Right HTML, Wrong JSON: Challenges in Replaying Archived Webpages Built with Client-Side Rendering

Michele C. Weigle, Michael L. Nelson

Web Science and Digital Libraries (WS-DL)
Department of Computer Science
Old Dominion University

Sawood Alam, Mark Graham

Wayback Machine
Internet Archive

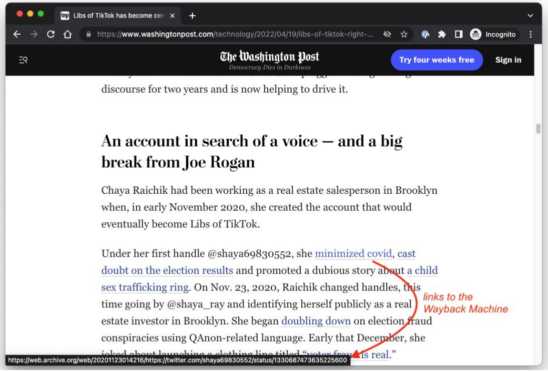
mweigle@odu.edu (@weiglemc, @WebSciDL) | June 27, 2023 | JCDL 2023 | Santa Fe, NM

Web archives have been used for journalistic and legal evidence

2004
“… conference addressing, understanding and preventing homosexuality.”

Last Week Tonight, Mar 18, 2018

Taylor Lorenz, "Meet the woman behind Libs of TikTok", *Washington Post*, April 2022, https://www.washingtonpost.com/technology/2022/04/19/libs-of-tiktok-right-wing-media/


Libs of TikTok has become cen
https://www.washingtonpost.com/technology/2022/04/19/libs-of-tiktok-right-...
Incognito
The Washington Post
Democracy Dies in Darkness
Try four weeks free
Sign in

discourse for two years and is now helping to drive it.

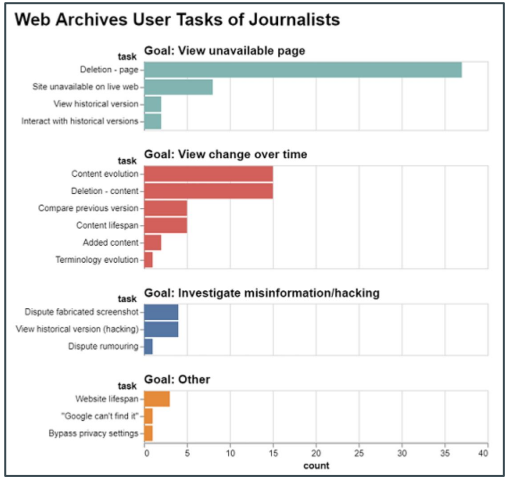
An account in search of a voice — and a big break from Joe Rogan

Chaya Raichik had been working as a real estate salesperson in Brooklyn when, in early November 2020, she created the account that would eventually become Libs of TikTok.

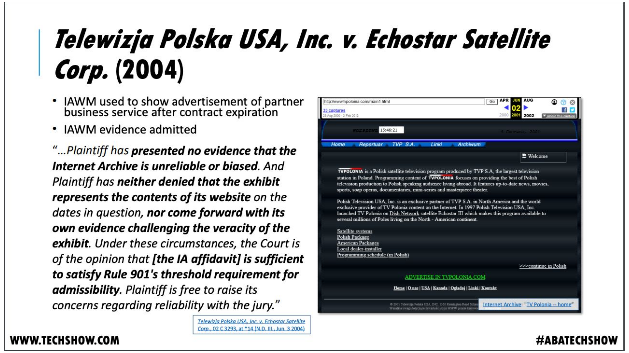
Under her first handle @shaya69830552, she minimized covid, cast doubt on the election results and promoted a dubious story about a child sex trafficking ring. On Nov. 23, 2020, Raichik changed handles, this time going by @shaya_ray and identifying herself publicly as a real estate investor in Brooklyn. She began doubling down on election fraud conspiracies using QAnon-related language. Early that December, she joked about launching a clothing line titled “voter fraud is real.”

https://web.archive.org/web/20201123014216/https://twitter.com/shaya69830552/status/1330687473635225600

links to the Wayback Machine



Lesley Frew, Web Archiving in Popular Media II: User Tasks of Journalists, https://ws-dl.blogspot.com/2022/08/2022-08-04-web-archiving-in-popular.html



Telewizja Polska USA, Inc. v. Echostar Satellite Corp. (2004)

- IAWM used to show advertisement of partner business service after contract expiration
- IAWM evidence admitted

“…*Plaintiff has* ***presented no evidence that the Internet Archive is unreliable or biased****. And Plaintiff has* ***neither denied that the exhibit represents the contents of its website*** *on the dates in question,* ***nor come forward with its own evidence challenging the veracity of the exhibit****. Under these circumstances, the Court is of the opinion that* ***[the IA affidavit] is sufficient to satisfy Rule 901's threshold requirement for admissibility****. Plaintiff is free to raise its concerns regarding reliability with the jury.*”

[Telewizja Polska USA, Inc. v. Echostar Satellite Corp., 02 C 3293, at *14 (N.D. Ill., Jun. 3 2004)]

TVPOLONIA is a Polish satellite television program produced by TVP S.A. the largest television station in Poland. Programming content of TVPOLONIA focuses on providing the best of Polish television production to Polish speaking audience living abroad. It features up-to-date news, movies, sports, soap operas, documentaries, mini-series and masterpiece theater.

Polish Television USA, Inc. is an exclusive partner of TVP S.A. in North America and the world exclusive provider of TV Polonia content on the Internet. In 1997 Polish Television USA, Inc. launched TV Polonia on Dish Network satellite Echostar III which makes this program available to several millions of Poles living on the North - American continent.

Satellite systems
Polish Package
American Packages
Local dealer-installer
Programming schedule (in Polish)

>>>continue in Polish

ADVERTISE IN TVPOLONIA.COM

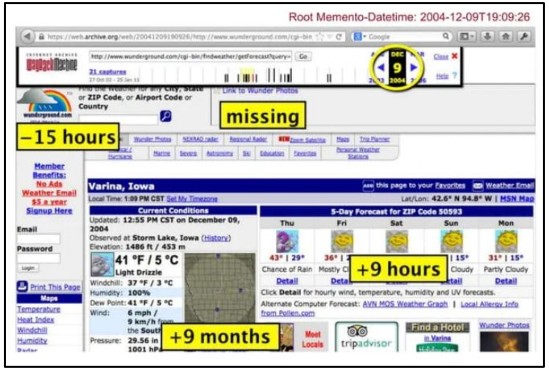
Home | O nas | USA | Kanada | Ogladaj | Linki | Kontakt

Internet Archive: "TV Polonia -- home"

WWW.TECHSHOW.COM #ABATECHSHOW

Nicholas Taylor and Joe Dugan, "Evidentiary Use of the Temporal Web", Mar 2022, https://nullhandle.org/pdf/2023-03-02_evidentiary_use_of_the_temporal_web.pdf#page=74
Nicholas Taylor, "A Brief Primer on Using Web-Archived Evidence", June 2023, https://nullhandle.org/blog/2023-06-07-a-brief-primer-on-using-web-archived-evidence.html

However, webpages replayed from web archives are not always exactly what was observed



Root Memento-Datetime: 2004-12-09T19:09:26

missing

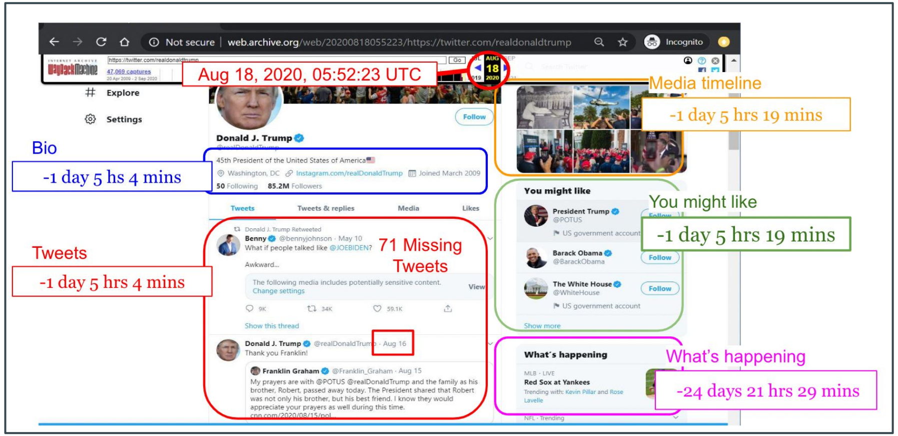
−15 hours

+9 hours

+9 months

S. Ainsworth, M.L. Nelson, H. Van de Sompel, "Only One Out of Five Archived Web Pages Existed as Presented", ACM Hypertext 2015, https://ws-dl.blogspot.com/2015/12/2015-12-08-evaluating-temporal.html

K. Garg, H.R. Jayanetti, S. Alam, M.C. Weigle, and M.L. Nelson, “Replaying Archived Twitter: When your bird is broken, will it bring you down?,” JCDL 2021, https://arxiv.org/abs/2108.12092



Aug 18, 2020, 05:52:23 UTC

Bio

-1 day 5 hs 4 mins

Tweets

-1 day 5 hrs 4 mins

71 Missing Tweets

Media timeline

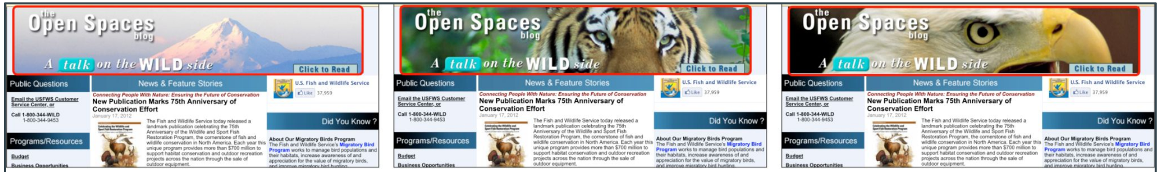
-1 day 5 hrs 19 mins

You might like

-1 day 5 hrs 19 mins

What’s happening

-24 days 21 hrs 29 mins



On 2018-02-28

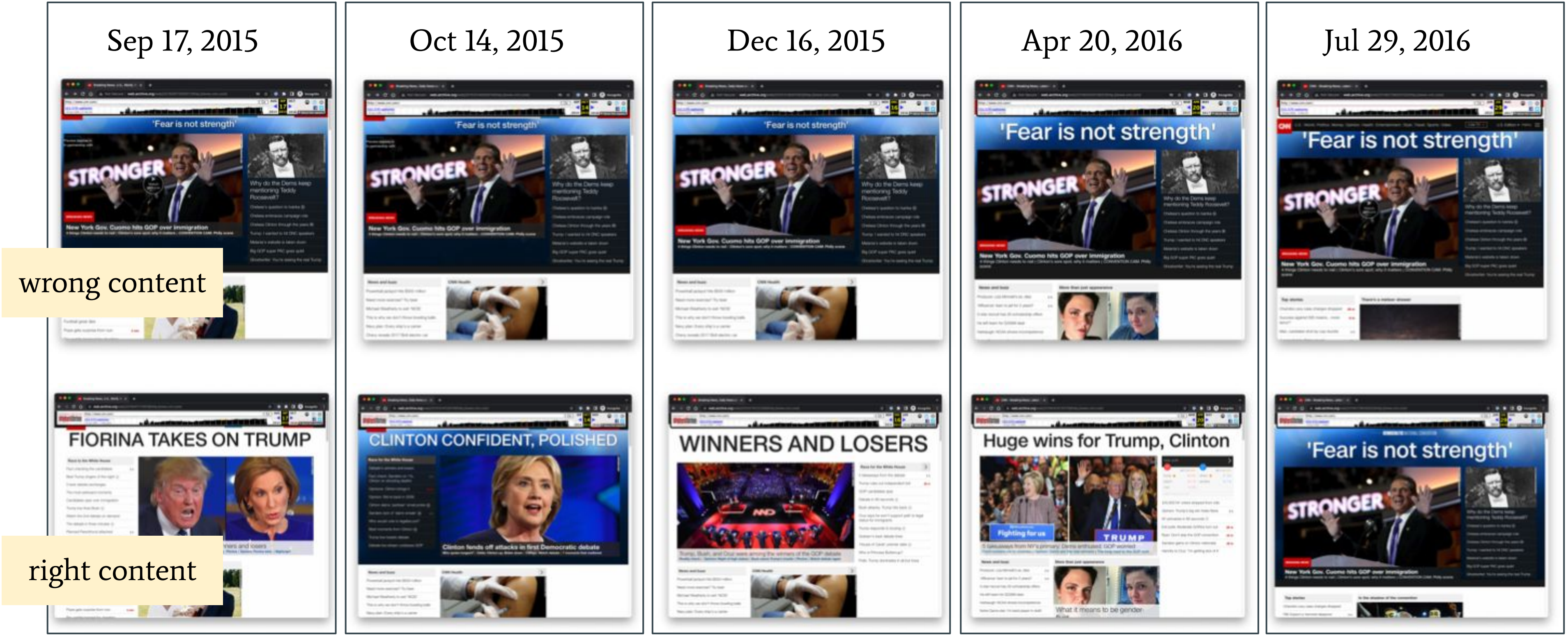
On 2018-07-08

On 2018-08-25

same memento (capture), replayed on different dates

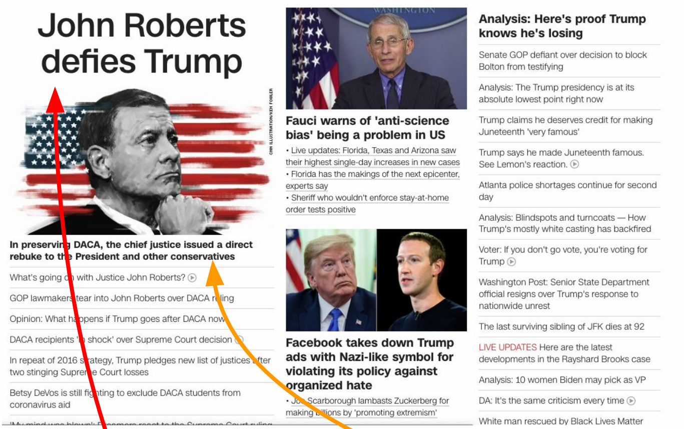
M. Aturban, M. Klein, H. Van de Sompel, S. Alam, M.L. Nelson, and M.C. Weigle, “Hashes are not suitable to verify fixity of the public archived web”, *PLOS ONE*, June 2023, https://doi.org/10.1371/journal.pone.0286879

We found temporal violations in some CNN.com mementos



CNN.com's front page uses client-side rendering, loading many JSON zone-manager.izl files that contain HTML code

Snippet from [https://wayback.archive-it.org/4887/20200618234850if_/https://www.cnn.com/data/ocs/section/index.html:**homepage1-zone-1**/views/zones/common/zone-manager.izl](https://wayback.archive-it.org/4887/20200618234850if_/https://www.cnn.com/data/ocs/section/index.html:homepage1-zone-1/views/zones/common/zone-manager.izl), which builds the "Hero" headlines.



John Roberts defies Trump

In preserving DACA, the chief justice issued a direct rebuke to the President and other conservatives

What's going on with Justice John Roberts?

GOP lawmakers tear into John Roberts over DACA ruling

Opinion: What happens if Trump goes after DACA now

DACA recipients 'in shock' over Supreme Court decision

In repeat of 2016 strategy, Trump pledges new list of justices after two stinging Supreme Court losses

Betsy DeVos is still fighting to exclude DACA students from coronavirus aid

Fauci warns of 'anti-science bias' being a problem in US

- Live updates: Florida, Texas and Arizona saw their highest single-day increases in new cases
- Florida has the makings of the next epicenter, experts say
- Sheriff who wouldn't enforce stay-at-home order tests positive

Facebook takes down Trump ads with Nazi-like symbol for violating its policy against organized hate

- Joe Scarborough lambasts Zuckerberg for making billions by 'promoting extremism'

Analysis: Here's proof Trump knows he's losing

Senate GOP defiant over decision to block Bolton from testifying

Analysis: The Trump presidency is at its absolute lowest point right now

Trump claims he deserves credit for making Juneteenth 'very famous'

Trump says he made Juneteenth famous. See Lemon's reaction.

Atlanta police shortages continue for second day

Analysis: Blindspots and turncoats — How Trump's mostly white casting has backfired

Voter: If you don't go vote, you're voting for Trump

Washington Post: Senior State Department official resigns over Trump's response to nationwide unrest

The last surviving sibling of JFK dies at 92

LIVE UPDATES Here are the latest developments in the Rayshard Brooks case

Analysis: 10 women Biden may pick as VP

DA: It's the same criticism every time

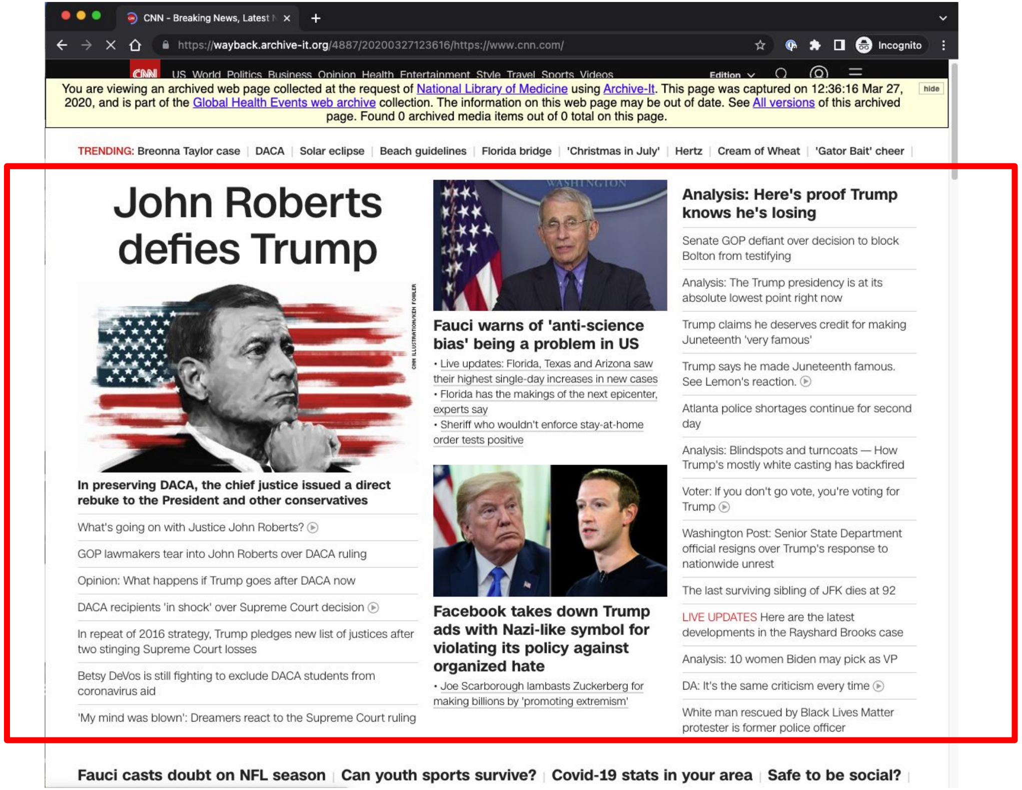
White man rescued by Black Lives Matter

```
{"izlData":{"cards":[]},"html":"<section class=\"zn zn-homepage1-zone-1 zn-left-flui
zn-left-fluid-share zn-has-multiple-containers zn-has-3-containers\" data-eq-pts=\"x
1100\" id=\"homepage1-zone-1\" data-vr-zone=\"zone-0-0\" data-zone-label=\"Hero\" da
class=\"l-container zn-top__label\"></div><div class=\"l-container zn-top__banner\">
class=\"zn-top__background\"></div></div><div class=\"l-container\"><div class=\"zn_
zn__column--idx-0\"><ul class=\"cn cn-list-hierarchical-xs cn--idx-0 cn-container_17
data-layout=\"list-hierarchical-xs\" data-vr-zone=\"home-top-col1\"><li><article cla
cd--vertical cd--has-siblings cd--has-media cd--media__image cd--has-banner\"
data-vr-contentbox=\"/2020/06/18/politics/john-roberts-supreme-court-daca-legacy/index.html\" data-eq-pts=\"xsmall: 0, small: 300,
medium: 460, large: 780, full16x9: 1100\" data-section-name=\"politics\"><a
href=\"/2020/06/18/politics/john-roberts-supreme-court-daca-legacy/index.html\" class=\"link-banner\" ><h2 class=\"banner-text
screaming-banner-text banner-text-size--char-26\" data-analytics=\"_list-hierarchical-xs_article_\">John Roberts defies
Trump</h2></a><div class=\"cd__wrapper\" data-analytics=\"_list-hierarchical-xs_article_\"><div class=\"media\"><a
href=\"/2020/06/18/politics/john-roberts-supreme-court-daca-legacy/index.html\" >
...
<span class=\"cd__headline-text\"><strong>In preserving DACA, the chief justice issued a direct rebuke to the President and other
conservatives</strong></span><span
```

Each zone-manager.izl file builds a separate section of the CNN.com front page

homepage1-zone-1

(Hero)


CNN - Breaking News, Latest N
https://wayback.archive-it.org/4887/20200327123616/https://www.cnn.com/
Incognito
CNN
US World Politics Business Opinion Health Entertainment Style Travel Sports Videos
Edition
hide

You are viewing an archived web page collected at the request of National Library of Medicine using Archive-It. This page was captured on 12:36:16 Mar 27, 2020, and is part of the Global Health Events web archive collection. The information on this web page may be out of date. See All versions of this archived page. Found 0 archived media items out of 0 total on this page.

TRENDING: Breonna Taylor case | DACA | Solar eclipse | Beach guidelines | Florida bridge | 'Christmas in July' | Hertz | Cream of Wheat | 'Gator Bait' cheer |

John Roberts defies Trump

In preserving DACA, the chief justice issued a direct rebuke to the President and other conservatives

What's going on with Justice John Roberts?

GOP lawmakers tear into John Roberts over DACA ruling

Opinion: What happens if Trump goes after DACA now

DACA recipients 'in shock' over Supreme Court decision

In repeat of 2016 strategy, Trump pledges new list of justices after two stinging Supreme Court losses

Betsy DeVos is still fighting to exclude DACA students from coronavirus aid

'My mind was blown': Dreamers react to the Supreme Court ruling

Fauci warns of 'anti-science bias' being a problem in US

- Live updates: Florida, Texas and Arizona saw their highest single-day increases in new cases
- Florida has the makings of the next epicenter, experts say
- Sheriff who wouldn't enforce stay-at-home order tests positive

Facebook takes down Trump ads with Nazi-like symbol for violating its policy against organized hate

- Joe Scarborough lambasts Zuckerberg for making billions by 'promoting extremism'

Analysis: Here's proof Trump knows he's losing

Senate GOP defiant over decision to block Bolton from testifying

Analysis: The Trump presidency is at its absolute lowest point right now

Trump claims he deserves credit for making Juneteenth 'very famous'

Trump says he made Juneteenth famous. See Lemon's reaction.

Atlanta police shortages continue for second day

Analysis: Blindspots and turncoats — How Trump's mostly white casting has backfired

Voter: If you don't go vote, you're voting for Trump

Washington Post: Senior State Department official resigns over Trump's response to nationwide unrest

The last surviving sibling of JFK dies at 92

LIVE UPDATES Here are the latest developments in the Rayshard Brooks case

Analysis: 10 women Biden may pick as VP

DA: It's the same criticism every time

White man rescued by Black Lives Matter protester is former police officer

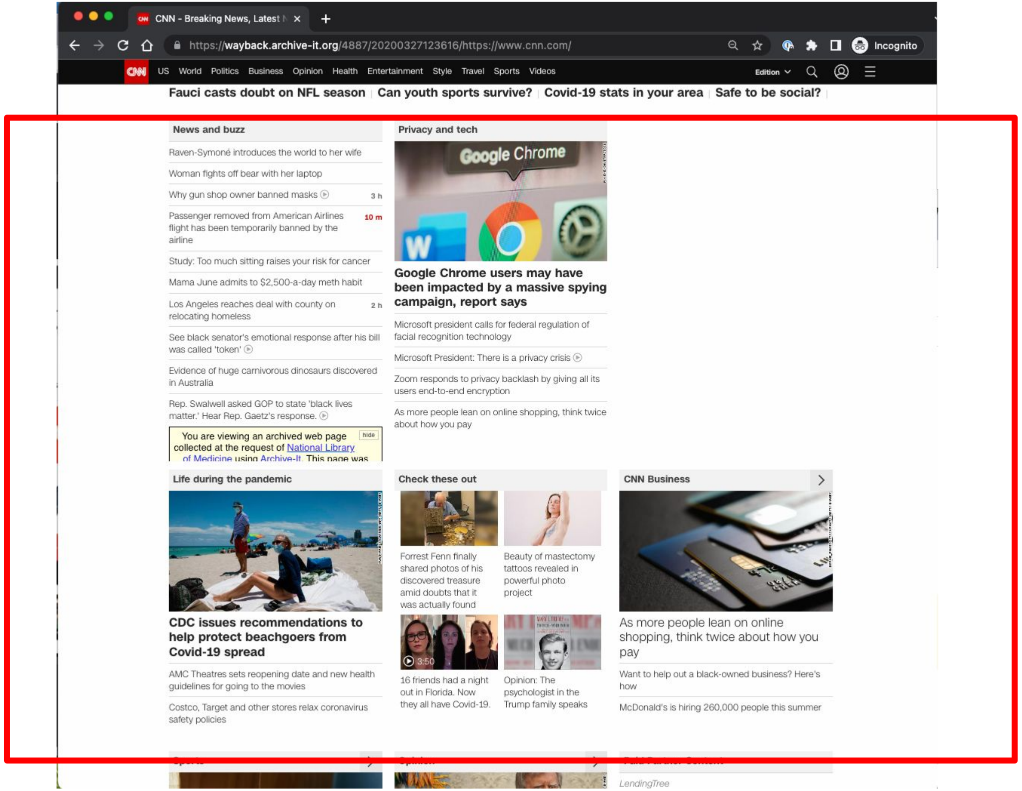
Fauci casts doubt on NFL season | Can youth sports survive? | Covid-19 stats in your area | Safe to be social? |

https://www.cnn.com/data/ocs/section/index.html:**homepage1-zone-1**/views/zones/common/zone-manager.izl

Each zone-manager.izl file builds a separate section of the CNN.com front page

homepage2-zone-1

(News and buzz, Privacy and tech, Life during the pandemic, CNN Business)


CNN - Breaking News, Latest N
https://wayback.archive-it.org/4887/20200327123616/https://www.cnn.com/
Incognito
US World Politics Business Opinion Health Entertainment Style Travel Sports Videos
Edition

Fauci casts doubt on NFL season | Can youth sports survive? | Covid-19 stats in your area | Safe to be social?

News and buzz

- Raven-Symoné introduces the world to her wife
- Woman fights off bear with her laptop
- Why gun shop owner banned masks 3 h
- Passenger removed from American Airlines flight has been temporarily banned by the airline 10 m
- Study: Too much sitting raises your risk for cancer
- Mama June admits to $2,500-a-day meth habit
- Los Angeles reaches deal with county on relocating homeless 2 h
- See black senator's emotional response after his bill was called 'token'
- Evidence of huge carnivorous dinosaurs discovered in Australia
- Rep. Swalwell asked GOP to state 'black lives matter.' Hear Rep. Gaetz's response.

You are viewing an archived web page collected at the request of National Library of Medicine using Archive-It. This page was hide

Privacy and tech

Google Chrome users may have been impacted by a massive spying campaign, report says

- Microsoft president calls for federal regulation of facial recognition technology
- Microsoft President: There is a privacy crisis
- Zoom responds to privacy backlash by giving all its users end-to-end encryption
- As more people lean on online shopping, think twice about how you pay

Life during the pandemic

CDC issues recommendations to help protect beachgoers from Covid-19 spread

- AMC Theatres sets reopening date and new health guidelines for going to the movies
- Costco, Target and other stores relax coronavirus safety policies

Check these out

- Forrest Fenn finally shared photos of his discovered treasure amid doubts that it was actually found
- Beauty of mastectomy tattoos revealed in powerful photo project
- 3:50 16 friends had a night out in Florida. Now they all have Covid-19.
- Opinion: The psychologist in the Trump family speaks

CNN Business

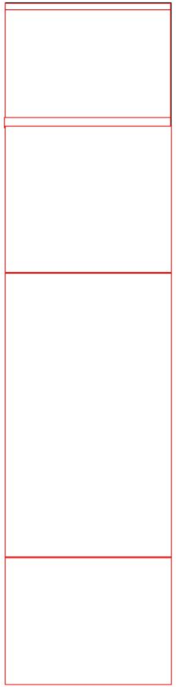
As more people lean on online shopping, think twice about how you pay

- Want to help out a black-owned business? Here's how
- McDonald's is hiring 260,000 people this summer

LendingTree

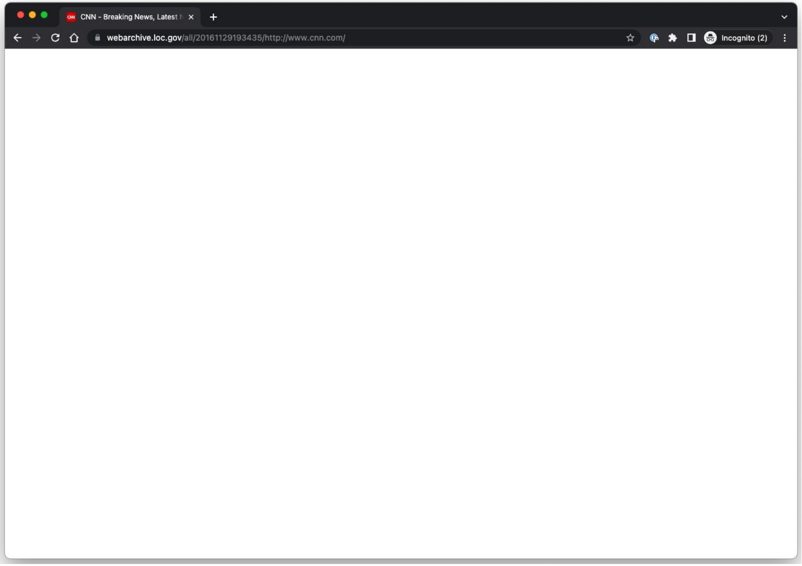
https://www.cnn.com/data/ocs/section/index.html:**homepage2-zone-1**/views/zones/common/zone-manager.izl

Without executing JavaScript, conventional web archive crawlers can't archive the content (just the placeholder)



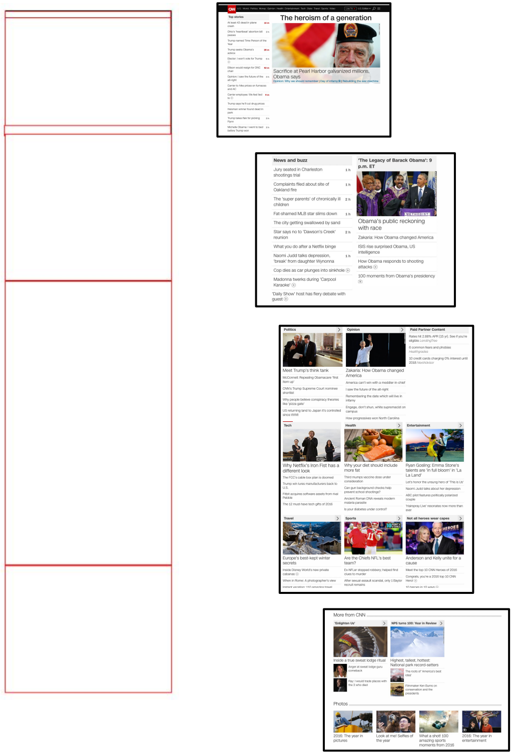
Justin F. Brunelle, Michele C. Weigle, and Michael L. Nelson, “Archival Crawlers and JavaScript: Discover More Stuff but Crawl More Slowly,” JCDL 2017.

Archived CNN.com, Library of Congress (uses conventional crawler), Nov 29, 2016

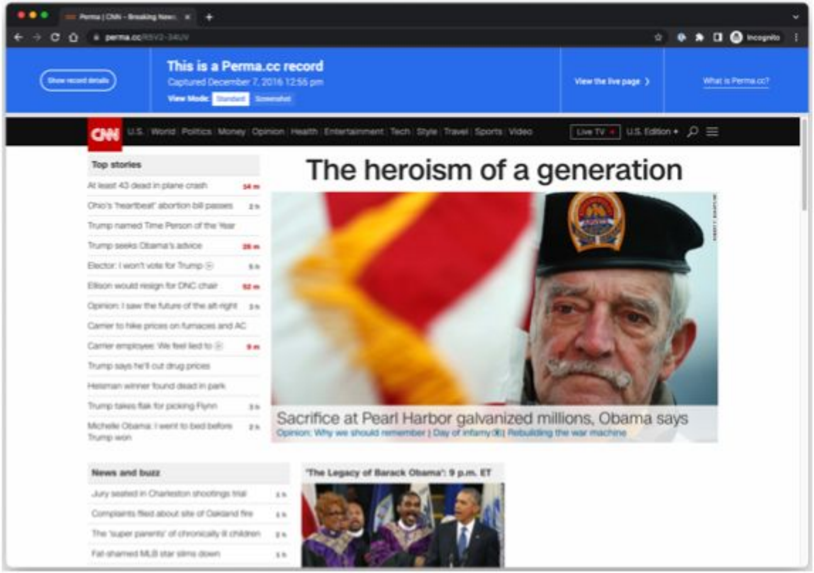

CNN - Breaking News, Latest N
webarchive.loc.gov/all/20161129193435/http://www.cnn.com/
Incognito (2)

https://webarchive.loc.gov/all/20161129193435/http://www.cnn.com/

Browser-based crawlers can capture both the placeholder and the content



Archived CNN.com, Perma.cc (uses browser-based crawler), Dec 7, 2016



Perma | CNN - Breaking News
perma.cc/R5V2-34UV
Incognito
Show record details
View the live page
What is Perma.cc?
U.S. World Politics Money Opinion Health Entertainment Tech Style Travel Sports Video
Live TV
U.S. Edition

This is a Perma.cc record

Captured December 7, 2016 12:56 pm

View Mode: Standard Screenshot

Top stories

At least 43 dead in plane crash 34 m

Ohio's 'heartbeat' abortion bill passes 2 h

Trump named Time Person of the Year

Trump seeks Obama's advice 25 m

Elector: I won't vote for Trump 5 h

Ellison would resign for DNC chair 52 m

Opinion: I saw the future of the alt-right 3 h

Carrier to hike prices on furnaces and AC

Carrier employee: We feel lied to 9 m

Trump says he'll cut drug prices

Heisman winner found dead in park

Trump takes flak for picking Flynn 3 h

Michelle Obama: I went to bed before Trump won 2 h

The heroism of a generation

Sacrifice at Pearl Harbor galvanized millions, Obama says

Opinion: Why we should remember | Day of infamy | Rebuilding the war machine

News and buzz

Jury seated in Charleston shootings trial 1 h

Complaints filed about site of Oakland fire 1 h

The 'super parents' of chronically ill children 2 h

Fat-shamed MLB star slims down 1 h

'The Legacy of Barack Obama': 9 p.m. ET

https://perma.cc/R5V2-34UV

Mixing the results of conventional and browser-based crawling can cause temporal violations

Archive-It

- 2020 - subscribers can use either Heritrix (conventional crawler) or Brozzler (browser-based crawler)
- subscribers can upload captures created with other tools (like Conifer, a browser-based crawler)

IA Wayback Machine

- Jun 2016 - incorporating captures from perma.cc
- Oct 2019 - Save Page Now service begins using Brozzler
- Large crawls still use Heritrix

https://blogs.harvard.edu/perma/2016/06/27/perma-v0-67-collaboration-internet-archive/
https://blog.archive.org/2019/10/23/the-wayback-machines-save-page-now-is-new-and-improved/

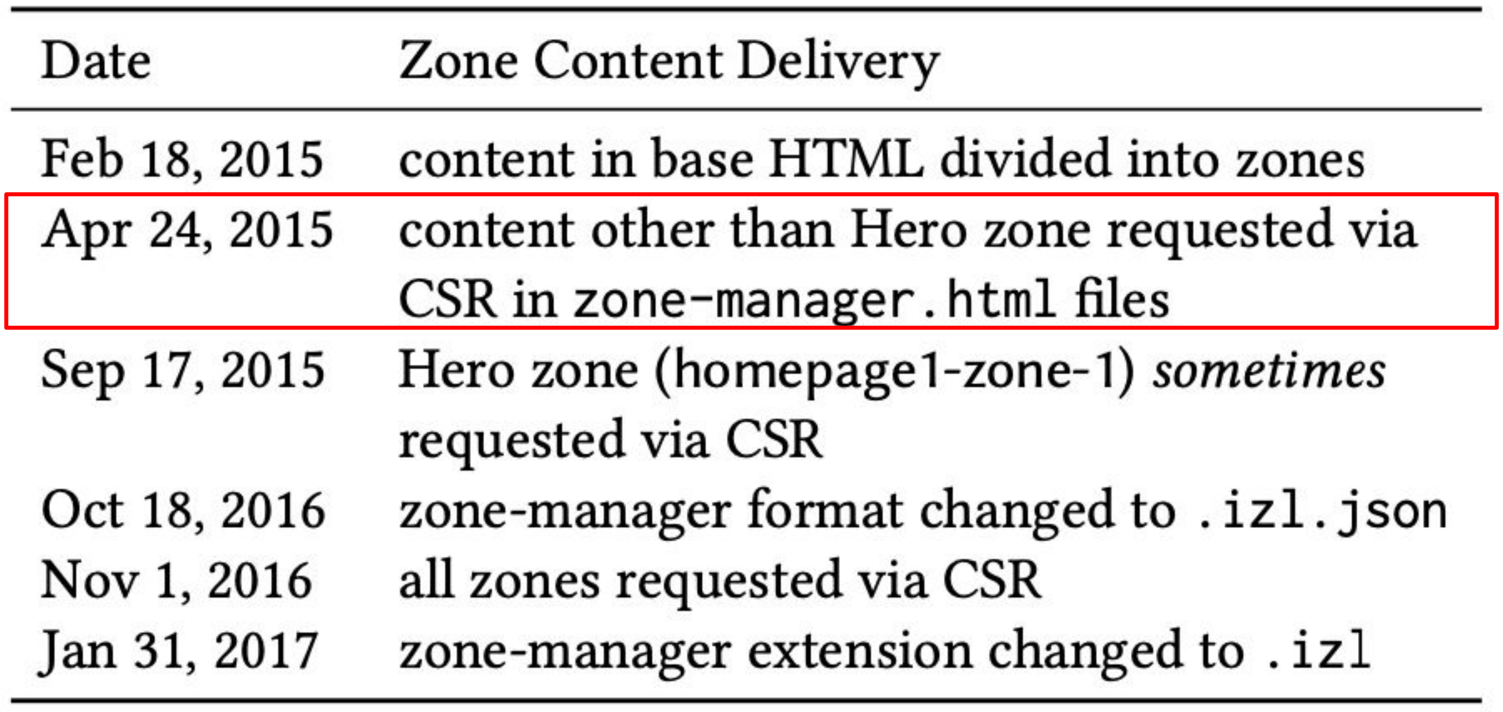
When did CNN start using client-side rendering (CSR)?

- Used Internet Archive's CDX API to identify 200,000 CNN.com mementos with HTTP 200 OK since 2010

- Used selenium-wire headless browser to replay over 1100 CNN.com mementos from Internet Archive's Wayback Machine (IAWM) between 2013-2020 and logged requests for zone-manager files

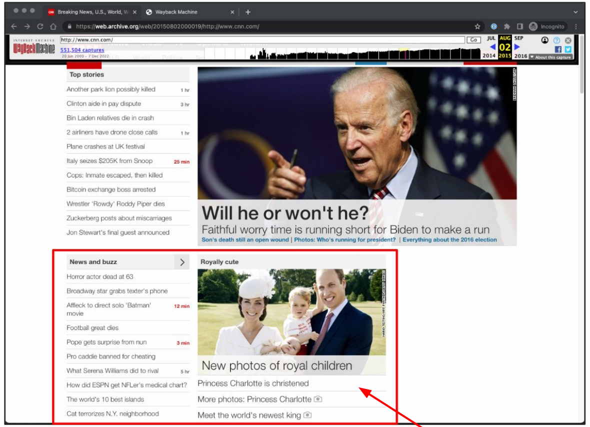
https://github.com/internetarchive/wayback/blob/master/wayback-cdx-server/README.md
https://pypi.org/project/selenium-wire/

Apr 2015: Everything but Hero zone requested via CSR



Date	Zone Content Delivery
Feb 18, 2015	content in base HTML divided into zones
Apr 24, 2015	content other than Hero zone requested via CSR in `zone-manager.html` files
Sep 17, 2015	Hero zone (homepage1-zone-1) *sometimes* requested via CSR
Oct 18, 2016	zone-manager format changed to `.izl.json`
Nov 1, 2016	all zones requested via CSR
Jan 31, 2017	zone-manager extension changed to `.izl`

Hero story in base HTML, but temporal violations in 2nd level content



Top stories

- Another park lion possibly killed 1 hr
- Clinton aide in pay dispute 3 hr
- Bin Laden relatives die in crash
- 2 airliners have drone close calls 1 hr
- Plane crashes at UK festival
- Italy seizes $205K from Snoop 25 min
- Cops: Inmate escaped, then killed
- Bitcoin exchange boss arrested
- Wrestler 'Rowdy' Roddy Piper dies
- Zuckerberg posts about miscarriages
- Jon Stewart's final guest announced

Will he or won't he?

Faithful worry time is running short for Biden to make a run

Son's death still an open wound | Photos: Who's running for president? | Everything about the 2016 election

News and buzz

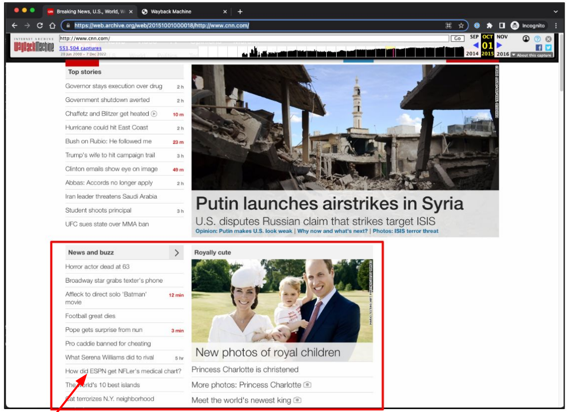
- Horror actor dead at 63
- Broadway star grabs texter's phone
- Affleck to direct solo 'Batman' movie 12 min
- Football great dies
- Pope gets surprise from nun 3 min
- Pro caddie banned for cheating
- What Serena Williams did to rival 5 hr
- How did ESPN get NFLer's medical chart?
- The world's 10 best islands
- Cat terrorizes N.Y. neighborhood

Royally cute

New photos of royal children

- Princess Charlotte is christened
- More photos: Princess Charlotte
- Meet the world's newest king

Aug 2, 2015



Top stories

- Governor stays execution over drug 2 h
- Government shutdown averted 2 h
- Chaffetz and Blitzer get heated 10 m
- Hurricane could hit East Coast 2 h
- Bush on Rubio: He followed me 23 m
- Trump's wife to hit campaign trail 3 h
- Clinton emails show eye on image 49 m
- Abbas: Accords no longer apply 2 h
- Iran leader threatens Saudi Arabia
- Student shoots principal 3 h
- UFC sues state over MMA ban

Putin launches airstrikes in Syria

U.S. disputes Russian claim that strikes target ISIS

Opinion: Putin makes U.S. look weak | Why now and what's next? | Photos: ISIS terror threat

News and buzz

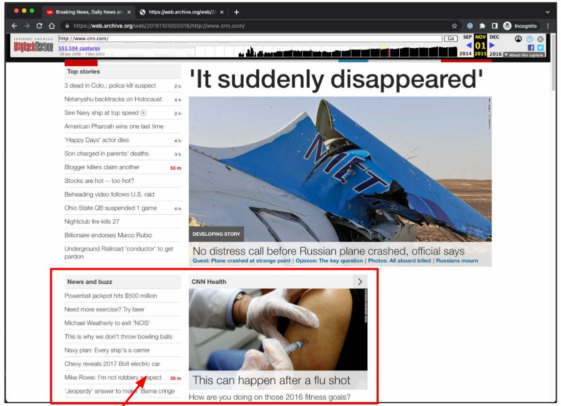
- Horror actor dead at 63
- Broadway star grabs texter's phone
- Affleck to direct solo 'Batman' movie 12 min
- Football great dies
- Pope gets surprise from nun 3 min
- Pro caddie banned for cheating
- What Serena Williams did to rival 5 hr
- How did ESPN get NFLer's medical chart?
- The world's 10 best islands
- Cat terrorizes N.Y. neighborhood

Royally cute

New photos of royal children

- Princess Charlotte is christened
- More photos: Princess Charlotte
- Meet the world's newest king

Oct 1, 2015



Top stories

- 3 dead in Colo.; police kill suspect 2 h
- Netanyahu backtracks on Holocaust 4 h
- See Navy ship at top speed 2 h
- American Pharoah wins one last time
- 'Happy Days' actor dies 4 h
- Son charged in parents' deaths 3 h
- Blogger killers claim another 53 m
- Stocks are hot -- too hot?
- Beheading video follows U.S. raid
- Ohio State QB suspended 1 game 4 h
- Nightclub fire kills 27
- Billionaire endorses Marco Rubio
- Underground Railroad 'conductor' to get pardon

'It suddenly disappeared'

No distress call before Russian plane crashed, official says

Quest: Plane crashed at strange point | Opinion: The key question | Photos: All aboard killed | Russians mourn

News and buzz

- Powerball jackpot hits $500 million
- Need more exercise? Try beer
- Michael Weatherly to exit 'NCIS'
- This is why we don't throw bowling balls
- Navy plan: Every ship's a carrier
- Chevy reveals 2017 Bolt electric car
- Mike Rowe: I'm not robbery suspect 39 m
- 'Jeopardy' answer to make 'Bama cringe

CNN Health

This can happen after a flu shot

How are you doing on those 2016 fitness goals?

Nov 1, 2015

from Jul 10, 2015

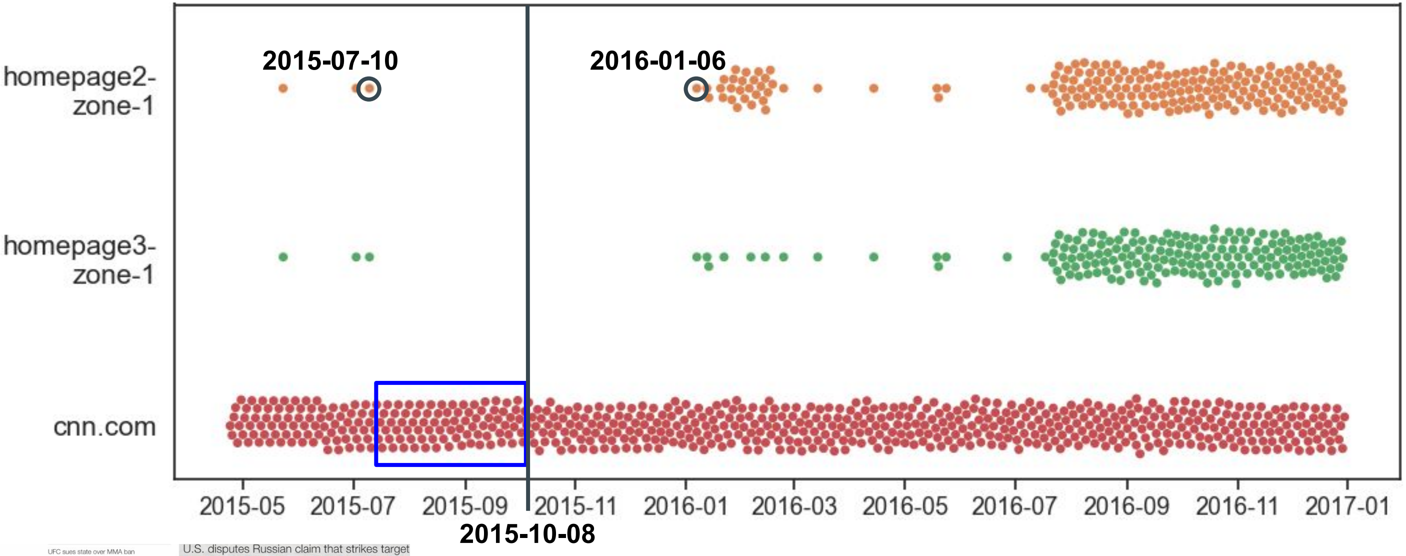
from Jan 6, 2016

https://web.archive.org/web/20150802000019/http://www.cnn.com/

https://web.archive.org/web/20151101000016/http://www.cnn.com/

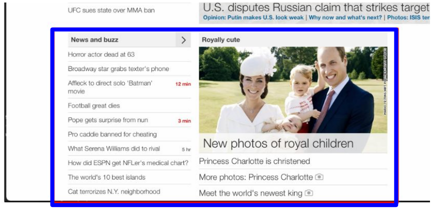
https://web.archive.org/web/20151001000018/https://www.cnn.com/

Every replay between July 10, 2015 and Oct 8, 2015 will use same homepage2-zone-1



mementos for zone-manager files (captured via browser-based crawler)

mementos for CNN.com base HTML



UFC sues state over MMA ban

U.S. disputes Russian claim that strikes target

Opinion: Putin makes U.S. look weak | Why now and what's next? | Photos: ISIS ter

News and buzz

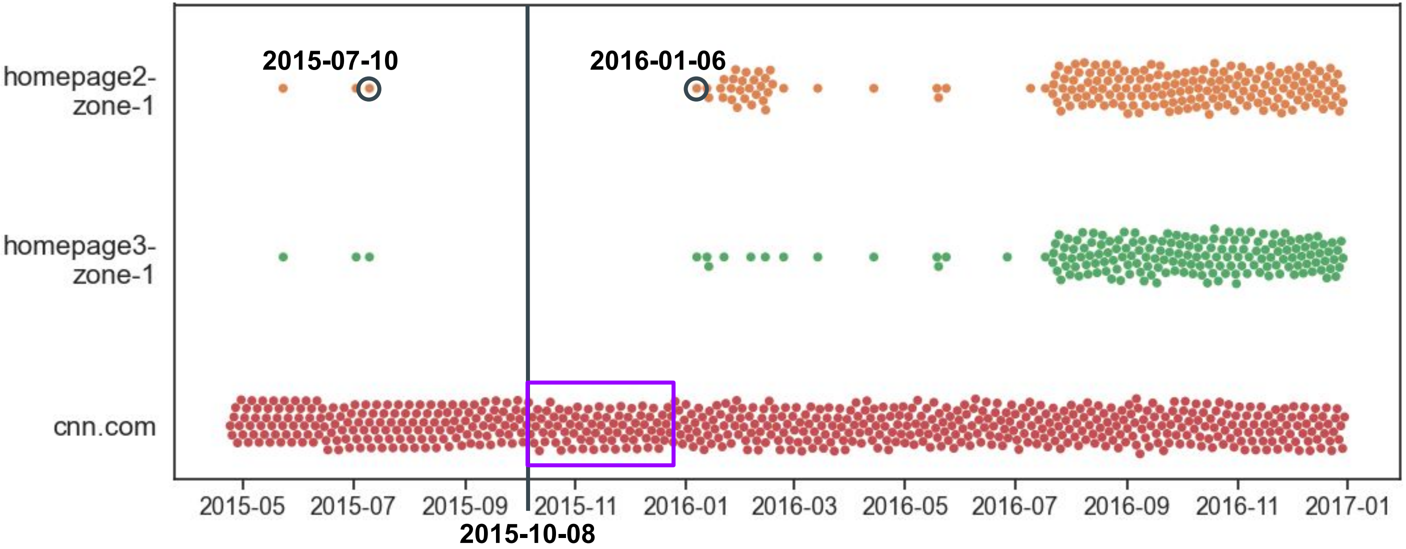
- Horror actor dead at 63
- Broadway star grabs texter's phone
- Affleck to direct solo 'Batman' movie 12 min
- Football great dies
- Pope gets surprise from nun 3 min
- Pro caddie banned for cheating
- What Serena Williams did to rival 5 hr
- How did ESPN get NFLer's medical chart?
- The world's 10 best islands
- Cat terrorizes N.Y. neighborhood

Royally cute

New photos of royal children

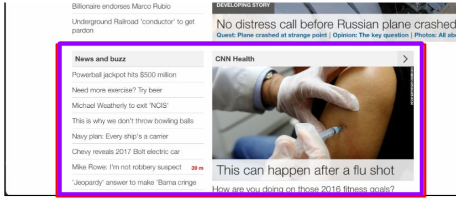
- Princess Charlotte is christened
- More photos: Princess Charlotte
- Meet the world's newest king

Every replay between Oct 8, 2015 and Jan 6, 2016 will use same homepage2-zone-1



mementos for zone-manager files (captured via browser-based crawler)

mementos for CNN.com base HTML



Billionaire endorses Marco Rubio

Underground Railroad 'conductor' to get pardon

DEVELOPING STORY

No distress call before Russian plane crashed

Quest: Plane crashed at strange point | Opinion: The key question | Photos: All ab

News and buzz

- Powerball jackpot hits $500 million
- Need more exercise? Try beer
- Michael Weatherly to exit 'NCIS'
- This is why we don't throw bowling balls
- Navy plan: Every ship's a carrier
- Chevy reveals 2017 Bolt electric car
- Mike Rowe: I'm not robbery suspect 39 m
- 'Jeopardy' answer to make 'Bama cringe

CNN Health >

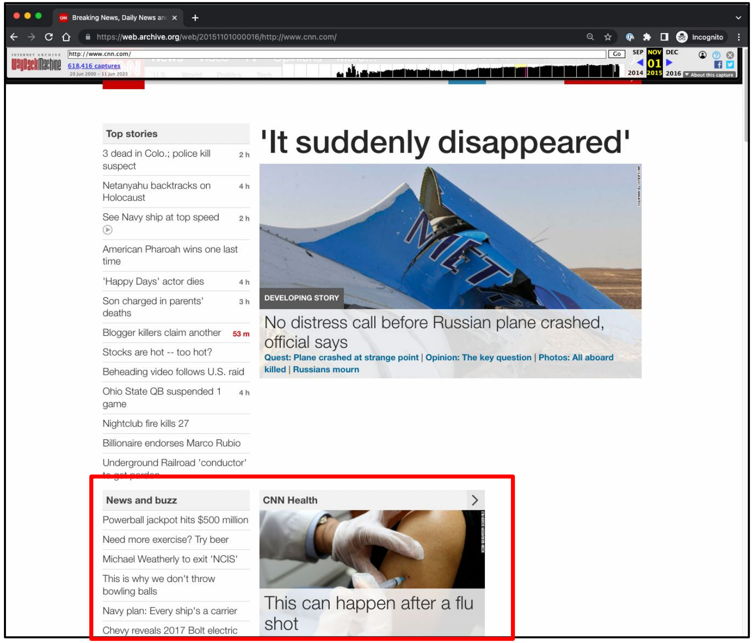
This can happen after a flu shot

How are you doing on those 2016 fitness goals?

Temporal violations are especially hard to detect on CNN.com

No date displayed on front page, even though individual articles do contain the date.

Nov 1, 2015



Top stories

- 3 dead in Colo.; police kill suspect 2 h
- Netanyahu backtracks on Holocaust 4 h
- See Navy ship at top speed 2 h
- American Pharoah wins one last time
- 'Happy Days' actor dies 4 h
- Son charged in parents' deaths 3 h
- Blogger killers claim another 53 m
- Stocks are hot -- too hot?
- Beheading video follows U.S. raid
- Ohio State QB suspended 1 game 4 h
- Nightclub fire kills 27
- Billionaire endorses Marco Rubio
- Underground Railroad 'conductor' to get pardon

'It suddenly disappeared'

DEVELOPING STORY

No distress call before Russian plane crashed, official says

Quest: Plane crashed at strange point | Opinion: The key question | Photos: All aboard killed | Russians mourn

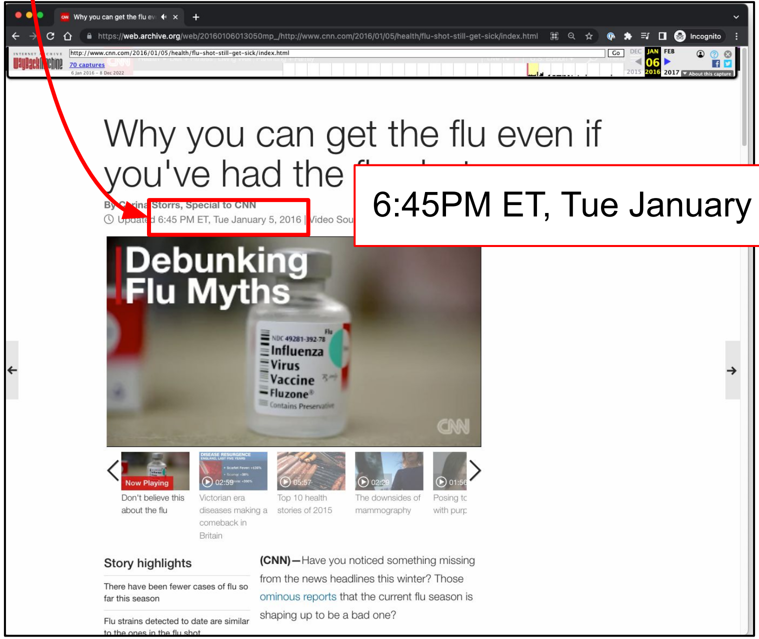
News and buzz

- Powerball jackpot hits $500 million
- Need more exercise? Try beer
- Michael Weatherly to exit 'NCIS'
- This is why we don't throw bowling balls
- Navy plan: Every ship's a carrier
- Chevy reveals 2017 Bolt electric

CNN Health

This can happen after a flu shot

https://web.archive.org/web/20151101000016/http://www.cnn.com/



Why you can get the flu even if you've had the fl

By Carina Storrs, Special to CNN

Updated 6:45 PM ET, Tue January 5, 2016 | Video Sou

6:45PM ET, Tue January 5, 2016

- Don't believe this about the flu
- Victorian era diseases making a comeback in Britain
- Top 10 health stories of 2015
- The downsides of mammography
- Posing tc with purp

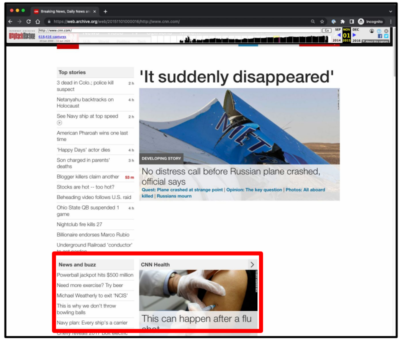
Story highlights

- There have been fewer cases of flu so far this season
- Flu strains detected to date are similar to the ones in the flu shot

(CNN)—Have you noticed something missing from the news headlines this winter? Those ominous reports that the current flu season is shaping up to be a bad one?

https://web.archive.org/web/20160106013050mp_/http://www.cnn.com/2016/01/05/health/flu-shot-still-get-sick/index.html

IAWM's "About this capture" lists temporal differences



Top stories

3 dead in Colo.; police kill suspect 2 h

Netanyahu backtracks on Holocaust 4 h

See Navy ship at top speed 2 h

American Pharoah wins one last time

'Happy Days' actor dies 4 h

Son charged in parents' deaths 3 h

Blogger killers claim another 53 m

Stocks are hot -- too hot?

Beheading video follows U.S. raid

Ohio State QB suspended 1 game 4 h

Nightclub fire kills 27

Billionaire endorses Marco Rubio

Underground Railroad 'conductor'

'It suddenly disappeared'

DEVELOPING STORY

No distress call before Russian plane crashed, official says

Quest: Plane crashed at strange point | Opinion: The key question | Photos: All aboard killed | Russians mourn

News and buzz

Powerball jackpot hits $500 million

Need more exercise? Try beer

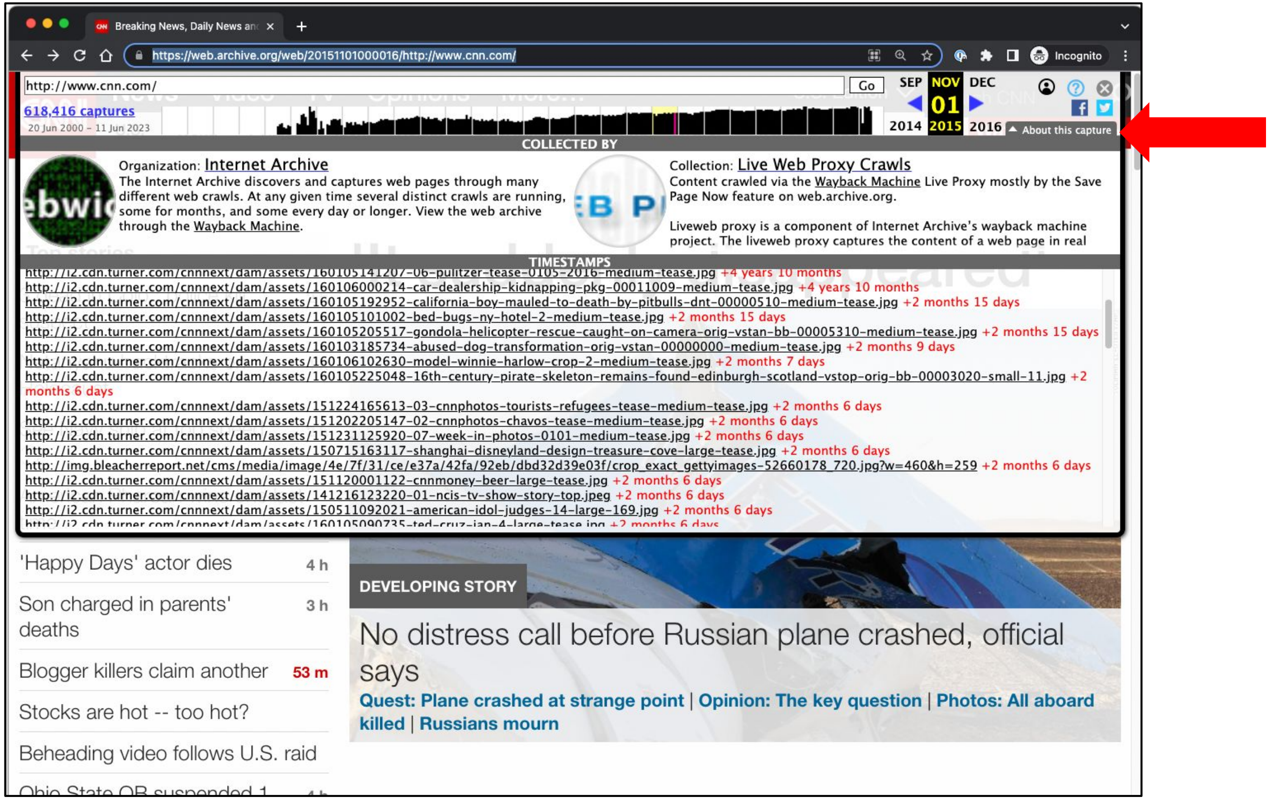
Michael Weatherly to exit 'NCIS'

This is why we don't throw bowling balls

Navy plan: Every ship's a carrier

CNN Health

This can happen after a flu shot



Breaking News, Daily News and
https://web.archive.org/web/20151101000016/http://www.cnn.com/
Incognito
http://www.cnn.com/
Go
618,416 captures
20 Jun 2000 – 11 Jun 2023
SEP
NOV
DEC
01
2014
2015
2016
About this capture

COLLECTED BY

Organization: Internet Archive

The Internet Archive discovers and captures web pages through many different web crawls. At any given time several distinct crawls are running, some for months, and some every day or longer. View the web archive through the Wayback Machine.

Collection: Live Web Proxy Crawls

Content crawled via the Wayback Machine Live Proxy mostly by the Save Page Now feature on web.archive.org.

Liveweb proxy is a component of Internet Archive's wayback machine project. The liveweb proxy captures the content of a web page in real

TIMESTAMPS

http://i2.cdn.turner.com/cnnnext/dam/assets/160105141207-06-pulitzer-tease-0105-2016-medium-tease.jpg +4 years 10 months
http://i2.cdn.turner.com/cnnnext/dam/assets/160106000214-car-dealership-kidnapping-pkg-00011009-medium-tease.jpg +4 years 10 months
http://i2.cdn.turner.com/cnnnext/dam/assets/160105192952-california-boy-mauled-to-death-by-pitbulls-dnt-00000510-medium-tease.jpg +2 months 15 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160105101002-bed-bugs-ny-hotel-2-medium-tease.jpg +2 months 15 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160105205517-gondola-helicopter-rescue-caught-on-camera-orig-vstan-bb-00005310-medium-tease.jpg +2 months 15 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160103185734-abused-dog-transformation-orig-vstan-00000000-medium-tease.jpg +2 months 9 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160106102630-model-winnie-harlow-crop-2-medium-tease.jpg +2 months 7 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160105225048-16th-century-pirate-skeleton-remains-found-edinburgh-scotland-vstop-orig-bb-00003020-small-11.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/151224165613-03-cnnphotos-tourists-refugees-tease-medium-tease.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/151202205147-02-cnnphotos-chavos-tease-medium-tease.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/151231125920-07-week-in-photos-0101-medium-tease.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/150715163117-shanghai-disneyland-design-treasure-cove-large-tease.jpg +2 months 6 days
http://img.bleacherreport.net/cms/media/image/4e/7f/31/ce/e37a/42fa/92eb/dbd32d39e03f/crop_exact_gettyimages-52660178_720.jpg?w=460&h=259 +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/151120001122-cnnmoney-beer-large-tease.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/141216123220-01-ncis-tv-show-story-top.jpeg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/150511092021-american-idol-judges-14-large-169.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160105090735-ted-cruz-ian-4-large-tease.jpg +2 months 6 days

'Happy Days' actor dies 4 h

Son charged in parents' deaths 3 h

Blogger killers claim another 53 m

Stocks are hot -- too hot?

Beheading video follows U.S. raid

Ohio State QB suspended 1

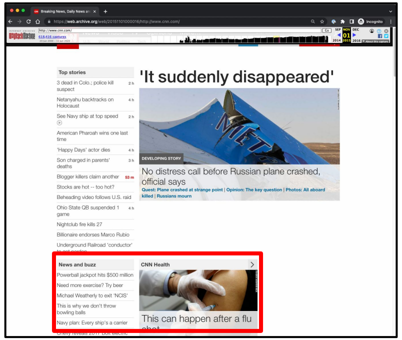
DEVELOPING STORY

No distress call before Russian plane crashed, official says

Quest: Plane crashed at strange point | Opinion: The key question | Photos: All aboard killed | Russians mourn

https://web.archive.org/web/20151101000016/http://www.cnn.com/

But "About this capture" does not include zone-manager files



Top stories

'It suddenly disappeared'

3 dead in Colo.; police kill suspect 2 h

Netanyahu backtracks on Holocaust 4 h

See Navy ship at top speed 2 h

American Pharoah wins one last time

'Happy Days' actor dies 4 h

Son charged in parents' deaths 3 h

Blogger killers claim another 53 m

Stocks are hot -- too hot?

Beheading video follows U.S. raid

Ohio State QB suspended 1 game 4 h

Nightclub fire kills 27

Billionaire endorses Marco Rubio

Underground Railroad 'conductor'

DEVELOPING STORY

No distress call before Russian plane crashed, official says

Quest: Plane crashed at strange point | Opinion: The key question | Photos: All aboard killed | Russians mourn

News and buzz

Powerball jackpot hits $500 million

Need more exercise? Try beer

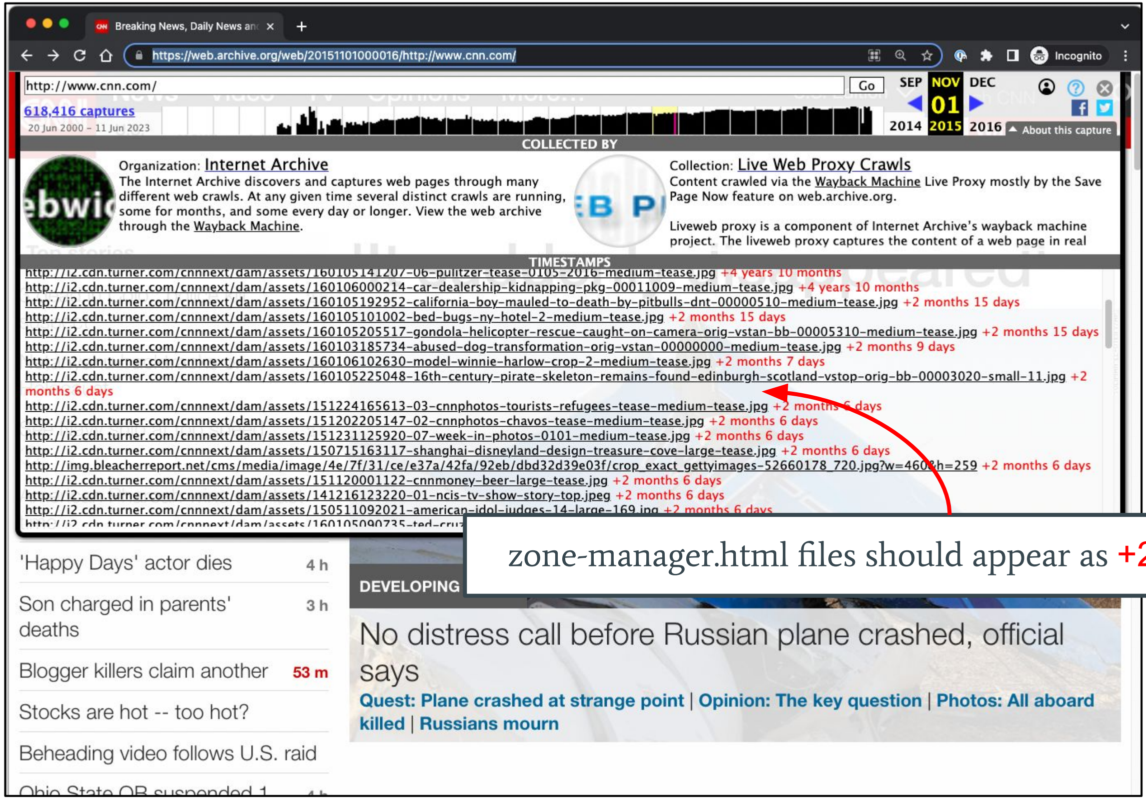
Michael Weatherly to exit 'NCIS'

This is why we don't throw bowling balls

Navy plan: Every ship's a carrier

CNN Health

This can happen after a flu shot



Breaking News, Daily News anc ×
+
https://web.archive.org/web/20151101000016/http://www.cnn.com/
Incognito
http://www.cnn.com/
Go
618,416 captures
20 Jun 2000 – 11 Jun 2023
SEP NOV DEC
01
2014 2015 2016
About this capture

COLLECTED BY

Organization: Internet Archive

The Internet Archive discovers and captures web pages through many different web crawls. At any given time several distinct crawls are running, some for months, and some every day or longer. View the web archive through the Wayback Machine.

Collection: Live Web Proxy Crawls

Content crawled via the Wayback Machine Live Proxy mostly by the Save Page Now feature on web.archive.org.

Liveweb proxy is a component of Internet Archive's wayback machine project. The liveweb proxy captures the content of a web page in real

TIMESTAMPS

http://i2.cdn.turner.com/cnnnext/dam/assets/160105141207-06-pulitzer-tease-0105-2016-medium-tease.jpg +4 years 10 months
http://i2.cdn.turner.com/cnnnext/dam/assets/160106000214-car-dealership-kidnapping-pkg-00011009-medium-tease.jpg +4 years 10 months
http://i2.cdn.turner.com/cnnnext/dam/assets/160105192952-california-boy-mauled-to-death-by-pitbulls-dnt-00000510-medium-tease.jpg +2 months 15 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160105101002-bed-bugs-ny-hotel-2-medium-tease.jpg +2 months 15 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160105205517-gondola-helicopter-rescue-caught-on-camera-orig-vstan-bb-00005310-medium-tease.jpg +2 months 15 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160103185734-abused-dog-transformation-orig-vstan-00000000-medium-tease.jpg +2 months 9 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160106102630-model-winnie-harlow-crop-2-medium-tease.jpg +2 months 7 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160105225048-16th-century-pirate-skeleton-remains-found-edinburgh-scotland-vstop-orig-bb-00003020-small-11.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/151224165613-03-cnnphotos-tourists-refugees-tease-medium-tease.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/151202205147-02-cnnphotos-chavos-tease-medium-tease.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/151231125920-07-week-in-photos-0101-medium-tease.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/150715163117-shanghai-disneyland-design-treasure-cove-large-tease.jpg +2 months 6 days
http://img.bleacherreport.net/cms/media/image/4e/7f/31/ce/e37a/42fa/92eb/dbd32d39e03f/crop_exact_gettyimages-52660178_720.jpg?w=460&h=259 +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/151120001122-cnnmoney-beer-large-tease.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/141216123220-01-ncis-tv-show-story-top.jpeg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/150511092021-american-idol-judges-14-large-169.jpg +2 months 6 days
http://i2.cdn.turner.com/cnnnext/dam/assets/160105090735-ted-cruz

'Happy Days' actor dies 4 h

Son charged in parents' deaths 3 h

Blogger killers claim another 53 m

Stocks are hot -- too hot?

Beheading video follows U.S. raid

Ohio State QB suspended 1

DEVELOPING

No distress call before Russian plane crashed, official says

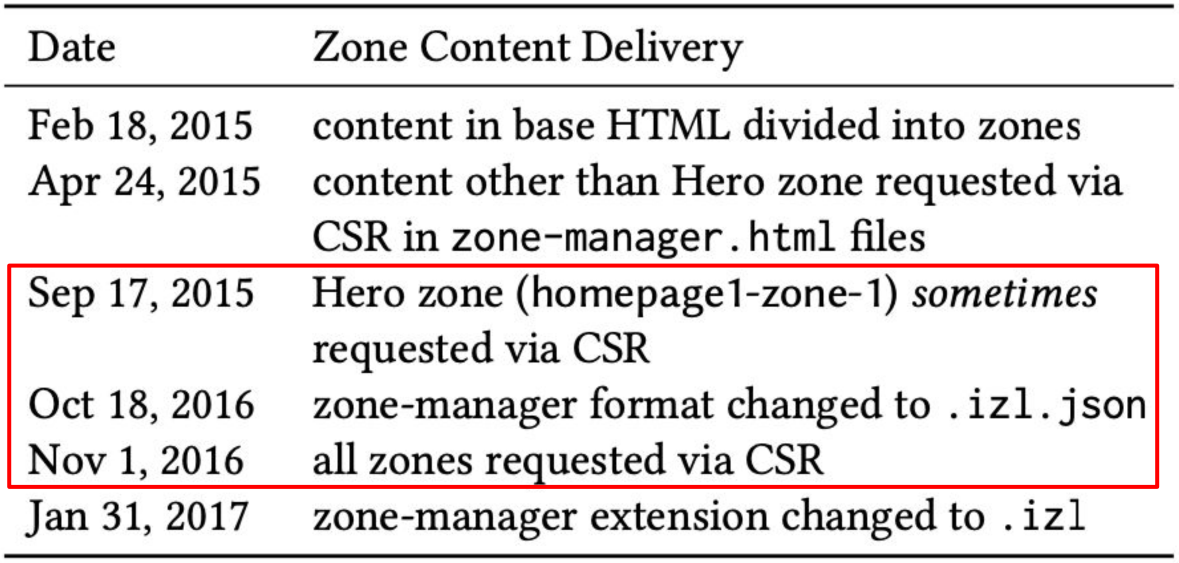
Quest: Plane crashed at strange point | Opinion: The key question | Photos: All aboard killed | Russians mourn

zone-manager.html files should appear as +2 months 6 days

https://web.archive.org/web/20151101000016/http://www.cnn.com/

Sep 2015 - Nov 2016: Hero zone sometimes requested via CSR

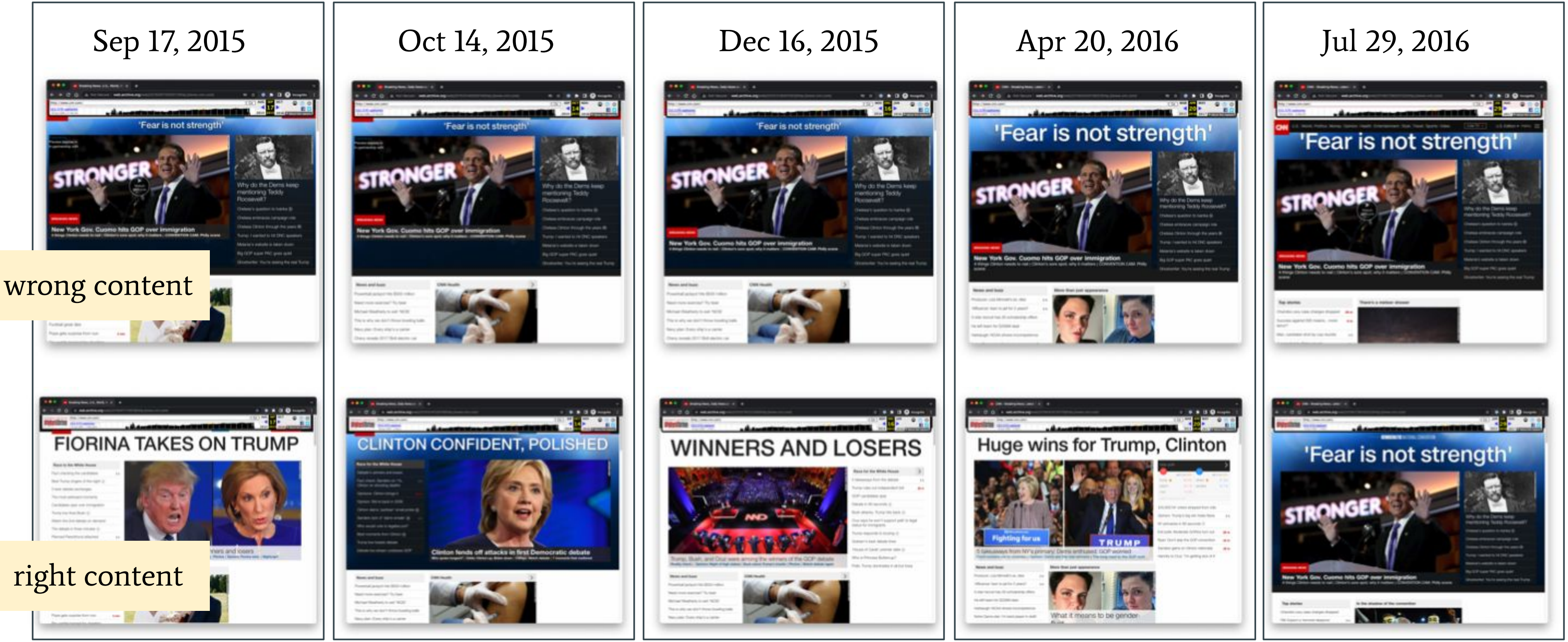
We downloaded raw HTML (using "id_") of 1316 mementos between 2010-2022 to investigate when Hero story delivered via base HTML vs. CSR.



Date	Zone Content Delivery
Feb 18, 2015	content in base HTML divided into zones
Apr 24, 2015	content other than Hero zone requested via CSR in `zone-manager.html` files
Sep 17, 2015	Hero zone (homepage1-zone-1) *sometimes* requested via CSR
Oct 18, 2016	zone-manager format changed to `.izl.json`
Nov 1, 2016	all zones requested via CSR
Jan 31, 2017	zone-manager extension changed to `.izl`

https://web.archive.org/web/20130806040521/http://faq.web.archive.org/page-without-wayback-code/

Hero story temporal violations



Sep 17, 2015

Oct 14, 2015

Dec 16, 2015

Apr 20, 2016

Jul 29, 2016

wrong content

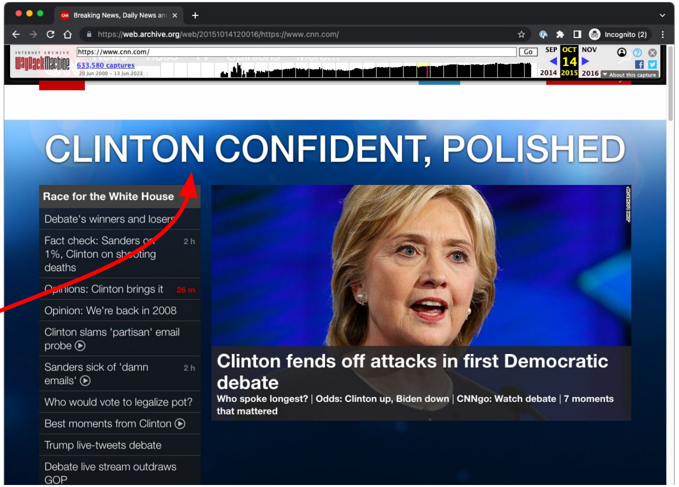
right content

'Fear is not strength'
STRONGER
New York Gov. Cuomo hits GOP over immigration
FIORINA TAKES ON TRUMP
CLINTON CONFIDENT, POLISHED
Clinton fends off attacks in first Democratic debate
WINNERS AND LOSERS
Huge wins for Trump, Clinton
Fighting for us
TRUMP
What it means to be gender

Hero stories *sometimes* in the base HTML

If the base HTML contains a `<section>` with `id="homepage1-zone-1"`, then Hero content is in the HTML.

```
<section class="zn zn-homepage1-zone-1
zn-left-fluid-right
      -stack zn--idx-0 t-light zn-loaded zn-left-fluid-
      right-stack zn-has-multiple-containers zn-has-3-
      containers" data-eq-pts="xsmall: 0,
      medium: 460, large: 780, full16x9: 1100"
      id="homepage1-zone-1" data-vr-zone="zone-0-0"
      data-zone-label="Hero" data-containers="3">
[...]
<h2 class="banner-text js-screaming-banner-text
screaming-banner-text">CLINTON CONFIDENT, POLISHED</h2>
[...]
```



CLINTON CONFIDENT, POLISHED

Race for the White House

- Debate's winners and losers
- Fact check: Sanders on 1%, Clinton on shooting deaths 2 h
- Opinions: Clinton brings it 26 m
- Opinion: We're back in 2008
- Clinton slams 'partisan' email probe
- Sanders sick of 'damn emails' 2 h
- Who would vote to legalize pot?
- Best moments from Clinton
- Trump live-tweets debate
- Debate live stream outdraws GOP

Clinton fends off attacks in first Democratic debate

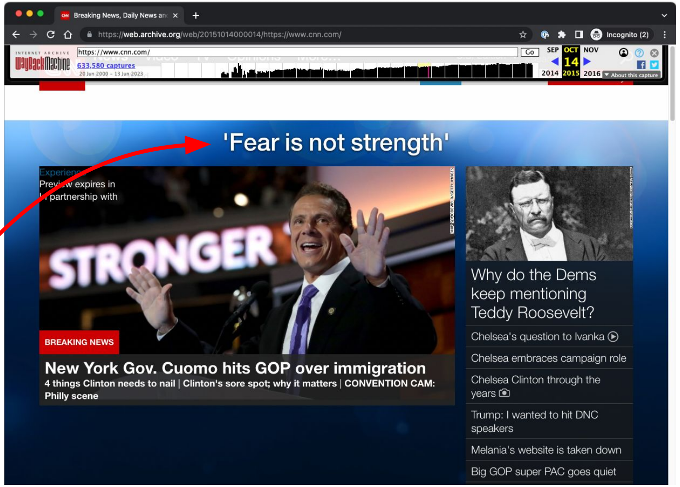
Who spoke longest? | Odds: Clinton up, Biden down | CNNgo: Watch debate | 7 moments that mattered

Oct 14, 2015

[https://web.archive.org/web/**20151014**120016/https://www.cnn.com/](https://web.archive.org/web/20151014120016/https://www.cnn.com/)

Hero stories *sometimes* loaded via CSR

If there is no `id="homepage1-zone-1"` in the base HTML, then Hero content is in the zone-manager file.



'Fear is not strength'

Experience
Preview expires in
partnership with

BREAKING NEWS

New York Gov. Cuomo hits GOP over immigration

4 things Clinton needs to nail | Clinton's sore spot; why it matters | CONVENTION CAM: Philly scene

Why do the Dems keep mentioning Teddy Roosevelt?

- Chelsea's question to Ivanka
- Chelsea embraces campaign role
- Chelsea Clinton through the years
- Trump: I wanted to hit DNC speakers
- Melania's website is taken down
- Big GOP super PAC goes quiet

Oct 14, 2015

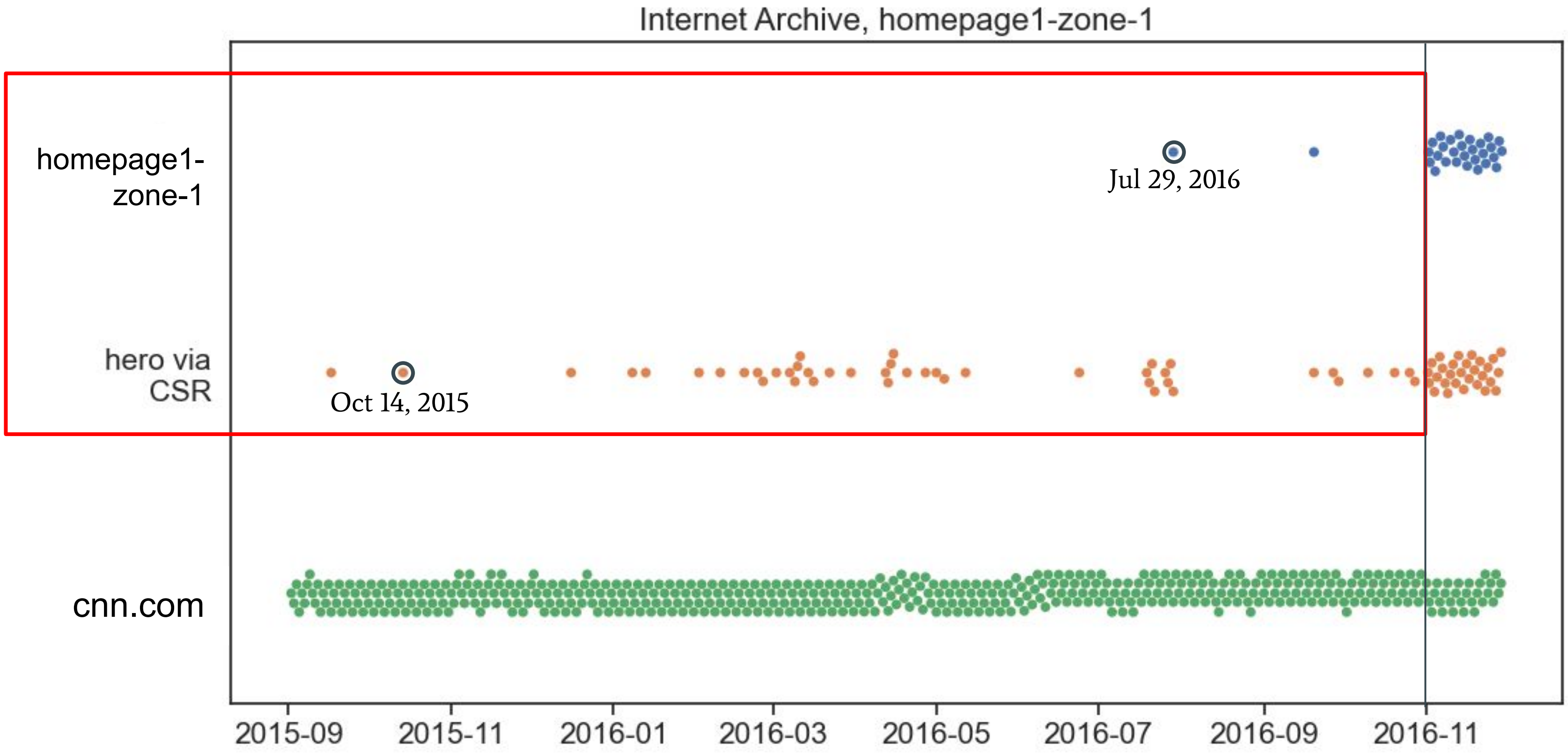
https://web.archive.org/web/**20151014**000014/https://www.cnn.com/

```
[...]
<h2 class="banner-text banner-text--maximized
banner-text-size--char-23">'Fear is not
strength'</h2>
[...]
```

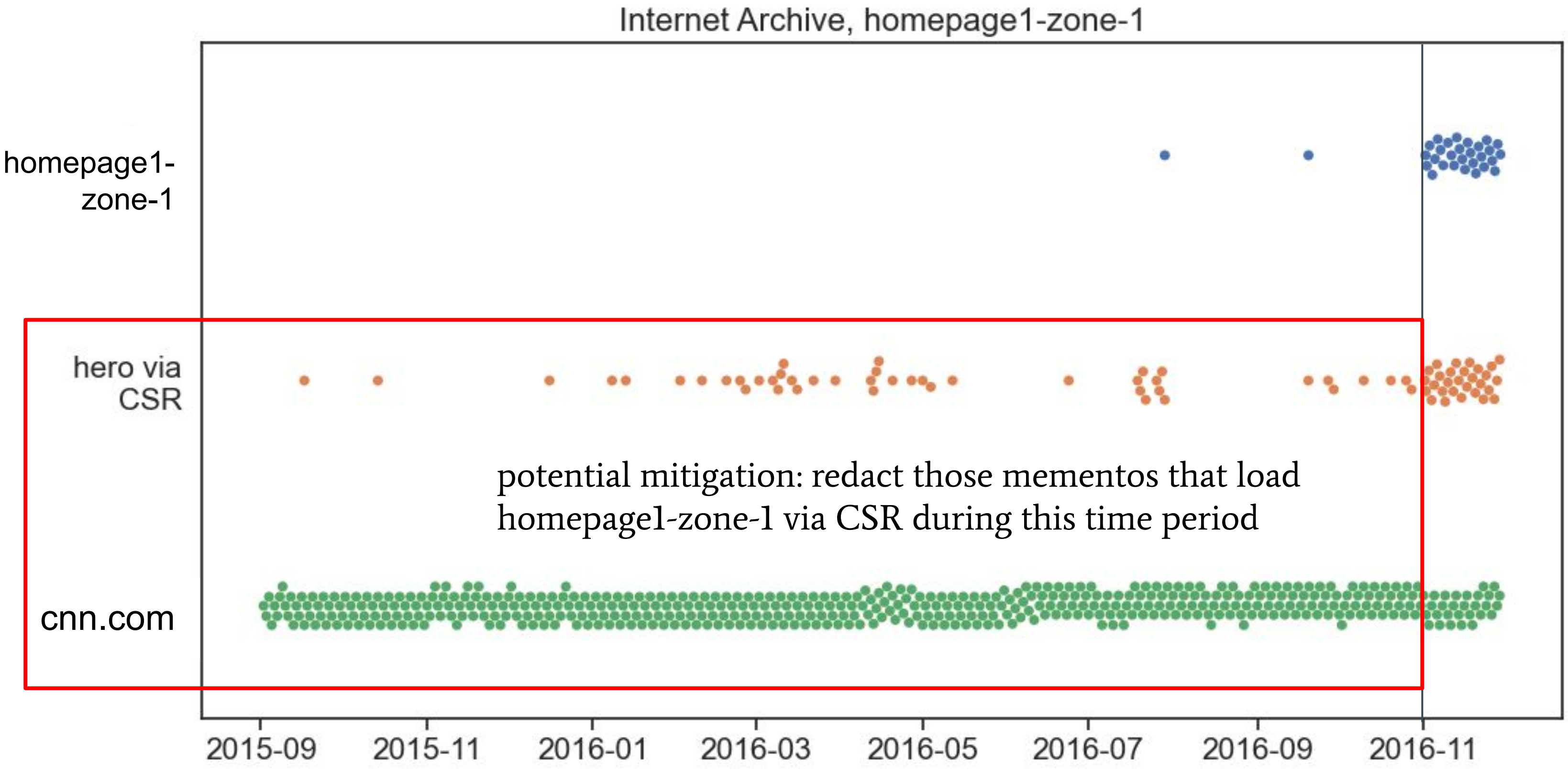
Content loaded from
https://web.archive.org/web/**20160729**003156/http://www.cnn.com/data/ocs/section/index.html:**homepage1-zone-1**/views/zones/common/**zone-manager.html**

verified by examining raw HTML of 1069 mementos between Apr 24, 2015 - Nov 1, 2016 and comparing to the requests logged by the headless browser

Hero story violation if requested via CSR and no hero memento

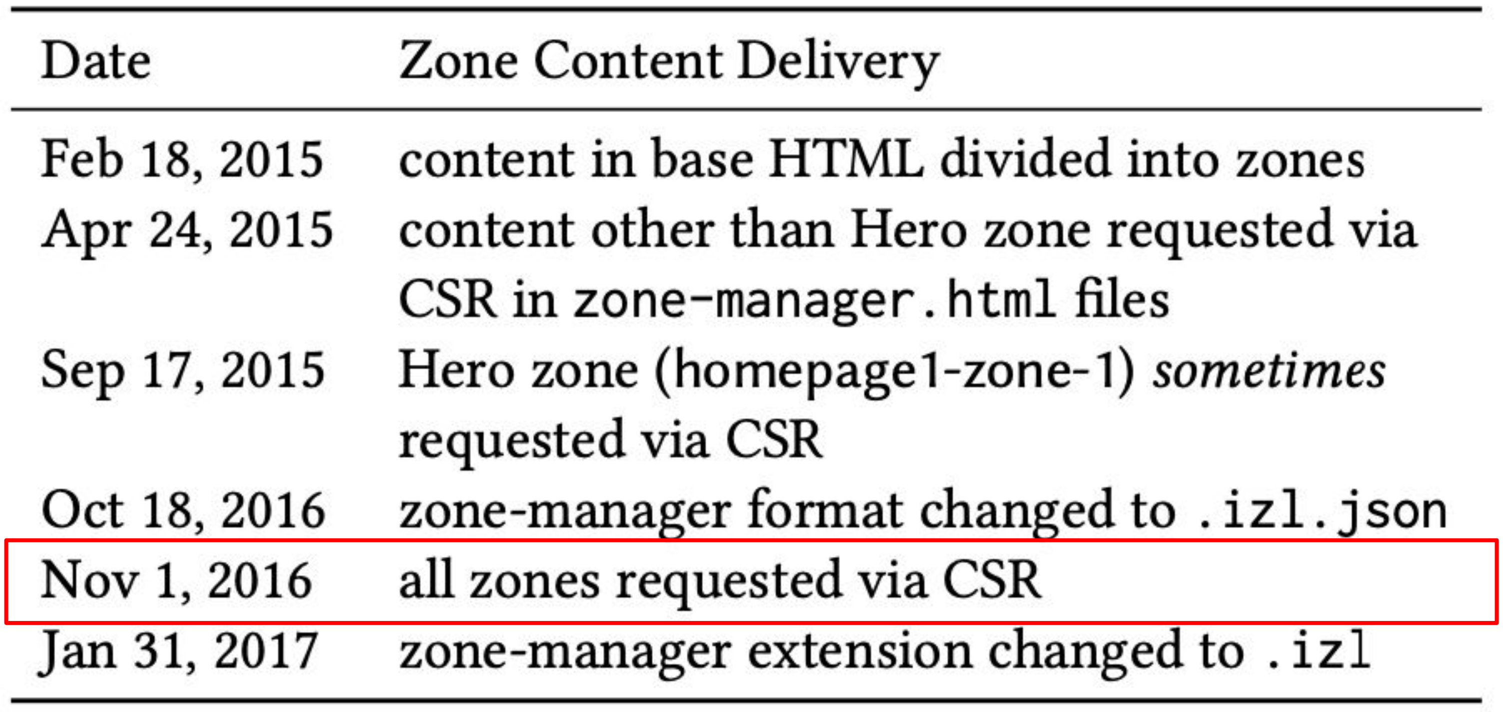


But, each day had mementos with the Hero story in the HTML



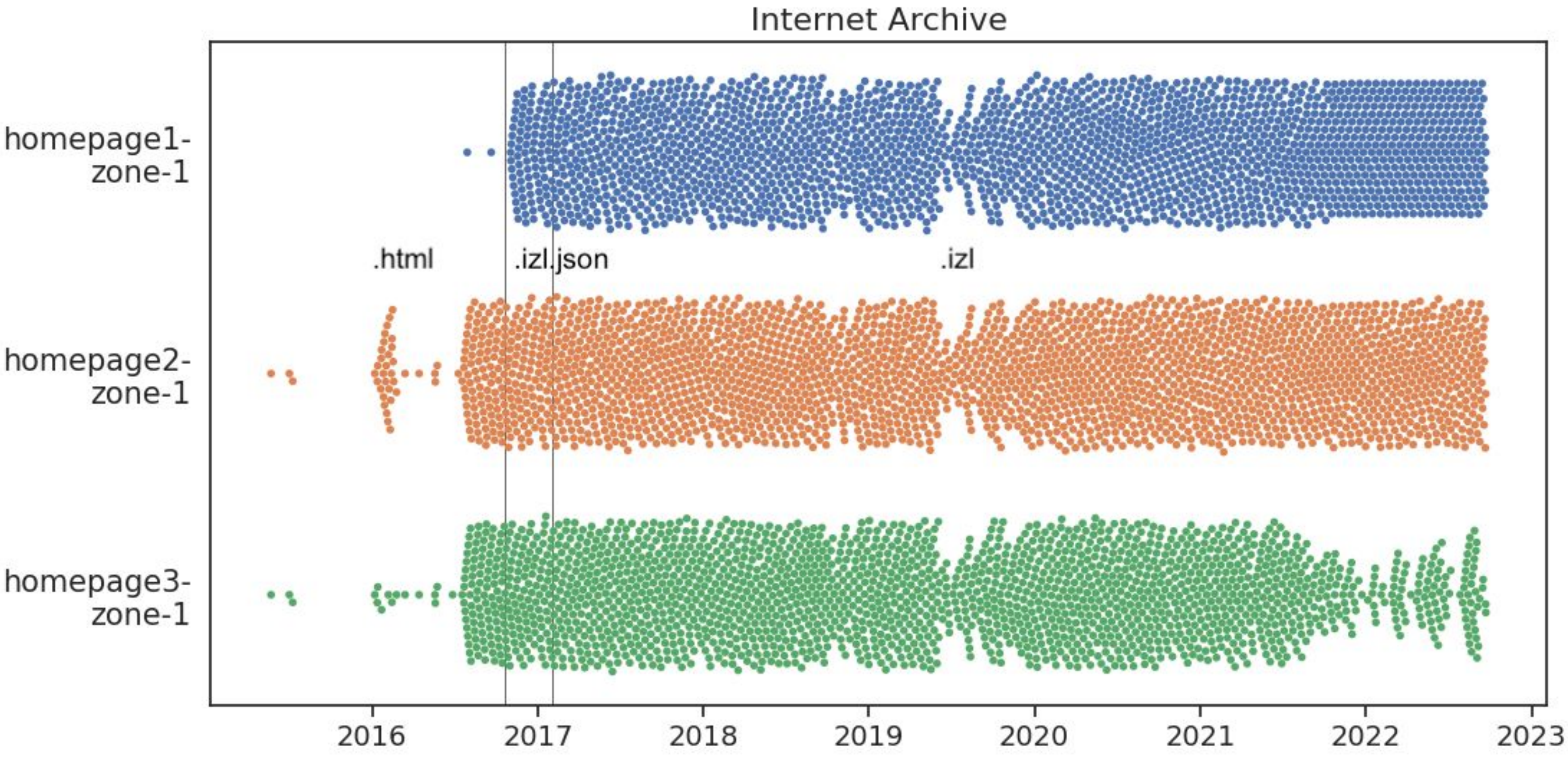
potential mitigation: redact those mementos that load homepage1-zone-1 via CSR during this time period

Nov 1, 2016: All content requested via CSR



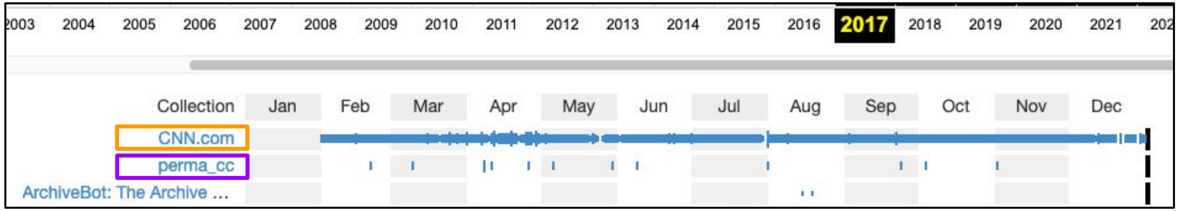
Date	Zone Content Delivery
Feb 18, 2015	content in base HTML divided into zones
Apr 24, 2015	content other than Hero zone requested via CSR in `zone-manager.html` files
Sep 17, 2015	Hero zone (homepage1-zone-1) *sometimes* requested via CSR
Oct 18, 2016	zone-manager format changed to `.izl.json`
Nov 1, 2016	all zones requested via CSR
Jan 31, 2017	zone-manager extension changed to `.izl`

Starting in 2017, good coverage of mementos of the top 3 CNN zones

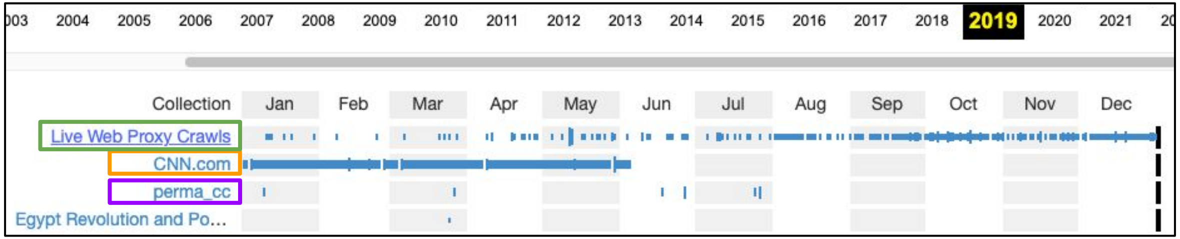


Contributions of homepage1-zone-1 zone-manager files from browser-based AIT collections, perma.cc, and Save Page Now

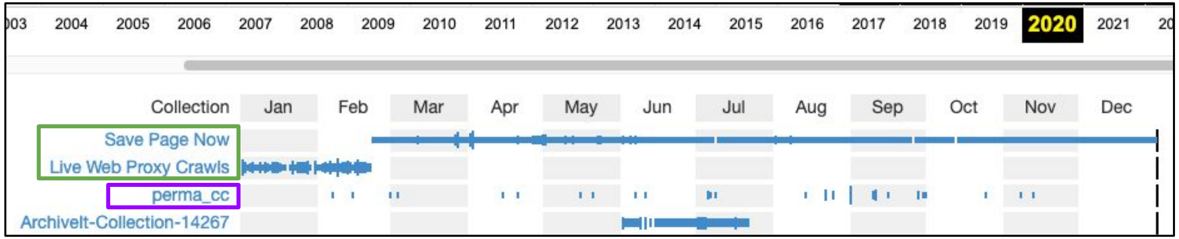
CNN.com - Archive-It collection 7678, Jul 2016 - Sep 2019



perma.cc

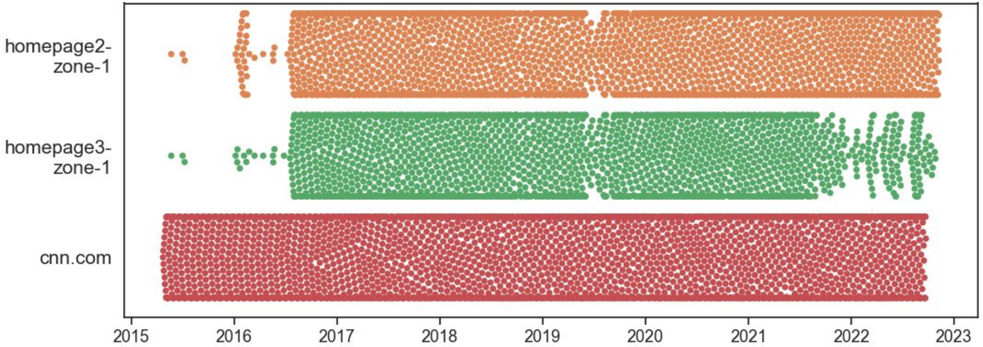


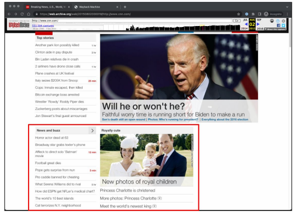
Live Web Proxy Crawls, Save Page Now



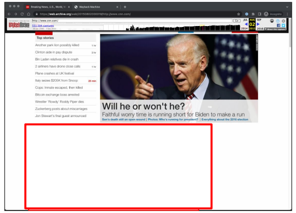
https://web.archive.org/web/collections/20170901000000*/https://www.cnn.com/data/ocs/section/index.html:**homepage1-zone-1**/views/zones/common/zone-manager.izl

How many days would have missing content if IAWM limited the allowed distance between replayed HTML and JSON?



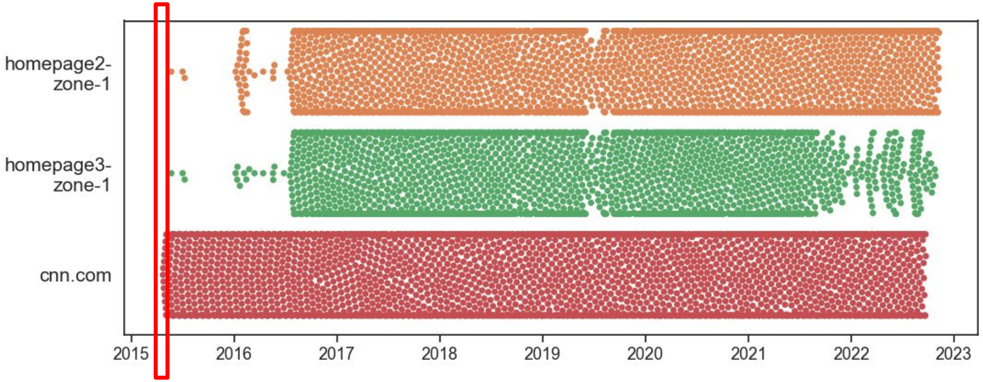

Will he or won't he?
Faithful worry time is running short for Biden to make a run
New photos of royal children

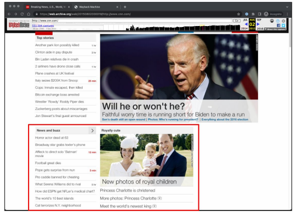
vs.


Will he or won't he?
Faithful worry time is running short for Biden to make a run

	Total			Affected	
Date Range	Mementos	Days	Threshold	Mementos	Days
Apr 24, 2015– May 23, 2015 (none)	788	29	1 hr	786	29
			2 hrs	785	29
			6 hrs	780	28
			24 hrs	754	27
			48 hrs	724	26
May 23, 2015– Jul 21, 2016 (sparse)	17,017	425	1 hr	16,863	425
			2 hrs	16,738	425
			6 hrs	16,244	410
			24 hrs	15,142	370
			48 hrs	14,178	342
Jul 21, 2016– Sep 23, 2022 (dense)	154,917	2256	1 hr	95,009	1826
			2 hrs	67,992	1425
			6 hrs	13,893	190
			24 hrs	1812	37
			48 hrs	526	8

48 hrs: Only 3/29 days would have 2nd level and lower content





Breaking News, U.S., World, W | Wayback Machine

https://web.archive.org/web/20150802000019/http://www.cnn.com/

http://www.cnn.com/ | Go | JUL AUG SEP | 02 | 2014 2015 2016

551,504 captures

Top stories

- Another park lion possibly killed 1 hr
- Clinton aide in pay dispute 3 hr
- Bin Laden relatives die in crash
- 2 airliners have drone close calls 1 hr
- Plane crashes at UK festival
- Italy seizes $205K from Snoop 25 min
- Cops: Inmate escaped, then killed
- Bitcoin exchange boss arrested
- Wrestler 'Rowdy' Roddy Piper dies
- Zuckerberg posts about miscarriages
- Jon Stewart's final guest announced

Will he or won't he?

Faithful worry time is running short for Biden to make a run

Son's death still an open wound | Photos: Who's running for president? | Everything about the 2016 election

News and buzz

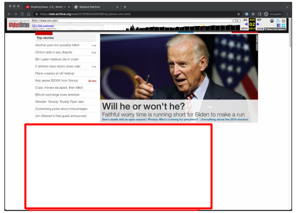
- Horror actor dead at 63
- Broadway star grabs texter's phone
- Affleck to direct solo 'Batman' movie 12 min
- Football great dies
- Pope gets surprise from nun 3 min
- Pro caddie banned for cheating
- What Serena Williams did to rival 5 hr
- How did ESPN get NFLer's medical chart?
- The world's 10 best islands
- Cat terrorizes N.Y. neighborhood

Royally cute

New photos of royal children

- Princess Charlotte is christened
- More photos: Princess Charlotte
- Meet the world's newest king

vs.



Top stories

- Another park lion possibly killed 1 hr
- Clinton aide in pay dispute 3 hr
- Bin Laden relatives die in crash
- 2 airliners have drone close calls 1 hr
- Plane crashes at UK festival
- Italy seizes $205K from Snoop 25 min
- Cops: Inmate escaped, then killed
- Bitcoin exchange boss arrested
- Wrestler 'Rowdy' Roddy Piper dies
- Zuckerberg posts about miscarriages
- Jon Stewart's final guest announced

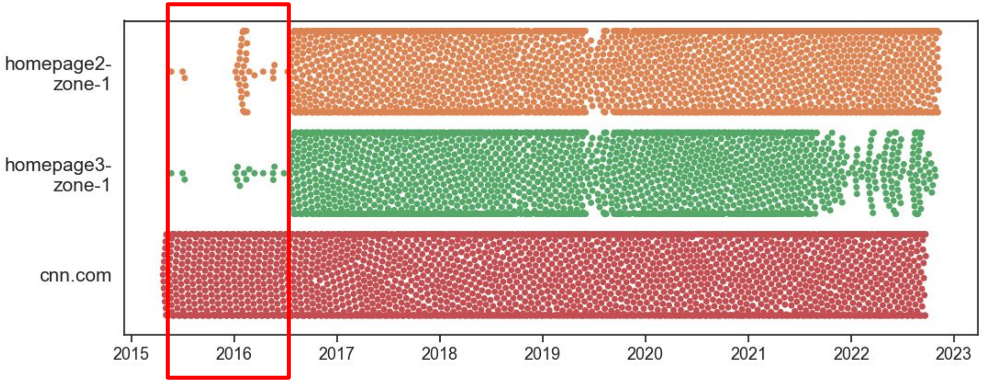
Will he or won't he?

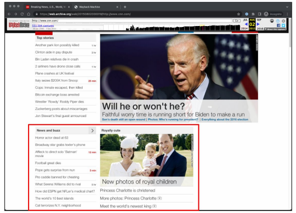
Faithful worry time is running short for Biden to make a run

Son's death still an open wound | Photos: Who's running for president? | Everything about the 2016 election

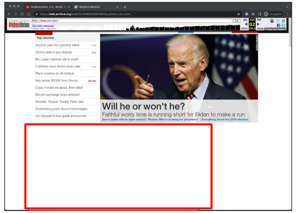
	Total			Affected	
Date Range	Mementos	Days	Threshold	Mementos	Days
Apr 24, 2015– May 23, 2015	788	29	1 hr	786	29
			2 hrs	785	29
			6 hrs	780	28
(none)			24 hrs	754	27
			48 hrs	724	26
May 23, 2015– Jul 21, 2016	17,017	425	1 hr	16,863	425
			2 hrs	16,738	425
			6 hrs	16,244	410
(sparse)			24 hrs	15,142	370
			48 hrs	14,178	342
Jul 21, 2016– Sep 23, 2022	154,917	2256	1 hr	95,009	1826
			2 hrs	67,992	1425
			6 hrs	13,893	190
(dense)			24 hrs	1812	37
			48 hrs	526	8

48 hrs: Only 83/425 days would have 2nd level and lower content



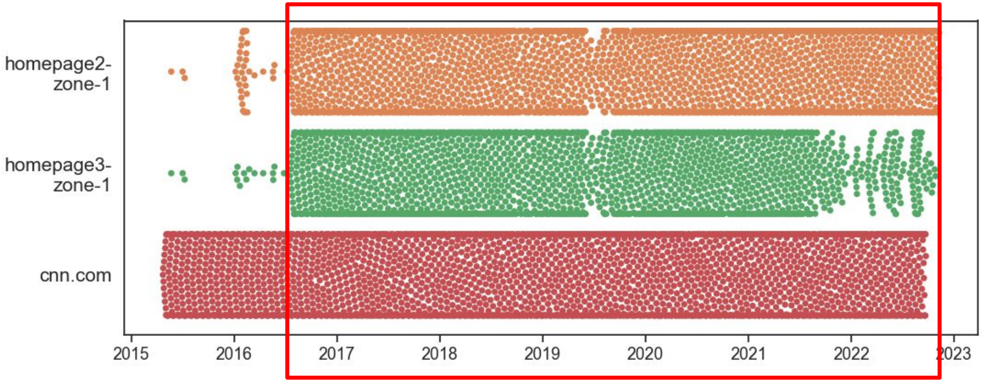

Will he or won't he?
Faithful worry time is running short for Biden to make a run
New photos of royal children

vs.

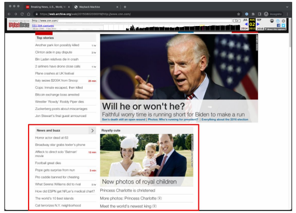

Will he or won't he?
Faithful worry time is running short for Biden to make a run

Date Range	Total Mementos	Total Days	Threshold	Affected Mementos	Affected Days
Apr 24, 2015– May 23, 2015 (none)	788	29	1 hr	786	29
			2 hrs	785	29
			6 hrs	780	28
			24 hrs	754	27
			48 hrs	724	26
May 23, 2015– Jul 21, 2016 (sparse)	17,017	425	1 hr	16,863	425
			2 hrs	16,738	425
			6 hrs	16,244	410
			24 hrs	15,142	370
			48 hrs	14,178	342
Jul 21, 2016– Sep 23, 2022 (dense)	154,917	2256	1 hr	95,009	1826
			2 hrs	67,992	1425
			6 hrs	13,893	190
			24 hrs	1812	37
			48 hrs	526	8

48 hrs: Only 8 days would be *missing* 2nd level and lower content



Date Range	Total Mementos	Total Days	Threshold	Affected Mementos	Affected Days
Apr 24, 2015– May 23, 2015	788	29	1 hr	786	29
			2 hrs	785	29
			6 hrs	780	28
(none)			24 hrs	754	27
			48 hrs	724	26
May 23, 2015– Jul 21, 2016	17,017	425	1 hr	16,863	425
			2 hrs	16,738	425
			6 hrs	16,244	410
(sparse)			24 hrs	15,142	370
			48 hrs	14,178	342
Jul 21, 2016– Sep 23, 2022	154,917	2256	1 hr	95,009	1826
			2 hrs	67,992	1425
			6 hrs	13,893	190
(dense)			24 hrs	1812	37
			48 hrs	526	8


Breaking News, U.S., World, W
Wayback Machine
https://web.archive.org/web/20150802000019/http://www.cnn.com/
http://www.cnn.com/
551,504 captures
Go
JUL
AUG
SEP
02
2014
2015
2016

Top stories

- Another park lion possibly killed 1 hr
- Clinton aide in pay dispute 3 hr
- Bin Laden relatives die in crash
- 2 airliners have drone close calls 1 hr
- Plane crashes at UK festival
- Italy seizes $205K from Snoop 25 min
- Cops: Inmate escaped, then killed
- Bitcoin exchange boss arrested
- Wrestler 'Rowdy' Roddy Piper dies
- Zuckerberg posts about miscarriages
- Jon Stewart's final guest announced

Will he or won't he?

Faithful worry time is running short for Biden to make a run

Son's death still an open wound | Photos: Who's running for president? | Everything about the 2016 election

News and buzz

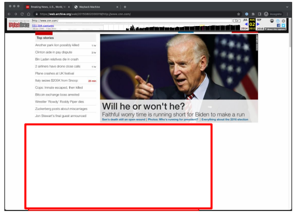
- Horror actor dead at 63
- Broadway star grabs texter's phone
- Affleck to direct solo 'Batman' movie 12 min
- Football great dies
- Pope gets surprise from nun 3 min
- Pro caddie banned for cheating
- What Serena Williams did to rival 5 hr
- How did ESPN get NFLer's medical chart?
- The world's 10 best islands
- Cat terrorizes N.Y. neighborhood

Royally cute

New photos of royal children

- Princess Charlotte is christened
- More photos: Princess Charlotte
- Meet the world's newest king

vs.


Top stories

- Another park lion possibly killed 1 hr
- Clinton aide in pay dispute 3 hr
- Bin Laden relatives die in crash
- 2 airliners have drone close calls 1 hr
- Plane crashes at UK festival
- Italy seizes $205K from Snoop 25 min
- Cops: Inmate escaped, then killed
- Bitcoin exchange boss arrested
- Wrestler 'Rowdy' Roddy Piper dies
- Zuckerberg posts about miscarriages
- Jon Stewart's final guest announced

Will he or won't he?

Faithful worry time is running short for Biden to make a run

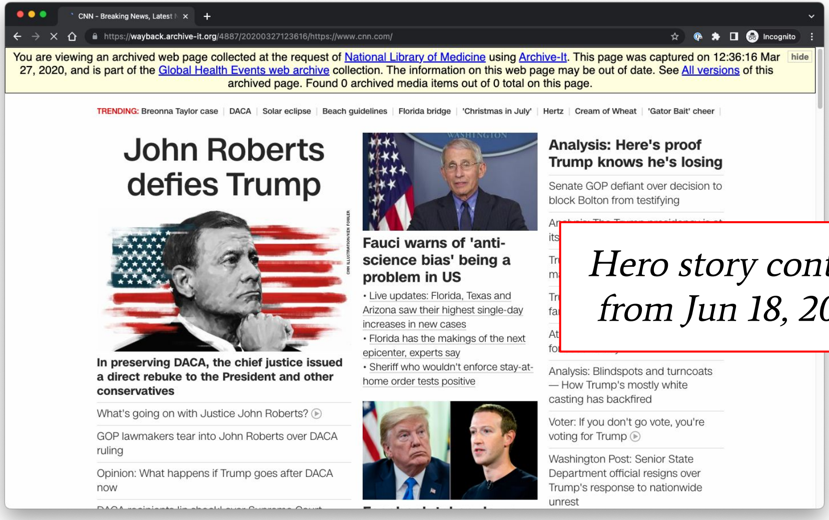
Son's death still an open wound | Photos: Who's running for president? | Everything about the 2016 election

Archive-It collections can be sparse, could have continuing issues

Mar 27, 2020

❌ Archive-It collection 4887
National Library of Medicine's Global Health Events
https://archive-it.org/collections/4887

✅ Internet Archive Wayback Machine (IAWM)


CNN - Breaking News, Latest
https://wayback.archive-it.org/4887/20200327123616/https://www.cnn.com/
Incognito
hide

You are viewing an archived web page collected at the request of National Library of Medicine using Archive-It. This page was captured on 12:36:16 Mar 27, 2020, and is part of the Global Health Events web archive collection. The information on this web page may be out of date. See All versions of this archived page. Found 0 archived media items out of 0 total on this page.

TRENDING: Breonna Taylor case | DACA | Solar eclipse | Beach guidelines | Florida bridge | 'Christmas in July' | Hertz | Cream of Wheat | 'Gator Bait' cheer |

John Roberts defies Trump

In preserving DACA, the chief justice issued a direct rebuke to the President and other conservatives

What's going on with Justice John Roberts?

GOP lawmakers tear into John Roberts over DACA ruling

Opinion: What happens if Trump goes after DACA now

Fauci warns of 'anti-science bias' being a problem in US

- Live updates: Florida, Texas and Arizona saw their highest single-day increases in new cases
- Florida has the makings of the next epicenter, experts say
- Sheriff who wouldn't enforce stay-at-home order tests positive

Analysis: Here's proof Trump knows he's losing

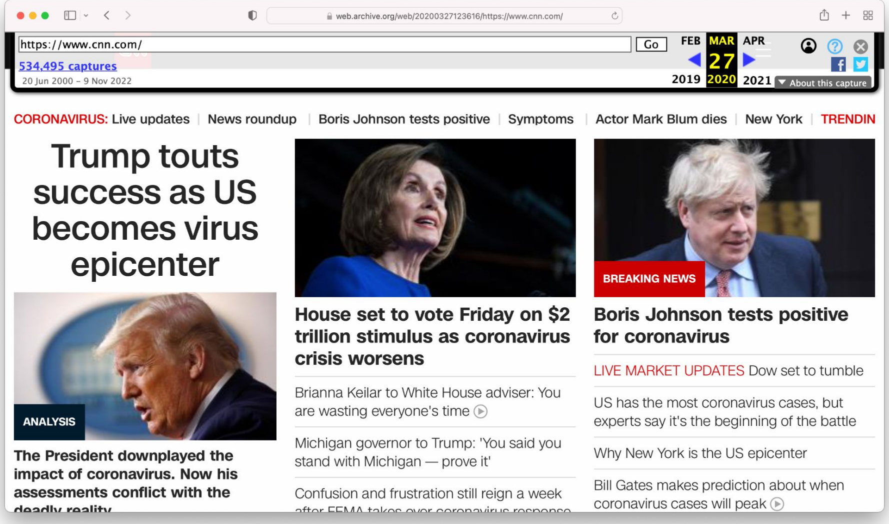
Senate GOP defiant over decision to block Bolton from testifying

Analysis: Blindspots and turncoats — How Trump's mostly white casting has backfired

Voter: If you don't go vote, you're voting for Trump

Washington Post: Senior State Department official resigns over Trump's response to nationwide unrest

Hero story content from Jun 18, 2020



CORONAVIRUS: Live updates | News roundup | Boris Johnson tests positive | Symptoms | Actor Mark Blum dies | New York | TRENDIN

Trump touts success as US becomes virus epicenter

ANALYSIS

The President downplayed the impact of coronavirus. Now his assessments conflict with the

House set to vote Friday on $2 trillion stimulus as coronavirus crisis worsens

Brianna Keilar to White House adviser: You are wasting everyone's time

Michigan governor to Trump: 'You said you stand with Michigan — prove it'

Confusion and frustration still reign a week

BREAKING NEWS

Boris Johnson tests positive for coronavirus

LIVE MARKET UPDATES Dow set to tumble

US has the most coronavirus cases, but experts say it's the beginning of the battle

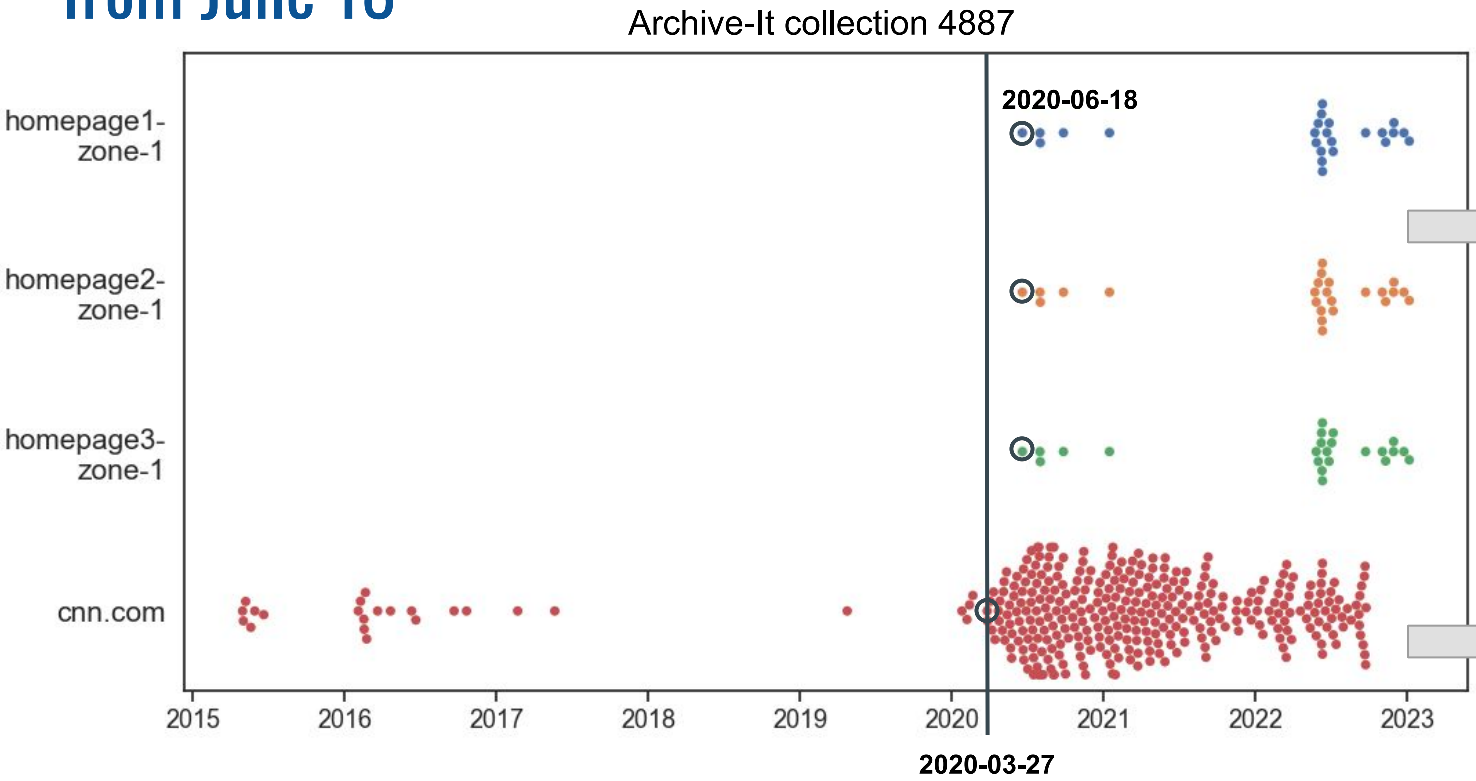
Why New York is the US epicenter

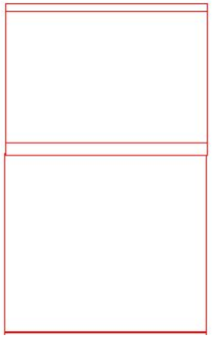
Bill Gates makes prediction about when coronavirus cases will peak

https://wayback.archive-it.org/4887/**20200327123616**/https://www.cnn.com/

https://web.archive.org/web/**20200327123616**/https://www.cnn.com/

March 27 replay contains placeholder from March 27 and content from June 18





Recommendations

- Use browser-based crawling for websites with heavy CSR, like CNN.com
- Develop crawling methods that can detect when CSR being used, nominate sites for browser-based crawls
- Be aware of implications of mixing mementos from conventional and browser-based crawls, audit replay of previous captures
- Develop replay methods that can highlight temporal violations on the page

- Potential mitigations
  - redact mementos between Sep 2015-Nov 2016 with temporal violations in the Hero story
  - increase the MIME types/file extensions included in "About this Capture"
  - include alerts for (or don't load) JSON files with a temporal gap greater than some threshold
  - allow Archive-It collections to borrow certain resources (like JSON) from other collections

Summary

- Since 2015, CNN.com has used client-side rendering (CSR) to build its front page
- CNN.com's CSR model conflicts with standard web archiving models (i.e., load content from closest memento)
- For webpages built with CSR (not only CNN.com), the mix of captures from conventional (Heritrix) and browser-based crawlers can result significant temporal violations upon replay
- We found over 15,000 mementos of CNN.com in the Internet Archive's Wayback Machine with temporal violations greater than 2 days
- We have documented the issue, but more work should be done in detecting these types of problems in real-time