Our Experiences for Web Archives Researcher Access

27 October 2017
Buffalo, NY
MARAC 2017
Why does a historian care so much about web archives?
Scarcity
Abundance

Scarcity
You can't study the 1990s and beyond without web archives!
And the 1990s are 25 years old. The Web is 25 years old.
Current Access Method

This won't be enough!
So what can we do?
Ask Nich Worby for data.
Canadian Political Parties & Political Interest Group Collection

- 50 Websites
- All major political parties
- Minor political parties
- Political interest groups
- Collected quarterly between 2005 & present.
How could we provide better access?
A prototype web archive exploration UI based on a Solar back-end that has been populated using...

Shine

Freestyle SOLR-powered web archive exploration UI. https://github.com/kwatts/shine
Groups Portal

The Canadian Political Parties and Political Interest

Welcome to the Web Archives for Historical Research. Political Parties Portal. Before diving in, we encourage you to visit our About page.
We also wanted to explore derivative datasets for researcher access with our new project.
Next Step

• The Archives Unleashed Cloud

stay tuned!
- workshops - with travel funding
- set of regional events and other DH work.

• under development - moving to a Blacklight-based system for community involvement;

• A GUI-based workbench for uploading WARCS or using Internet Archive APIs to extract derivatives from your collections
to do text analysis, networks,

• set of regional events and workshops - with travel funding
Compute Canada

Spark Cluster (Project Blacklight)
Search Portal

Archives Unleashed Toolkit

Archives Unleashed

Spark Cluster

Altiscale

Analysis if desired etc. For additional local results, derived data, scholars can download

Data Ingestion from Archive-IT,

Archives Unleashed Cloud

Using local resources or via Archives Unleashed Toolkit

Scholars can directly use the
Visit us at archivesunleashed.org

AUT PROJECT

Our goal is to make petabytes of historical internet content accessible to scholars and others interested in researching the recent past.
Thanks very much!