

# Engaging the Public with Web Archives

Providing Access to 10 Years of Political History  
with WebArchives.ca

UNIVERSITY OF  
**WATERLOO**



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



The background of the image shows a perspective view of a library or archive. There are several rows of bookshelves. The shelves are filled with books, many of which have white labels on their spines. Some of the labels are clearly visible, showing numbers like '73', '374', and '375'. The lighting is somewhat dim, creating a sense of a quiet, organized space. The overall color palette is muted, with a lot of greys and browns from the bookshelves and books, contrasted with the white labels.

**Why should we  
care about web  
archives?**

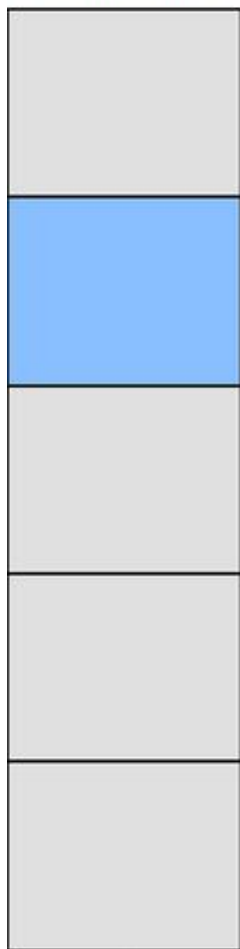
**First, more data than  
ever before is being  
preserved...**

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



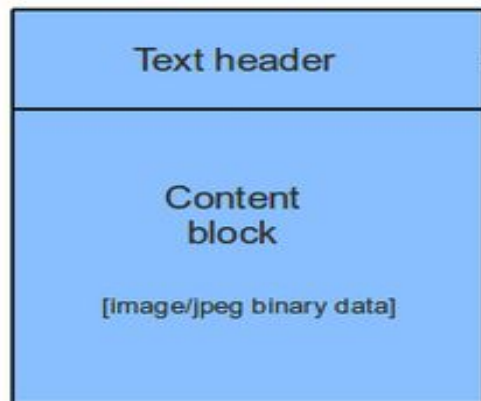
**WARC (ISO 28500:2009)**

# WARC file



...etc.

## WARC record



```
WARC/1.0
WARC-Type: resource
WARC-Target-URI: file://var/www/htdocs/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
```

You are viewing an archived web page, collected at the request of [Internet Archive Global Events](#) using [Archive-It](#). This page was captured on 5:07:27 Dec 03, 2011, and is part of the [Occupay Movement 2011/2012](#) collection. The information on this web page may be out of date. See [All versions](#) of this archived page. [Videos](#) [Metadata](#)



Welcome [login](#) | [signup](#)  
Language [en](#) [es](#) [fr](#)

[News](#) [LiveStream](#) [#HowToOccupy](#) [Forum](#) [Chat](#) [User Map](#) [NYCGA](#) [About](#) [Donate](#) [Twitter](#) [Facebook](#) [RSS](#)

## Farmers Join Occupy Wall Street, Calling for Food Justice

Posted 5 hours ago on Dec. 2, 2011, 6:21 p.m. EST by [OccupayWallSt](#)



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 [OccupayWallSt](#)



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 occupy movement: the liberation of vacant bank-owned homes for those in need. The banks not

**General Inquiries:**  
[general@occupywallst.org](mailto:general@occupywallst.org)  
**Press Inquiries:**  
[press@occupywallst.org](mailto: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.



Wharf  
1006 Rth

RIC GRILL  
STEAK SEAFOOD  
& CHOP HOUSE

Look Here

81

IDLE NO MORE

IDLE  
NO MORE

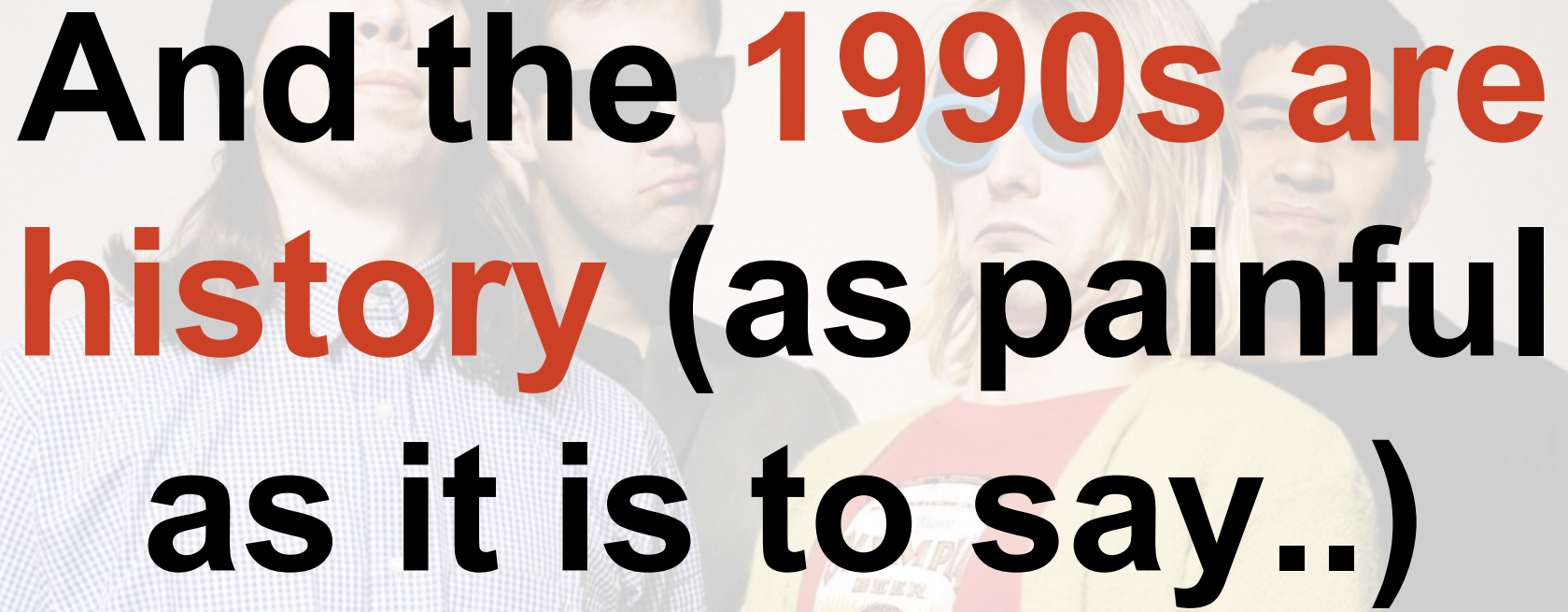




~~Scarcity~~  
Abundance

The background features the text '1990s' in a large, stylized font at the top. Below it, the faces of two people are visible, one on the left and one on the right, both appearing to be smiling or speaking. The overall color palette is light blue and pink.

**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!**



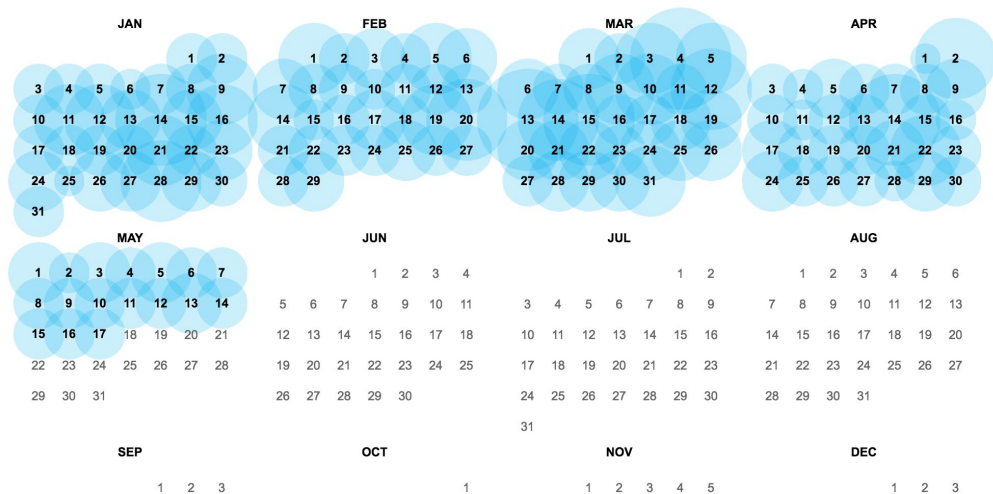
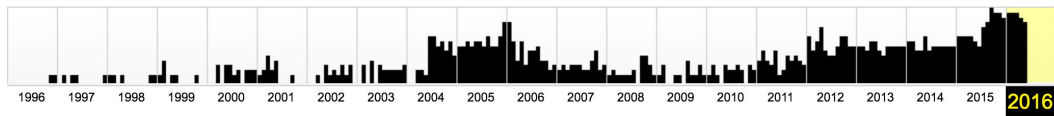
http://liberal.ca

BROWSE HISTORY

<http://liberal.ca>

Saved **2,451 times** between November 9, 1996 and May 17, 2016.

**PLEASE DONATE TODAY.** Your generosity preserves knowledge for future generations. Thank you.



Liberal Party of Canada - H x

web.archive.org/web/19961109132852/http://www.liberal.ca/english/index.html

INTERNET ARCHIVE  
Wayback Machine

6 captures  
9 Nov 96 - 25 Jun 04

http://www.liberal.ca/english/menu.html

Go

OCT NOV 9 APR  
1995 1996 2002

Close X  
Help ?



October 1996

Welcome Cybernauts!

You have arrived at the home page of the Liberal Party of Canada. Perhaps you know who we are?

In fact, you may have read about us in your local newspaper, heard about us on your local radio or seen stories about us on your TV (especially if you're from Canada). We do get lots of exposure in the traditional media...unfortunately, it's all one-way communication. We Liberals are excited about the potential of the World Wide Web...the potential for interactive communication with you!

We are hoping that you will find our site interesting...we are certainly trying to make it so. And we are using your [suggestions](#) to improve as we grow. So please keep those e-mails coming!

On behalf of our leader Prime Minister Jean Chrétien, our President [Senator Daniel Hays](#) and our National Executive, welcome...enjoy your visit...and come back again!

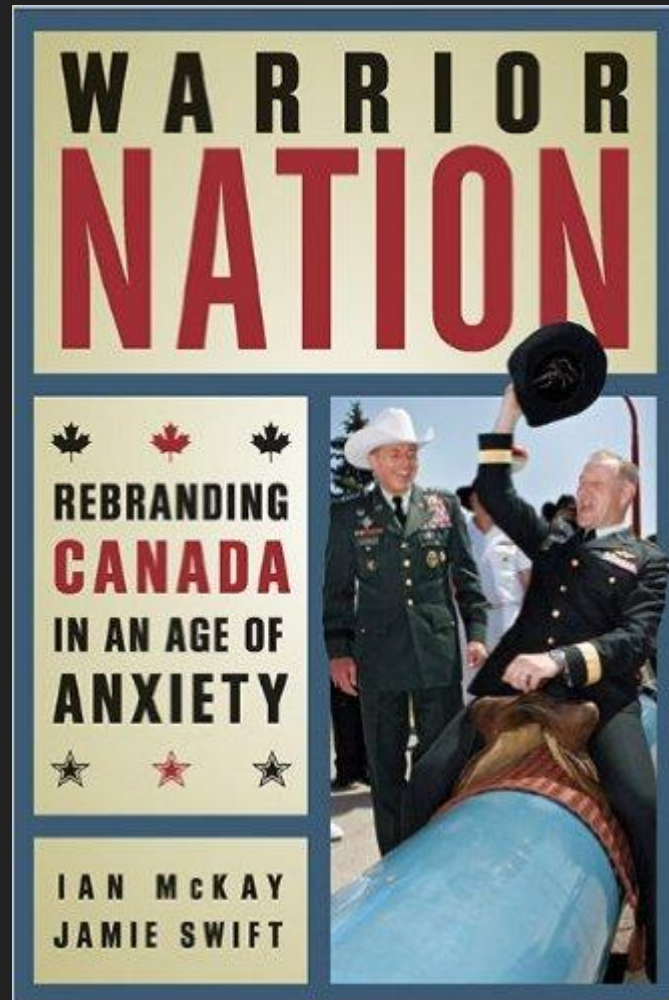
Yours in surfing,

The WebMaster at LPC

This site is authorized by the Federal Liberal Agency of Canada

# **The Case Study:** Canadian Politics

Ian McKay and  
Jamie Swift,  
*Warrior Nation*





# Pivotal Changes in Canadian Politics

- Militarization of Canadian society?
- Change from 'natural governing party' of Liberals to Conservatives
- Major policy changes on foreign policy, environment, etc.
- Sweeping change to how we understand our history
- How to measure?



# Canadian Political Parties & Political Interest Group Collection

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

The screenshot shows the Liberal Party of Canada website. At the top is a navigation menu with links: HOME, THE TEAM, THE PARTY, ISSUES, MEDIA CENTRE, YOUR RIDING, and DONATE. Below the menu is a large banner image of Prime Minister Paul Martin and other officials, with the text "ADDRESS BY PRIME MINISTER PAUL MARTIN". To the right of the banner is a "Stay Informed" section with a search bar and a "GO" button. Below the banner is a "TAKE ACTION TODAY!" section with three buttons: "Volunteer", "Join Today!", and "Donate Now". The main content area features a letter from the Prime Minister to Honourable Members, Ladies and Gentlemen, dated September 29, 2005. The letter discusses the appointment of Adrienne Clarkson and John Ralston Saul as Honourable Members. To the right of the letter is a "Commissions" section listing four commissions: Young Liberals of Canada, National Women's Liberal Commission, Aboriginal Peoples' Commission, and Senior Liberals Commission. At the bottom of the page is a footer with navigation links: Home | News | Your Riding | Issues | Contact Us | français, and a copyright notice: "This website is the property of the Liberal Party of Canada and may not be reproduced in whole or in part without express written permission. © Liberal Party of Canada 2005. All rights reserved. Authorized by the registered agent for the Liberal Party of Canada. Privacy Policy".

HOME THE TEAM THE PARTY ISSUES MEDIA CENTRE YOUR RIDING DONATE

**Stay Informed**  
GO

**Top Stories**  
**September 29, 2005**  
Statement by the Prime Minister on the retirement of John Hamm, Premier of Nova Scotia

**September 28, 2005**  
Charity Barbecue Raises \$125,000 for Hurricane Katrina Victims

**September 27, 2005**  
Address by Prime Minister Paul Martin at the installation of the new Governor General

[Complete List of Stories](#)

**Commissions**  
 Young Liberals of Canada  
 National Women's Liberal Commission  
 Aboriginal Peoples' Commission  
 Senior Liberals Commission

Your Excellencies, Honourable Members, Ladies and Gentlemen:

Let me begin by expressing, on behalf of all Canadians, our appreciation to the Right Honourable Adrienne Clarkson and John Ralston Saul. With warmth, intelligence, and wit, they have honoured this high office and made an indelible contribution to our nation.

Over the course of six years, Madame Clarkson recognized achievement, decorated bravery, bore witness to tragedy and grief, and encouraged the disadvantaged. She welcomed foreign visitors and eloquently explained before audiences abroad what it is that makes Canada special. She took great interest in our cities and towns, and especially the north. She traveled to more than 200 communities across Canada; in some of them, it was the first-ever visit by a representative of the Crown.

[Full Story](#)

[Home](#) | [News](#) | [Your Riding](#) | [Issues](#) | [Contact Us](#) | [français](#)

This website is the property of the Liberal Party of Canada and may not be reproduced in whole or in part without express written permission. © Liberal Party of Canada 2005. All rights reserved. Authorized by the registered agent for the Liberal Party of Canada. [Privacy Policy](#)

# The Current Interface..

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

The screenshot shows the Internet Archive website interface. At the top, there are navigation links: HOME, EXPLORE, LEARN MORE, and CONTACT US. On the right, there is a login button and a tagline: "The leading web archiving service for collecting and accessing cultural heritage on the web. Built at the Internet Archive".

The main content area displays search results for "stephen harper". The search bar contains "stephen harper" and a "Search" button. Below the search bar, it states: "The following results were found for the term(s): stephen harper". A note indicates: "No metadata results for stephen harper, but there are up to 1233056 matches within the page text."

The search results are displayed in a list format. The first result is titled "Stephen Harper | Facebook". The URL is "http://www.facebook.com/pages/Stephen-Harper/9106562109". The text was captured on May 02, 2009. The snippet includes: "Facebook Sign up for Facebook to connect with Stephen Harper. Information Country: Canada Currently... Stephen Harper | Showing 10 photos Most Recent | Edit Pictures YouTube Box 10 of 13 See all PM on Wolf... the Prime Minister 11:28am Dec 22 | 30 Comments Create a Page Report Page Stephen Harper Wall Info Boxes Notes Stephen Harper + Fans Just Stephen Harper Just Fans Stephen Harper Celebrating... Stephen Harper Launched the Apprenticeship Completion Grant. \$2000 to eligible apprentices. http://tinyurl.com/cqzvyv April 9 at 11:47am Stephen Harper 'Lest we forget.' Statement on the 92nd anniversary of the battle of Vimy Ridge. http://bit.ly/ERb11 April 9 at 11:25am Stephen Harper Announced new... Content: text/html Size: 108 KB More Results from facebook.com".

The second result is titled "Stephen Harper (pmharper) on Twitter". The URL is "http://twitter.com/PMHarper". The text was captured on Aug 03, 2010. The snippet includes: "Stephen Harper (pmharper) on Twitter Skip past navigation On a mobile phone? Check out m.twitter.com ! Skip to navigation Skip to sign in form Have an account? Sign in Username or email Password Remember me Forgot password? Forgot username? Already using Twitter on your phone? Get short, timely messages from Stephen Harper. Twitter is a rich source of instantly updated information. It's easy to stay updated on an incredibly wide variety of topics. Join today and follow @pmharper - Get updates via SMS by texting follow pmharper to 40404 in the United States Codes for other countries Two-way (sending and receiving) short codes: Country Code For customers of Australia 0198089488 Telstra Canada... Account Name Stephen Harper Location Ottawa, Ontario Web http://www.conser... Bio Prime Minister of... Content: text/html Size: 46 KB More Results from twitter.com".

At the bottom of the search results, there is a link: "The Walrus » The Man Behind Stephen Harper » Tom Flanagan » politics". The URL is "http://www.walrusmagazine.com/articles/the-man-behind-stephen-harper-tom-flanagan/". The text was captured on Aug 03, 2010.

**Enter Open Source  
Software:  
Shine**

**What is it?**

**UKWA++**

# Play Framework; MVC w/Scala

<https://www.playframework.com/>

# Apache Solr

<http://lucene.apache.org/solr/>



**14M+ Solr docs!**

# webarchive-discovery

<https://github.com/ukwa/webarchive-discovery>

# Shine

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



News - Canada's NDP

wayback.archive-it.org (27/201411072244:7)http://www.ndp.ca/news/archive/2014-03/special/modules/breaking\_news\_...

You are viewing an archived web page, collected at the request of [University of Toronto](#) using [Archive-It](#). This page was captured on 22:44:57 Nov 07, 2014, and is part of the [Canadian Political Parties and Political Interest Groups](#) collection. The information on this web page may be out of date. See [All versions](#) of this archived page.

Home Tom Mulcair Party News Photos and Videos Take Action **DONATE**

# News



**All News Types**

- Press Releases
- Statements
- Speeches
- Reality Checks

**Recent News**

- October 2014
- September 2014
- August 2014

**2014 03 31**  
**The NDP still fighting to save Canada Post**  
In order to bring the drastic restructuring at Canada Post back to the spotlight, the NDP held a press conference this morning to condemn the Conservatives' new price hikes for stamps, which will bring the cost to a record \$1 and affect consumers across the country.

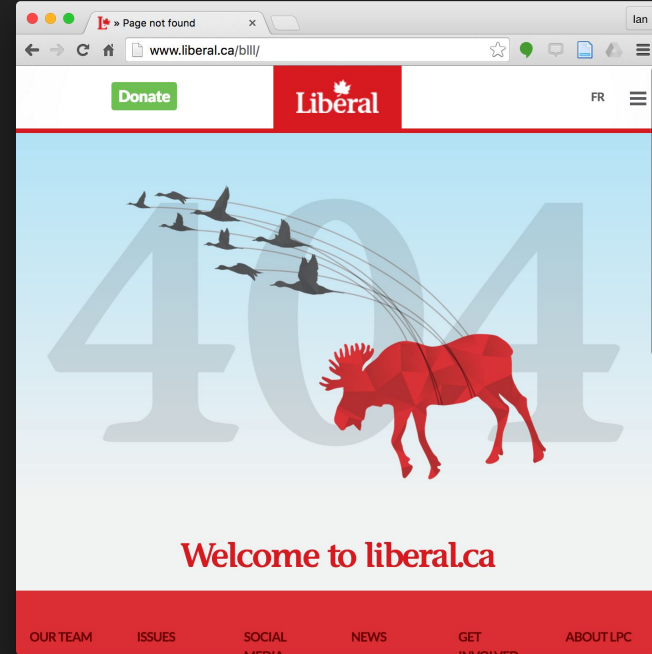
**2014 03 31**  
**Conservatives leaving municipalities behind**  
The New Building Canada plan will in no way reverse the infrastructure and public transit deficiencies that communities are facing throughout the country. The NDP is urging the Conservative government to take immediate action to ensure that the plan meets the needs of municipalities.

**2014 03 28**  
**Official Opposition statement on Earth Hour**  
Official Opposition Environment Critic Megan Leslie (Halifax) made the following statement on Earth Hour:  
  
"On Saturday night (8:30-9:30 PM local time), New Democrats will join millions of people and businesses around the world in turning off our lights to raise awareness about the need to tackle

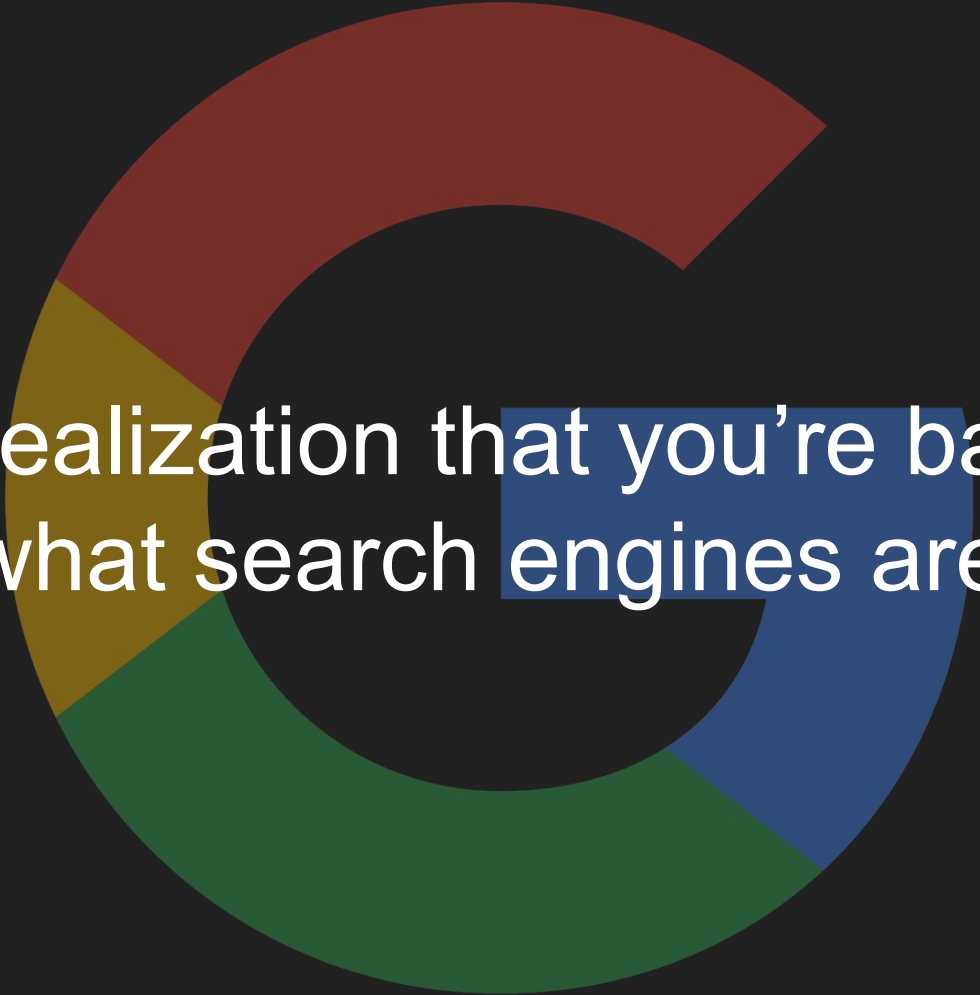
Weblog by wayback.archive-it.org

# Five Things We Learned

- Political parties delete content
- User-generated comments were more common in political parties
- Absences can be more informative than presences
- We can see the rise/fall of prominent people
- Enabling user access is truly transformative



...and the fascinating thing about all this is



...the realization that you're basically  
doing what search engines are doing.



Wait! I'm not done!

**What if we made Shine  
talk to 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=c0sEV985ScI",
      "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_/to
tal_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 in0
Don'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 nowCoffe
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 May
or 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 ro
b 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 nowCan
adian Lifestyle is Shit, Bedbugs, Mold in Toronto; Major Nourhighighl 's Proofs 0^U 0±U†0^U by CanadianLifeStyle469 views4:36Play nextPlay nowToronto Mayor Rob Ford on Jimmy Kimmel Live PART 1by Jim
my 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 buil
ding in Torontoby kegfiwuegfu hsegrfcyweg10,900 views2:34Play nextPlay now200 Wellesley Apartment Fire Victims Struggle To Resettle TCHCby 200Wellesley2,242 views3:45Play nextPlay nowCouncillor Fletc
her 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 views2:40P
lay 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 GL
obe and Mail31,295 viewsLoading more suggestions...Show moreLanguage:EnglishCountry:WorldwideSafety:OffHelpLoading...Loading...Loading...AboutPress & BlogsCopyrightCreators & PartnersAdvertisingDeve
lopersYouTubeTermsPrivacyPolicy & 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%3Ddayton%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\[@name%3D%27crawl\\_years%27\]](http://voyant-tools.org/?input=http://rho.library.yorku.ca:8080/solr/select%3Fq%3Ddayton%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])



Jack Layton | Leadership(ndp.ca)

Jack Layton | Leadership

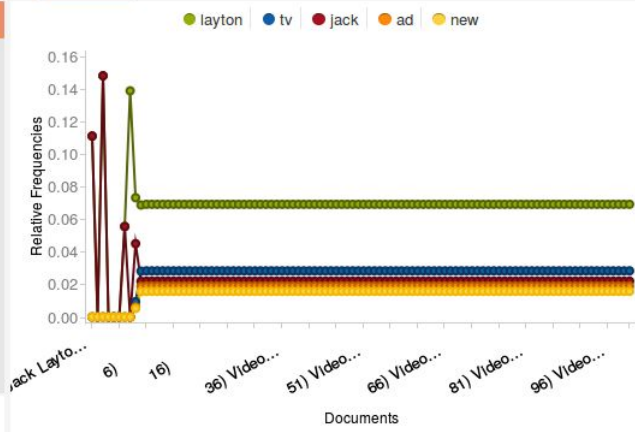
Jack Layton, Leaders Tour - Tournée du Chef - Jack Layton Jack Layton, Leaders Tour - Tournée du Chef - Jack Layton(davidsuzuki.org)

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)



Scale Terms

Navigation controls

Scale Frequencies

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: 0; 0; 0; 0; 0

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

# Thanks and Acknowledgements

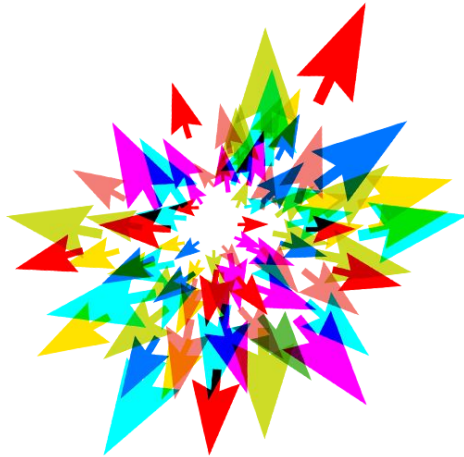


Social Sciences and Humanities  
Research Council of Canada

Conseil de recherches en  
sciences humaines du Canada

Canada

**compute** | **calcul**  
canada | canada



MINISTRY OF RESEARCH AND INNOVATION  
MINISTÈRE DE LA RECHERCHE ET DE L'INNOVATION

# Contact

Ian Milligan: @ianmilligan1

[i2milligan@uwaterloo.ca](mailto:i2milligan@uwaterloo.ca)

Nick Ruest: @ruebot

[ruestn@yorku.ca](mailto:ruestn@yorku.ca)