



# Beyond the Affidavit: Towards Better Standards for Web Archive Evidence

Nicholas Taylor  
[nullhandle.org](http://nullhandle.org)

IIPC Web Archiving Conference  
Session 7: Legal and Ethical Considerations  
3 May 2023

"Origami" by Helgi Halldórsson under CC BY 2.0 

# a historical webpage(?)

http://www.wunderground.com/cgi-bin/findweather/getForecast?query=50593

34 captures  
27 Oct 2002 - 10 Jun 2017

Go AUG DEC MAR  
09 2003 2004 2006

About this capture

wunderground.com

Find the Weather for any City, State or ZIP Code, or Airport Code or Country

Features: Wunder Photos NEXRAD radar Regional Radar NEW Zoom Satellite Maps Trip Planner

Tropical / Hurricane Marine Severe Astronomy Ski Education Favorites Personal Weather Stations

Member Benefits:  
No Ads  
Weather Email  
\$5 a year  
Signup Here

Email taylor@gmail.com  
Password  
Login  
Print This Page

Maps

Temperature  
Heat Index  
Windchill  
Humidity  
Radar  
Dew Point  
Wind  
Visibility  
Visible Satellite  
Satellite  
Fronts  
Snow Depth  
Jet Stream  
Flight Rules  
Health Maps

UV Forecast  
Air Quality  
Peak Air Quality Index

International  
United States  
Canada  
Mexico  
Central America  
South America  
Northern Asia

Varina, Iowa

Local Time: 1:09 PM CST Set My Timezone

Add this page to your Favorites Weather Email

Lat/Lon: 42.6° N 94.8° W | MSN Map

Current Conditions

Updated: 12:55 PM CST on December 09, 2004  
Observed at Storm Lake, Iowa (History)  
Elevation: 1486 ft / 453 m

41 °F / 5 °C Light Drizzle

Windchill: 37 °F / 3 °C  
Humidity: 100%  
Dew Point: 41 °F / 5 °C  
Wind: 6 mph / 9 km/h from the South  
Pressure: 29.56 in / 1001 hPa  
Visibility: 5.0 miles / 8.0 kilometers  
UV: 1 out of 12  
Clouds Overcast 6500 ft / 1981 m (AGL):  
Show full METARS (help)

Local Radar  
Regional Radar  
Local Satellite  
Ski Conditions  
Trip Planner  
Personal Weather Stations

5-Day Forecast for ZIP Code 50593

Thu	Fri	Sat	Sun	Mon
43°   29°	36°   22°	47°   34°	41°   15°	31°   15°
Chance of Rain	Mostly Cloudy	Mostly Cloudy	Partly Cloudy	Partly Cloudy
<a href="#">Detail</a>				

Click Detail for hourly wind, temperature, humidity and UV forecasts.

Alternate Computer Forecast: AVN MOS Weather Graph | Local Allergy Info from Pollen.com

Meet Locals  
read their Hotel Tips and Travel Guides for Varina, IA

Unbiased reviews and advice for Varina hotels and other cities

in Varina  
Best Rates GUARANTEED  
Discount Hotels

Wunder Photos

Show Varina or Iowa

Now  
Dense fog will persist along and west of Highway 169 through the early afternoon. Locations from Algona to Estherville can expect visibilities to remain around a quarter mile or less with localized areas in low lying spots under one eighth of a mile.

Nowcast as of 11:42 am CST on December 9, 2004

Forecast for Pocahontas County

Updated: 11:20 am CST on December 9, 2004  
No Active Advisories (US Severe Weather)

This Afternoon  
Cloudy with a chance of light rain showers. High in the lower 40s. West wind 10 to 15 mph. Chance of rain 30 percent.  
[Detail](#)

Tonight  
Light rain likely through the evening. Light snow likely after midnight. Little or no snow around 30. Northwest wind 15 to 20 mph. Chance of snow 30 percent.

Scott G. Ainsworth, et al.: "Only One Out of Five Archived Web Pages Existed as Presented"

Internet Archive: "Weather Underground: Varina, Iowa Forecast"



# a historical webpage(?)

Find the Weather for any City, State or ZIP Code, or Airport Code or Country

Features: Wunder Photos NEXRAD radar Regional Radar NEW Zoom Satellite Maps Trip Planner

Tropical / Hurricane Marine Severe Astronomy Ski Education Favorites Personal Weather Stations

**Varina, Iowa**

Local Time: 1:09 PM CST Set My Timezone

Current Conditions

Updated: 12:55 PM CST on December 09, 2004  
Observed at Storm Lake, Iowa (History)  
Elevation: 1486 ft / 453 m

**41 °F / 5 °C**  
Light Drizzle

Windchill: 37 °F / 3 °C  
Humidity: 100%  
Dew Point: 41 °F / 5 °C  
Wind: 6 mph / 9 km/h from the South  
Pressure: 29.56 in / 1001 hPa  
Visibility: 5.0 miles / 8.0 kilometers  
UV: 1 out of 12  
Clouds: Overcast 6500 ft / 1981 m (AGL):  
Show full METARS (help)

Local Radar Regional Radar Local Satellite Ski Conditions Trip Planner Personal Weather Stations

Add this page to your Favorites Weather Email

Lat/Lon: 42.6° N 94.8° W | MSN Map

5-Day Forecast for ZIP Code 50593

Thu	Fri	Sat	Sun	Mon
43°   29°	36°   22°	47°   34°	41°   15°	31°   15°
Chance of Rain	Mostly Cloudy	Mostly Cloudy	Partly Cloudy	Partly Cloudy
<a href="#">Detail</a>				

Click Detail for hourly wind, temperature, humidity and UV forecasts.

Alternate Computer Forecast: AVN MOS Weather Graph | Local Allergy Info from Pollen.com

Meet Locals  
read their Hotel Tips and Travel Guides for Varina, IA

Unbiased reviews and advice for Varina hotels and other cities

Best Rates GUARANTEED Discount Hotels

Wunder Photos

Show -Varina or -Iowa

Now  
Dense fog will persist along and west of Highway 169 through the early afternoon. Locations from Algona to Estherville can expect visibilities to remain around a quarter mile or less with localized areas in low lying spots under one eighth of a mile.

Forecast for Pocahontas County

Updated: 11:20 am CST on December 9, 2004  
No Active Advisories (US Severe Weather)

This Afternoon  
Cloudy with a chance of light rain showers. High in the lower 40s. West wind 10 to 15 mph. Chance of rain 30 percent.  
[Detail](#)

Tonight  
Light rain likely through the evening. Light snow likely after midnight. Little or no snow around 30. Northwest wind 15 to 20 mph. Chance of snow 80 percent.

Scott G. Ainsworth, et al.: "Only One Out of Five Archived Web Pages Existed as Presented"

Internet Archive: "Weather Underground: Varina, Iowa Forecast"



# a historical webpage(?)

Find the Weather for any City, State or ZIP Code, or Airport Code or Country

Features: Wunder Photos NEXRAD radar Regional Radar NEW Zoom Satellite Maps Trip Planner Tropical / Hurricane Marine Severe Astronomy Ski Education Favorites Personal Weather Stations

Member Benefits:  
No Ads  
Weather Email  
\$5 a year  
Signup Here

Email: taylor@gmail.com  
Password: \*\*\*\*\*  
Login  
Print This Page  
Maps  
Temperature  
Heat Index  
Windchill  
Humidity  
Radar  
Dew Point  
Wind  
Visibility  
Visible Satellite  
Satellite  
Fronts  
Snow Depth  
Jet Stream  
Flight Rules  
Health Maps  
UV Forecast  
Air Quality  
Peak Air Quality  
Index  
International  
United States  
Canada  
Mexico  
Central America  
South America  
Northern Asia

Varina, Iowa

Local Time: 1:09 PM CST Set My Timezone

41 °F / 5 °C Light Drizzle

Windchill: 37 °F / 3 °C  
Humidity: 100%  
Dew Point: 41 °F / 5 °C  
Wind: 6 mph / 9 km/h from the South  
Pressure: 29.56 in / 1001 hPa  
Visibility: 5.0 miles / 8.0 kilometers  
UV: 1 out of 12  
Clouds: Overcast 6500 ft / 1981 m (AGL)  
Show full METARS (help)

43° | 29° 36° | 22° 47° | 34° 41° | 15° 31° | 15°

Chance of Rain Mostly Cloudy Mostly Cloudy Partly Cloudy Partly Cloudy

Detail Detail Detail Detail Detail

Click Detail for hourly wind, temperature, humidity and UV forecasts.

5-Day Forecast for ZIP Code 50593

Now Dense fog will persist along and west of Highway 169 through the early afternoon. Locations from Algona to Estherville can expect visibilities to remain around a quarter mile or less with localized areas in low lying spots under one eighth of a mile.

Forecast for Pocahontas County

No Active Advisories (US Severe Weather)

This Afternoon Cloudy with a chance of light rain showers. High in the lower 40s. West wind 10 to 15 mph. Chance of rain 30 percent.  
Detail

Tonight Light rain likely through the evening. Light snow likely after midnight. Little or no snow around 30. Northwest wind 15 to 20 mph. Chance of snow 80 percent.

in Varina  
Best Rates GUARANTEED  
Discount Hotels

Show Varina or Iowa

Wunder Photos

Talk to State Farm agents near you:

Conditions Nearby

City	Temp	Conditions	Updated
Storm Lake	41 °F / 5 °C	Light Drizzle	12:55 PM CST
Fort Dodge	41 °F / 5 °C	Partly	12:55 PM CST

Scott G. Ainsworth, et al.: "Only One Out of Five Archived Web Pages Existed as Presented"

Internet Archive: "Weather Underground: Varina, Iowa Forecast"



# ahistorical webpage

INTERNET ARCHIVE  
**Wayback Machine**

<http://www.wunderground.com/cgi-bin/findweather/getForecast?query=50593> Go AUG DEC MAR  
34 captures 09 2003 2004 2006 About this capture

27 Oct 2002 - 10 Jun 2017

COLLECTED BY Alexa Crawls EA

Organization: Alexa Crawls  
Starting in 1996, Alexa Internet has been donating their crawl data to the Internet Archive. Flowing in every day, these data are added to the Wayback Machine after an embargo period.

Features: Wunder Photos Radar Regional Radar Weather Satellite Maps Trip Planner

**TIMESTAMPS**

<http://icons.wunderground.com/data/wximagednew/a/abby/29-banner.jpg> +9 years 5 months  
<http://icons.wunderground.com/data/wximagednew/d/d70dcave/0-thumb.jpg> +9 years 18 days  
<http://banners.wunderground.com/cgi-bin/statefarmbanner?zip=50593&width=150> +1 year 3 months  
[http://www.wunderground.com/data/ids/DMX19\\_thumb.jpg](http://www.wunderground.com/data/ids/DMX19_thumb.jpg) +9 months 7 days  
<http://icons.wunderground.com/graphics/smash/blue/localRadar.gif> -20 days 1 minute  
<http://icons.wunderground.com/graphics/smash/blue/weatherStation.gif> -20 days 52 seconds  
<http://icons.wunderground.com/graphics/highlightr.gif> -20 days 48 seconds  
[http://icons.wunderground.com/graphics/smash/blue\\_noWarning.gif](http://icons.wunderground.com/graphics/smash/blue_noWarning.gif) -20 days 43 seconds  
[http://icons.wunderground.com/graphics/smash/red\\_satellite.gif](http://icons.wunderground.com/graphics/smash/red_satellite.gif) -20 days 38 seconds  
<http://icons.wunderground.com/graphics/highlightr.gif> -20 days 37 seconds  
<http://icons.wunderground.com/graphics/lowlift.gif> -20 days 18 seconds  
[http://icons.wunderground.com/graphics/smash/blue\\_add3.gif](http://icons.wunderground.com/graphics/smash/blue_add3.gif) -20 days 16 seconds  
[http://icons.wunderground.com/graphics/smash/blue\\_ski.gif](http://icons.wunderground.com/graphics/smash/blue_ski.gif) -20 days 13 seconds  
[http://icons.wunderground.com/graphics/smash/blue\\_envelope2.gif](http://icons.wunderground.com/graphics/smash/blue_envelope2.gif) -20 days 11 seconds  
[http://icons.wunderground.com/graphics/smash/blue\\_tripPlanner.gif](http://icons.wunderground.com/graphics/smash/blue_tripPlanner.gif) -20 days 4 seconds  
<http://icons.wunderground.com/graphics/lowlightr.gif> -20 days 2 seconds  
[http://icons.wunderground.com/graphics/smash/blue\\_radar.gif](http://icons.wunderground.com/graphics/smash/blue_radar.gif) -19 days 23 hours

Add this page to your Favorites Weather Email Lat/Lon: 42.6° N 94.8° W | MSH Forecast

**5-Day Forecast for ZIP Code 50593**

Thu	Fri	Sat	Sun	Mon
43°   29° Chance of Rain	36°   22° Mostly Cloudy	47°   34° Mostly Cloudy	41°   15° Partly Cloudy	31°   15° Partly Cloudy
<a href="#">Detail</a>	<a href="#">Detail</a>	<a href="#">Detail</a>	<a href="#">Detail</a>	<a href="#">Detail</a>

Click Detail for hourly wind, temperature, humidity and UV forecasts.

**Alternate Computer Forecast:** AVN MOS Weather Graph | Local Allergy Info from Pollen.com

**Meet Locals**  
read their Hotel Tips and Travel Guides for Varina, IA

**in Varina**  
Unbiased reviews and advice for Varina hotels and other cities

**Best Rates GUARANTEED Discount Hotels**

**Wunder Photos**

**Show**  
Varina or Iowa

**Now**  
Dense fog will persist along and west of Highway 169 through the early afternoon. Locations from Algona to Estherville can expect visibilities to remain around a quarter mile or less with localized areas in low lying spots under one eighth of a mile.

**Nowcast as of 11:42 am CST on December 9, 2004**

**Forecast for Pocahontas County**

Updated: 11:20 am CST on December 9, 2004

No Active Advisories (US Severe Weather)

**This Afternoon**  
Cloudy with a chance of light rain showers. High in the lower 40s. West wind 10 to 15 mph. Chance of rain 30 percent.  
[Detail](#)

**Tonight**  
Light rain likely through the evening. Light snow likely after midnight. Little or no snow around 30. Northwest wind 15 to 20 mph. Chance of snow 30 percent.

**Health Maps**

[Heat Index](#)  
[Windchill](#)  
[Humidity](#)  
[Radar](#)  
[Dew Point](#)  
[Wind](#)  
[Visibility](#)  
[Visible Satellite](#)  
[Satellite](#)  
[Fronts](#)  
[Snow Depth](#)  
[Jet Stream](#)  
[Flight Rules](#)

[Talk to State Farm agents near you:](#)

**Conditions Nearby**

City	Temp	Conditions	Updated
Storm Lake	41°F / 5°C	Light Drizzle	12:55 PM CST
Fort Dodge	41°F / 5°C	Partly Cloudy	12:55 PM CST

Scott G. Ainsworth, et al.: "Only One Out of Five Archived Web Pages Existed as Presented"

Internet Archive: "Weather Underground: Varina, Iowa Forecast"



# ahistorical webpage

INTERNET ARCHIVE  
**Wayback Machine**

http://www.wunderground.com/cgi-bin/findweather/getForecast?query=50593

34 captures | Go AUG DEC MAR  
27 Oct 2002 - 10 Jun 2017 | < 09 > 2003 2004 2006 | About this capture

COLLECTED BY Alexa Crawls EA  
Organization: Alexa Crawls  
Starting in 1996, Alexa Internet has been donating their crawl data to the Internet Archive. Flowing in every day, these data are added to the Wayback Machine after an embargo period.

Features Wunder Photos NEXTAD Radar Regional Radar Weather Radar Maps Trip Planner

TIME STAMPS

Collection: Alexa Crawls EA  
Crawl data donated by Alexa Internet. This data is currently not publicly accessible

http://icons.wunderground.com/data/wximagenew/a/abby/29-banner.jpg +9 years 5 months  
http://icons.wunderground.com/data/wximagenew/d/d70dave/0-thumb.jpg +9 years 18 days  
http://banners.wunderground.com/cgi-bin/statefarmbanner?zip=50593&width=150 +1 year 3 months  
http://www.wunderground.com/data/nids/DMX19\_thumb\_t.jpg +9 months 7 days

Add this page to your Favorites Weather Email Lat/Lon: 42.6° N 94.8° W | MSN

5-Day Forecast for ZIP Code 50593

Thu Fri Sat Sun Mon

http://www.wunderground.com/data/nids/DMX19\_thumb\_t.jpg +9 months 7 days

[http://icons.wunderground.com/data/wximagenew/a/abby/29-banner.jpg +9 years 5 months](http://icons.wunderground.com/data/wximagenew/a/abby/29-banner.jpg)  
[http://icons.wunderground.com/data/wximagenew/d/d70dave/0-thumb.jpg +9 years 18 days](http://icons.wunderground.com/data/wximagenew/d/d70dave/0-thumb.jpg)  
[http://banners.wunderground.com/cgi-bin/statefarmbanner?zip=50593&width=150 +1 year 3 months](http://banners.wunderground.com/cgi-bin/statefarmbanner?zip=50593&width=150)  
[http://www.wunderground.com/data/nids/DMX19\\_thumb\\_t.jpg +9 months 7 days](http://www.wunderground.com/data/nids/DMX19_thumb_t.jpg)

Radar  
Dew Point  
Wind  
Visibility  
Visible Satellite  
Satellite  
Fronts  
Snow Depth  
Jet Stream  
Flight Rules  
Health Maps  
UV Forecast  
Air Quality  
Peak Air Quality Index  
International  
United States  
Canada  
Mexico  
Central America  
South America  
Northern Asia

9 km/h from the South  
Pressure: 29.56 in / 1001 hPa  
Visibility: 5.0 miles / 8.0 kilometers  
UV: 1 out of 12  
Clouds Overcast 6500 ft / 1981 m (AGL):  
Show full METARS (help)

Local Radar  
Regional Radar  
Local Satellite  
Ski Conditions  
Trip Planner  
Personal Weather Stations

read their Hotel Tips and Travel Guides for Varina, IA  
Unbiased reviews and advice for Varina hotels and other cities  
Best Rates GUARANTEED Discount Hotels

Show -Varina or Iowa

Now  
Nowcast as of 11:42 am CST on December 9, 2004  
Dense fog will persist along and west of Highway 169 through the early afternoon. Locations from Algona to Estherville can expect visibilities to remain around a quarter mile or less with localized areas in low lying spots under one eighth of a mile.

Forecast for Pocahontas County  
Updated: 11:20 am CST on December 9, 2004  
No Active Advisories (US Severe Weather)

This Afternoon  
Cloudy with a chance of light rain showers. High in the lower 40s. West wind 10 to 15 mph. Chance of rain 30 percent.  
Detail

Tonight  
Light rain likely through the evening. Light snow likely after midnight. Little or no snow around 30. Northwest wind 15 to 20 mph. Chance of snow 60 percent.

Scott G. Ainsworth, et al.: "Only One Out of Five Archived Web Pages Existed as Presented"

Internet Archive: "Weather Underground: Varina, Iowa Forecast"



# ahistorical webpage

Text on webpage and associated infographic indicate current conditions as raining. Infographic captured on **12 December 2004 at 8:54a CST.**

Timestamp in text on webpage corroborates time of capture (i.e., **9 December 2004 at 1:09p CST**).

Radar graphic reflects weather conditions (clear) at the time of its capture: **12 September 2005 at 5:34p CDT.**

Internet Archive: "Weather Underground: Varina, Iowa Forecast"

Scott G. Ainsworth, et al.: "Only One Out of Five Archived Web Pages Existed as Presented"

The screenshot shows the Weather Underground homepage for Varina, Iowa. It features a search bar at the top with the placeholder "Find the Weather for any City, State or ZIP Code, or Airport Code or Country". Below the search bar is a navigation menu with links to "Features", "Wunder Photos", "NEXRAD radar", "Regional Radar", "Zoom Satellite", "Map", "Tropical / Hurricane", "Marine", "Severe", "Astronomy", "Ski", "Education", "Favorites", and "Personal Weather Stations". On the left side, there's a sidebar for "Member Benefits" including "No Ads", "Weather Email \$5 a year", and a "Signup Here" button. There are also links for "Email", "Password", "Login", "Print This Page", "Maps", and "Temperature". The main content area displays the following information:

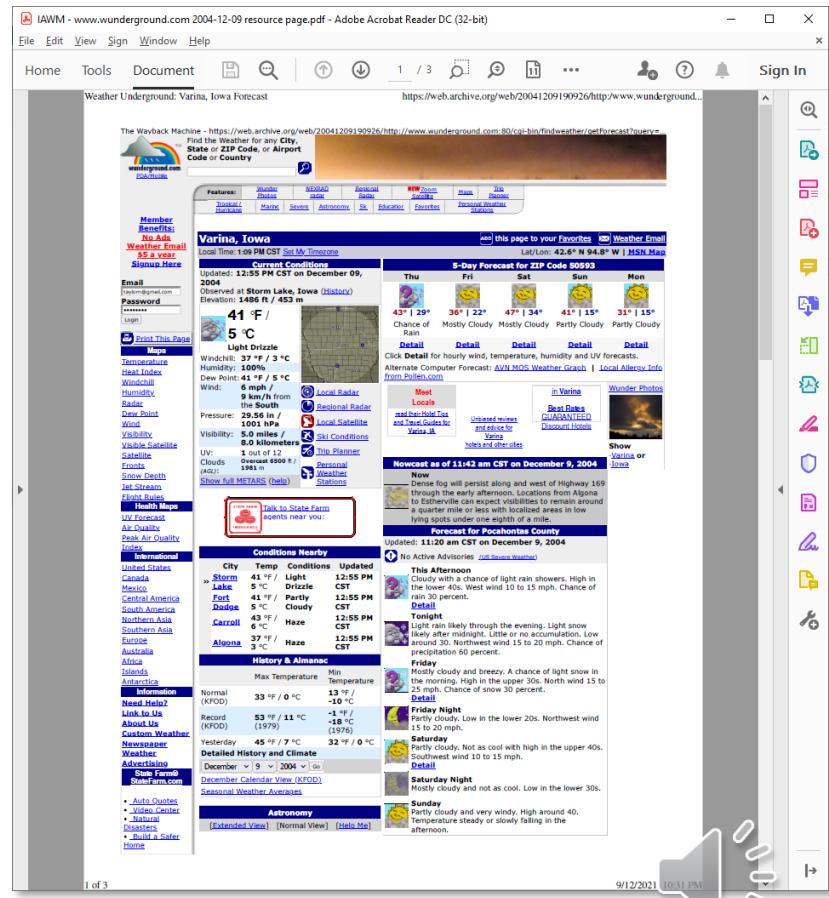
- Varina, Iowa**
- Local Time: 1:09 PM CST [Set My Timezone](#)
- Updated: 12:55 PM CST on December 09, 2004
- Observed at Storm Lake, Iowa ([History](#))
- Elevation: 1486 ft / 453 m
- Current Conditions**: 41 °F / 5 °C, Light Drizzle, Windchill: 32 °F / 3 °C
- 5-day Forecast for ZIP Code 50593**:
  - Thu: 43° | 29°
  - Fri: 36° | 22°
  - Sat: 47° | 34°
  - Sun: 41° | 15°
  - Mon: 31° | 15°
- Click Detail** for hourly wind, temperature, humidity and UV forecasts.
- Alternate Computer Forecast**: AVN MOS Weather Graph | Local Allergy Info from Pollen.com
- Radar graphic**: Shows a clear sky over the region.
- Conditions Nearby**:

City	Temp	Conditions	Updated
Storm Lake	41 °F / 5 °C	Light Drizzle	12:55 PM CST
Fort Dodge	41 °F / 5 °C	Partly	12:55 PM CST
- Local Services**:
  - Local Radar
  - Regional Radar
  - Local Satellite
  - Ski Conditions
  - Trip Planner
  - Personal Weather Stations
- Now**: Dense fog will persist along and west of Highway 169 through the early afternoon. Locations from Algona to Estherville can expect visibilities to remain around a quarter mile or less with localized areas in low lying spots under one eighth of a mile.
- Forecast for Pocahontas County**: Updated: 11:20 am CST on December 9, 2004
- This Afternoon**: Cloudy with a chance of light rain showers. High in the lower 40s. West wind 10 to 15 mph. Chance of rain 30 percent.
- Tonight**: Light rain likely through the evening. Light snow likely after midnight. Little or no snow accumulation expected. Low around 30. Northwest wind 15 to 20 mph. Chance of snow 60 percent.



# misleading PDF

- IAWM evidence submitted as PDF screenshots
- temporal variations no longer apparent
- may be taken for granted as historical facsimile



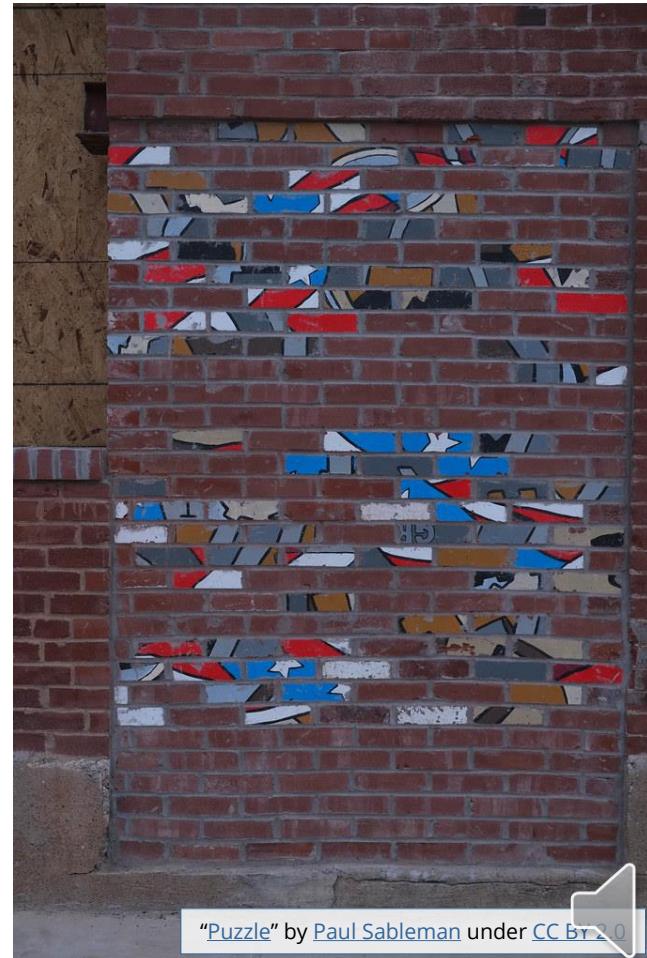
# why does this matter for evidentiary use?

- cases may turn on temporal attributes of embedded mementos, specifically
- bears on arguments about authenticity – i.e., that evidence is what it purports to be
- IAWM evidence generally coming in for less scrutiny now than it should
- if discrepancy becomes apparent, will likely prompt more scrutiny of IAWM evidence than it should



# why less scrutiny?

- legal professionals don't understand web archives
- questions of authentication + admissibility foregrounded
- too neatly addressed w/ IA standard affidavit



"Puzzle" by Paul Sableman under CC BY 2.0



# flawed approach

- correct admissibility determinations for wrong reasons
- declining to admit probative evidence
- IAWM evidence should generally be admitted + treated as authentic
- seems to be consistent w/ IA's perspective



"Wrong Way Around" by [Colin and Sarah Northway](#) under [CC BY 2.0](#)

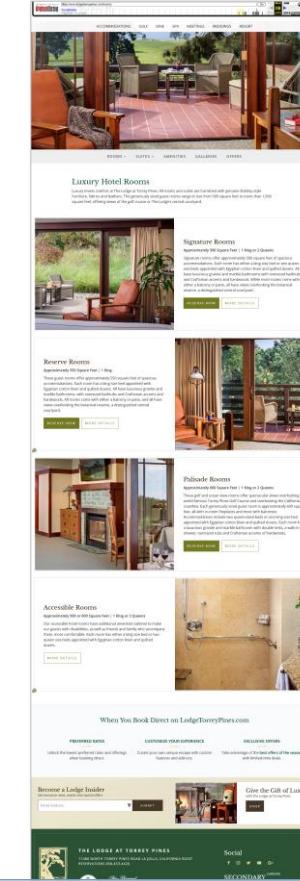


# *Rutherford v. Evans Hotels, LLC (2020)*

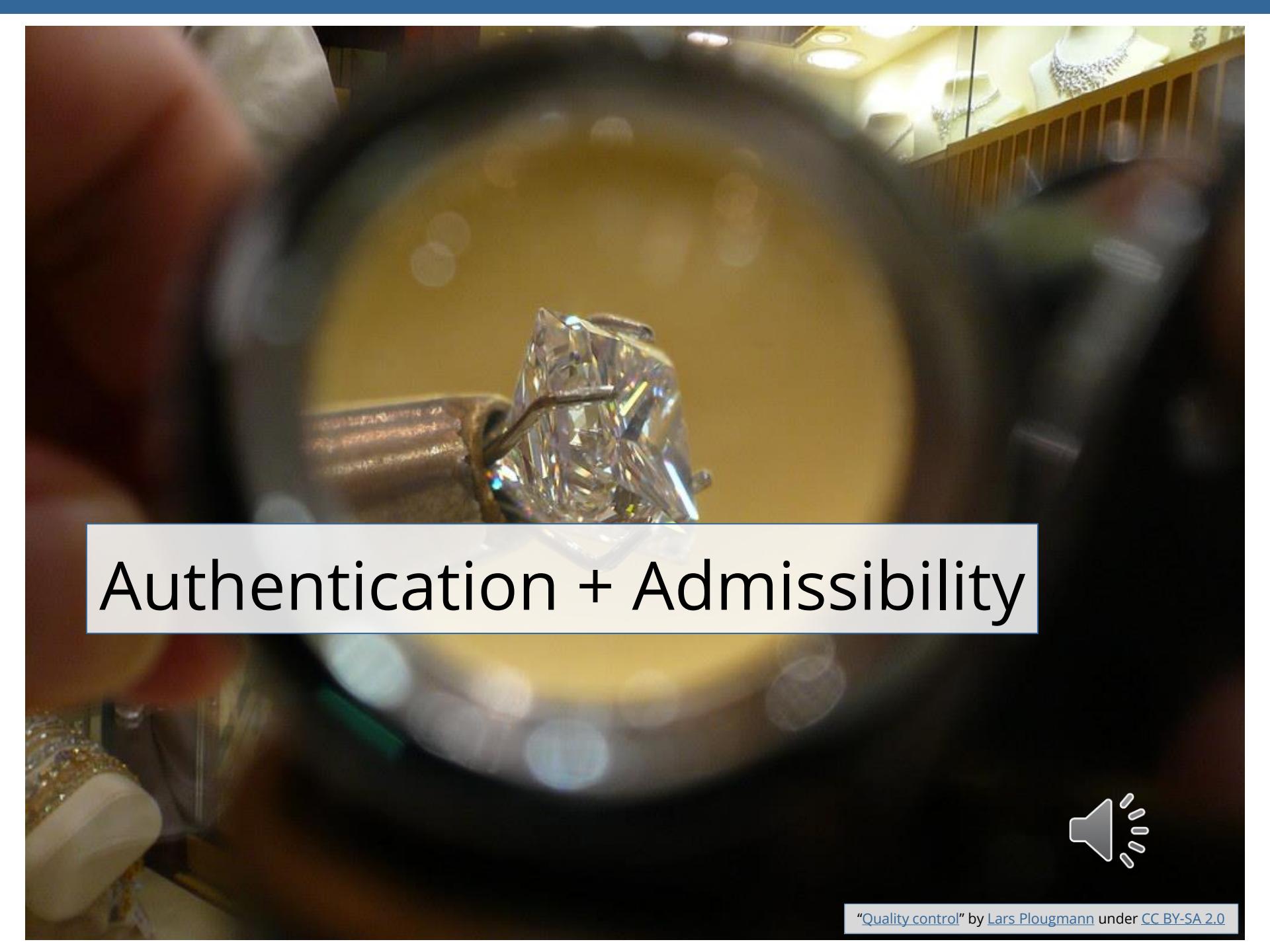
- [PageFreezer](#) archives used to show lack of compliance of online hotel reservation system w/ ADA
- [PageFreezer](#) archives denied admission

*"[T]he Court has found **no cases addressing the admissibility of, let alone referencing, 'PageFreezer' archives.** Coupled with the differences between PageFreezer and the Internet Archive—including that the Internet Archive offers a free, publicly available service whereas PageFreezer offers a paid service in anticipation of litigation—the Court declines to admit the PageFreezer archives as analogous to those from the Wayback Machine admitted in other litigation."*

*Rutherford v. Evans Hotels, LLC, No. 18-CV-435 JLS (MSB),  
2020 WL 5257868, at \*1 (S.D. Cal. Sept. 3, 2020)*



*Internet Archive: "La Jolla Resort - Hotels in La Jolla CA | Lodge Torrey Pines"*



# Authentication + Admissibility



# authentication + admissibility overview

- IAWM cited in 300+ U.S. federal court cases
- harder to authenticate than other web content
- strategies include:
  - judicial notice
  - IA affidavit
  - witness w/ personal knowledge
  - expert witness
  - mutual stipulation



# authentication via judicial notice

- judge exercises discretion to admit uncontested info
- efficient but inconsistent + jurisdiction-sensitive
- more likely to prevail in absence of specific assertions of unreliability



["Earthrise"](#) by Bill Anders under [Public Domain](#)



# authentication via witness w/ personal knowledge

- testimony by witness w/  
personal knowledge  
that evidence is what it's  
claimed to be
- requires a witness w/  
actual, incidental  
personal knowledge of  
historical web content
- less effective when  
instrumentally entering  
into record a  
comparison performed  
by a random person



# authentication via expert witness

- testimony by an expert w/ specialized knowledge
- reliable but more involved + less-used mechanism
- useful when interpretation, explanation, or qualification needed



"Test tubes and other recipients in chemistry lab" by Old Photo Profile under CC BY 2.0

# authentication via IA affidavit

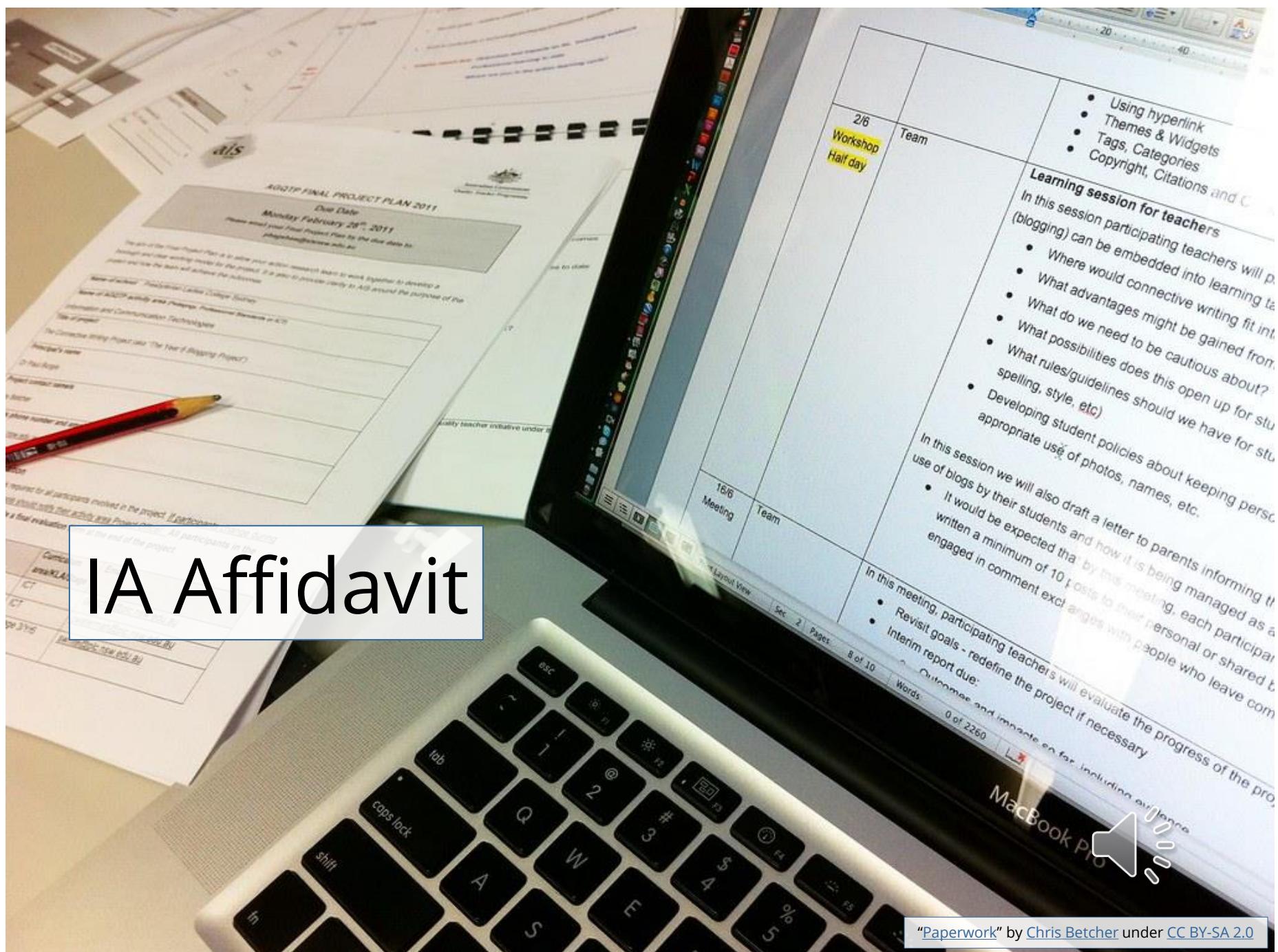
- reliable + efficient
- has only failed when incorrectly applied
- works best for simple cases
- but less ideal when conclusions require interpretation



"Paperwork" by Damian Gadal under CC BY 2.0



# IA Affidavit



# IA affidavit guidance

## **Do I really need an affidavit from the Internet Archive?**

No. Please consider alternatives to an affidavit from the Internet Archive. Judicial notice and stipulation to a document's authenticity are two typical and straightforward options that might be used instead of an affidavit. Since our resources are limited, we urge you to pursue these alternatives before coming to us with authentication requests.



[Internet Archive: "Legal FAQ"](#)

# IA affidavit guidance

## Do I really need an affidavit from the Internet Archive?

No. Please consider alternatives to an affidavit from the Internet Archive. Judicial notice and stipulation to a document's authenticity are two typical and straightforward options that might be used instead of an affidavit. Since our resources are limited, we urge you to pursue these alternatives before coming to us with authentication requests.

## Does the Internet Archive's affidavit mean that the printout was actually the page posted on the Web at the recorded time?

The Internet Archive's affidavit only affirms that the printed document is a true and correct copy of our records. It remains your burden to convince the finder of fact what pages were up when.



[Internet Archive: "Legal FAQ"](#)

# IA affidavit

INTERNET ARCHIVE G25»

ABOUT BLOG PROJECTS HELP DONATE CONTACT JOBS VOLUNTEER PEOPLE

SIGN UP | LOG IN UPLOAD Search

## News [more]

Hooniverse: Wayback Machine Allows a Peek into Defunct Detroit Automaker Websites

Laughing Squid: An Amazing Collection of Pulp Magazines Going Back 75 Years Is Available Online at The Internet Archive

Far Out Magazine: Over 100,000 historic vinyl records are being digitised and made available to stream online for free

GigaZine: ウェブ上の情報を記録・保存する「インターネット・アーカイブ」の存続をひっそりとめかしているものとは?

Actualitte: Plongez dans l'art japonais de la fin du XIXe siècle grâce à ce magazine numérisé

Library Journal: Better World Libraries, Internet Archive Partner, Acquires Better World Books

Open Culture: The Internet Archive Is Digitizing & Preserving Over 100,000 Vinyl Records: Hear 750 Full Albums Now

Against The Grain: ATG Newsflash: For the Love of Literacy—Better World Books and the Internet Archive Unite to Preserve Millions of Books

Research Information: Better World Books affiliates with Internet Archive

Wired: The Internet Archive Is Making Wikipedia More Reliable

## Standard Affidavit

1. I am a Records Request Processor at the Internet Archive, located in San Francisco, California. I make this declaration of my own personal knowledge.

2. The Internet Archive is a website that provides access to a digital library of Internet sites and other cultural artifacts in digital form. Like a paper library, we provide free access to researchers, historians, scholars, and the general public. The Internet Archive has partnered with and receives support from various well-known institutions and libraries, including the Library of Congress.

3. The Internet Archive has created a service known as the Wayback Machine. The Wayback Machine makes it possible to browse more than 450 billion pages stored in the Internet Archive's web archive. Visitors to the Wayback Machine can search archives by URL (i.e., a website address). If archived records for a URL are available, the visitor will be presented with a display of available dates. The visitor may select one of those dates, and begin browsing an archived version of the Web. Links on archived files in the Wayback Machine point to other archived files (whether HTML pages or other file types), if any are found for the URL indicated by a given link. For instance, the Wayback Machine is designed such that when a visitor clicks on a hyperlink on an archived page that points to another URL, the visitor will be served the archived file found for the hyperlink's URL with the closest available date to the initial file containing the hyperlink.

4. The archived data made viewable and browseable by the Wayback Machine is obtained by use of web archiving software that automatically stores copies of files available via the Internet, each file preserved as it existed at a particular point in time.

5. The Internet Archive assigns a URL on its site to the archived files in the format `http://web.archive.org/web/[Year in yyyy][Month in mm][Day in dd][Time code in hh:mm:ss]/[Archived URL]` aka an "extended URL". Thus, the extended URL `http://web.archive.org/web/19970126045828/http://www.archive.org/` would be the URL for the record of the Internet Archive home page HTML file (`(http://www.archive.org/)`) archived on January 26, 1997 at 4:58 a.m. and 28 seconds (1997/01/26 at 04:58:28). A web browser may be set such that a printout from it will display the URL of a web page in the printout's footer. The date indicated by an extended URL applies to a preserved instance of a file for a given URL, but not necessarily to any other files linked therein. Thus, in the case of a page constituted by a primary HTML file and other separate files (e.g., files with images, audio, multimedia, design elements, or other embedded content) linked within that primary HTML file, the primary HTML file and the other files will each have their own respective extended URLs and may not have been archived on the same dates.

6. Attached hereto as Exhibit A are true and accurate copies of browser printouts of the Internet Archive's records of the archived files for the URLs and the dates specified in the footer of the printout or an attached coversheet in the case of records for which a browser does not provide a ready option to print a URL in the footer, e.g., in the case of a PDF file.

7. I declare under penalty of perjury that the foregoing is true and correct.

Internet Archive: "Legal: Affidavit"



# IA affidavit

...when a visitor clicks on a hyperlink on an archived page that points to another URL, the visitor will be served the archived file found for the hyperlink's URL with the closest available date to the initial file containing the hyperlink.



# IA affidavit

...when a visitor clicks on a hyperlink on an archived page that points to another URL, the visitor will be served the archived file found for the hyperlink's URL with the closest available date to the initial file containing the hyperlink.

The archived data made viewable and browseable by the Wayback Machine is obtained by use of web archiving software that automatically stores copies of files available via the Internet, each file preserved as it existed at a particular point in time.



# IA affidavit

...when a visitor clicks on a hyperlink on an archived page that points to another URL, the visitor will be served the archived file found for the hyperlink's URL with the closest available date to the initial file containing the hyperlink.

The archived data made viewable and browseable by the Wayback Machine is obtained by use of web archiving software that automatically stores copies of files available via the Internet, each file preserved as it existed at a particular point in time.

... in the case of a page constituted by a primary HTML file and other separate files (e.g., files with images, audio, multimedia, design elements, or other embedded content) linked within that primary HTML file, the primary HTML file and the other files will each have their own respective extended URLs and may not have been archived on the same dates. ...



# IA affidavit

...when a visitor clicks on a hyperlink on an archived page that points to another URL, the visitor will be served the archived file found for the hyperlink's URL with the closest available date to the initial file containing the hyperlink.

The archived data made viewable and browseable by the Wayback Machine is obtained by use of web archiving software that automatically stores copies of files available via the Internet, each file preserved as it existed at a particular point in time.

... in the case of a page constituted by a primary HTML file and other separate files (e.g., files with images, audio, multimedia, design elements, or other embedded content) linked within that primary HTML file, the primary HTML file and the other files will each have their own respective extended URLs and may not have been archived on the same dates. ...

Attached hereto as Exhibit A are true and accurate copies of browser printouts of the Internet Archive's records of the archived files for the URLs and the dates specified in the footer...



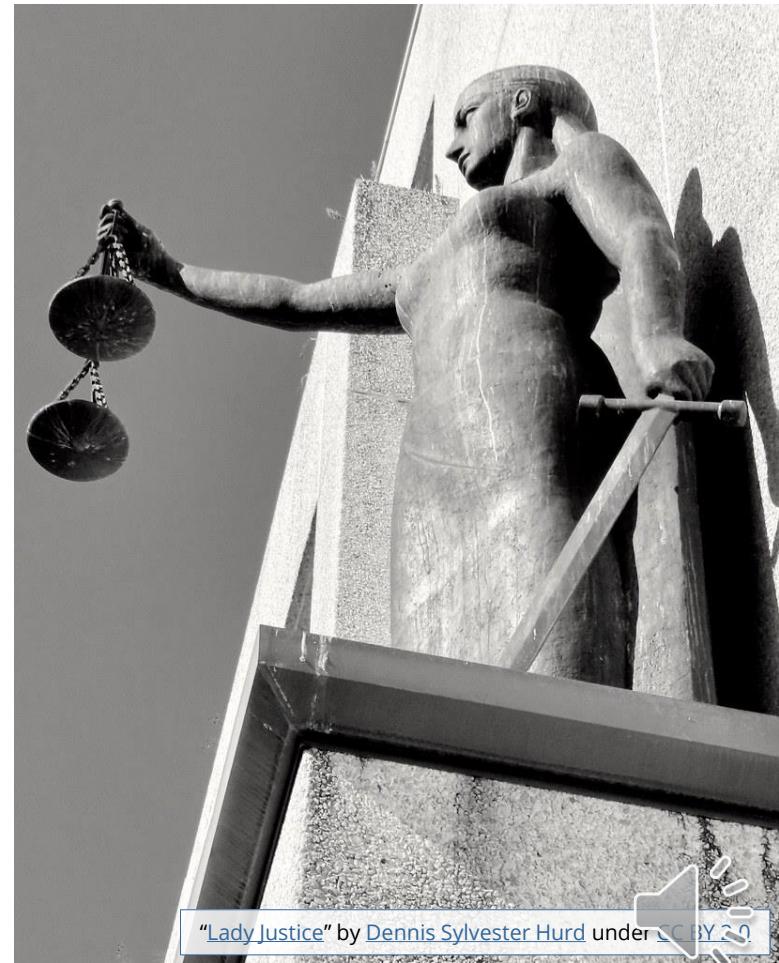
# Extant Frameworks



["Gaps between teeth - Explored"](#) by [Clint Budd](#) under [CC BY 2.0](#)

# judicial precedents

- no explicit guidance from judiciary
- treatment of IAWM evidence is court- or judge-specific
- judicial presumptions of the authenticity of IAWM evidence vary widely
- universal deference to IA affidavit



"Lady Justice" by Dennis Sylvester Hurd under CC BY 2.0

# e-discovery community guidance

- reflected in Sedona Conference Working Group publications
- some correct:
  - reliability of IA affidavit
- some okay:
  - courts only occasionally take judicial notice of IAWM evidence
  - WARC standard addresses authentication issues

The Sedona Conference Journal

Volume 22

Forthcoming 2021

The Sedona Conference Commentary  
on ESI Evidence & Admissibility,  
Second Edition

The Sedona Conference



October 2020

Final Post-Public-Comment Version

Recommended Citation:

The Sedona Conference, *Commentary on ESI Evidence & Admissibility, Second Edition*, 22 SEDONA CONF. J. 83 (forthcoming 2021).

Copyright 2020, The Sedona Conference

For this and additional publications see: <https://thesedonaconference.org/publications>  
**Sedona Conference: "The Sedona Conference Commentary  
on ESI Evidence & Admissibility, Second Edition"**



# e-discovery community guidance

- reflected in Sedona Conference Working Group publications
- some questionable or wrong:
  - conflation of capture + replay technologies
  - possible to delete or change archived information
  - IAWM “not accepted as a forensic evidence collection method”
  - IAWM provides for “an identical replica of the website”

The Sedona Conference Journal

Volume 22

Forthcoming 2021

The Sedona Conference Commentary  
on ESI Evidence & Admissibility,  
Second Edition

The Sedona Conference

October 2020

Final Post-Public-Comment Version



Recommended Citation:

The Sedona Conference, *Commentary on ESI Evidence & Admissibility, Second Edition*, 22 SEDONA CONF. J. 83 (forthcoming 2021).

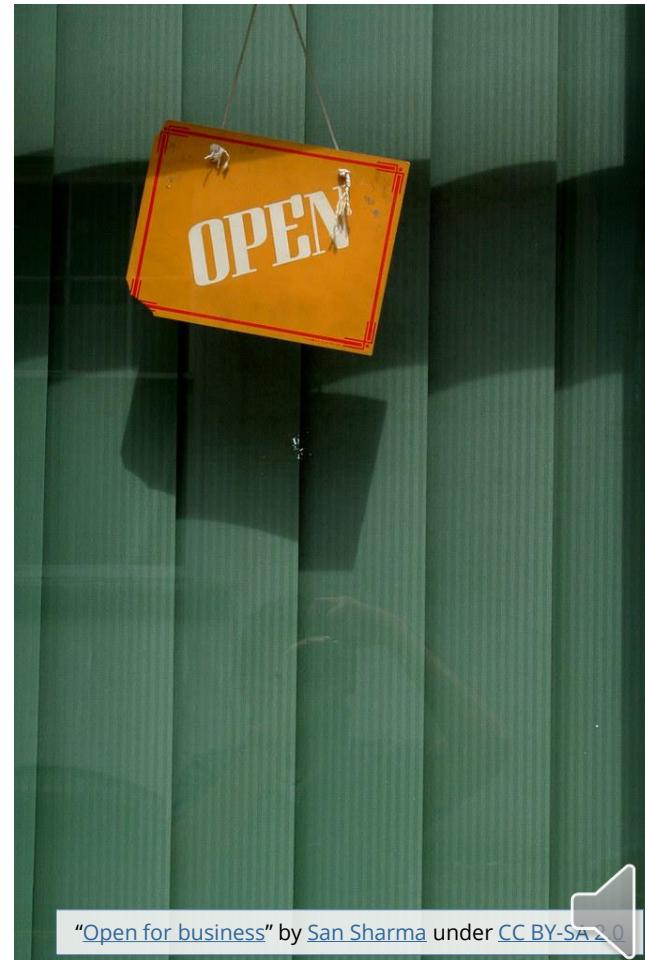
Copyright 2020, The Sedona Conference

For this and additional publications see: <https://thesedonaconference.org/publications>  
**Sedona Conference: "The Sedona Conference Commentary  
on ESI Evidence & Admissibility, Second Edition"**



# commercial archiving companies service marketing

- timestamp + capture metadata
- real-time collection
- completeness + fidelity via browser-based capture
- third-party-platform-independent storage
- accompanied by affidavits or expert testimony



["Open for business"](#) by [San Sharma](#) under [CC BY-SA 2.0](#)



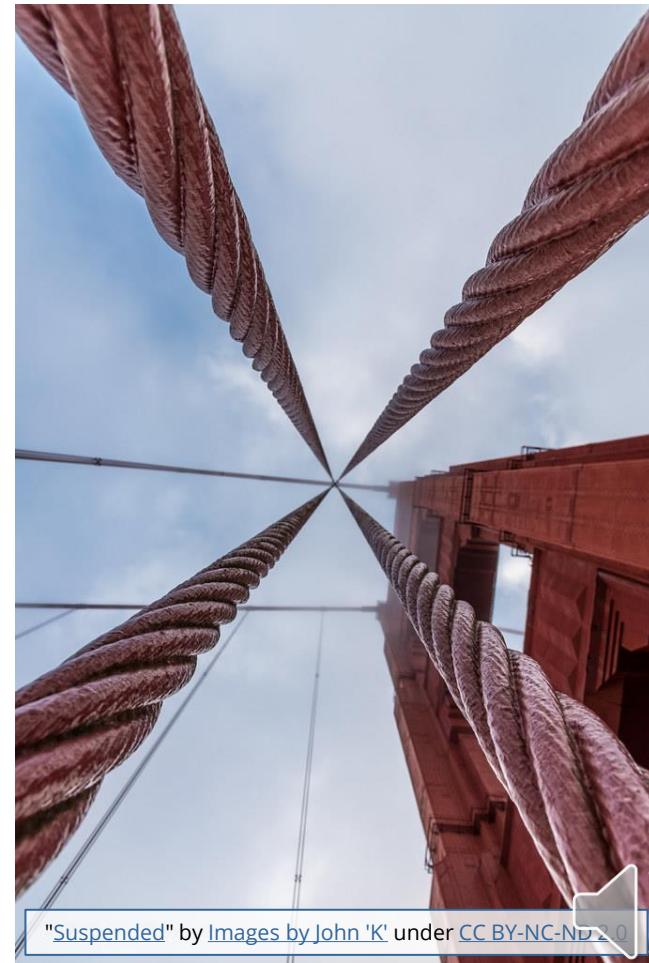


# Trustworthy Web Archives



# what makes web archives trustworthy?

- mostly in place:
  - capture atomic resources
  - capture http conversation
  - timestamped captures
  - archiving / attesting org
  - auditable / OSS tech stack
  - community standards
- could be improved:
  - measures supporting temporal coherence
  - measures supporting robustness / reproducibility
  - signing / chain of custody
  - corroboration w/ independent sources



"Suspended" by [Images by John 'K'](#) under [CC BY-NC-ND 2.0](#)

# capture atomic resources

- unlike screenshots, web archiving granularly captures a webpage's constituent resources
- enables examination of files' distinct provenance
- affords higher fidelity reproduction of original presentation



"T2i Testing" by Chris Potako under CC BY 2.0



# capture http conversation

- capture agents converse w/ web resource hosts in http
- not just the artifact is captured
- archiving conditions + context also recorded
- this is key provenance information



# timestamped captures

- archived web resources are situated in both place + time
- live web presents as perpetual present
- temporal dimensions can sometimes be (less authoritatively) discerned

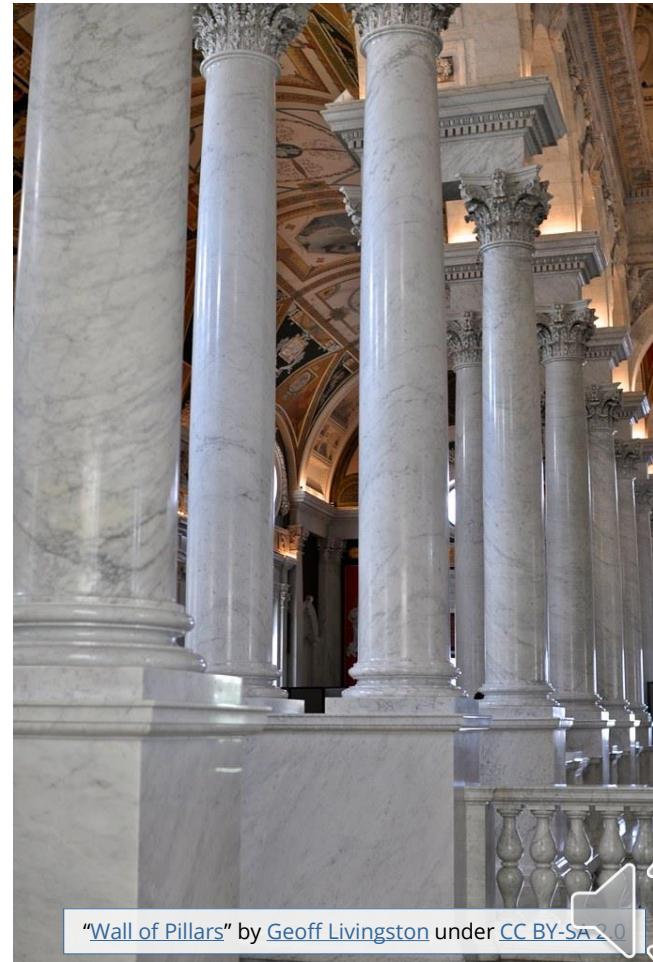


"South Side Sidewalk Co 9-25-51" by Thirteen of Clubs under [CC BY-SA 2.0](#)



# archiving / attesting org

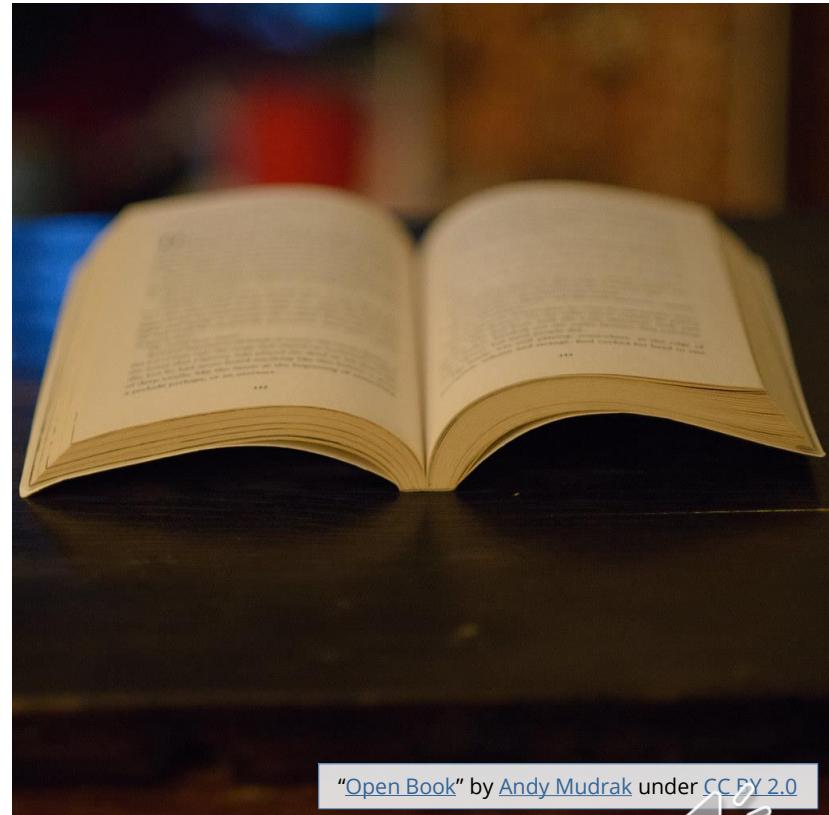
- emergently overweighted in U.S. federal courts' treatment of web archive evidence
  - particularly with respect to the IA affidavit
- not that it shouldn't matter
- should weigh in favor of authentication



["Wall of Pillars"](#) by Geoff Livingston under CC BY-SA 2.0

# auditable / OSS tech stack

- proprietary tools / methods can't be audited
- OSS tools could be
- gaps b/t live + archived versions may be more understandable w/ OSS tool stack



["Open Book"](#) by [Andy Mudrak](#) under [CC BY 2.0](#)



# community standards

- reflect community consensus as to needs for long-term digital preservation
- WARC, CDX, Memento, WACZ, etc.



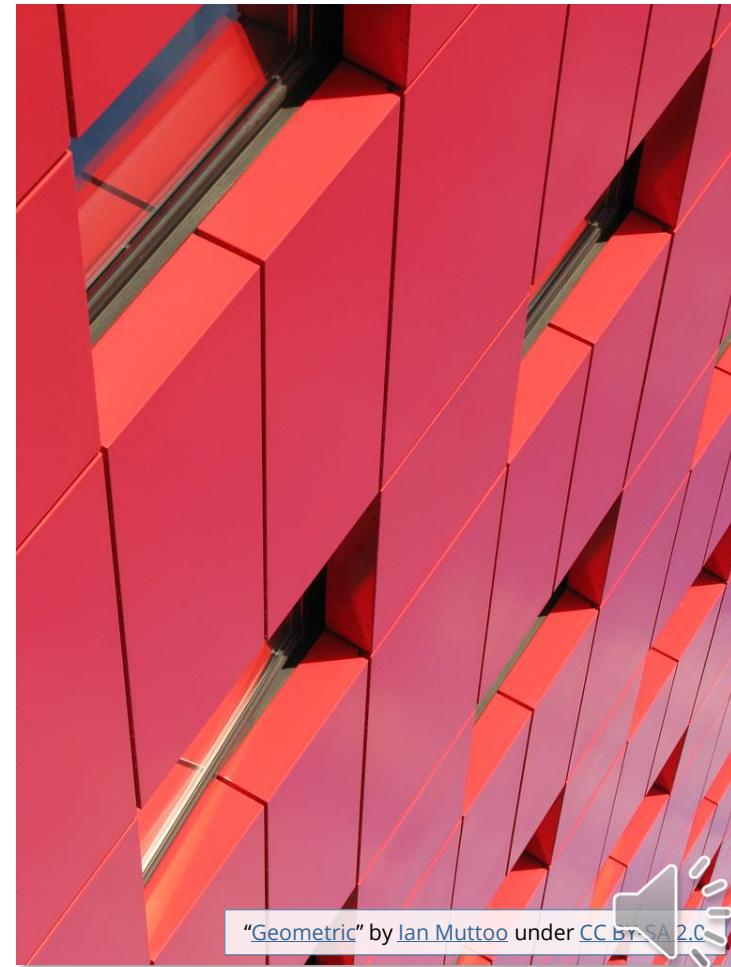
# measures supporting temporal coherence

- IAWM [About this capture feature](#) + [Time Travel](#) help visualize temporal topography of composite mementos
- however, no production systems help visualize temporally violative elements per se or allow them to be toggled
- browser-based crawling can mitigate, but can it scale?



# measures supporting robustness / reproducibility

- sub-optimal that composite memento replay yields variable hashes
- can we robustify our archive access services?
- can we characterize archive-related variations?
- can we rationalize the resources bundle that replay engines deliver?



"Geometric" by Ian Muttoo under CC BY-SA 2.0

# signing / chain of custody

- a technical complement to legal attestation
- a best practice for more conventional digital forensics



"Certified" by Damian Gadal under [CC BY 2.0](#)

# corroboration w/ independent sources

- Lots of Copies Keep Stuff Safe (LOCKSS)
- not just redundancy, but also independence
  - independent web archives
  - independently hosted copies
  - independently captured copies
  - heterogeneous technical stacks
- corroboration vice sameness





# Where Does that Leave Us?



"electric luminaria" by [grendelkhan](#) under [CC BY-SA 2.0](#)



**THANK YOU**

**EXIT**



"Thank You" by Brian Crawford under CC BY-SA 2.0