Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory
by Darryl Mead

Abstract
This case study uses routine activity theory to contextualise the method used by an external bad actor to create fake links within the Internet Archive for the Web site Yourbrainonporn.com. It then discusses the social media campaign which occurred two years later using screenshots of these fake links accessed via the Wayback Machine to defame the site owner. An organised disinformation campaign on social media began attacking the site owner of Yourbrainonporn.com (a pornography recovery Web site) for allegedly, accidentally, posting evidence on his own site of him searching for and hosting hardcore pornography. In fact, the list of purportedly incriminating links did not point to any content, but the defamers’ intentions seemed to have always been to set up a smear campaign against a particular site and its author. Options are discussed for the Internet Archive to provide improved guardianship and to educate the public to minimise harm from this type of social media attack based on screenshots of fake URLs.

Contents
1. Introduction
2. Materials and methods
3. Study
4. Discussion
5. Conclusion

1. Introduction

1.1. The Internet Archive

The Internet Archive is vast in scale. It was established in 1996 and its Wayback Machine has now archived over 815 billion pages from the World Wide Web, making it the largest archive ever created (Internet Archive, 2023). Along with its huge scale has come a reputation for reliability. If you can find material in the Internet Archive, then people are likely to accept that at some point in the past it appeared on the Web. Public access to the Internet Archive is often provided through the Wayback Machine, a software interface launched in October 2001 (Wikipedia, 2021b).

1.2. Routine activity theory

Inspired by the application of routine activity theory (Cohen and Felson, 1979) to crime prevention in general
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

(Choi, 2008; Kleemans, *et al.*, 2012; Leukfeldt and Yar, 2016), and to cyberbullying in particular (Arntfield, 2015; Reyns, 2017; Kaluarachchi, *et al.*, 2020), this paper offers a case study of a pornography industry-connected cyber-attack. In this example a bad actor had created the illusion of content being archived through the Wayback Machine, which later served as the basis of what appeared to be a sophisticated disinformation campaign calculated to damage the reputation of a high-profile Web site and its owner. The social engineering-style cyberattack (Imperva, 2022) allowed archived addresses, without accompanying content, to be employed as the basis of an intentional, social media smear campaign. The forensics involved in responding to the campaign are discussed and relevant data are analysed.

Routine activity theory is a criminological theory of predatory victimisation that explains how patterns of routine activities create opportunities for criminal activity. For a crime to occur, according to this theory, there must be three elements: a motivated offender, a suitable target, and the absence of capable guardians (Schreck, 2017; Schaefer, 2021). The theory posits that crime is more likely to occur when these three elements converge in time and space. The space can be real or virtual.

A motivated offender is an individual who is not only capable of committing criminal activity but is also willing to do so. The phrase ‘suitable target’ was made deliberately wide, so that it encompasses both people and their property — real and virtual. It can also refer to the vulnerability of people to becoming a target. Capable guardians are individuals, groups or systems whose presence or proximity discourages offenders from committing crime. Capable guardians, as a category, can also refer to intentionally implemented characteristics or systems that “harden targets,” that is, environmental (real and virtual) adjustments that raise the cost and difficulty of specific victimization targets. These intentional characteristics or systems are deterrent in nature.

Routine activity theory looks at crime from an offender’s point of view, and a crime will only be committed if a likely offender thinks that a target is suitable and a capable guardian is absent or relatively and episodically incapable, and the cost of the effort outweighs the benefit to the offender as a rational actor.

1.3. Routine activity theory as a lens on the pornography industry

At present, commercial pornographers operate an industry within a very weak international regulatory framework. Numerous companies and individuals within the pornography industry have been criticised for engaging in unsavoury conduct, including but not limited to sexual assault, human trafficking, sexual coercion, distribution of child sexual abuse material, distribution of revenge porn and other non-consensual content and making pornographic material readily accessible to children (Grant, 2015; Kristof, 2020; Newman, 2023; Ladden-Hall, 2022; Ritter, 2021; Belanger, 2023; Debusmann, 2023; Associated Press, 2023).

Before the Web, the marketplace for online pornography was negligible. By 1998 the online pornography market had grown to nine million consumers per year (Cooper, *et al.*, 2000). Jump forward 25 years and the number of people accessing pornography over the Internet is now billions every month (Website SEO Checker, 2023). The industry is highly profitable, but potentially fragile. It may be vulnerable to a wide range of legal, regulatory and social challenges. Ever since the arrival of *Playboy* in the 1950s the industry has used the full public relations toolkit to overcome all obstacles to sustained growth and profit (Jacquet, 2022; Oreskes and Conway, 2012).

In the attack chronicled in this case study, associates and activist supporters of the pornography industry appear to have been acting in the role of a motivated offender.

When motivated offenders within the industry began to campaign against a range of potential threats, they actually had quite a limited range of suitable targets with whom they could engage. There was an established movement which had long opposed pornography use on religious, moral, political or anti-exploitation grounds (National Center on Sexual Exploitation [NCOSE], 2023). Equally vocal have been the second-wave feminists, who since the 1970s opposed pornography on, generally speaking, grounds based in its lack of gender equality and the promotion of attitudes encouraging violence against women (Dines, 2010).

Shortly after the arrival of the smart phone in 2007, a new movement of voices questioning the desirability of pornography consumption appeared from consumers themselves. In setting up the Yourbrainonporn.com Web site in 2010, Gary Wilson (1956–2021) became a leader in documenting research on physical health and mental well-being issues which accompanied unlimited access to free, streaming Internet pornography. As
Yourbrainonporn.com began to build a substantial user base, it moved onto the radar of pornography industry supporters, and other individuals who wished to suppress or otherwise undermine the research and health-based messages propagated by Mr. Wilson. From 2013, Gary Wilson became a suitable target, both as a person and as a Web site. Over a period of eight years Wilson was subject to a wide, varied and sustained range of aggression from pornography industry associates and supporters. These included false reports to law enforcement agencies, unfounded accusations of academic misconduct, social media attacks, trademark and copyright infringements, a baseless restraining order request (which a judge promptly dismissed; the request had been filed by one involved in the Internet Archive attack), and a variety of de-platforming attempts (Yourbrainonporn.com, 2021d).

This paper focuses on an unusual and sophisticated attack of a type that has not been previously reported in the literature. Mr. Wilson’s importance and significance as an industry suitable target is emphasised by the fact that multiple individuals worked over a period of years in an attempt to fundamentally undermine his credibility. The attack was an attempt to reduce the impact that Mr. Wilson was having in illuminating research on the physical and mental health implications of consumers interacting with the pornographic industry’s products.

“If crime needs the convergence in time and space of a motivated offender, a suitable target, and absence of a capable guardian (the ‘crime triangle’), this means that crime can be prevented by keeping motivated offenders away from suitable targets at specific points in time and space or by increasing the presence of capable guardians.” [1]

This paper concludes by examining the opportunities available to the Internet Archive, and the Wayback Machine to better adopt the role of capable guardian to minimise or mitigate future cyberattacks of this type.

Ireland (2021) looked at the concept of capable guardians working on “target hardening”, highlighting what measures and technologies that Internet users might adopt to minimize their likelihood of future victimization. The same concept applies to the technology and systems deployed by the Internet Archive. The Internet Archive’s managers can use an understanding of the types of potential attacks that the Wayback Machine can be targeted with to develop prevention and mitigation strategies.

The Internet Archive could decrease the potential risk or impact of the type of cyberattack reported in this case study through incremental improvements to the interface of the Wayback Machine. Ideally in the future, the type of intentional campaign that seems to have occurred in this study could be prevented as a part of the standard ingestion or maintenance processes employed by Archive.org. Well before submitting this paper, these vulnerabilities were disclosed in writing to the Wayback Machine’s managers. Changes to operating procedures at the Internet Archive noted later in this paper may (or may not) have partly been influenced by this dialogue with the Internet Archive.

1.4. State of research

Academic literature on the Wayback Machine used in this paper was sourced through keyword searches for ‘Internet Archive’ and ‘Wayback Machine’ in Google Scholar, Scopus, the Web of Knowledge and Semantic Scholar. Given the huge cultural importance of the Internet Archive, there are surprisingly few research articles focusing on the organisation and its software.

The Internet Archive has established itself as a credible source of information, giving consumers access to the past (ALNoamany, et al., 2014), which can have a large influence on emerging present day news cycles (Lepore, 2015). The role of Internet archivists has been examined using ethnographic approaches (Ogden, et al., 2017). Quality and veracity of the archival record of the Web has been discussed, highlighting its strengths and limitations (Nguyen and Weber, 2015; Gaffney and Matias, 2018; Ogden, et al., 2023).

Murphy, et al. (2007) discussed the potential of the Wayback Machine as a tool for investigating validity across the fields of e-commerce, information systems and computer-mediated communication.

As a disinterested third-party with extensive independent record keeping, the Internet Archive has become a source of legal evidence. Eltgroth (2009) looked at the potential for the Wayback Machine to provide
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

The Internet Archive has established a role for its content to be admissible as legal documents, provided:

“the original page accurately represented material placed on the originating site by the site’s owner or operator. With this additional step, reliable evidence from the Wayback Machine can become as easily admitted as any other Internet-derived proof.” [2]

The Internet Archive has been used to determine dates establishing precedent in patent litigation and intellectual property rights law. Evidence from the Wayback Machine is widely accepted in courts around the world (Pearce and Charlton, 2009; Quarles and Crudo, 2014).

Lerner, et al. (2017) were among the first to recognize the technical vulnerability of the Internet Archive’s Wayback Machine for a range of attacks. Their academic advice on archival safety was published after the likely attackers in this study had already set their trap. Lerner, et al. (2017) focused on mapping the incentives and techniques available to those who desired to maliciously manipulate content within the site. They asked: “How might attackers attempt to rewrite history? How might they intentionally cause clients who view the archive to see archived websites with content, appearance, and behavior that are different from the actual website at archival time?” [3]. Their work demonstrated potential vulnerabilities rather than providing proof that those vulnerabilities had been used to create disinformation. They emphasised that they were not aware of any attacks like the ones hypothesised in their paper being used in practice. They addressed attacks requesting that the Archive capture specific URLs [4]. The particular vulnerabilities of sites using only HTTP, rather than encrypted HTTPS, were also highlighted (Stock, et al., 1997). At the time of the campaign, the Web site in this study was still on HTTP, though there is no evidence to suggest that that created a particular vulnerability to the type of attack described. As unencrypted HTTP is now largely considered defunct; it is not relevant to this story (Radić, 2023). Archive.org is now employing encryption for all content on its public-facing site.

All of the interventions covered by Lerner and her colleagues ultimately involved changing content within the Archive, using a wide range of techniques to rewrite history. They described attacks where external actors planted material for later use [5]. In this case study, an unknown party injected URLs that all resolved to “404 Page not found”. Then, years later, malicious defamers (who may or may not have been the link-placers) used images of the URL lists as the basis of their external hostile social media campaign. The great majority of consumers of Internet content do not check the basis of every claim that they encounter, making it very likely that this claim would be accepted at face value. It was a lightweight but effective rewriting of history without having to inject substantive content into the Internet Archive. The presence of substantive content was simply inferred. This did not make it any less effective as the basis of a defamation campaign.

Acker and Chaiet (2020), focusing on the COVID-19 pandemic, specifically discussed the ways that the Internet Archive’s Wayback Machine was being weaponized to propagate and preserve health misinformation circulating on platforms like Facebook and Twitter/X. The attack on Gary Wilson predated the creation of disinformation included in their data, but their wider perspectives on the role of the Wayback Machine are relevant and insightful (Acker and Chaiet, 2020).

Thanks to a boom in “fake news” accusations, misinformation and digitally-spread disinformation over the past decade, fact-checking is a growing business on the Internet. The Internet Archive is actively engaging with at least nine leading fact-checking organisations (Graham, 2020; Binder, 2020). However, attention generally falls on major news providers, political movements, pundits and other material judged to be of high value, or substantial significance to society, or to the Internet Archive. There are also smaller fact checking Web sites, such as Snopes.com, but to date they have not shown any interest in engaging with pornography-industry stories beyond those covered by mainstream media (Snopes.com, 2023).

The emergence of ChatGPT and Bard in 2023, and the prediction by many, including Bill Gates, that their emergence was of profound and still developing consequence (Gerken, 2023), produced problems for the Internet Archive. Yet it also offered potential solutions, such as creating effective AI tools to scan and fact-check ingested material. OpenAI utilises information gathered from the open Web to train AI tools, including ChatGPT (Brown, et al., 2020). As such, online astroturfing in the form of smear campaigns has an ability to feed popular
AI tools with disinformation.

As an experiment, I asked ChatGPT about an aspect of the attack discussed in this paper (see section 3.2.3). ChatGPT’s answer contained fabricated information, calling to mind the account of George Washington University law professor Jonathan Turley, regarding whom ChatGPT “hallucinated” (Ji, et al., 2023) an accusation of sexual harassment (Weiss, 2023).

In reality, it is early for the fact-checking process to operate at scale; only a very tiny percentage of material being added to archive.org will be subject to cross-checking and labelling. At the level of overall digital information generation, this makes sense, but it exposes smaller Web sites and their participants to a greater risk of malicious activity by bad actors (Takata, et al., 2018).

Untrustworthy Web sites are now common (Mintal, et al., 2021). An ecosystem of blacklists has grown up, helping Internet security professionals focus their energies on obvious threats in their field (Zhao, et al., 2019). Separately, many bad actors are now using mechanisms designed to allow users to report sites that they disagree with as ‘unreliable sites’, leading to shadow bans and reduced prominence in search results. Research has also shown that the Wayback Machine can be used in a positive way to construct better maps of the information universe (Ben-David and Amram, 2018).

Twitter (relaunched as Elon Musk’s X) has been historically used for propaganda, cyberattacks and disinformation campaigns on an industrial scale (Bastos and Farkas, 2019). Until 2022 the company employed mechanisms such as its Trust and Safety Council in a purported effort to reduce bad-faith conduct (BBC.com, 2022). In 2022, the company transitioned to ownership by Elon Musk. The dynamic pace of change under Musk placed an analysis of Twitter policies and processes for damping down disinformation beyond the scope of this paper.

As a registered, non-profit organisation focused on the preservation of the ever-changing Internet, the Internet Archive could be considered as well positioned to create more tools to combat online disinformation. To undertake to do so in an effective way would require a strategic choice within archive.org. Success would depend on accessing sufficient funding and talent to keep up with developments in AI across text-based and multimedia content.

2. Materials and methods

This case study is based around an incident that began with the archiving of over 100 “fake URLs” (URLs that have never resolved to any existing content) with evocative paths suggesting they linked to pornography, all purportedly to be found on Yourbrainonporn.com — a Web site dedicated to recovery from problematic pornography use. The Internet Archive first collected some of them at least as early as 3 March 2016. New URLs in the sequence were then intermittently registered in the Archive, with the latest being collected on 28 February 2017. In August 2019 screenshots listing the fictional URLs were used in an intentional, hostile smear campaign. I look at the process from the point of view of a Web site under attack. Subsequent analysis indicated the identity of at least some of the apparent bad actors, but there was not enough information to identify, definitively, all the potentially involved motivated offenders.

Background material for the study appears in Appendices A to E. To collect the data a search was made on the Wayback Machine at https://web.archive.org/web/*/yourbrainonporn.com/*.

Appendix A — Snapshots of Yourbrainonporn.com from the Wayback Machine — provides screenshots of all URLs including the element ‘//’ immediately after ‘.com’, which is prominent in all the fake URLs. This element automatically sorts in sequence ahead of alpha-numeric characters. Appendix B — Distinguishing fake links from real links on Yourbrainonporn.com — then discusses the analysis to distinguish between legitimate and fake URLs containing the element ‘//’ after the ‘lsquo;com’. Appendix C — Yourbrainonporn.com Fake Mormon URLs — describes the process used to take the data from the live screens in Appendix A and process it in a MS Excel spreadsheet to allow deeper analysis. Appendix D — Screenshots of the ‘Real Your Brain on Porn’ Twitter account —
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

Appendix E — Spreading disinformation in other publicly visible places — provides examples of peripheral elements of the cyberbullying campaign.

3. Study

3.1. The target Web site

The target site of the disinformation campaign was https://yourbrainonporn.com. It was created in 2010 by author Gary Wilson, who had taught anatomy, physiology and pathology for many years at vocational schools, as well as anatomy and physiology labs at Southern Oregon University (Cowell, 2013).

The Web site mapped out the interaction between consumption of pornography from the Internet and its possible effects on physical and mental health. This was done by reference to academic research and through reports of users and former users of pornography. By the time of Mr. Wilson’s death in May 2021, the site had grown to more than 12,000 pages and cited over 900 peer-reviewed studies. It attracts a wide audience, currently receiving about 4.75 million users per year, for a global traffic ranking of #32,880 (SimilarWeb, 2022a). As public visibility of the site rose, its creator became a target of sustained personal and academic attacks from individuals who did not agree with Wilson’s evidence-based approach revealing the risks of Internet pornography use. The apparent campaign documented in this study can be seen in the context of a much wider programme of pushback against many organisations and individuals who suggest that there are potential risks associated with the use of digital pornography.

Gary Wilson became a suitable target for pushback, receiving attacks from many angles in a sustained and complex campaign to undermine his credibility (Hess, 2022). This included branding him a “pseudoscientist” and falsely accusing him of a wide range of antisocial behaviours ranging from stalking to academic misrepresentation. As a defensive tactic, Mr. Wilson began to document comprehensively many of the attacks on Yourbrainonporn.com (Yourbrainonporn.com, 2021a). Gary Wilson’s status as a suitable target for a pornography-industry-connected actor was further demonstrated by his subsequent success in the Los Angeles County Superior Court on 6 August 2020, which ruled in his favour. The judge determined that a baseless legal filing targeting Wilson was a strategic lawsuit against public participation (SLAPP) (Yourbrainonporn.com, 2020).

In addition to creating Yourbrainonporn.com, in 2012 Gary Wilson gave a TEDx talk in Glasgow, Scotland, called “The great porn experiment” (Wilson, 2012) which at the time of writing had been viewed over 16 million times on YouTube. Building upon this effort, in 2014 Wilson wrote a popular book (Wilson, 2014) and in 2016 he authored a peer-reviewed paper, recommending more research on pornography use (Wilson, 2016).

Also in 2016, Wilson co-authored with seven U.S. Navy doctors another peer-reviewed paper in this field. This paper, Park, et al. (2016) has been widely cited in the academic literature (Scopus lists 86 citations, Web of Science 69 and Google Scholar 234). There had been over 180,800 full text views as of 24 January 2023. Behavioral Sciences lists this as the most viewed paper of all of the 1,626 papers that it has published since the journal was founded in 1996 (MDPI, 2023). However, this success was achieved in the face of sustained efforts by an individual reviewer who attempted to suppress the paper and its authors in a wide range of ways, including repeatedly contacting the Committee on Publication Ethics demanding its retraction and reporting six of the Navy doctors who co-authored it to their medical boards for professional malpractice. The journal’s publisher MDPI resisted these attacks, and subsequently published a small correction where the only material change was to remove the academic editor’s name from the paper (Park, et al., 2018). The same individual who attempted to block Wilson’s paper was a primary individual propagating a social media defamation campaign described in this paper.

The Internet Archive first collected the homepage of Yourbrainonporn.com on 8 December 2010. After a decade, the homepage appears in the Internet Archive more than 1,700 times and in November 2021 more than 330,000
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

3.2. Analysing the attack

3.2.1. Why was the theme of ‘Mormon porn’ chosen as the subject of the Wayback Machine attack?

I believe it is likely that the attackers carefully chose the concept of ‘Mormon pornography’ for the URLs that were screenshot from the Wayback Machine because of its potential for a very high negative impact on Gary Wilson’s reputation, if people believed the campaign to be based on truth. While the field of people opposing unrestricted use of pornography is diverse, some leaders and activists within organisations have a strong religious faith, including members of the Church of Jesus Christ of Latter-day Saints. People within this Church are also often referred to as “Mormons” in popular culture (Weaver, 2018). By contrast, the late Gary Wilson had been an atheist all his life (West, 2018). Creating disinformation that linked Wilson’s purported salacious behaviour to the Latter-day Saint religious faith and practices would potentially be divisive, and perhaps even introduce an element of apparent religious hate speech into Mr. Wilson’s health-focused information service.

“Mormon pornography” is an existing genre, with its own page in Wikipedia (Wikipedia.org, 2021a). An unfiltered Google search for the same term in November 2021 returned just over 9,000 results, along with a warning that “some results may be explicit” (Google.co.uk, 2021). By portraying Wilson as a consumer or purveyor of Mormon porn, the attackers could have believed that such a revelation might have sown distrust and undermined his credibility within the pornography-harm-awareness community.

Themes in the fake links targeted many elements central to the Latter-day Saint faith or culture, including families, motherhood and the Church itself. The fake links included 61 unique URLs incorporating the word ‘Mormon’ as well as references to Utah, the U.S. state with the largest Latter-day Saint population, and to Brigham Young University, the world’s largest LDS-affiliated academic institution. The use of the word ‘Mormon’ itself, rather than ‘LDS’ or other phrases, appears to be controversial within the Latter-day Saint community (Weaver, 2018).

3.2.2. How and when were the links added to the Internet Archive?

There are two probable ways that links could have been added to the Internet Archive. The first is through the ‘Save Page Now’ box on the front page of https://archive.org/web/, which allows a user to submit any URL (real or not) to the Archive. So long as the URL can connect to a server and receive a response, and the site allows crawlers, even if that response is an error, the page will be archived immediately. Ogden, et al. (2023) provides a deep analysis of the ‘Save Page Now’ tool.

With this method, new URLs would have to have been manually added over an 11-month span that they appeared. This, along with the three-year window from the initial fake URL to the social media campaign, strongly implies a deliberate plan rather than an opportunistic attack.

The second way is publishing a new Web page with a list of fake links or publishing a list of fake links to an existing Web page. These “fake links” could have been published online intentionally (for the purpose of creating defamation fodder) or non-intentionally (for example, the very unlikely scenario of a person mistakenly writing “Yourbrainonporn.com” instead of a different intended URL). The Internet Archive uses a Web crawler that regularly scans known Web sites to collect new archival snapshots, and to find URLs for Web pages that have not been previously archived (Kumar, et al., 2017; Mirtaheri, et al., 2014). The crawler might eventually find the published URL list and attempt to archive each new URL that it discovers. As noted in the previous method, the URLs mined this way would be archived even if the URLs discovered led to error pages.

While it’s possible that all URLs were published in a single batch using this method, the 11-month span seems excessive for a crawler to finish processing and archiving a single page of URLs. In my opinion it is more likely that either the page was periodically updated, or that multiple pages of links were published in smaller batches.

3.2.3. Generating the social media storm

This study is based around an incident which began with the creation of fake links in 2016 and evolved into a
full-scale disinformation campaign in 2019. It began with a tweet from the currently-suspended @BrainOnPorn Twitter account associated with the imposter, trademark-infringing Web site RealYourBrainOnPorn.com. The attacker’s Twitter (X) account and a corresponding press release were initially promoted by Pornhub, one of the world’s most popular pornography Web sites (SimilarWeb.com, 2022b).

One thing immediately stands out: the image in Figure D3 showing the tweet that launched the incident depicts the Wayback Machine record of Yourbrainonporn.com. It shows the list of URLs captured. However, the Wayback Machine’s procedure also involves saving a snapshot of a Web site’s HTML and assets (including images) at the URLs that it captures. This detail is crucial. The tweet thread features only a screenshot of the URL list; it does not include any screenshots or links to the implied page content. Nor does it include a URL of the address from which the screenshot was taken (https://web.archive.org/web/*/http://yourbrainonporn.com/*).

Another thing that stands out is that all of the suspect URLs that the Wayback Machine crawled go to “404 Page Not Found” (e.g., https://web.archive.org/web/*/http://www.yourbrainonporn.com/hot-blonde-mormon-feet/). There are at most two or three attempts to crawl each page it before it appears that the Wayback Machine decides that it’s a non-existent URL and terminates the collection process.

Finally, it was strange, for example, that http://www.yourbrainonporn.com//hot-blonde-mormon-feet/ appeared before http://www.yourbrainonporn.com/brain-activation-patterns-associated-cue-reactivity-and-craving-abstinent-problem-gamblers-heavy in the list. URLs are sorted alphabetically, so “brain” should always appear before “hot”.

However, if you look closely at the pornographic URLs in Figures A1–A3, you’ll see that they all have two slashes (//) instead of the usual single slash, between ‘.com’ and the rest of the URL, e.g., http://www.yourbrainonporn.com//hot-blonde-mormon-feet/. First, the presence of the second slash is not an expected address component, making it difficult to notice. It is easy to overlook without careful examination. Second, non-alphanumeric symbols automatically rank higher in most string sorting algorithms. This simple stratagem would ensure that the pornographic URLs appeared legitimate and yet remained near the beginning of the Wayback Machine’s list for Yourbrainonporn.com (which had some 10,000 captured URLs by 20 November 2022). This meant that visitors to the Wayback Machine seeking pages from Yourbrainonporn.com would almost certainly stumble upon these apparently pornographic, but empty, URL addresses. The double slash also grouped the suspect URLs together for easy screenshotting, which was essential for grabbing images to use in any subsequent social media smear campaign. A detailed discussion of the sorting issue can be found in Appendices B and C.

How did so many non-existent pages end up being catalogued by the Wayback Machine? Several possible scenarios are proposed.

The first is that someone created, and perhaps periodically updated, a Web page featuring a list of pornographic URLs. The Wayback Machine’s regular crawling activities then gathered and archived elements of that URL list at different dates.

The second seems less likely, but is also possible. Like most Internet indexing software, the Wayback site allows anyone to submit URLs for its engine to crawl. Those who managed the campaign may have used this functionality to submit carefully crafted, non-existent URLs over multiple months to create the purported records.

Last, while it is conceivable that the Wayback Machine was hacked, or an Internet Archive employee was involved, both of these options seem so unlikely as to be safely dismissed. This is especially so when the other attack opportunities are readily available to any individual on the Internet.

It’s worth noting that while the malicious, now defunct @BrainOnPorn Twitter (X) account claimed to exist in order to share pornography research and to advocate for the health benefits of using pornography, it regularly veered into tweeting hundreds of defamatory statements about Wilson, anyone associated with him, and anyone raising awareness about the potential adverse effects of Internet pornography use. This conduct by the @BrainOnPorn Twitter account operator(s) escalated to posting disparaging or defamatory statements targeting Wilson and others upwards of over 170 times in just one day alone. Some of those unsupported accusations are
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

The unusual feature of this study is not that Wilson was being attacked again, but that the attackers created and presented fresh “evidence” which might suggest Wilson was secretly a pornography user whose private activities were supposedly at odds with his work in the pornography recovery community.

The potential power of this imputed accusation can be considered through the lens of neutralization theory where “acts that violate norms or go against beliefs can carry with them guilt and shame” (Copes, 2014). Neutralization theory subsequently included the concept of ‘drift’, where an individual could indulge in a temporary period of irresponsibility or episodic relief from moral constraint (Matza, 1964).

The attackers implicitly suggested that Mr. Wilson was employing one or more of the techniques of neutralization to provide himself with episodic relief from moral constraint. The unsavoury suggestion was that he was possibly drifting back and forth between delinquent and conventional behaviour (Sykes and Matza, 1957). Sykes and Matza put forward five neutralization techniques: denial of responsibility, denial of injury, denial of victims, appeal to higher loyalties and condemnation of condemners. In this case the most applicable technique of neutralization for potentially undermining Wilson’s credibility was the invocation of ‘condemning the condemner’ (Copes, 2014).

Twitter later disabled the @BrainOnPorn account after it posted personal information about Wilson himself (including his residential address) and members of Wilson’s family (including photographs and financial information). However, the account operator(s) appeared to have created another new Twitter account, @ScienceOfPorn in March 2021. This account subsequently posted negative comments about Gary Wilson in October 2021 (ScienceOfPorn 2021). The corresponding Web site connected to the @BrainOnPorn Twitter handle, RealYourBrainOnPorn.com, was transferred to Gary Wilson as part of a legal settlement after a trademark infringement dispute (U.S. Patent and Trademark Office, 2019).

Returning ChatGPT, the following query was put into the AI tool: “Was the YourBrainOnPorn.com creator accused of hosting any Mormon-themed pornographic material on his website?” The response:

“Gary Wilson, the creator of YourBrainOnPorn.com, was accused of hosting Mormon-themed pornographic material on his website. However, it is important to note that these accusations were made by a former colleague of Wilson’s, and there is no concrete evidence to support the claim. Wilson himself has vehemently denied the accusation, and there is no evidence that he intentionally hosted any sexually explicit content on his website. It is possible that the accusation was made as part of a broader campaign to discredit Wilson and his work on pornography addiction.” [6]

First, ChatGPT users sometimes receive results that differ from one another in response to similar queries. That does not negate the points made below, as explained.

Second, the ChatGPT response above may appear to be phrased relatively neutrally at first glance, but it nevertheless defames Mr. Wilson. Puzzlingly, it ascribes the Mormon-porn accusations to a (fictitious) “former colleague of Wilson’s” rather than to those with a long history of targeting Wilson, and who openly employed the fake URLs in an associated social media defamation campaign. None was a former colleague of Wilson.

The false statement that Wilson was accused by somebody he had met or known, rather than by online trolls, has stronger defamatory imputation. Instead, ChatGPT might have either replied that it was unable to answer, or answered definitively that no evidence demonstrated that Mormon-themed pornography was ever present on Yourbrainonporn.com, without further elaboration.

ChatGPT’s account of the Mormon-porn smear campaign, with or without embellishment, tends to show that the
disinformation seeds of attackers on social media can “sprout” in these AI tools, causing potential harm. The results are then delivered as “fact”, sometimes seasoned with the AI’s inventions \textit{du jour}.

This is a concerning development. Wilson is not the only victim of defamatory statements served up by ChatGPT (Verma and Oremus, 2023; Nelson, 2023; Tifft, 2023). While ChatGPT includes a disclaimer stating that it may output false information, people still use it for a wide variety of applications (Shendruk, 2023). It enjoys a certain amount of public trust and is one of the fastest growing software tools ever created (Hu, 2023).

Commenting on the subject of ChatGPT’s credibility, UCLA law professor Eugene Volokh noted:

“In libel cases, the threshold ‘key inquiry is whether the challenged expression, however labelled by defendant, would reasonably appear to state or imply assertions of objective fact.’ OpenAI has touted ChatGPT as a reliable source of assertions of fact, not just as a source of entertaining nonsense. Its current and future business model rests entirely on ChatGPT’s credibility for producing reasonable accurate summaries of the facts.” (Weiss, 2023)

In short, by not confining itself to fact-based responses, a tool like ChatGPT can unintentionally spread and augment smear campaigns. The fact that ChatGPT’s answers may become more accurate over time does not negate the harm or “evidence” provided by a flawed answer. Experts have sounded the alarm about how the emergence of artificial intelligence could lead to dangerous consequences (Metz, 2023). In this way, tools with the presumption of truth can be employed to distort reality. A recent U.S. federal court case where Chat GPT generated fictitious legal citations led to lawyers being fined by the judge (Mangan, 2023). Similar distortions have plagued Wikipedia (Grabowski and Klein, 2023; Yourbrainonporn.com, 2021c).

4. Discussion

I respect and value the work of the Internet Archive and the Wayback Machine, as do the technical experts with whom I consulted. In this study, the Wayback Machine worked as intended, as the fake “Mormon pornography” URLs did not resolve to any content on Yourbrainonporn.com. However, this campaign used a technique that is relatively unknown. Explaining its nuances to the public is difficult without appearing defensive. This unfortunate reality of online communication limits victims’ options for recourse.

The key thing is the recognition by the attackers of their potential power to wield Internet Archive data as a tool to disrupt the positive impressions that Mr. Wilson had cultivated with his audience over time (Goffman, 1959). It was a deliberate attempt to introduce pseudo-evidence to create stigma. The goal was to degrade his status as a reliable information source [7]. If the attackers succeeded in convincing people that stigma attached to Mr. Wilson was real, it is an “attribute that is deeply discrediting” [8] and it would have a powerful negative impact on his message that pornography use was not risk free.

By acting as a capable guardian, the Internet Archive is in the best position to effectively mitigate these types of attacks and stop them from becoming commonplace. Preventative measures seem called for. Here are some potential avenues that would mitigate the impact of future attacks.

Given that anyone can publish lists of fake URLs online, or anonymously submit any URL to the Wayback Machine, there is nothing pre-emptive that individuals or organizations can do themselves to prevent an attack of this type from occurring. The ingestion page available to the public even has a checkbox that allows the individual requesting the archiving of a page to make 404 errors collectible. This means that it is theoretically possible to create a massive archival record of URLs without content.

Loading a Web page involves submitting a request and receiving a response. The response includes an HTTP
status code with it (W3.org, 2021). The 200-code range (200, 204, etc.) indicated successful request and response. The 400-range indicated that there was an error with the request (e.g., 404 means the requested resource does not exist). The 500-range means the request was correct, but there was an error generating a response.

Legitimate URLs will occasionally return 400 and 500 range responses. The Internet Archive maintains multiple snapshots of a single URL and will therefore encounter these status ranges for some snapshots. However, if all snapshots of a given URL are in the 400 or 500 range, the Wayback Machine could prefix them with a warning emoji, change their typeface or trigger some other visual identifier (e.g., ). The Internet Archive would also want to publish a page or notation explaining what this means in accessible terms. Such a flag would make it easier for victims of this attack to identify and share facts about suspect URLs of this type.

Another potential solution would be to ensure that submitted and crawled URLs resolve to an HTTP status code in the 200 range before allowing them to be archived or appear in a search. This would still allow 4** and 5** status code snapshots to exist if they occurred naturally for valid URLs, while preventing bad actors from having fake URLs archived maliciously. However, this would likely not be ideal because the ability of the Wayback Machine to prove that a page doesn’t exist could be useful in certain situations.

For completeness, I include a third, less preferable option: to automatically place archived URLs that contain only error snapshots at the end of the list. That prevents bad actors from easily abusing the order of sorting records to put fake links at the top of the list, while still allowing content to exist and be searchable. This is not a fully effective solution, as a screenshot of all these at the end would still create a usable list for social media purposes. Malefactors could then mischaracterise the content in any way they please ... unless the links are labelled as ‘404’ errors on the Internet Archive results page.

My principal analysis of material held in the Internet Archive was conducted in November 2021. When I revisited the data within the Internet Archive again in November 2022, there were substantial changes in the visibility of fake links, which may make this type of attack less likely in future. Over the 12-month period between my successive reviews, for the Web site Yourbrainonporn.com the number of archived URLs visible to the public fell from 100,000 to 10,000. The site currently has about 12,400 pages. I manually checked the 10,000 page addresses now listed and noted that the Internet Archive is no longer including any URLs which both include the structure ‘//’ and which then resolved to pages without content. This means that a simple inquiry of the content of ‘Yourbrainonporn.com’ no longer shows Mormon-themed links. However, if the exact addresses of the fake links are put into the Internet Archive as a request, it will still serve them up, as shown in Figures A5 and A6 generated on 21 November 2022.

Finally, in this study, the attackers relied upon individuals not knowing that the existence of a written URL does not prove that the URL is legitimate or actually points to actual content. The Internet Archive and unaffiliated technical professionals could do a better job of communicating with the public. As the Internet is a relatively new tool in the grand scheme of human civilisation, more awareness about its vulnerabilities could help curb the spread of disinformation.

5. Conclusion

Routine activity theory provides a helpful framework for conceptualizing the roles of the motivated offenders, suitable targets and capable guardians in this case study. While motivated offenders remain only vaguely visible, Gary Wilson’s status as a suitable target was confirmed. The need for the Internet Archive to conceive itself as delivering the role of the capable guardian has also been suggested.

The Internet Archive’s credibility can be co-opted to manufacture legitimacy for false and/or misleading claims using simple techniques available to anyone online. There are ways to mitigate and prevent this kind of abuse without sacrificing the transparency or openness of the Internet Archive. Complete solutions require both technical and educational components. However, most of these mitigations can only be effectively implemented
by the Internet Archive itself. Victims of this kind of attack are left with limited options on their own.

Within the ingestion mechanism of the Wayback Machine, there is scope for the identification of ‘//' or similar suspect elements within URLs. This identification could be used to create software features to flag this type of potentially fake link. Ideally they should be flagged as 404 errors.

About the author

Darryl Mead is Chair of The Reward Foundation in Edinburgh, Scotland.
E-mail: darryl [at] rewardfoundation [dot] org

Declaration of interests

This research received no external funding. The late Gary Wilson was the author of Your brain on porn: Internet pornography and the emerging science of addiction, a book described in this paper. The Reward Foundation, a registered non-profit (Scottish Charitable Incorporated Organisation SC044948), receives the author royalties from Gary Wilson’s book. Darryl Mead is now a volunteer at this non-profit. Mead’s wife Mary Sharpe supplies management services to The Reward Foundation on a contract basis. The Reward Foundation has been targeted with various forms of online harassment, defamatory statements, and administrative reports by pornography industry-connected activists. Mead has participated in surveys, mailings, and public consultations on problematic pornography use, government policy for sex education, age verification, digital archiving and other topics. He has provided advice to public bodies, governments and legislative review programmes around the world. He has performed grant reviews for institutions in the U.K. and New Zealand. He has published peer-reviewed articles on pornography use. He has delivered continuing professional education events in Scotland, England, Northern Ireland, Ireland, Croatia, Germany, Turkey, Japan, South Korea, Australia and the U.S.

Notes

6. OpenAI’s ChatGPT 3.5 AI language model, personal communication, 4 April 2023, see Figure E4.

References


B. Debusmann, Jr., 2023. “Pornhub blocks Utah users ahead of age verification law,” *BBC* (2 May), at
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory


Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory


Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory


doi: [https://doi.org/10.1109/BigDataCongress.2015.118](https://doi.org/10.1109/BigDataCongress.2015.118), accessed 18 December 2021.


doi: [https://doi.org/10.1145/3091478.3091506](https://doi.org/10.1145/3091478.3091506), accessed 18 December 2021.


doi: [https://doi.org/10.3390/bs8060055](https://doi.org/10.3390/bs8060055), accessed 18 December 2021.

doi: [https://doi.org/10.3390/bs6030017](https://doi.org/10.3390/bs6030017), accessed 18 December 2021.

D. Pearce and B.G. Charlton. 2009, “Plagiarism of online material may be proven using the Internet Archive Wayback Machine (archive.org),” *Medical Hypotheses*, volume 73, number 6, p. 875.  


doi: [https://doi.org/10.4324/9781315117249](https://doi.org/10.4324/9781315117249), accessed 20 June 2023.

Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory


Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory


Appendix A: Snapshots of Yourbrainonporn.com from the Wayback Machine
This section provides screenshots of all URLs including the element ‘//’ inserted in the suspect URLs immediately after ‘.com’ rather than the standard ‘/’. The second ‘/’ in the element has a higher sort precedence than standard alpha-numeric characters, and it is visually similar enough to ‘/’ as to be difficult to spot upon casual review. This simple ordering trick formed a critical aspect of this apparent attack, allowing all of the bogus URL addresses to remain grouped together. This later facilitated creation of screenshots of pornographic URL lists. It made the search easy to replicate and always produced a list which brought the fake links to the first page of the Internet Archive’s collection of URLs for Yourbrainonporn.com.

Appendix A first presents screenshots of the initial four pages returned for a query in the Wayback Machine for the string https://web.archive.org/web/*/yourbrainonporn.com/*/*. The search was made on 6 November 2021. They include all of the URLs which have been identified for this study as fake links, along with some legitimate content.

It offers the first three pages of snapshots, each of 50 records, labelled as Figures A1 to A3. Figure A4 is the fourth page and includes the next 21 records.

The results are in the default order, sorted by URL, so the first explicitly pornography related URL containing the ‘//’ element was 17th on the list in Figure A1.

 Altogether 131 fake links were identified using methods discussed in Appendix B. Of these 107 containing the word ‘Mormon’ in the URL. The last fake link happens to fall on the last line of page 3. As a result, Figure A4 begins immediately with records showing real pages, starting with pages with addresses beginning with the numeral ‘1’. These have been sorted by the Wayback Machine so that they appear immediately after the records with ‘//’.

Figure A5 shows the re-run search on 21 November, 2022 for the example of http://www.yourbrainonporn.com//closeup-of-tight-mormon-pussy/. This returned the archived image shown in Figure A6 demonstrating that the page ‘could not be found’ when it was added to the Internet Archive on 3 March 2016.

A year later, in November 2022, none of the ‘Mormon porn’ links still appear in the default list of search results. Figure A7 shows the current sequence where all of the URLs containing the structure ‘//’ are still visible. The sequence is identical to that shown in Figure A1 down to the entry for http://www.yourbrainonporn.com//advertising/. The 131 ‘Mormon porn’ links have then been removed by the Internet Archive and the sequence picks up from http://www.yourbrainonporn.com/100-benefits-quitting-pornere.
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

Figure A1.
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

<table>
<thead>
<tr>
<th>URL</th>
<th>NME TYPE</th>
<th>FROM</th>
<th>TO</th>
<th>CAPTURES</th>
<th>DUPLICATES</th>
<th>UNIQUES</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.yourbrainporn.com/hot-married-mormon-girl/">http://www.yourbrainporn.com/hot-married-mormon-girl/</a></td>
<td>text/html</td>
<td>Mar 12, 2016</td>
<td>Mar 17, 2016</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainporn.com/hot-mormon-shower-voyeur-apture/">http://www.yourbrainporn.com/hot-mormon-shower-voyeur-apture/</a></td>
<td>text/html</td>
<td>Feb 12, 2017</td>
<td>Feb 12, 2017</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainporn.com/hot-mormon-shower-voyeur-apture/feed/">http://www.yourbrainporn.com/hot-mormon-shower-voyeur-apture/feed/</a></td>
<td>text/html</td>
<td>Apr 26, 2017</td>
<td>Apr 26, 2017</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainporn.com/hot-teasing-mormon-mill/">http://www.yourbrainporn.com/hot-teasing-mormon-mill/</a></td>
<td>text/html</td>
<td>Apr 11, 2016</td>
<td>Apr 11, 2016</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainporn.com/hot-wife-showing-her-perfect-tits/feed/">http://www.yourbrainporn.com/hot-wife-showing-her-perfect-tits/feed/</a></td>
<td>text/html</td>
<td>Apr 12, 2016</td>
<td>Apr 17, 2016</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainporn.com/married-wife-fingering-pussy/">http://www.yourbrainporn.com/married-wife-fingering-pussy/</a></td>
<td>text/html</td>
<td>Mar 9, 2016</td>
<td>Mar 1, 2016</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure A2.
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

<table>
<thead>
<tr>
<th>URL</th>
<th>Type</th>
<th>Date</th>
<th>Page</th>
<th>Visits</th>
<th>Links</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.yourbrainon.com/nude-mormon-wife-selfie/">http://www.yourbrainon.com/nude-mormon-wife-selfie/</a></td>
<td>text/html</td>
<td>Apr 11, 2015</td>
<td>1</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainon.com/petite-tan-mormon-wife/">http://www.yourbrainon.com/petite-tan-mormon-wife/</a></td>
<td>text/html</td>
<td>Feb 28, 2016</td>
<td>1</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainon.com/petite-tan-mormon-wife/feed/">http://www.yourbrainon.com/petite-tan-mormon-wife/feed/</a></td>
<td>text/html</td>
<td>Feb 28, 2016</td>
<td>1</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainon.com/reai-mormon-pussy-closeup/">http://www.yourbrainon.com/reai-mormon-pussy-closeup/</a></td>
<td>text/html</td>
<td>Mar 25, 2016</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainon.com/removal-request/">http://www.yourbrainon.com/removal-request/</a></td>
<td>text/html</td>
<td>Mar 5, 2016</td>
<td>1</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainon.com/siexy-blende-mormon-selfie/feed/">http://www.yourbrainon.com/siexy-blende-mormon-selfie/feed/</a></td>
<td>text/html</td>
<td>Apr 18, 2016</td>
<td>1</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainon.com/siexy-sweaty-mormon-boobs/">http://www.yourbrainon.com/siexy-sweaty-mormon-boobs/</a></td>
<td>text/html</td>
<td>Feb 28, 2016</td>
<td>1</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainon.com/squishy-mormon-tits/">http://www.yourbrainon.com/squishy-mormon-tits/</a></td>
<td>text/html</td>
<td>May 31, 2016</td>
<td>1</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainon.com/thick-mormon-teen-squeezing-tits/">http://www.yourbrainon.com/thick-mormon-teen-squeezing-tits/</a></td>
<td>text/html</td>
<td>Apr 18, 2016</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><a href="http://www.yourbrainon.com/thick-mormon-teen-squeezing-tits/feed/">http://www.yourbrainon.com/thick-mormon-teen-squeezing-tits/feed/</a></td>
<td>text/html</td>
<td>Apr 18, 2016</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Showing 1 to 15 of 130,000 entries

First Previous 1 2 3 4 5 ... 2000 Next Last

Figure A3.
<table>
<thead>
<tr>
<th>URL</th>
<th>TYPE</th>
<th>FROM</th>
<th>TO</th>
<th>CAPTURES</th>
<th>DUPLICATES</th>
<th>UNIQUES</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.yourbrainerpopcom/100-benefits-quitting-pornere">http://www.yourbrainerpopcom/100-benefits-quitting-pornere</a></td>
<td>unk</td>
<td>May 3, 2021</td>
<td>May 3, 2021</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure A4.
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

Figure A5.

Figure A6.
Appendix B: Distinguishing fake links from real links on Yourbrainonporn.com

To understand the parameters of the attack, I needed to be able to distinguish fake links from the real links held for Yourbrainonporn.com within the Wayback Machine. In this appendix, I report on tests used to distinguish between real links and fake URLs.

Data collected from the Wayback Machine sample at: https://web.archive.org/web/*/yourbrainonporn.com/* was reported in Appendix A. This was the precise URL the attackers quoted in their screenshots. This appendix details the approach employed when analysing this dataset to distinguish between legitimate links having the element ‘//’ after the ‘.com’ and those where the link was a fake.

Links that include the element ‘//’ after the ‘.com’, but which appear to relate to legitimate components of Yourbrainonporn.com were then subject to simple forensic examination. The Internet Archive has collected the 16 URLs at the beginning of the sequence containing http://www.yourbrainonporn.com// which are legitimate. The unusual structure of these URLs has not resulted from insertions by bad actors.

As an example, Figure B1 shows a search within the Wayback Machine for the first ‘//’ example on the list, http://www.yourbrainonporn.com//677/. When tested in the Wayback Machine it returns the following null result.
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

Figure B1: Wayback Machine search result for http://www.yourbrainonporn.com/677/.
These last two URLs were collected more than 60 times each by the Wayback Machine.

The only other ‘//’ URL which does not include the word ‘Mormon’ or some form of reference to pornographic imagery was http://www.yourbrainonporn.com/removal-request/ which again resolves to http://www.yourbrainonporn.com/removal-request/ and does not return a result.

All of the links shown in Figure A4 involve URLs beginning with numerals after the ‘.com’ and were legitimately collected from Yourbrainonporn.com by the Wayback Machine.

Appendix C: Analysing the Yourbrainonporn.com Fake Mormon URLs

A spreadsheet was used to analyse information from the live screens in Appendix A, Figures A1 to A4. The data was processed successively through four different spreadsheets. Processing was carried out in November 2021 using Microsoft Excel for Mac, version 16.54. The analysis can be summarised as follows.

In the first sheet the links were listed as found by the search using the query string https://web.archive.org/web/*/yourbrainonporn.com/*, preserving their original order as they appeared in the Wayback Machine. This sheet was named ‘Links in original order’.

To create the second sheet, this data was duplicated onto a sheet named ‘Sorted by date first archived’. The data transfer process rendered the dates in columns C and D as text. These were then manually converted to machine-readable dates. The records were then sorted by date order when first archived by the Wayback Machine. Based on the analysis in the described in Appendix B, non-pornographic links were identified and removed. The non-pornographic links appeared among the others presumably due to authentic typing errors adding a superfluous ‘/’. The result was 131 suspect fake links.

In the third sheet the records were sorted according to the date of the most recent snapshot created by the Wayback Machine.

Finally, this third sheet was duplicated to create the fourth sheet labelled ‘Mormon links sorted by URL’. Of the 131 links identified as ‘fake’, 55 also include the string ‘/feed/’. These URLs were identified as the crawler explicitly trying to access an RSS feed for each of the non-existent pages. This is likely a side effect of how the Wayback machine crawls WordPress sites and were not added by the attacker. They are simply a by-product generated by the attack mechanism. After eliminating these URLs, the total number of fake links inserted was found to be 76. Of these 61 included the word ‘Mormon’ and several others mentioned Utah or Brigham Young University.

The Wayback Machine reports that the entire body of fake URLs was first archived between 3 March 2016 and 28 February 2017.

With the URLs sorted alphabetically, a simple visual analysis of the writing style strongly suggested that the suspect links were all written as a single batch. This opinion is based on their structure where words and ideas were added successively in cascades to make longer and more complex URLs. Structural elements are repeated again and again. Note: all of these URLs have always gone to “error” (404) pages, indicating that actual pages had never existed. Samples demonstrating this are shown at the end of Appendix D.

Appendix D: Screenshots of the “Real Your Brain on Porn” Twitter account

This appendix provides a series of screenshots chronicling tweets from the Twitter account of the impostor Web site RealYourBrainOnPorn.com. That site utilized Yourbrainonporn.com’s trademarks for over a year (Trademarkia.com, 2022) until Yourbrainonporn.com was awarded RealYourBrainOnPorn.com’s trademark-infringing URL in a legal settlement. One of the RealYourBrainOnPorn.com operators attempted to trademark the exact URL of Gary Wilson’s Yourbrainonporn.com site, which Gary Wilson subsequently opposed with the U.S. Patent and Trademark Office. Eventually, after Gary Wilson’s opposition was filed, the attempt to
trademark Gary Wilson’s URL was abandoned (Wilson, 2020). The Twitter handle of this infringing site was @BrainOnPorn. It was the account that was used to attack Yourbrainonporn.com in August 2019. The @BrainOnPorn account itself was established on 18 April 2019 (Figure D1) as a Twitter handle targeting Yourbrainonporn.com in particular, and the wider pornography-harm awareness movement in general. Twitter eventually banned it for violating Twitter policies.

Immediately after @BrainOnPorn’s launch, it appears that Pornhub was the first account to retweet @BrainOnPorn’s maiden tweet (Figure D2), suggesting a coordinated effort between Pornhub and the Real Your Brain on Porn account. It is possible that Pornhub was implicated in creating the account, or conferred privately with the Twitter account holder(s). Pornhub is one of the world’s three largest commercial pornography suppliers (SimilarWeb.com, 2022b).

Contemporaneous information was gathered during the social media attack in August 2019 when Mr. Wilson and his system administrator made screenshots of many elements of the social media disinformation campaign. A screenshot is supplied that publicises the attacker’s initial ‘discovery’ of the Mormon porn links for Yourbrainonporn.com within the Wayback Machine (Figure D3), along with numerous efforts to boost the spread of this particular piece of disinformation across Twitter (Figures D4–D6, E1) and elsewhere (Figure E3).

Several key assertions were made on the Real Your Brain on Porn account during the attack. In Figure D3 the description of the fake links asserted that “These cannot be backdated”. Another tweet in Figure D3 also claimed that “the links cannot be inserted manually, and cannot be back dated. He got caught with Mormon porn text ... We will make our computer engineer available to verify everything with a complete site scrape for media.” The same Figure also included a tweet from the site owner, falsely claiming that Mr. Wilson “appears on multiple white supremacist shows, now sending death threats”.

To understand, and then refute, the @BrainOnPorn claims that the records appearing in the Wayback Machine “could not be fake”, Mr. Wilson was able to demonstrate that anyone, too, could insert fake links (that went to 404 pages) containing the ‘//’ element after the ‘.com’ for pages preserved by the Internet archive. A detailed description of his method has been publicly documented (Yourbrainonporn.com, 2019).
Figure D1: The initial tweet from Real Your Brain on Porn sent on 18 April 2019.
Figure D2: Immediately after the launch, it appears that Pornhub was first account to “retweet” this (Twitter slang for sharing it with their own Twitter audience), suggesting a coordinated effort between Pornhub and the Real Your Brain on Porn account. Pornhub is among the world’s top three commercial pornography suppliers.

In the five months after this account opened, Real Your Brain on Porn posted a range of content disparaging Mr. Wilson and the authentic Yourbrainonporn.com site.

**The Twitter storm**

The campaign against Yourbrainonporn.com began on Thursday 22 August 2019 and ran into the weekend. Content of the attack as shown in Figures D3 to D6. A linked, parallel campaign on a ‘burner’ Twitter account is shown in Figures E1 and E2.
Figure D3: The attack was launched at 7:08 pm (Pacific time) on 22 August 2019.
This week Mormon porn links were found on @YourBrainOnPorn site. He screamed "hacker", "conspiracy", & "optical illusion". A lesson in crawlers!

Search: web.archive.org/web/*/yourbrain...

Crawlers try to follow anything that looks like a link. Our super porn-free site!

Figure D4: A follow-up tweet the next day, rewording the accusation that Wilson had mistakenly uploaded porn and claiming it “could not be ‘hacker’”.
This:
/yourpornhere/
is format of a relative link on yourbrainonporn.com. Crawler tried to follow. No doubt Wilson, sole author, did not intend to upload his porn, so just text. Hundreds of porn links between 2016/2017 on “young Mormons”. None recent, could not be “hacker”.

Figure D5: Another follow-up tweet the next day, rewording the accusation that Wilson had mistakenly uploaded porn and claiming it “None recent, could not be ‘hacker’”. 
Figure D6: Example of a fake Mormon porn link in the Wayback Machine (accessed 8 November 2021).

Figure D7: However, requesting this content from the Wayback Machine demonstrated that no content had been archived (accessed 8 November 2021).
I tested a sample of the 131 fake links listed in Appendix B. In all cases the Wayback Machine returned a consistent record of ‘Page not found’, a 404 error. This pattern of the archived links lacking content in the archive was repeated in every case.

Appendix E: Spreading disinformation in other publicly visible places

The defamatory messages did not appear only on ‘Real Your Brain on Porn’s’ @BrainOnPorn Twitter account. Another Twitter account (https://twitter.com/CorrectingWils1) appears to have been set up specifically to post the same material. The @CorrectingWils1 account tagged many of the exact same Twitter accounts that @BrainOnPorn was tagging in dozens of similar tweets. Twitter banned the @CorrectingWils1 account after it was reported.
CorrectingWilson @CorrectingWils1 · 1h
Replying to @JohnFOubert
Are you interested in continuing to present @YourBrainOnPorn when "Mormon" porn content is all over his anti-porn website? See for yourself:
web.archive.org/web/*/yourbrain...

Denounce any time you believe what you preach.
thick-mormon-teen-squeezing-t**s
nude-virgin-mormon-collage.png

CorrectingWilson @CorrectingWils1 · 2h
Does @SASH123 still work with the anti-porn activist @YourBrainOnPorn who has Mormon porn links all over his website? See for yourself:
web.archive.org/web/*/yourbrain...

kinky-mormon-girl-in-utah
married-mormon-milf-with-huge-t**ts
molly-mormon-wife-naked
mormon-t**ts-in-garments
And so on

CorrectingWilson @CorrectingWils1 · 2h
Replying to @FightTheNewDrug and @YourBrainOnPorn
Are you prepared to defend the 'Mormon' porn all over Wilson's website? See for yourself:
web.archive.org/web/*/yourbrain...

He said it's 'optical illusion'. It's not. This came directly from his website:
creamy-hairy-mormon-teen-p*ssy
cute-innocent-mormon-teen
cute-young-mormon-butt
**Figure E1**: A ‘burner’ Twitter account set up by a bad actor to spread carefully focused disinformation.
Figure E2: The ‘burner’ Twitter account was reported and Twitter eventually suspended it.
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory

--- Forwarded message ---
From: Real Your Brain On Porn <sciencearousal@gmail.com>
Date: Thu, Aug 22, 2019 at 1:02 PM
Subject: NCOS speaker posts "mormon" p*rn
To: <PUBLIC@ncose.com>

One of you frequent speakers, Gary Wilson, has a public collection of "mormon porn texts" on his website yourbrainonporn.com

You can view the content using this search:
web.archive.org/web/*/yourbrainonporn.com/*
Filter/search for "Mormon" to show the pornography content. He has not provided any reason other than he accidentally uploaded his own porn searches to his anti-porn website.

He may have joined the movement to get access to those girls. He has not denied that the text was on his website, he only claims that it is some conspiracy against him. This is what predators do.

**Figure E3**: Defamation also extended beyond social media to e-mail messages.

---

Was the YourBrainOnPorn.com creator accused of hosting any Mormon-themed pornographic material on his website?

Gary Wilson, the creator of YourBrainOnPorn.com, was accused of hosting Mormon-themed pornographic material on his website. However, it is important to note that these accusations were made by a former colleague of Wilson’s, and there is no concrete evidence to support the claim. Wilson himself has vehemently denied the accusation, and there is no evidence that he intentionally hosted any sexually explicit content on his website. It is possible that the accusation was made as part of a broader campaign to discredit Wilson and his work on pornography addiction.

**Figure E4**: Defamation was provided in ChatGPT 3.5 on 4 April 2023.

---

**Editorial history**

Received 8 March 2023; revised 26 June 2023; revised 24 September 2023; accepted 25 September 2023.

---

This paper is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).
Creating disinformation: Archiving fake links on the Wayback Machine viewed through the lens of routine activity theory
by Darryl Mead.
First Monday, volume 28, number 10 (October 2023).
doi: https://dx.doi.org/10.5210/fm.v28i10.13154