

The background features a repeating pattern of circular buttons with the word 'VOTE' in blue capital letters. The buttons have a patriotic color scheme of red, white, and blue, with some featuring stars. A white silhouette of the state of Michigan is centered on the page, serving as a frame for the text.

# **Web Archiving Local Election Websites Project**

Presentation for GODORT of Michigan Spring Meeting

June 7, 2024

Julia Ezzo, Government Information & Political Science Librarian

Eli Landaverde, Special Collections LGBTQ+ Librarian

Sarah Mainville, Media Preservation Librarian

# Project Description

Starting in 2019 with a grant from the MSU Libraries, this project focus on the capture and preservation of websites for candidates and significant proposals for Ingham County.



# Why?

Candidate websites are highly ephemeral and often taken down immediately after an election

How can researchers study candidate platforms (especially losing candidates) and how their campaigns are run if the way they disseminate this information isn't preserved?

Many politicians for major offices start their careers in local politics—how can one study their progression if their early campaign materials are lost?

**Ed Busch**, Electronic Records Archivist

**Julia Ezzo**, Government Information & Political  
Science Librarian

**Eli Landaverde**, Special Collections LGBTQ+  
Librarian

**Jessica Martin**, African Studies Librarian &  
Adjunct Assistant History Professor

**Lydia Tang**, Special Collections Archivist

# Initial Project Members

# Grant Funding

~\$3,500 for expansion of MSU Library/University Archives' Archive-It subscription, student labor, and a focus group event

Focus group pulled in (predominately) faculty stakeholders to discuss what sorts of sites they'd find most useful



# Initial Phase

Capture of candidate websites and proposal sites for the 2018 and 2019 election cycles for Ingham County, as well as some sites for Flint and Detroit

Crawls and quality assurance completed by grant-funded graduate assistant

Spreadsheet of candidates with websites is created approximately a month before an election (including primaries, special elections, and general election)

Crawls and QA

URLs to the Archive-It pages are added to a LibGuide to help users easily find and view specific pages

**Ongoing**

# Michigan Politics Web Crawling Project

Collected by: Michigan State University

Archived since: Aug, 2019

Description: This collection contains a large number of Michigan election websites. The collection is part of an Michigan State University project to collect and archive political websites in order to preserve the historical record of Michigan politics and the application of knowledge to the study of Michigan politics.

Subject: Politics & Elections -- Michigan, Michigan -- Politics and government, Judges -- Michigan, Michigan -- Politics and government -- City councils -- Michigan, Courts -- Michigan, Michigan -- Politics and government -- Municipal government -- Michigan, Michigan -- Politics and government -- Mayors -- Michigan, Referendum -- Michigan

Rights: Educational use only, no other permissions. This resource is held by the author and is provided here for educational purposes. It may not be reproduced or distributed in any format without written permission.

Language: English

URL: <https://www.archive-it.org/collections/12662>

Collector: Michigan State University Libraries

## Access

# Archive-It



## Michigan Politics Web Crawling Project

Collected by: Michigan State University

Archived since: Aug, 2019

Description: This collection is mostly focused on the Greater Lansing political and election websites. The collection is part of an MSU Libraries pilot grant to archive highly ephemeral Greater Lansing elections and political websites in order to support the Library's mission of preservation, creation, transmission and application of knowledge.

Subject: Politics & Elections, Government - US States, Society & Culture, Michigan -- Politics and government, Judges -- Michigan, Local government -- Michigan, City council members -- Michigan, City councils -- Michigan, County government -- Michigan, Elections -- Michigan, Municipal government -- Michigan, Political campaigns -- Michigan, School board members -- Michigan, Mayors -- Michigan, Referendum -- Michigan

Rights: Educational use only, no other permissions given. Copyright to this resource is held by the author and is provided here for educational purposes. It may not be reproduced or distributed in any format without written permission.

Language: English

URL: <https://www.archive-it.org/collections/12662>

Collector: Michigan State University Libraries

Genre Form: Web archives

Sort By: [Title \(A-Z\)](#) | [Title \(Z-A\)](#) | [URL \(A-Z\)](#) | [URL \(Z-A\)](#)

Title: Canis Arbrouet NW

URL: <http://arbrouetfornwschools.org/>

Description: Official campaign website to elect Canis Arbrouet to Northwest Community School Board of Education.

Captured 2 times between October 22, 2020 and October 23, 2020

Subject: [Michigan -- Politics and government](#), [School boards -- Michigan -- Jackson](#), [Political campaigns -- Michigan](#), [Elections -- Michigan](#)

Group: [School Board](#)

Language: [English](#)

Coverage: [North America -- United States -- Michigan -- Jackson](#)

Format: [Web archives](#)

Rights: Educational use only, no other permissions given. Copyright to this resource is held by the author and is provided here for educational purposes. It may not be reproduced or distributed in any format without written permission.

Collector: [Michigan State University Libraries](#)

URL: <http://arbrouetfornwschools.org/>

Contributor: [Canis Arbrouet](#)

---

Title: Jim Jamo for Ingham County Circuit Court

URL: <http://jimjamojudge.com/>

Description: Official campaign website to elect Jim Jamok as Ingham County Circuit Court Judge.

Captured once on October 15, 2019

Subject: [Michigan -- Politics and government](#), [Elections -- Michigan](#), [Political campaigns -- Michigan](#), [Judges -- Michigan](#)

Group: [Judges](#)

Language: [English](#)

Coverage: [North America -- United States -- Michigan -- Ingham County](#)

Format: [Web archives](#)

Rights: Educational use only, no other permissions given. Copyright to this resource is held by the author and is provided here for educational purposes. It may not be reproduced or distributed in any format without written permission.

# LibGuide

## Michigan Politics Web Archive: About

This guide provides information and access points to selected local election websites that were crawled by the MSU Libraries and Archives.

[About](#)[City Clerk](#)[City Council](#)[County Clerk](#)[County Commissioner](#)[Drain Commissioner](#)[Judges](#)[Mayors](#)[Meetings](#)[Proposals](#) ▾[Prosecuting Attorney](#)[Register of Deeds](#)[Representative](#)[School Board](#)[Sheriff](#)[Supervisor](#)[Treasurer](#)[Trustee](#)

### Project Description

In 2019, the MSU Libraries funded a pilot grant to archive the websites of local elections. Using a service called [Archive-It](#), we began preserving websites for the 2018 and 2019 election cycle for the Greater Lansing area, with limited coverage of Flint and Detroit. While there are many formal efforts and mandates to preserve national and established statewide governmental records, local elections are extremely ephemeral but can provide important useful information.

This web archive is not a comprehensive collection for all candidates and ballot measures, as having a web site is not required for a campaign. Additionally, at the time of the web crawls, some candidate and ballot websites were no longer active and thus could not be preserved.

We continue to capture and preserve Ingham County election websites for primary and general elections.

[Access the web archives](#)

### Why is web archiving important?

Just as we have archives and museums to preserve our history, important history is made and documented on the web. Just as easily as it is created and shared, it can also be altered or completely deleted. The [Internet Archive](#) is one of a handful of organizations that take snapshots of websites called "crawls." If you have a URL, you can search the Internet Archive's [Wayback Machine](#) to see how a website looked years ago and to see how it has evolved over time.

### Questions, Comments, or Feedback

If you have questions, comments, or feedback about this project and the archived links, please feel free to send them to Julia Ezzo at [julia@msu.edu](mailto:julia@msu.edu).

# Michigan Politics Web Archive: City Council

This guide provides information and access points to selected local election websites that were crawled by the MSU Libraries and Archives.

- About
- City Clerk
- City Council**
- County Clerk
- County Commissioner
- Drain Commissioner
- Judges
- Mayors
- Meetings
- Proposals ▾
- Prosecuting Attorney
- Register of Deeds
- Representative
- School Board
- Sheriff
- Supervisor
- Treasurer
- Trustee

**2018 Election**

**Mason**

- [Marlon Brown](#)
- [Jon Droscha](#)

**2019 Election**

**East Lansing**

- [Erik Altmann](#)
- [Lisa Babcock](#)
- [Jessy Gregg](#)
- [Mark Meadows](#)

**Lansing**

**At Large**

- [Yanice Jackson-Long](#)
- [Dan Ross](#)

**2020 Election**

**Mason**

- [Angela Madden](#)
- [Rita Vogel](#)

**Williamston**

- [Tammy Gilroy](#)

**2021 Election**

**East Lansing**

**At Large (4-Year Term)**

- [George Brookover](#)
- [Adam DeLay](#)
- [Chuck Grigsby](#)

**At Large (Partial Term)**

- [Ron Bacon](#)

**2022 Election**

**Lansing**

- [Brian Daniels](#)

**2023 Election**

**Lansing**

**At-Large**

- [Tamera Carter](#)
- [Missy Lilje](#)
- [Trini Lopez Pehlivanoglu](#)
- [Farhan Sheikh-Omar](#)
- [Jody Washington](#)
- [Nicklas W. Zande](#)

**1st Ward**

- [Ryan Kost](#)
- [D. Taft](#)



# **How the Sausage is Made**

Archive-It is a subscription service that provides tools and technical support to organizations seeking to preserve digital content.

Data captured via Archive-It is stored in data centers owned and operated by the Internet Archive. This data can be downloaded, stored, and shared independently by subscribers.

No technical expertise needed.

# Archive-It

# It Takes a Village

## Collaborative effort

- Special Collections and University Archives & Historical Collections ran independent web archiving projects, and both had their own Institutional accounts.

## Different approaches

- Special Collections was more of an experimental endeavor, UAHC was already a well-honed machine.

## Learning as you go

- Documentation and training.

# 1. Adding URLs to Archive-It.

The screenshot shows the Archive-It web interface. At the top, there is a navigation bar with the Archive-It logo and menu items: Home, Collections, Crawls, Archives, and ARS. On the right side of the navigation bar, there are links for 'Help Center' and 'Welcome, Eli Landaverde'. Below the navigation bar, a breadcrumb trail reads 'Home / Collections / Michigan Politics Web Crawling Project / Seeds'. The main heading is 'Michigan Politics Web Crawling Project', with a pencil icon for editing. To the right of the heading, it says 'Created: Aug 15, 2019 by buschedw' and 'Updated: Sep 1, 2023 by Eli'. Below the heading, there are tabs for 'Overview', 'Seeds' (which is selected), 'Crawls', 'Collection Scope', 'Metadata', 'Wayback QA', and 'Upload WARC's'. Under the 'Seeds' tab, it says 'Seed List (1 to 100 of 395 Seeds)' and has a pagination control '< 1 2 ... 4 >'. There is a search box with the placeholder text 'Type to filter Seeds by URL or Group Name' and a 'Download Seed List' button. Below the search box, there are several buttons: 'Run Crawl', 'Edit Settings', 'Add Metadata', 'Edit Groups', 'Delete Seeds', 'Add Rules', 'Manage Seed Groups', and 'Add Seeds'. At the bottom, there is a table with the following columns: 'Seed URL', 'Groups', 'Status', 'Frequency', 'Type', 'Access', 'Last Crawl', 'Crawls', and 'Wayback'. The table contains five rows of seed data.

<input type="checkbox"/> Seed URL	Groups	Status	Frequency	Type	Access	Last Crawl	Crawls	Wayback
<input type="checkbox"/> <a href="https://www.erikaltmann.com/">https://www.erikaltmann.com/</a>	City Council	Active	One-Time	Standard	Public	Feb 28 2024	1	<a href="#">Wayback &gt;</a>
<input type="checkbox"/> <a href="https://www.wardellforel.com/">https://www.wardellforel.com/</a>	City Council	Active	One-Time	Standard	Public	Nov 01 2023	1	<a href="#">Wayback &gt;</a>
<input type="checkbox"/> <a href="https://www.kerryesingh.com/">https://www.kerryesingh.com/</a>	City Council	Active	One-Time	Standard	Public	Nov 01 2023	1	<a href="#">Wayback &gt;</a>
<input type="checkbox"/> <a href="https://ramirezforel.com/">https://ramirezforel.com/</a>	City Council	Active	One-Time	Standard	Public	Nov 01 2023	1	<a href="#">Wayback &gt;</a>
<input type="checkbox"/> <a href="https://www.noelgarciaforcitycouncil.com/">https://www.noelgarciaforcitycouncil.com/</a>	City Council	Active	One-Time	Standard	Public	Nov 01 2023	1	<a href="#">Wayback &gt;</a>

## 2. Set parameters for what to capture

The screenshot shows the Archive-IT interface. At the top, there is a navigation bar with the Archive-IT logo and menu items: Home, Collections, Crawls, Archives, and ARS. On the right side of the navigation bar, there are links for 'Help Center' and 'Welcome, Eli Landaverde'. Below the navigation bar, the main content area displays the URL 'https://www.noelgarciaforcitycouncil.com/' with a pencil icon for editing. To the right of the URL, there are status indicators: 'Active', 'Public', and 'One-Time'. Below the URL, it shows 'Created: Nov 1, 2023' and 'Updated: Nov 9, 2023'. On the right, it says 'In collection Michigan Politics Web Crawling Project'. A horizontal menu below the URL includes 'Settings', 'Metadata', 'Groups', 'Sharing', 'Crawling History', 'Notes', and 'Seed Scope' (which is currently selected). Under the 'Seed Scope' tab, there is a section titled 'Add Seed Scope Rule'. A dropdown menu is open, showing options: 'Select Seed Scope Rule Type...', 'Add Data Limit', 'Ignore Robots.txt', 'Block URL If...', and 'Expand Scope to Include URL If...'. Below the dropdown, there is a 'Scope Rule List (1 Rule)' section with a search box labeled 'Type to Search...'. At the bottom, there is a table with columns: 'Applies To', 'Rule', 'Url Match', 'Value', and 'Controls'. The table contains one row: 'For seed https://www.noelgarciaforcitycouncil.com/' under 'Applies To', 'ignore robots.txt' under 'Rule', and a toggle switch under 'Controls'.

ARCHIVE-IT

Home Collections Crawls Archives ARS

Help Center Welcome, Eli Landaverde

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

Active Public One-Time

Created: Nov 1, 2023 Updated: Nov 9, 2023

In collection [Michigan Politics Web Crawling Project](#)

Settings Metadata Groups Sharing Crawling History Notes **Seed Scope**

**Add Seed Scope Rule**

- ✓ Select Seed Scope Rule Type...
- Add Data Limit
- Ignore Robots.txt
- Block URL If...
- Expand Scope to Include URL If...

Scope Rule List (1 Rule)

Type to Search...

Applies To	Rule	Url Match	Value	Controls
For seed <a href="https://www.noelgarciaforcitycouncil.com/">https://www.noelgarciaforcitycouncil.com/</a>	ignore robots.txt			<input checked="" type="checkbox"/>

# 3. Schedule a test crawl

The screenshot shows the Archive-IT web interface. At the top, there is a navigation bar with 'Home', 'Collections', 'Crawls', 'Archives', and 'ARS'. A user is logged in as 'Welcome, Eli Landaverde'. The breadcrumb trail indicates the current location: 'Home / Collections / Michigan Politics Web Crawling Project / Seeds'. The main content area is titled 'Michigan Politics Web Crawling Project' and shows a 'Seed List (1 to 100 of 39)'. A 'Run Crawl' dialog box is open in the center, allowing configuration for a test crawl. The dialog includes options for 'Crawl Type' (One-Time Crawl or Test Crawl), 'Doc. Limit' (Whole Number, e.g. 10000), 'Data Limit' (Limit in GB, e.g. 20.91), 'Time Limit' (1 Day), 'Crawl PDFs Only' (checkbox), and 'Crawling Technology' (Standard or Brozzler). A summary states '1 selected seed will be crawled.' and there are 'Cancel' and 'Crawl' buttons.

ARCHIVE-IT

Home Collections Crawls Archives ARS Help Center Welcome, Eli Landaverde

Home / Collections / Michigan Politics Web Crawling Project / Seeds

## Michigan Politics Web Crawling Project

Updated: Sep 1, 2023 by Eli

Overview **Seeds**

Seed List (1 to 100 of 39)

Type to filter Seeds by U

Run Crawl Edit Seeds

Seed URL

- <https://www.erikal...>
- <https://www.warde...>
- <https://www.kerrye...>
- <https://ramirezfore...>
- <https://www.noelg...>

### Run Crawl

Please select options below for a test crawl or one-time crawl of the selected seeds.

Crawl Type  One-Time Crawl  Test Crawl

Doc. Limit  Documents

Data Limit  GB

Time Limit

Crawl PDFs Only  (Available for Standard crawls)

Crawling Technology  Standard  Brozzler

**1 selected seed will be crawled.**

Cancel Crawl

< 1 2 ... 4 >

Download Seed List

Groups Add Seeds

Crawl	Crawls	Wayback
2024	1	Wayback >
2023	1	Wayback >
2023	1	Wayback >
2023	1	Wayback >
2023	1	Wayback >

# 4. QA & patching

The screenshot shows the Archive-IT web interface. At the top, there is a navigation bar with the Archive-IT logo and links for Home, Collections, Crawls, Archives, and ARS. On the right, there are links for Help Center and a user greeting: 'Welcome, Eli Landaverde'. The main content area is titled 'Michigan Politics Web' and includes tabs for Overview, Seeds, and Crawls. A 'Missing Documents' section is visible, with a table listing 21 missing documents. A modal dialog box titled 'Run Patch Crawl' is open in the center, prompting the user to select options for a patch crawl. The dialog includes a 'Cancel' button and a 'Crawl' button. The background table has columns for 'Missing Document', 'Doc. Missing From', 'Possible Reason', 'File Type', 'Size', and 'Crawl Date'. The first row shows a document from 'http://frog.wix.com/bpm?\_msid=b5...' that is 'Invalid (Unknown)' and was crawled on 'Nov 1, 2023'. The dialog box also indicates that 21 selected documents will be crawled and that the 'Ignore Robots.txt' option is checked.

**Run Patch Crawl**

Please select options below for a patch crawl of the selected missing documents.

**Ignore Robots.txt**  capture documents


blocked by robots.txt

**21 selected Documents will be crawled.**

Cancel Crawl

Missing Document	Doc. Missing From	Possible Reason	File Type	Size	Crawl Date
<input checked="" type="checkbox"/> <a href="http://frog.wix.com/bpm?_msid=b5...">http://frog.wix.com/bpm?_msid=b5...</a>	<a href="https://www.noelgarciaforcitycouncil.com/">https://www.noelgarciaforcitycouncil.com/</a>	Invalid (Unknown)		0 by...	Nov 1, 2023
Filter Match: "b=356&url=https://www.noelgarciaforcitycouncil.com/&v=1.13060"					
<input checked="" type="checkbox"/> <a href="http://frog.wix.com/bpm?_msid=b5...">http://frog.wix.com/bpm?_msid=b5...</a>	<a href="https://www.noelgarciaforcitycouncil.com/">https://www.noelgarciaforcitycouncil.com/</a>	Invalid (Unknown)		0 by...	Nov 1, 2023
Filter Match: "b=578&url=https://www.noelgarciaforcitycouncil.com/&v=1.13060"					
<input checked="" type="checkbox"/> <a href="http://frog.wix.com/bpm?_msid=b5...">http://frog.wix.com/bpm?_msid=b5...</a>	<a href="https://www.noelgarciaforcitycouncil.com/">https://www.noelgarciaforcitycouncil.com/</a>	Invalid (Unknown)		0 by...	Nov 1, 2023
Filter Match: "=3602&url=https://www.noelgarciaforcitycouncil.com/&v=1.13060"					
<input checked="" type="checkbox"/> <a href="http://frog.wix.com/pulse?_brandId...">http://frog.wix.com/pulse?_brandId...</a>	<a href="https://www.noelgarciaforcitycouncil.com/">https://www.noelgarciaforcitycouncil.com/</a>	Invalid (Unknown)		0 by...	Nov 1, 2023
Filter Match: "rc=72&url=https://www.noelgarciaforcitycouncil.com/&webvitals"					
<input checked="" type="checkbox"/> <a href="http://frog.wix.com/pulse?_brandId...">http://frog.wix.com/pulse?_brandId...</a>	<a href="https://www.noelgarciaforcitycouncil.com/">https://www.noelgarciaforcitycouncil.com/</a>	Invalid (Unknown)		0 by...	Nov 1, 2023
Filter Match: "rc=72&url=https://www.noelgarciaforcitycouncil.com/&webvitals"					

# 4. Metadata

[Home](#) [Collections](#) [Crawls](#) [Archives](#) [ARS](#) [Help Center](#) [Welcome, Eli Landaverde](#)

Created: Nov 1, 2023 Updated: Nov 9, 2023 In collection [Michigan Politics Web Crawling Project](#)

[Settings](#) **Metadata** [Groups](#) [Sharing](#) [Crawling History](#) [Notes](#) [Seed Scope](#)

[Edit](#)

[Import from WorldCat](#)

**Title** Noel Garcia for City Council

---

**Subject**

Michigan -- Politics and government

Elections -- Michigan

Political campaigns -- Michigan

City council members -- Michigan -- East Lansing

---

**Description** Official campaign website to elect Noel Garcia to East Lansing City Council.

---

**Contributor** Noel Garcia for East Lansing City Council

---

**Format** Web archives

---

**Coverage** North America -- United States -- Michigan -- East Lansing

A black and white photograph of a snail. The snail's shell is replaced by a clock face with numbers 1 through 12 and hands. The snail is moving slowly across a wooden surface, leaving a trail of mucus. The text "It's time consuming!" is overlaid in white on the clock face.

**It's time consuming!**

# Challenges

Out-of-date training manual

Social media

Ephemeral nature of the site's purpose

Content could be updated frequently on a site, but we only crawl a few times and might miss changes

User interface

Updating documentation

Metadata clean-up and streamlining the process

Capture PDFs for problematic websites

Long-term preservation storage of WARC files

Any suggestions or feedback for the future?

**The Future?**



Thank you!

Any questions?

Contact info:

Julia Ezzo: [julia@msu.edu](mailto:julia@msu.edu)

Eli Landaverde: [elandav@msu.edu](mailto:elandav@msu.edu)

Sarah Mainville: [Mainvil3@msu.edu](mailto:Mainvil3@msu.edu)