreReportNeed to report the video?Sign in to report inappropriate content.Sign inStatistics161 views1Like this video?Sign in to make your opinion count.Sign in0Don't like this video?Sign in to make your opinion count.Sign inLoading...Loading...Loading...Sign inRatings have been disabled for this video.Rating is available when the video has been rented.This feature is not available right now. Please try again later.Uploaded on Feb 1, 2011A look at a capitalist as mayor of torontoCategoryComedyLicenseStandard YouTube LicenseShow moreShow lessLoading...40:51Play nextPlay nowThe Rob Ford Story - the fifth estateby CBC News5,330 views24:11Play nextPlay nowThe Funniest MAYOR ROB FORD Compilation!by Aiden McLean949,863 views30:44Play nextPlay nowCoffee Run - Rob Fordby deadmau5815,474 views9:19Play nextPlay nowJon Stewart DESTROYS Obama and Mayor Rob Ford 11/14/2013 |by richardnyavor35,226 views6:16Play nextPlay nowJon Stewart's Crack Smoking Mayor Of Toronto Rob Ford segment 11 14 2013by Peter Tsirlis70,575 views5:36Play nextPlay nowDel Grande Speaks on Appointmentby Matt Elliott178 views4:45Play nextPlay nowJon Stewart on Rob Ford Smoking Crack Again (May 2014) New Videoby LittleDragon7789,240 views5:01Play nextPlay nowRob Ford's Brother Doug: Top 10 WTFord Momentsby Truth Mashup6,791 views12:58Play nextPlay nowctv news kitchener rob ford mayor of toronto crack cocaine nov 4 2013by derek otto260 views5:42Play nextPlay nowToronto Mayor Rob Ford on Jimmy Kimmel Live PART 3by Jimmy Kimmel Live953,911 views6:01Play nextPlay nowCanadian Lifestyle is Shit, Bedbugs, Mold in Toronto; Major Nourhaghighi 's Proofs Ø³Ù˚Ø±Ù†Ø³Ù˚by CanadianLifeStyle469 views4:36Play nextPlay nowToronto Mayor Rob Ford on Jimmy Kimmel Live PART 1by Jimmy Kimmel Live2,247,246 views2:51Play nextPlay nowI Have Bed Bugs...by Raya55,292 views0:08Play nextPlay nowCranked cruiserby pippyconstable78 views4:45Play nextPlay now500 Dawes road-the worst building in Torontoby kegfiwuegfu hsegfcyweg10,900 views2:34Play nextPlay now200 Wellesley Apartment Fire Victims Struggle To Resettle TCHCby 200Wellesley2,242 views3:45Play nextPlay nowCouncillor Fletcher and the Toronto Bed Bug Projectby PaulaFletcher30158 views2:30Play nextPlay nowToronto Community Housing Corporation Public Scams And Corruption TCHC Part 26 of 78by 200Wellesley1,070 views2:40Play nextPlay nowFast Ford Nation! Tasha Hearts ROB FORDby Tasha James33,791 views1:25:41Play nextPlay nowToronto 2014 mayoral debate with Rob Ford, John Tory, Olivia Chow and David Soknackiby The Globe and Mail31,295 viewsLoading more suggestions...Show moreLanguage:EnglishCountry:WorldwideSafety:OffHelpLoading...Loading...Loading...AboutPress & BlogsCopyrightCreators & PartnersAdvertisingDevelopers+YouTubeTermsPrivacyPolicy & SafetySend feedbackTry something new!Loading...Working...Sign in to add this to Watch LaterAdd to"],
        "title":"Ford bedbug - YouTube"}]
  }}
[nruest@rho:nruest]$
```

http://voyant-tools.org/?input=http://rho.library.yorku.ca:8080/solr/select%3Fq%3Dlayton%26start%3D0%26rows%3D250%26fl%3Durl,crawl_date,content,title,crawl_year,domain%26wt%3Dxml%26indent%3Dtrue&xmlDocumentsXpath=//doc&xmlContentXpath=//arr[@name%3D%27content%27]&xmlTitleXpath=//str[@name%3D%27title%27]&xmlAuthorXpath=//str[@name%3D%27domain%27]&xmlPubDateXpath=//str[@name%3D%27crawl_years%27]

voyant-tools.org/?corpus=bf2139b6aac42babdbd5dfbf2c49bfb4&input=http://rho.library.yorku.ca:8080/solr/select?q=layton&start=0&row

Search

Voyant Tools

Cirrus | Terms | Links

Scale | Terms:

Reader

Jack Layton | Leadership(ndp.ca)

Jack Layton | Leadership

Jack Layton, Leaders Tour - Tournée du Chef - Jack Layton Jack Layton, Leaders Tour - Tournée du Chef - Jack Layton(davidsuzuki.org)

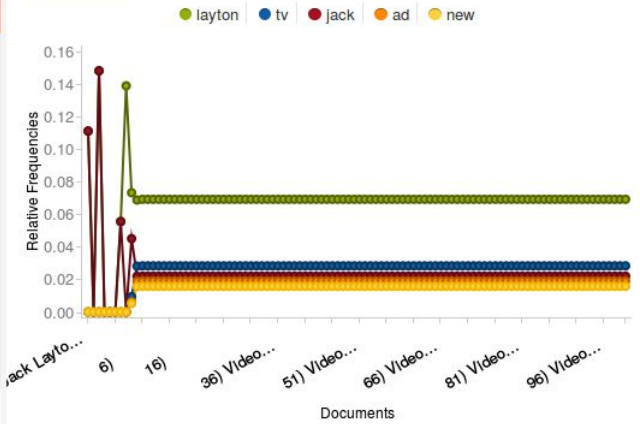
Jack Layton, Leaders Tour - Tournée du Chef - Jack Layton Jack Layton, Leaders Tour - Tournée du Chef - Jack Layton

Jack Layton, Leaders Tour - Tournée du Chef - Jack Layton Jack Layton, Leaders Tour - Tournée du Chef - Jack Layton(davidsuzuki.org)

Jack Layton, Leaders Tour - Tournée du Chef - Jack Layton Jack Layton, Leaders Tour - Tournée du Chef - Jack Layton

INSPIRE awards honour Jack Layton - YouTube(youtube.com)

Trends | Document Terms



Scale | Frequencies

Summary | Documents | Phrases

This corpus has 250 documents with 72,685 total words and 255 unique word forms. Created 2 seconds ago.

Document Length:

- Longest: Video Gallery | NDP (320); Video Gallery | NDP (320); Video Gallery | NDP (320); Video Gallery | NDP (318); Video Gallery | NDP (318)
- Shortest: (0); (0); (0); (0); (0)

Vocabulary Density:

- Highest: INSPIRE awards honour… (1.000); Layton at Central Station… (1.000); Jack Layton | Leadership (1.000); Comments on: People's… (0.667); Video Gallery | NDP (0.603)
- Lowest: (); (); (); (); ()

items:

Contexts

Document	Left	Term	Right
1) Jack …	Jack	lay…	| Leadership
7) Jack …	Jack	lay…	, Leaders Tour - Tournée du Chef
7) Jack …	Tour - Tournée du Chef - Jack	lay…	Jack Layton, Leaders Tour - Tournée
7) Jack …	du Chef - Jack Layton Jack	lay…	, Leaders Tour - Tournée du Chef
7) Jack …	Tour - Tournée du Chef - Jack	lay…	
8) Jack …	Jack	lay…	, Leaders Tour - Tournée du Chef
8) Jack …	Tour - Tournée du Chef - Jack	lay…	Jack Layton, Leaders Tour - Tournée
8) Jack …	du Chef - Jack Layton Jack	lay…	Leaders Tour - Tournée du Chef

5,037 context | expand | Scale

warcbase

Again :-)

Future

- The future is now!
- Ways of categorizing websites?
- Include a third set category (eg. by archiving institution)?
- Expand the number of collections for analysis.
- Refine the tool with better UX so it could be turned into a usable website.

Links!

- https://uwaterloo.ca/web-archive-group/
- https://github.com/web-archive-group/
- https://github.com/ianmilligan1/
- https://github.com/ruebot
- http://dataverse.scholarsportal.info/dvn/dv/wahr

Contact

Nick Ruest: @ruebot
ruestn@yorku.ca
Ian Milligan: @ianmilligan1
i2milligan@uwaterloo.ca