

*Think, before you link*  
**Recent research and guidance on link rot  
risk**

DP Clinic June 2026

Garth Stewart

22 June 2026

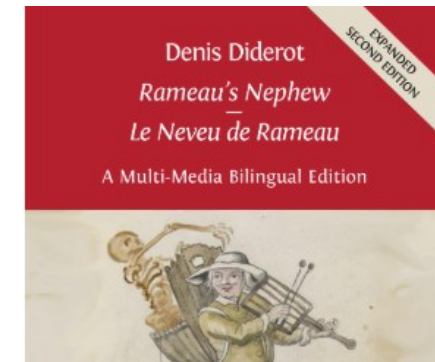


Digital**Preservation**Coalition

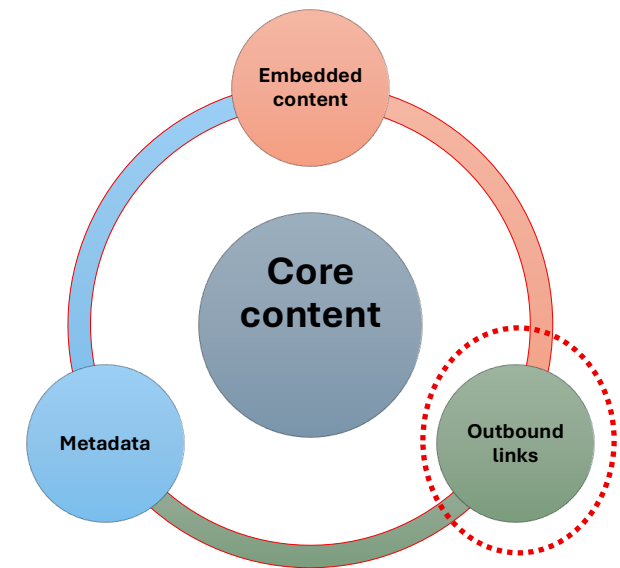


# The challenge

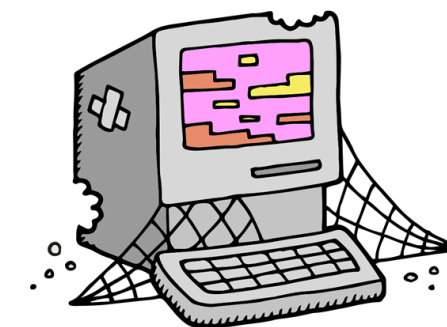
- Web links are an essential (*de facto?*) means for linking and sharing digital information
- User (client) → URL (link) → content served to user via HTTP → plays in browser → happy user 😊
- BUT links very prone to break, or ‘rot’, over time



*“the median lifespan of a URL is now 2.3 years” (Nelson)*



# Defining link rot



‘When a hyperlink breaks, it means that when a user client attempts to access that information via its URL, they will receive an error message indicating that the information is not available...

...the authority to the website has moved or changed the location of that resource, and now we can't access it from that original location. ‘Link rot’ has occurred.

User (client) → URL → content not found or irrelevant content served to user via HTTP → error message or wrong content → unhappy user ☹️

# Two main effects on access

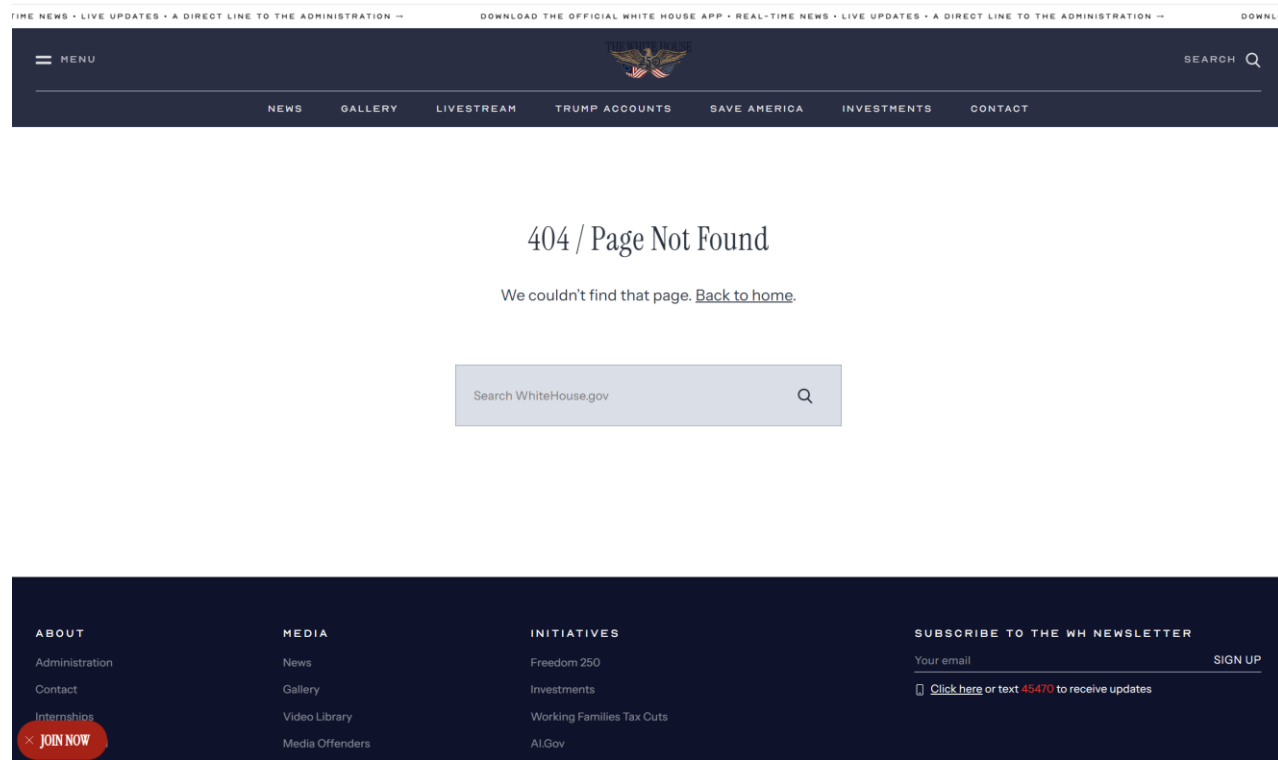
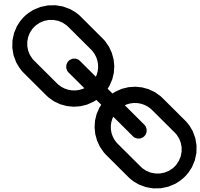
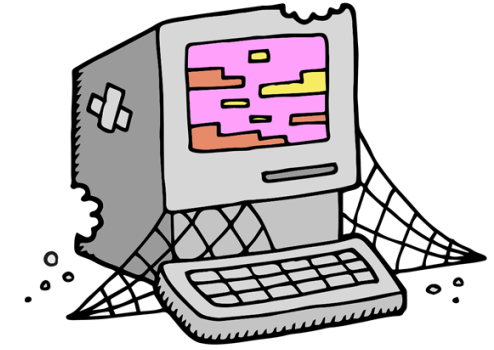


Figure 6: screenshot showing where original URL for a former poetry website now leads to gambling content - evidence of 'cyber squatting'

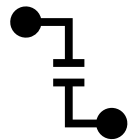
# User trust becomes a casualty

‘When someone follows a link and it breaks, they generally lose confidence in the owner of the server. They are also frustrated – emotionally and practically from accomplishing their goal.’  
(Berners-Lee, 1998)

# The times we live in: “a petri dish for link rot”



## Link rot causes

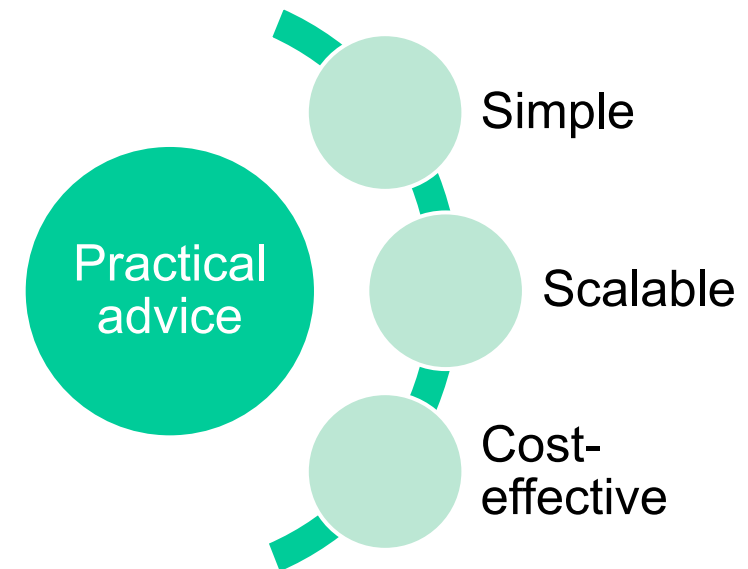
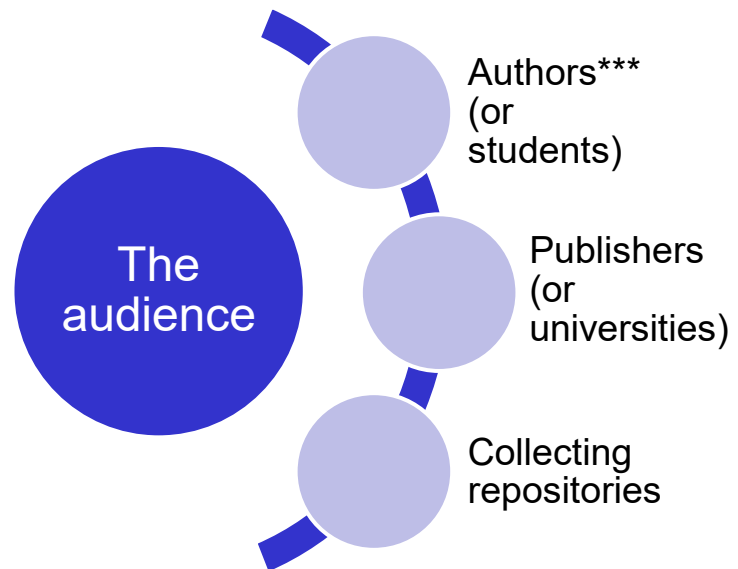


- Organizational change, rebrand or closure
- Neglect
- Content or website superseded
- Cyber attack and server failure
- Bots traffic (particularly for AI LLM crawlers)
- Dependency on third party platforms
- Politics and geopolitics
- Human error e.g. redirection incorrectly setup
- Privacy, data protection, and copyright
- Imposition of paywalls and/or sign in to content once freely open



- Fragile websites
- Forgotten resources
- Frustrated users
- Link rot is (almost?) inevitable...
- Tricky for web archiving

# Building resilience against link rot



# Practical Steps



## Step 1: Think before you link, determine ‘acceptable loss’

- Checklists for stakeholders: risk-based approach, evidential value, user experience, is the content archived already?
- Informational value assessment
- Package associated content locally if you can

## Step 2: Cite resilient links

- What makes a resilient link?: short, simple, self-explanatory, sourced
- Link to web archive URLs
- URL health checks pre-publication

# Practical Steps



## Step 3: Describe more, be accessible

- Provide simple text descriptions of linked content, especially AV → doing anything will be cheaper than link rot repair

## Step 4: Archive your links

- Use community tools such as Save Page Now and ArchiveWeb.Page

## Step 5: Monitor your links, do digital archaeology

- Check links and dip into web archive collections if link rot occurs

# Triage Table:

## Prioritise where and how to build resilience against link rot

	Critical links	Important links	Non-important links
Remain in situ within the eBook?	Yes, but augment with web archive URL	Yes, but augment with web archive URL if available	Yes/No – see Step 1 for further guidance
If available, package and preserve available files?	Yes	Possibly – assess value and resourcing	No
Add descriptive metadata?	Yes – as much as possible, total content if possible	Yes – as much as possible	If resource permits, but keep minimal
Create a web archive copy for online access?	Yes – ArchivePage and Browsertrix	Yes – ArchivePage, Browsertrix if necessary	No
Create a WARC or WACZ file for offline access?	Yes – SavePageNow and Browsertrix	Possibly – assess value and resourcing	No
Monitor over time?	Yes	Yes	No

Thank you!  
Any questions?  
How could we take this work further?

