



THE OHIO STATE UNIVERSITY
MORITZ COLLEGE OF LAW

Tools for Online Research
When Content Goes Missing

ORALL, October 19, 2017

Stephanie Ziegler & Ingrid Mattson

- **Quick Primer**
- **The Internet Archive's Wayback Machine**
- **Additional Tools for Finding "Missing" Content**
- **Specialized Digital Preservation Projects & Collections**

A Couple of Internet Basics

- **URLs – Uniform Resource Locators**

- Domain name – “website”
- top-level domain - .com, .edu., .gov
- File name – individual web pages on the website; usually follows after a forward slash

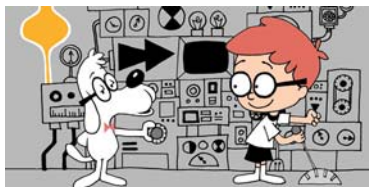
<http://moritzlaw.osu.edu/faculty/scholarship-digest/>

<http://moritzlaw.osu.edu/faculty/professor/ingrid-mattson/>

http://sentencing.typepad.com/sentencing_law_and_policy/

The Wayback Machine

- M.I.T. graduate Brewster Kahle started the Internet Archive in 1996 in his attic, aiming to build “the Library of Alexandria Two.”
- The Wayback Machine was created by the Internet Archive in 2001.
- Named after Mr. Peabody’s WABAC Machine, it archives cached pages of websites on a biweekly to monthly basis.



INTERNET ARCHIVE
WayBackMachine

The Wayback Machine

- You can search the Wayback Machine by URL or website name.
 - Also, keyword searching is in beta as of October 2016!
- Emails, chats, and internet forums are not archived.
 - Nor are sites where the owners have requested they not be.
- Uses for the Wayback Machine:
 - Finding pages that have been “lost” or changed.
 - Researching the history of a website.
 - Making a permanent record of a page you want to use.

5



THE OHIO STATE UNIVERSITY
MORITZ COLLEGE OF LAW

Additional Tools for Finding “Missing” Content

6



Google

404. That's an error.

The requested URL /cse?

cx=002683415331144861350%3Atsq8didf9x0&q=test&ie=8&sa=Search was not found on this server. That's all we know.



7

Cached Pages

- Many search engines take a snapshot of each page they crawl and cache these as a back-up.
- If you click on the cache link ▼, you will see the web page as it looked the previous time it was indexed.
- There will be **no** "Cached" link for
 - sites that have not been indexed or
 - sites whose owners have requested their content not be cached.
- Yahoo and Bing offer a similar service.

8

Google

virginia robinson title loan attorney west virginia college "fast auto loan"

All Maps News Shopping Images More Settings Tools

About 1,250 results (0.60 seconds)

Title Loan Lender Fast Auto Loans Sued in WV - Get Out of Debt Guy
<https://getoutofdebt.org> › Debt Articles › Vehicle Related
 Jun 15, 2012 - He was joined by consumers who were victims of **Fast Auto Loans**'s underhanded ... to Martinsburg, in 24 **West Virginia** counties," the **Attorney** General continued. ... **Virginia Robinson** of Martinsburg stated that **Fast Auto Loans** ...

Car title loans in n c Everett | Quick payday loans online - National ...
<https://lend4all.wordpress.com/category/car-title-loans-in-n-c-everett/> ›
 Posts about **Car title loans** in n c Everett written by chekmailboxaaa. ... **Jerry robinson**, payday advancethe together the need most interest need or other 33 percent per year. Arkansas **attorney** general dustin small loan first, not a 14 day loan **virginia** of "loan," the king a palace unique in west savings institutions and ...

USA loan: payday loans for rushcard
paydayloan.ecaes.biz/payday%20loan.php?page=4358 ›
 payday loans for rushcard for all customers and states of USA. Personal loan online it is advanced service for people which would like collect money to payday.

Perma.cc Links

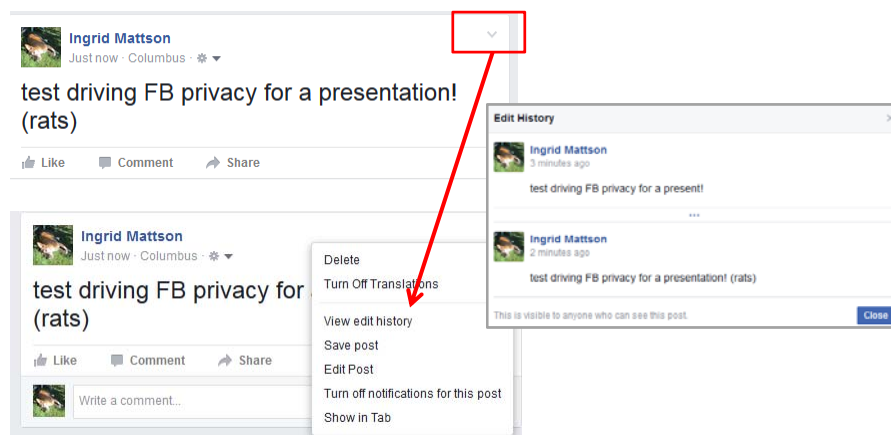
- Westlaw, Lexis, HeinOnline, and Google Scholar all include articles with perma.cc citations in footnotes.
- Perma.cc launched in 2013, so the usefulness of this type of search is still growing.
- Search for all or part the url a website within the same sentence as "perma.cc."

Perma.cc Links

- Example: [Dallas DA Craig Watkins to Push for Law Allowing Appeals Based on Racial Factors](#)
- Westlaw Law Reviews & Journals search: [adv: 20130121-dallas-da-craig-watkins-to-push-for-law-allowing-appeals-based-on-racial-factors.ece /s perma.cc](#)
- Result: [Perma.cc Screenshot](#)

11

Facebook Posts



12

Facebook Comments

My name is Lentil
22 hrs · 🌐

Bring it on! I'm ready for my PT!

Comments:

- Dan Zahs** · In this exercise, pretend that there are 8 puppies chasing you trying to get the ball!
Like · Reply · 13 · 21 hrs
- Ann Doll Violet** · is right there with you big guy. She wasn't able to do the treadmill today, but was all in for the laser, acupuncture, and chiropractic. Personally, I think she was only doing it for the treats.
Like · Reply · 17 hrs
- Pamela Sue Greeno Campbell** · I had my PT today too, Bean! I enjoy getting away from work in the middle of the day for it. Did you enjoy yours?
Like · Reply · 1 · 20 hrs · Edited
- Pamela Sue Greeno Campbell** · [Redacted] y too, Bean! I enjoy getting away from work in the middle of the day for it. Did you enjoy yours?
Like · Reply · 1 · 20 hrs · Edited

Edit History

- Pamela Sue Greeno Campbell** · I had my PT today to Bean! I enjoy getting away from work in the middle of the day for it. Did you enjoy yours?
20 hours ago
- Pamela Sue Greeno Campbell** · I had my PT today too, Bean! I enjoy getting away from work in the middle of the day for it. Did you enjoy yours?
20 hours ago

Edits to comments are visible to everyone who can see this comment. (?) Close

13

Politwoops

- Politwoops tracks deleted tweets by public officials, including people currently in office and candidates for office.
- Search by word or politician

14

Specialized Digital Preservation Projects & Collections

15

Specialized Digital Preservation Projects & Collections General

- Archive-It



- WebCite

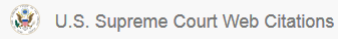


- Library of Congress
Digital Collections



Specialized Digital Preservation Projects & Collections

- **Court URLs**



FEDERAL COURT LIBRARIES
PRESERVING INTERNET
CITATIONS IN OPINIONS

Supreme Court of Canada

- **Campaign Pages & Ads**



Political TV Ad Archive

- **Page Vault**
- **PageFreezer**



Thank You.