

# Taking the witness stand: attestation and authentication for web archive evidence

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

[Starling Lab](#) Web-Archiving /  
Ukraine Accountability Workshop  
25 August 2022

"The Modern Court" by Michael Foley under [CC BY-NC-ND 2.0](#)

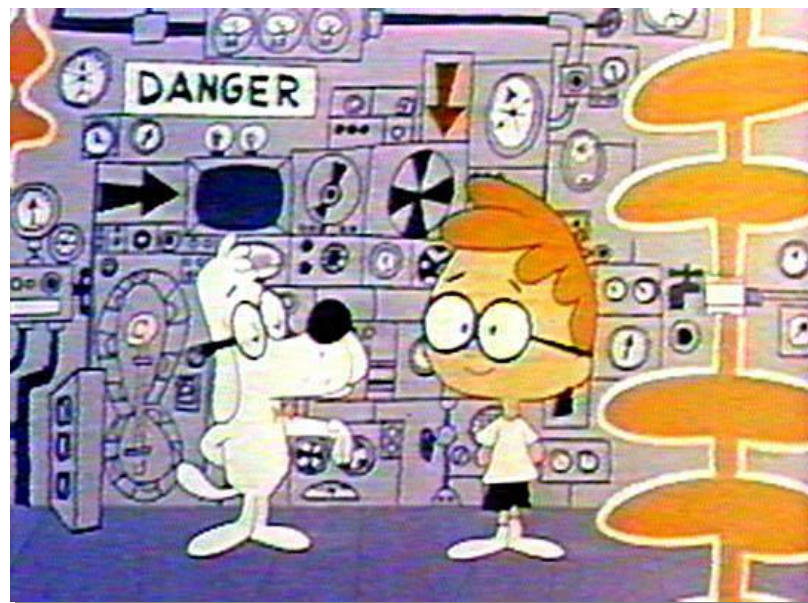
LEGAL PROFESSIONALS DON'T  
UNDERSTAND WEB ARCHIVES  
WELL ENOUGH FOR THAT TO BE A  
STABLE BASIS FOR WHAT SHOULD  
BE EVIDENTIARY BEST PRACTICE.



# evidentiary use of the Internet Archive Wayback Machine (IAWM)

- increasing use in litigation, since 2002
- breadth of arbitrary web content is great
- inconsistently admitted, however
- generalizability of precedents uncertain

INTERNET ARCHIVE  
**WayBackMachine**





# authentication + admissibility

- admissibility turns on reliability assessment
- IAWM reliability concerns:
  - disclaimers
  - archival incompleteness
  - temporal coherence
  - mixed provenance
- concerns raised generally vice specifically



"Dollar" by Images Money 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
- more focus needed on interpretation
- seems to be consistent w/ IA's perspective



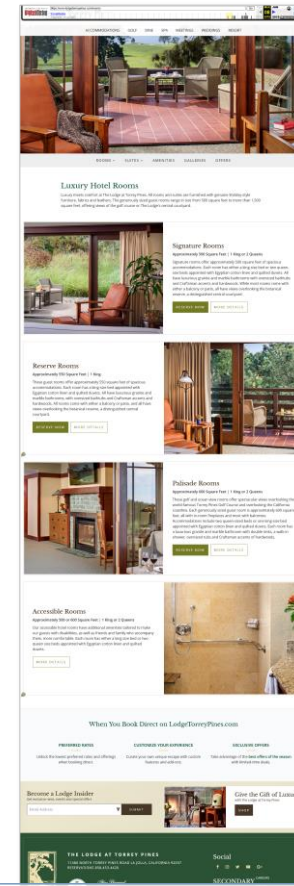
"Wrong Way Around" by [Colin and Sarah Northway](#) under [CC BY 2.0](#)

# 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”

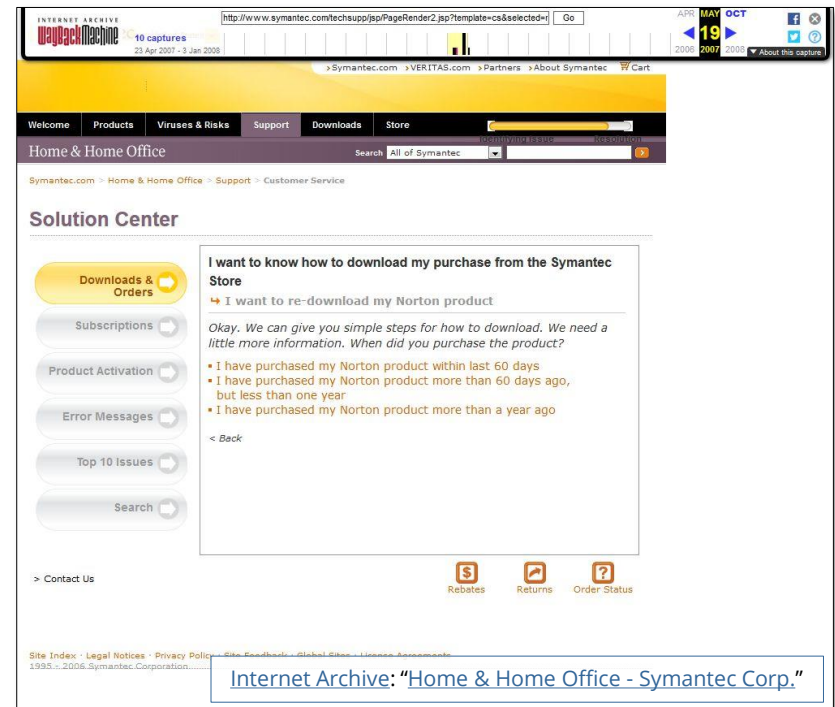


# Khoday v. Symantec Corp., et al. (2015)

- IAWM used to show probable user navigation paths to re-download software over period of time
- IAWM evidence admitted

***"Expert testimony that merely repeats information capable of easy comprehension by a jury is excludable...The Court finds that Taylor offers conclusions not readily available to the jury without the assistance of expert testimony. Whether legitimate alternatives for redownloading Norton products were available, at no cost, to customers during the relevant time period is central to the Plaintiffs' claims, making Taylor's proffered testimony highly relevant."***

[Khoday et al v. Symantec Corp. et al, 93 F.Supp.3d 1067, at \\*1084 \(D. Minn. 2015\)](#)

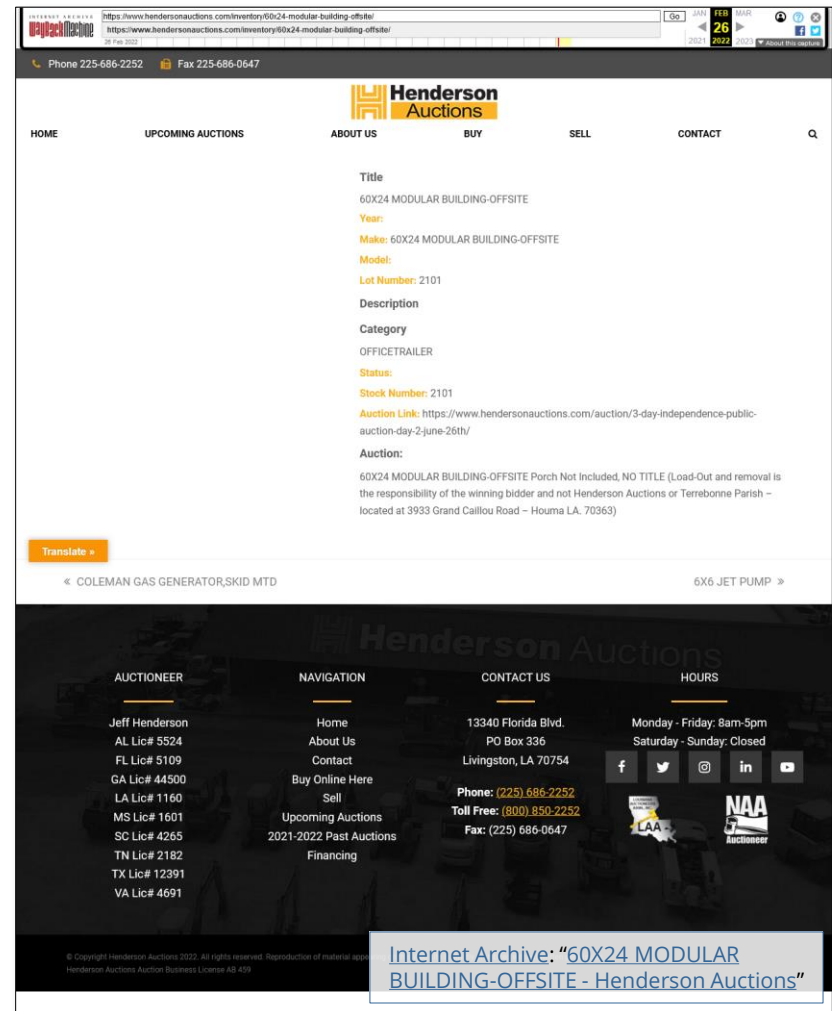


# Weinhoffer v. Davie Shoring, Inc. (2022)

- IAWM used to show online auction terms of sale
- district court's admission of IAWM evidence via judicial notice reversed

*"Here, there was no testimony to authenticate the archived webpage. Our sister circuits' decisions that the Wayback Machine is not self-authenticating are persuasive in the context of judicial notice. In sum, **the district court erred in taking judicial notice of the terms because a private internet archive falls short of being a source whose accuracy cannot reasonably be questioned** as required by Rule 201."*

*Weinhoffer v. Davie Shoring, Inc., No. 20-30568, at \*7 (5th Cir. Jan. 20, 2022)*

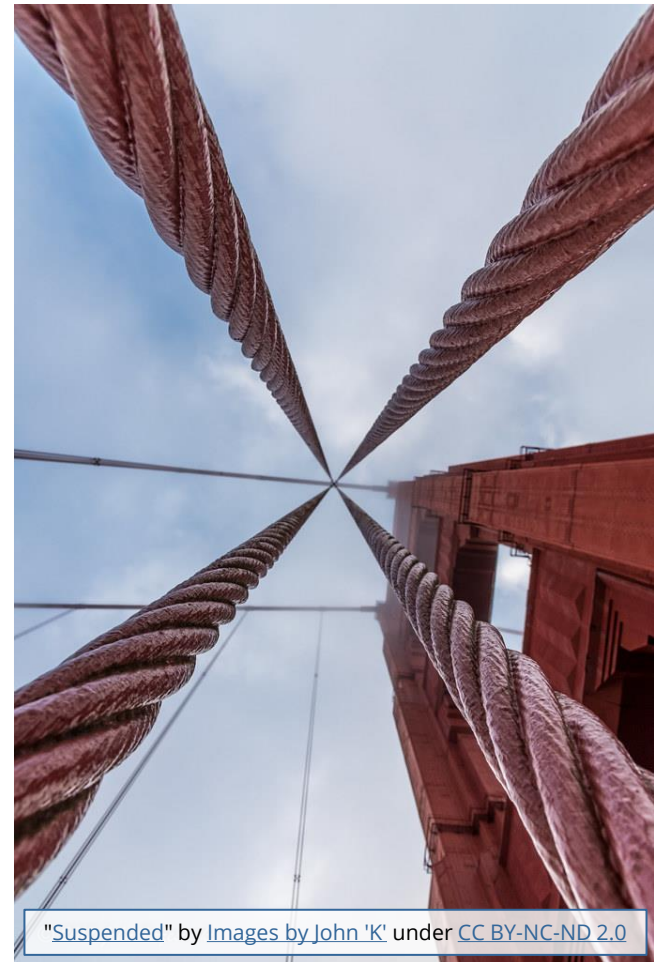


Internet Archive: "60X24 MODULAR BUILDING-OFFSITE - Henderson Auctions"



# 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





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