

Uncomfortable Untruths in the Archive:

The Irish Slaves Meme and the Creation of Controversial Social Media Archives

By

Claire O'Halloran

2201194

A thesis submitted to Leiden University in partial fulfilment of the
requirements for the degree of:

MA History: Archival Studies

Supervisor: Professor Marieke Bloembergen



**Universiteit
Leiden**

2019

Table of Contents

Introduction.....	5
Chapter 1: The Irish Slaves Meme.....	13
Chapter 2: Appraisal and Acquisition.....	33
Chapter 3: Harvesting and Preservation.....	45
Chapter 4: Arrangement and Description.....	57
Chapter 5: Ethics.....	69
Conclusion.....	81
References.....	85

List of Figures

Figure 1. Facebook post about Irish slaves being treated worse than others.....	17
Figure 2. Facebook post about Irish slaves being treated worse than others.....	18
Figure 3. Facebook post in reaction to Black Lives Matter.....	19
Figure 4. Facebook post with refrain of 'Irish lives matter.'	20
Figure 5. Facebook post highlighting the 'cover-up' of white slavery with refrain of 'White lives matter.'	21
Figure 6. Instagram post about the forgotten history of Irish slavery.....	22
Figure 7. Tweet shared on Facebook about reparations.....	23
Figure 8. Tweet shared on Facebook about reparations.....	23
Figure 9. Twitter post about reparations.....	24
Figure 10. Facebook post about reparations.....	24
Figure 11. Twitter post about reparations.....	25
Figure 12. Facebook post about the treatment of Irish slaves with 7,500 comments.....	43

Introduction

'The Irish were bred with the African slaves, you know? Even the Irish, we were slaves. At some point, you just have to get over it.' – Confederate flag rally attendee, Mississippi, USA.¹

In recent years, the terms 'fake news' and 'post-truth' have entered into common parlance. The prevalence of social media has done much to aid the spread of conspiracy theories and false or hateful information online. Historian Sophia Rosenfeld has argued that social media platforms have created a 'wild west climate' in which fact and fiction are muddled and popularity rather than reliability dictates the type of information that is presented to users.² The primary focus of this thesis is on the role archivists can play in preserving problematic and false information that has been created and spread on social media. 'Problematic' in this instance can mean controversial, xenophobic, or otherwise bigoted social media content containing false information or conspiracy theories. For instance, this content can include homophobic Tweets, memes spread by white supremacist Facebook pages, or conspiracy theories about Hillary Clinton shared during the 2016 United States presidential campaign.

In order to make my argument clear, this paper will focus on one particular case study of false and controversial information shared on social media: the so-called 'Irish slaves myth' or 'Irish slaves meme'.³ Deeper analysis of the meme and its online development will be given in Chapter 1. However, to briefly summarise, this is a myth which falsely claims that Irish indentured servants were treated as slaves in America and the Caribbean in the seventeenth

¹ Neely Tucker, 'In Mississippi, defenders of state's Confederate-themed flag dig in,' *The Washington Post*, 18 August, 2015. https://www.washingtonpost.com/lifestyle/style/mississippi-resists-calls-to-change-its-confederate-themed-flag/2015/08/18/cd590924-41d3-11e5-846d-02792f854297_story.html?utm_term=.586bed5e530b

² Sophia Rosenfeld, *Democracy and Truth: A Short History*, (Philadelphia: University of Pennsylvania Press, 2019) Ebook edition, Chapter 4, no page number.

³ I will be discussing memes and meme culture at length in this paper, therefore it would be prudent to provide a definition of a 'meme' here. The Merriam-Webster Dictionary describes a meme as 'an idea, behavior, style, or usage that spreads from person to person within a culture' or 'an amusing or interesting item (such as a captioned picture or video) or genre of items that is spread widely online especially through social media'. <https://www.merriam-webster.com/dictionary/meme>

and eighteenth centuries, equating indentured servitude with African chattel slavery. It has been argued that the rise in popularity of this narrative among American social media users coincided with the formation of the Black Lives Matter movement and is used to deny the existence of systemic racism against African-Americans while simultaneously promoting a narrative of white victimhood.⁴ Individuals and groups have shared memes, pictures and posts using hashtags such as #whitelivesmatter, #whiteslavery, #irishslavery, and #whitehistorymonth thousands of times. One of the central questions of this paper is whether this content merits a place in an archive based on its importance as a historical record. Proceeding from this question, the paper will then consider the issues that an archival institution must address if it were to create such an archive.

Archival scholarship relating to social media is still in its nascent stages, which makes this paper all the more topical and relevant. This paper will build on existing scholarship on issues surrounding social media archiving and the documentation of controversial elements of society. It will argue that the Irish slaves meme, despite its false and bigoted content, should be preserved as a snapshot of race relations and of the dissemination of racist ideas in the twenty-first century. However, archiving this material would not be a straightforward process. There are many special issues that an archival institution would have to consider when creating this archive, such as appraisal (determining the archival value of the material), preservation, arrangement, description, and ethics. This paper will study these issues as they apply to the creation of a hypothetical archive of the Irish slaves meme and it will attempt to offer some solutions. This experiment could lead the way for future social media archives of a controversial nature. Archival scholar Richard J. Cox argues that archivists need to be more assertive about the need to preserve controversial or problematic material as historical records.⁵

⁴ Bryan Fanning, 'Slaves to a Myth,' Dublin Review of Books, 1 November 2017. <https://www.drb.ie/essays/slaves-to-a-myth>

⁵ Richard J. Cox, *Archival Anxiety and the Vocational Calling*. (Duluth, Minn: Litwin Books, 2011), 248.

This paper will also examine other projects that have been undertaken to archive social media content, and will use them as a guideline for the proposed Irish slaves meme archive. It can be argued that preserving social media content is of vital importance to the historical record due to its ephemeral nature and the fact that significant social movements have begun on social media, such as Black Lives Matter or #MeToo. However, the overwhelming amount of data generated by social media users makes the creation of a general social media archive an extremely expensive, difficult, and time consuming endeavour; the failure of the Library of Congress's Twitter archive is the most obvious example of this.⁶ Despite these pitfalls, several social media archives have been set up with a narrower and more attainable scope.

Documenting the Now (DocNow) is a project established in 2016 with the aim of developing tools for collecting social media datasets from significant events to be used by scholars in their research.⁷ Other institutions have embraced the 'Archives 2.0' approach and created open source software through which users can add material to existing collections or even create their own collections of social media archives.⁸

Archiving the controversial or 'undesirable' elements of society is not a new phenomenon. For example, the Clark Historical Library in Michigan, USA caused outrage after its acquisition of Ku Klux Klan membership records in 1992.⁹ Archivist Karen M. Lamoree defends the inclusion of hate speech in archives by arguing that people in the future need to be fully informed about all elements of society, and despite the offensiveness of the material, remembering that such opinions existed or events occurred is of paramount importance.¹⁰ It would be doing a disservice to future historians and citizens not to document such material, despite how personally distasteful it may seem to archivists. As Cox powerfully argues, 'Archivists are not in

⁶ Axel Bruns, 'The Library of Congress Twitter Archive: A Failure of Historic Proportions,' Medium, 2 January 2018. <https://medium.com/dmrc-at-large/the-library-of-congress-twitter-archive-a-failure-of-historic-proportions-6dc1c3bc9e2c>

⁷ Julie Hale, 'Documenting the Now: Archiving Social Media for Generations to Come,' University Libraries, Washington University in Saint Louis, 20 January 2016, <https://library.wustl.edu/8497-2/>

⁸ E.g. Social Feed Manager, developed by George Washington University. <https://gwu-libraries.github.io/sfm-ui/about/overview>

⁹ Frank Boles, "'Just a Bunch of Bigots' A Case Study in the Acquisition of Controversial Material,' *Archival Issues* 19, no. 1, (1994): 53.

¹⁰ Karen M. Lamoree, 'Documenting the Difficult or Collecting the Controversial,' *Archival Issues* 20, no. 2 (1995): 150.

business to make people or our society feel good, but to provide critical evidence for understanding our past and how we have gotten to where we are today.¹¹

Projects have been undertaken in recent years to archive hate speech online; for example, the Internet Archive preserves Stormfront, a white supremacist message board. This paper seeks to demonstrate the value in adding the Irish slaves meme to this growing collection of problematic social media archives. The theme of this paper was partly inspired by a panel discussion titled 'Documenting Hate' which took place during the National Forum on Ethics and Archiving the Web in New York in March 2018.¹² Speakers on this panel highlighted the scholarly and larger societal benefits of archiving instances of white supremacist online hate speech while also considering the challenges that such an endeavour entails.¹³ This discussion, despite raising more questions than answers, provides much food for thought and demonstrates that there is a larger discussion around problematic social media archives to which this paper can be a contribution.

Methodology

The primary sources used in this paper will come from Facebook, Twitter, and Instagram. In compiling the material for potential inclusion in the Irish slaves meme archive, I carried out hashtag and keyword searches. The hashtags used were: #irishslaves, #irishslavery, #irishlivesmatter, #whitelivesmatter, #whitehistorymonth, and #reparations2020. Identical keyword searches were also conducted. These search terms were chosen due to their frequent use in the spread of the meme. Out of concern for privacy, only public posts will be shown in this paper and the usernames and avatars of users will be redacted. The material ranges in nature from text-only posts to photos and graphics on the subject of Irish slavery. There are some issues to consider when utilising social media posts as primary sources. Due to technological constraints, the posts are saved as screenshots. A disadvantage of this method is

¹¹ Cox, *Archival Anxiety*, 248.

¹² 'Ethics and Archiving the Web,' Rhizome, accessed 13 June 2019. <https://eaw.rhizome.org/>

¹³ 'Ethics and Archiving the Web,' Vimeo, 27 June 2018. Accessed 13 June 2019. <https://vimeo.com/277336010>

that it effectively freezes the material at the point of saving, thus further interaction after that moment (such as comments and shares) is not shown. However, this method does ensure that a copy of the post is saved even if the post is later deleted by the user, thus screenshotting is more reliable than using the 'save' function provided by Facebook and Instagram. One encounters difficulties when searching for this meme in 2019 due to the fact that a noticeable number of posts made about this myth in previous years have been deleted. Fortunately, between May 2015 and May 2016, Liam Hogan, a librarian and independent scholar based in Limerick, Ireland, compiled an online repository of hundreds of screenshots of the Irish slaves meme posts from Facebook and Twitter dating back to 2014.¹⁴ As an archive, it is flawed due to its lack of contextualisation and logical order. However, the collection is a useful source for the early development of the meme and some of its material will be cited in this paper.

This paper is structured according to the questions posed earlier. The first chapter will give a detailed background into the Irish slaves myth, considering its historical origins and online spread. In order to prove the historical value of preserving instances of this meme in an archive, this chapter will draw from the debate around social justice and inclusion in relation to archives, and the postmodern goal of presenting multiple perspectives or 'truths' in the archive. Furthermore, the Irish slaves meme will be placed in the context of similar cases of the preservation of controversial material in libraries or archives.

The proceeding four chapters will examine four issues which the archivist must overcome in the process creating the archive: acquisition and appraisal, harvesting and preservation, arrangement and description, and ethics. Chapter 2, on acquisition and appraisal, explores criteria for deciding which material will be included in the hypothetical Irish slaves meme archive. The chapter will build upon scholarship related to both traditional and digital appraisal methods. For example, as long ago as 1975, F. Gerald Ham asserted that it is the archivist's

¹⁴ "'But the Irish were slaves too'", Tumblr, accessed 13 June 2019. <https://theishwereslavestoo.tumblr.com/>

responsibility to ‘fill in the gap’ in the historical record, conducting appraisal from the viewpoint that archives should reflect wider society as fully as possible.¹⁵

Chapter 3 will focus on harvesting and preserving social media data. As already mentioned, scholarship in this area is in its early stages but it is possible to explore comparable methods with which the Irish slaves meme archive can preserve material from social media sites. Harvesting and preservation go hand in hand. Media studies scholar Niels Brügger asserts that simply collecting data is not web archiving; the process is not complete without planning for long-term preservation.¹⁶

Chapter 4 is a detailed consideration of the order and description of the Irish slaves meme archive. A case can be made for arranging the material chronologically and linking it thematically via metadata. It is here that description plays an important role in providing context to the material. For instance, Geoffrey Yeo recommends the development of descriptive systems that account for the wide range of perspectives and frames of reference that different users bring to an archive.¹⁷ This is especially pertinent to the controversial and emotionally-charged material that would be contained in the Irish slaves meme archive.

Chapter 5, the final chapter, is concerned with the myriad ethical issues that would arise in the creation of the Irish slaves meme archive. Among these issues are privacy concerns for social media users who contributed to the spread of the meme. Another key area of concern is limiting access to the archive to avoid it becoming a propaganda tool which would further contribute to the spread of the myth of Irish slavery. While most archival scholarship on the subject of access advocates its widening and democratisation, this paper will argue in favour of placing controls on who is allowed to access the material.

¹⁵ F. Gerald Ham, ‘The Archival Edge’ *The American Archivist* 38, no.1 (January 1975): 9.

¹⁶ Niels Brügger, *The Archived Web: Doing History in the Digital Age*, (Cambridge, MA : The MIT Press, 2018), 60.

¹⁷ Geoffrey Yeo, ‘Continuing Debates about Description,’ in *Currents of Archival Thinking*, eds. Heather MacNeil and Terry Eastwood, Second Edition. (Santa Barbara, California Denver, Colorado: Libraries Unlimited, an imprint of ABC-CLIO, LLC, 2017), 180.

The paper will finish with a conclusion which will tie together the various arguments presented in the previous chapters and indicate where the debate can go in the future. As this is a new and relatively-uncharted area of research, it is hoped that this paper will provide some useful suggestions for future projects aimed at preserving problematic or controversial social media data.

Chapter 1: The Irish Slaves Meme

Introduction

This chapter will provide a background to the origins and evolution of the Irish slaves myth, from book to meme form. Some examples of the meme will be given to illustrate the various contexts in which the myth is presented. A case will then be made for the preservation of the meme in an archive. This argument will draw on such archival theories as inclusion and social memory.

Origins and Spread of the Irish Slaves Meme

Before presenting an argument for the creation of an Irish slaves meme archive, it is necessary to describe the origins and nature of the meme. Scholars cite Irish journalist Sean O’Callaghan’s popular history book, *To Hell or Barbados: The Ethnic Cleansing of Ireland* (2000)¹ as a major contribution to the spread of the myth of Irish slavery in the Caribbean.² O’Callaghan had previously written investigative books about slavery around the world, such as *White Slave Trade* (1967) and *Yellow Slave Trade* (1968). *To Hell or Barbados* relies on ‘questionable sources,’ and conflates voluntary and involuntary forms of indentured servitude with chattel slavery.³ *White Cargo: The Forgotten History of Britain’s White Slaves in America* (2008) by British television documentarians Don Jordan and Michael Walsh echoes the conflation of servitude and slavery that can be observed in *To Hell or Barbados*.⁴ The authors argue that the term ‘slavery’ can be applied to anyone who was the property of another, whether temporarily

¹ Sean O’Callaghan, *To Hell or Barbados: The ethnic cleansing of Ireland* (Dingle: Brandon, 2000).

² Liam Hogan, Laura McAtackney, and Matthew C. Reilly, ‘The Irish in the Anglo-Caribbean: Servants or Slaves?’ *History Ireland* 24, no. 2 (April 2016): 19; Jerome S. Handler and Matthew C. Reilly, ‘Contesting “White Slavery” in the Caribbean: Enslaved Africans and European Indentured Servants in Seventeenth-Century Barbados’. *New West Indian Guide* 91 (2017): 31. <https://doi.org/10.1163/22134360-09101056>.

³ Handler and Reilly, ‘Contesting “White Slavery,”’ 47.

⁴ Don Jordan and Michael Walsh, *White Cargo: The Forgotten History of Britain’s White Slaves in America* (New York: New York University Press, 2008).

or for a lifetime.⁵ It is worth noting that none of the above authors are trained historians and their books lie very much in the realm of popular history. Therefore, their work is lacking the rigorous and systematic interrogation of sources that is typical of historians. In particular, O’Callaghan demonstrates a disregard for using robust evidence to back up his argument. At several points in his book, he admits that his assertions are based on assumptions due to a lack of sources.⁶ *To Hell or Barbados* was published by Brandon, an imprint of popular Irish general publisher, The O’Brien Press. More worryingly, *White Cargo* was published by the New York University Press, a company which purports to uphold scholarly values.⁷

Although the false history of Irish slavery has been debunked by scholars and journalists alike,⁸ it has been embraced on social media in recent years and predominantly appears in a racially-charged context:

In its most charged form, social media on both sides of the Atlantic reflect a highly racialized rhetoric that adopts the discourse of O’Callaghan and others to condemn or criticize calls for reparations to those of African descent whose ancestors were enslaved and who experience the repercussions of this past in the present.⁹

Liam Hogan has done extensive work on examining and debunking the Irish slaves meme.¹⁰ His Facebook research has shown that although the meme has been shared among Irish Facebook users (predominantly as a means of brushing aside Irish involvement in the transatlantic slave

⁵ Joyce Lau, ‘Master and Servant,’ *New York Times*, 27 April, 2008. <https://www.nytimes.com/2008/04/27/books/review/Lau-t.html>

⁶ O’Callaghan, *To Hell or Barbados*, 87-88.

⁷ ‘History,’ NYU Press, accessed 13 June 2019. <https://nyupress.org/about-nyupress/history/>

⁸ See footnote 2; Liam Hogan, ‘Open letter to Irish Central, Irish Examiner and Scientific American about their “Irish slaves” disinformation,’ Medium, 8 March 2016. <https://medium.com/@Limerick1914/open-letter-to-irish-central-irish-examiner-and-scientific-american-about-their-irish-slaves-3f6cf23b8d7f>; Liam Stack, ‘Debunking a Myth: The Irish Were Not Slaves, Too,’ *New York Times*, 17 March 2017. <https://www.nytimes.com/2017/03/17/us/irish-slaves-myth.html>

⁹ Handler and Reilly, ‘Contesting “White Slavery,”’ 31.

¹⁰ Liam Hogan, ‘All of my work on the “Irish slaves” meme (2015–’18),’ Medium, 12 March 2017. <https://medium.com/@Limerick1914/all-of-my-work-on-the-irish-slaves-meme-2015-16-4965e445802a>

trade), it is more widely shared in the United States.¹¹ For example, in March 2016 the ‘Irish Americans’ public Facebook page posted an image of the meme which has been shared over 580,000 times.¹² According to Hogan’s analysis, the Irish slaves meme was shared at least 4 million times from 2014 to 2016.¹³ Google trend data shows that the search term ‘Irish slaves’ reached its peak popularity in September 2016.¹⁴ Hogan links the evolution of the meme to the backlash against the Black Lives Matter movement as it has been used to ‘derail discussions about the legacy of slavery or ongoing anti-black racism’.¹⁵ Social policy scholar Bryan Fanning similarly links the evolution of the meme to the American alt-right movement:

[The Irish slaves meme] has been mobilised by the American alt-right, among others, to disavow legacies of racism and present-day racism while simultaneously promoting a white nationalist political agenda based on claims of white victimhood.¹⁶

As will be shown below, it is difficult to isolate the meme from its racial connotations.

The following examples come from both my own hashtag-based searches and the online repository created by Liam Hogan. The examples were selected due to their relevance to the most commonly-recurring themes identified during the course of research for this paper. They come from Facebook, Twitter, and Instagram. Figures 1 and 2 are common iterations of the meme which use the myth of Irish slavery to delegitimise claims of systemic racism against black Americans. The meme posits that Irish ‘slaves’ were treated worse than African slaves; in reality, it is impossible to accurately compare the treatment of indentured servants and slaves

¹¹ Alex Amend, ‘How the Myth of the "Irish slaves" Became a Favorite Meme of Racists Online,’ Southern Poverty Law Center, 19 April 2016. <https://www.splcenter.org/hatewatch/2016/04/19/how-myth-irish-slaves-became-favorite-meme-racists-online>

¹² Irish Americans, ‘Thoughts?’ Facebook, 24 March 2016. <https://www.facebook.com/Irish.Americans1/photos/a.1391286387806114.1073741828.1391282101139876/1691045911163492/?type=3&theater>

¹³ Liam Hogan, ‘Two years of the ‘Irish slaves’ myth: racism, reductionism and the tradition of diminishing the transatlantic slave trade,’ openDemocracy, 7 November 2016. <https://www.opendemocracy.net/en/beyond-trafficking-and-slavery/two-years-of-irish-slaves-myth-racism-reductionism-and-tradition-of-diminis/>

¹⁴ Google trend data, carried out on 13 June 2019. <https://trends.google.com/trends/explore?date=all&q=irish%20slaves>

¹⁵ Amend, ‘How the Myth of the "Irish slaves" Became a Favorite Meme of Racists Online.’

¹⁶ Fanning, ‘Slaves to a Myth.’

in the seventeenth century Caribbean given the dearth of surviving evidence.¹⁷ The memes also serve as a self-congratulation of Irish Americans by contrasting their successful social mobility to that of African Americans, which is a common trope of the white slavery myth.¹⁸ In this way, the Irish slaves meme can generally be seen as a marker of Irish or Irish American identity, which archivist Corrie Commisso argues is not unusual in the 'post-truth' era, in which 'objective facts are less influential in shaping public opinion than are appeals to emotion or personal belief.'¹⁹ Key to this identity is self-victimisation and consequent denial of Irish involvement in the African slave trade.²⁰ In a similar way, the meme is utilised as a means of condemning Black Lives Matter, as evidenced by the use of hashtags such as #whitelivesmatter and #irishlivesmatter (Figures 3-5). Another frequent refrain of those sharing the Irish slaves meme is that this history has been 'forgotten' or even 'hidden', continuing the victimisation narrative (Figures 5 and 6). In this way, the Irish slaves meme fulfils meme culture's potential to allow those who feel marginalised to create new (albeit false) narratives about themselves and their history.²¹ The meme also appears in the context of the debate around reparations for descendants of slaves in the United States. Statements made by various Democratic Party candidates in the 2020 US Presidential elections on the issue of reparations has re-ignited this debate.²² It has prompted some to ridicule the argument for giving reparations to descendants of African slaves and others to contend that reparations should also be awarded to descendants of 'Irish slaves' (Figures 7-11).

¹⁷ Handler and Reilly, 'Contesting "White Slavery,"' 36.

¹⁸ *Ibid.*, 48.

¹⁹ Corrie Commisso, 'The Post-Truth Archive: Considerations for Archiving Context in Fake News Repositories,' *Preservation, Digital Technology & Culture* 46, no.3 (2017): 100. <https://doi.org/10.1515/pdte-2017-0010>.

²⁰ Hogan, McAtackney, and Reilly, 'The Irish in the Anglo-Caribbean,' 22.

²¹ An Xiao Mina, *Memes to Movements: How the World's Most Viral Media is Changing Social Protest and Power*. (Boston: Beacon Press, 2019), 103.

²² Oliver Laughland, 'Reparations: Democrats renew debate over how to heal the legacy of slavery,' *The Guardian*, 21 March 2019. <https://www.theguardian.com/us-news/2019/mar/21/reparations-democrats-2020-slavery>



Fig 1. Facebook post about Irish slaves being treated worse than others.²³

²³ <https://theirishwereslavestoo.tumblr.com/post/143086722740>. Accessed 13 June 2019.



Fig 2. Facebook post about Irish slaves being treated worse than others.



Fig 3. Facebook post in reaction to Black Lives Matter.



Fig 4. Facebook post with refrain of 'Irish lives matter.'

The first slaves imported into the American colonies were 100 White children in 1619, four months before the arrival of the first shipment of Black slaves. Many were brought from Ireland, where the law held that it was "no more sin to kill an Irishman than a dog or any other brute".



A CHILDREN IN THE FERRY

King James II, followed by Charles I and Oliver Cromwell, sold over 500,000 Irish Catholics into slavery throughout the 1600's onto plantations in the West Indies Islands of Antigua, Montserrat, Jamaica, Barbados, as well as Virginia and New England. Irish slaves were less expensive than African, and treated with more cruelty & death.

In the 17th Century, from 1600 until 1699, there were many more Irish sold as slaves than Africans. There are records of Irish slaves well into the 18th Century. Many never made it off the ships. According to written record, in at least one incident 132 slaves, men, women, and children, were dumped overboard to drown because ships' supplies were running low. They were drowned because the insurance would pay for an "accident," but not if the slaves were allowed to starve.

White Slavery

History Denied, Covered Up, & Marginalized

WHITE LIVES MATTER **Whitelivesmatter** · Like This Page · February 14, 2019

#WhiteSlavery
#WhiteLivesMatter
!!HardWorkPaysOff!!
#NoReparationsForIrish

👍 146 💬 219 Comments (4/ Share)

👍 Like 💬 Comment ➦ Share

All Comments ▾

It goes to show some people would rather play the victim oppressed, than see reality, no matter what truth this are brought forward, there will always be a group of uneducated people to refute it.

Like Reply 3y

This picture is from the late 18th first of all and the 100 children were not the first 100 white children.

Write a comment...

Fig 5. Facebook post highlighting the 'cover-up' of white slavery with refrain of 'White lives matter.'

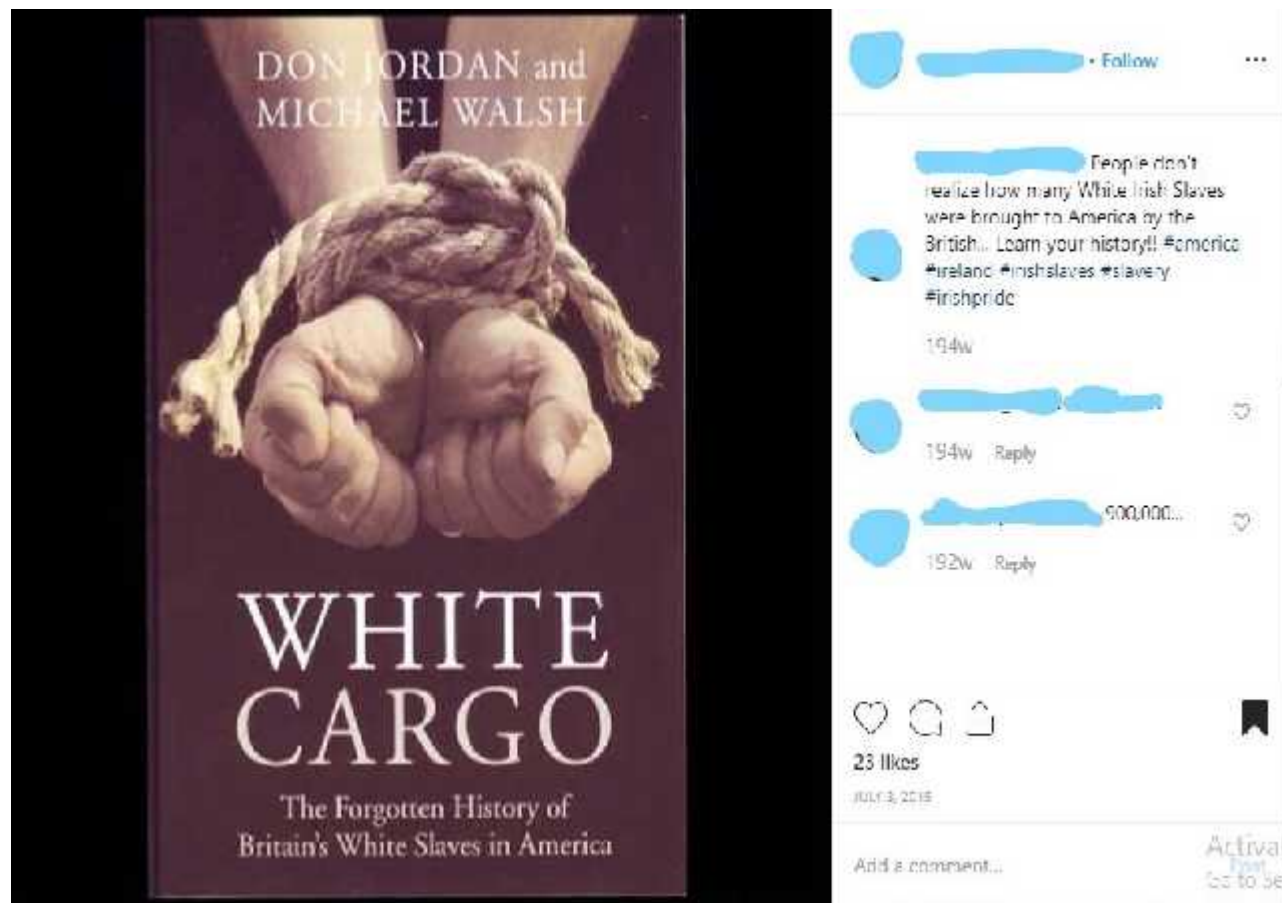


Fig 6. Instagram post about the forgotten history of Irish slavery.



Fig 7. Tweet shared on Facebook about reparations.

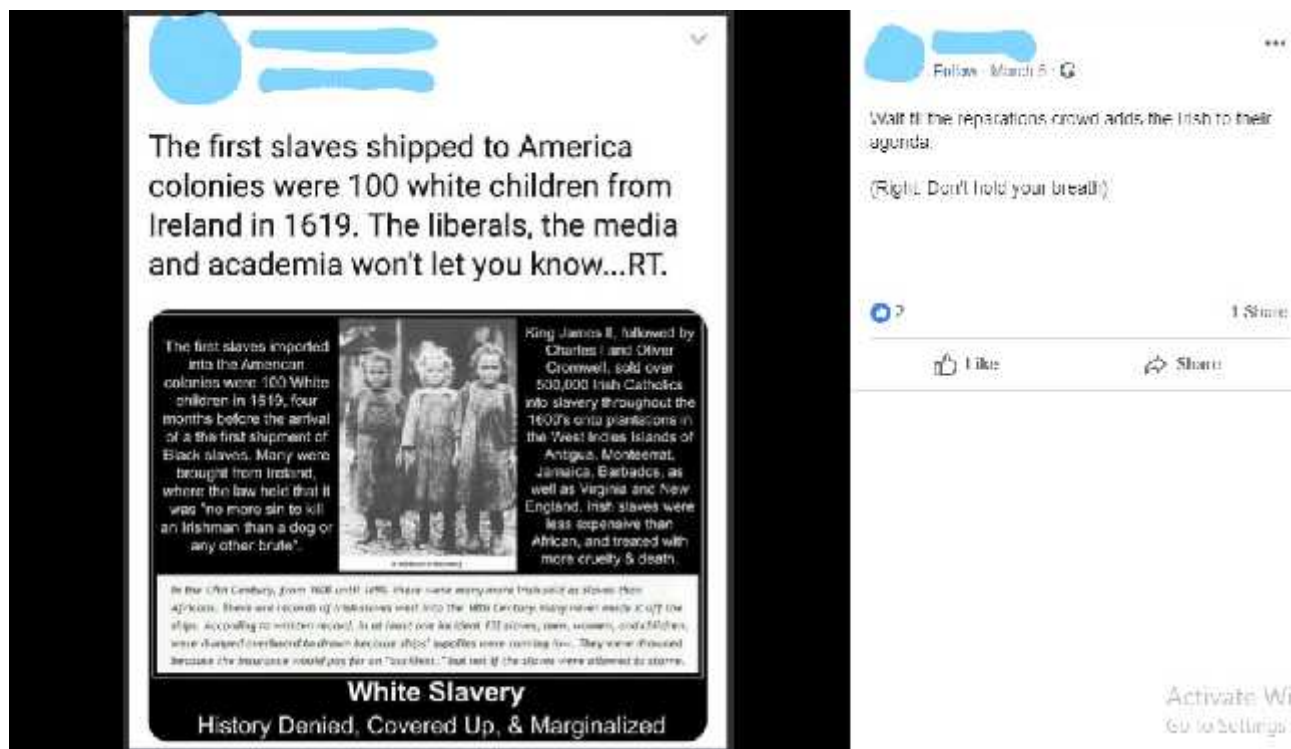


Fig 8. Tweet shared on Facebook about reparations.



Fig 9. Twitter post about reparations.



Fig 10. Facebook post about reparations.



Fig 11. Twitter post about reparations.

Archival Value of the Irish Slaves Meme

Archival scholar Terry Cook asserts that archivists are ‘active agents in constructing social and historical memory,’ deciding which voices will be heard and remembered or marginalised and forgotten.²⁴ Therefore, it can be argued that archivists have a responsibility to represent as many facets of society as possible through the records they preserve. Ham has posited that the archivist’s most important role is to ‘hold up a mirror for mankind’ in order to ‘[help] people understand the world they live in.’²⁵ If one considers that archivists have agency over the creation of social memory and have a duty to provide the fullest possible picture of society, it would follow that archives should be inclusive of all members of society. This idea is not new. In 1977, historian Howard Zinn famously criticised the passivity of American archivists in the face of government secrecy and urged them to ‘take the trouble to compile a whole new world of documentary material, about the lives, desires, needs, of ordinary people’ rather than only documenting the lives of the powerful elite.²⁶ This has been credited as the first public iteration of the notion of social justice in relation to archives, or the ‘activist archivist.’²⁷ The majority of archival scholars are now in favour of the concept of an activist, non-neutral archivist and archival social justice has become a hugely popular topic of research since the turn of the twenty-first century.²⁸

Randall Jimerson has been highly influential in the field of archival social justice.²⁹ He posits that archivists should embrace their power as shapers of societal memory and use this power in a morally positive way by ‘giving equal voice to those groups that too often have been

²⁴ Terry Cook, ‘Remembering the Future: Appraisal of Records and the Role of Archives in Constructing Social Memory’ in Francis X. Blouin, ed. *Archives, Documentation, and Institutions of Social Memory: Essays from the Sawyer Seminar*. (Ann Arbor, Mich: Univ. of Michigan Press, 2006), 232.

²⁵ Ham, ‘The Archival Edge,’ 13.

²⁶ Howard Zinn. ‘Secrecy, Archives, and the Public Interest’. *The Midwestern Archivist* II, no. 2 (1977): 25.

²⁷ David A. Wallace, ‘Archives and Social Justice,’ in *Currents of Archival Thinking*, eds. Heather MacNeil and Terry Eastwood, Second Edition, 279-280.

²⁸ Wallace, ‘Archives and Social Justice,’ 279 & 271.

²⁹ *Ibid.*, 283.

marginalized and silenced' through inclusivity and greater representation of minorities.³⁰

Another venerated activist archivist, Verne Harris, advocates hospitality and engaging with the 'other' in the archives, even if one fears this otherness.³¹ He concedes that this is not always an easy task:

How to invite in what is always beyond the limits of understanding? [...] How to invite in what one wishes to resist – the voices, for instance, of white supremacists, hard drug dealers, paedophiles, rapists, pimps, and so on and on and on?³²

One may be tempted to resist or turn away from the voices of those propagating the Irish slaves myth but this paper argues that the archivist has a responsibility to allow these voices to be preserved as part of the historical record. Regardless of the veracity of the claims they make, it can be argued that a significant amount of people believe in the myth and use the meme to share their grievances about several contemporary issues, such as reparations and African American activism. Mark Greene, a proponent of postmodern archival thinking, asserts that the power of the archive does not lie in its value as evidence of transactions, but in its capacity as a 'repository of meaning.'³³ Furthermore, he states that meaning and truth in archives are subjective notions and it is possible to find multiple meanings or 'truths' even in a document that contains unreliable information.³⁴ This shows that despite its lack of factual grounding, the Irish slaves meme has archival value due to its meaning for different audiences.

If the commonly accepted social justice aim of archivists is to better represent marginal views and minorities in society, why should this not be applied to controversial topics that may not have our sympathy? As the DocNow project aims to provide scholars with tools to research contemporary social movements born on social media, surely such tools would also be useful

³⁰ Randall Jimerson. 'Archives for All: Professional Responsibility and Social Justice'. *The American Archivist* 70, no. 2 (September 2007): 254. <https://doi.org/10.17723/aarc.70.2.5n20760751v643m7>.

³¹ Verne Harris, 'The Archival Sliver: Power, Memory, and Archives in South Africa' *Archival Science* 2, nos. 1–2 (2002): 86.

³² Harris, 'The Archival Sliver,' 86.

³³ Mark Greene, 'The Power of Meaning: The Archival Mission in the Postmodern Age,' *The American Archivist* 65 (Spring/Summer 2002): 55.

³⁴ Greene, 'The Power of Meaning,' 52.

for researching the counterpoint to these movements. In order to present a comprehensive overview of an issue, one must consider all sides. Lamoree advises archivists to focus their efforts on documenting an issue rather than a particular stance, group, or individual to create archives that represent ‘the complexity of society.’³⁵ Archives are an effective means of holding up a mirror to society in all its complexity. Only inviting in the marginalised with sympathetic causes simplifies this mirror image and limits the potential for scholarly research and human understanding. Cox criticises archivists for too often choosing to align themselves with the ‘feel-good heritage industry’, resulting in a loss of nuance and depth in the view of the past created by their records.³⁶

Archives, museums, and libraries have the potential to act as tools for us to confront uncomfortable truths about our society, both in the past and in the present. In 2012, the Jim Crow Museum of Racist Memorabilia was opened to the public. The museum exhibits historical items that depict African-Americans in a racist light to demonstrate the domination of anti-black images in American cultural life, from the slavery era to the present day. Dr. David Pilgrim, founder of the museum, states that the ultimate goal of the museum is to educate the public and to promote dialogue on the issue of racism.³⁷ Preserving examples of online racism in the form of the Irish slaves meme can fulfil the same aims. Rather than being a ‘shrine to racism,’ an archive of the meme would provide the evidence needed in order to have an informed and meaningful dialogue on the subject of modern-day racism.³⁸

The Irish slaves meme is also worth preserving as it touches upon several potential areas of research such as representations of white victimhood or (the repudiation of) white guilt on social media, the spread of conspiracy theories or false narratives via social media, or Irish

³⁵ Lamoree, ‘Documenting the Difficult,’ 150.

³⁶ Cox, *Archival Anxiety*, 248.

³⁷ Mike Householder, ‘New Michigan museum showcases racist artifacts,’ About the Jim Crow Museum, Ferris State University, 19 April 2012. <https://www.ferris.edu/HTMLS/news/jimcrow/more.htm#>

³⁸ Householder, ‘New Michigan museum.’

American (or to a lesser extent, Irish) identity in the twenty-first century. However, perhaps the slaves meme archive's most valuable contribution to scholarly research and society in general would be in demonstrating another facet of the debate around slavery, reparations, and race in the United States in the twenty-first century. In order for future generations to understand these complicated race relations, they must be provided with information from various quarters, no matter how controversial and hateful the material may be. Therefore, this material could form a vital component of societal memory due to its links to several hotly contested issues in contemporary society.

The question of collecting examples of the false narrative of the Irish slaves meme ties into that of the preservation of 'fake news' on social media, a highly topical one in the current political climate.³⁹ Alexandra Juhasz argues that fake news merits preservation:

Fake news—and the Internet's mountain of attempts to better see it, know it, defang, debunk, and stop it—should be carefully saved for no better reason than that it existed, and thus proved itself to be at once inordinately powerful within the attention economy of the Internet and also for associated material manifestations that occurred offline.⁴⁰

A similar line of reasoning can be followed in relation to the Irish slaves meme. Although the historical basis for the meme may be false, the narrative still exists and a notable number of people identify with its claims. Studying fake news can help to illustrate the myriad perspectives that can be held on an issue.⁴¹ Thus, the Irish slaves meme can provide a window into people's views on the range of topics outlined earlier.

The Irish slaves meme could also function as a tool for educating the public about the nature and dangers of online misinformation. It has been posited that libraries and archives have a

³⁹ Elle Hunt, 'What is fake news? How to spot it and what you can do to stop it,' *The Guardian*, December 17, 2016. <https://www.theguardian.com/media/2016/dec/18/what-is-fake-news-pizzagate>

⁴⁰ Alexandra Juhasz, 'A Preservationist's Guide to #100hardtruths- #fakenews: One Fake News Preserve,' *Preservation, Digital Technology & Culture* 46, no.3 (2017): 105. <https://doi.org/10.1515/pdtc-2017-0011>

⁴¹ Editor-in-Chief, 'Preserving Fake News,' *Preservation, Digital Technology & Culture* 46, no.3 (2017): 89. <https://doi.org/10.1515/pdtc-2017-9001>

responsibility to promote informational literacy in an age in which it is increasingly difficult to differentiate between truth and lies in the media. As Commisso argues, ‘the information professions have a vital role to play in equipping people to navigate the murky waters of an information ecosystem that is constantly changing and increasingly complicated.’⁴²

Memes and online fake news are emerging research topics. Viz Lab, begun in 2018 and still in its development stages, is a project which aims to collect and preserve memes that propagate false narratives in order to create ‘visual search and discovery tools to help contextualize and track memes as they travel and evolve online’.⁴³ Researchers are also beginning to appreciate the value of social media in general as a primary source. Cayce Myers and James F. Hamilton argue that social media will become key primary sources for future historians due to the ‘historical tradition of examining conflict coupled with the continued relevance of social history.’⁴⁴ Indeed, social media data is already proving to be a useful resource in the fields of sociology and media studies. For example, sociologists Aaron Panofsky and Joan Donovan analysed data from white nationalist online forum, Stormfront, to examine users’ construction of race identity based on genetic ancestry tests.⁴⁵ Therefore, the Irish slaves meme archive could provide valuable primary source material for current and future researchers.

Conclusion

This chapter has offered a contextual background to the development of the Irish slaves meme and has attempted to demonstrate its worth as an addition to the historical record, arguing that archival theories around social justice and documenting all aspects society should also be applied to controversial material or ‘fake news.’ The next chapter will focus on how the Irish

⁴² Commisso, ‘The Post-Truth Archive,’ 99.

⁴³ ‘Vizlab’, Harmony Labs, accessed 13 June 2019. <https://harmonylabs.org/vizlab>

⁴⁴ Cayce Myers & James F. Hamilton, ‘Social Media as Primary Source,’ *Media History* 20, no. 4 (2014): 432.

⁴⁵ Aaron Panofsky and Joan Donovan, ‘Genetic Ancestry Testing Among White Nationalists.’ *SocArXiv*. 17 August 2013. doi:10.31235/osf.io/7f9bc

slaves meme can be acquired and appraised by an archival institution: the first step in the process of creating an archival collection.

Chapter 2: Appraisal and Acquisition

Introduction

Appraisal can be defined as ‘the process of determining whether records and other materials have permanent (archival) value.’¹ Ham defines acquisition as ‘the process of acquiring records from any source by transfer, donation, or purchase, or the body of records so acquired.’² This chapter examines how these procedures could be conducted on the Irish slaves meme. There will first be an exploration of the most appropriate type of organisation to acquire the meme archive. Then, the appraisal of the meme will be examined within the framework of Terry Cook’s four archival paradigms. The meme’s applicability to an institution’s collection development policy will then be explored. The chapter will conclude with some final comments about appraisal issues that are specific to the Irish slaves meme.

Finding a Suitable Home for the Irish Slaves Meme Archive

The best approach to collecting and appraising the Irish slaves meme archive will depend on the type of organisation that will undertake this project. For instance, the project could be carried out by an independent group funded by grants or user donations. Alternatively, the archive could become part of a collection within a larger organisation, such as a university library or a cultural institution. There are advantages and disadvantages to each collecting body. Smaller scale initiatives are free from the constraints of scope and policy that govern larger institutions. Like community archives, these initiatives allow for greater participation from users and are therefore more likely to reflect the needs and interests of their audience.³ However, archiving web data is an expensive and time-consuming process which requires long-term maintenance.

¹ Richard Pearce-Moses, *A Glossary of Archival and Records Terminology*, (Chicago: Society of American Archivists, 2005), 22.

² F. Gerald Ham, *Selecting and Appraising Archives and Manuscripts*, (Chicago: Society of American Archivists, 1993), 2.

³ Rebecka Sheffield, ‘Community Archives,’ in *Currents of Archival Thinking*, eds. Heather MacNeil and Terry Eastwood, Second Edition, 364.

Many smaller, non-profit web archiving initiatives rely on short-term funding and this can be problematic for the long-term preservation of the archive.⁴ For example, Viz Lab was awarded a grant of \$50,000 by Democracy Fund in 2017 to begin its work.⁵ No information is given about whether funding will be provided for the project on an ongoing basis. This could affect the future of the archive as its wide scope may cause it to exceed its initial budget. Relying on project-based funding carries the risk that the archive will stagnate or become inaccessible if financial support ceases. For this reason, it would be more advantageous for a larger organisation to take responsibility for creating the Irish slaves meme archive. For example, a university library would offer a more reliable guarantee that the material will be safeguarded for future generations due to its relatively greater financial, technical, and personnel resources.⁶

An academic setting is perhaps the most suitable one for the meme archive given its controversial content. Archivists could mediate access and ensure that the material is presented in the fullest possible context. Issues of description and access will be more closely addressed in later chapters of this paper. Therefore, this paper proposes that an academic body, such as a university, a research institution, or a national archives, should be responsible for archiving the Irish slaves meme.

The Irish Slaves Meme in Cook's Four Paradigms

⁴ Colin Post, 'Building a Living, Breathing Archive: A Review of Appraisal Theories and Approaches for Web Archives,' *Preservation, Digital Technology & Culture* 46, no.2. (2017): 72. <https://doi.org/10.1515/pdte-2016-0031>

⁵ Josh Stearns, '20 Projects Receive Funding to Combat Misinformation and Build a More Trustworthy Public Square,' Democracy Fund, 22 June 2017. <https://www.democracyfund.org/blog/entry/20-projects-receive-funding-to-combat-misinformation-and-build-a-more-trust>

⁶ For example, the main library of the London School of Economics and Political Science asserts in its Collection Development Policy that any material it acquires is kept on a long-term basis. Anna Towlson, 'Collection Development Policy for Archives, Museum Objects and Ephemera,' LSE Library, June 2015. <http://www.lse.ac.uk/library/assets/documents/Archives-Collection-Development-Policy-PDF.pdf>

The Irish slaves meme is an unusual kind of primary source. It does not come from one creator; it does not provide a record of the functions of an organisation; and the information contained in the documents is false. As a result, traditional theories of appraisal (as will be shown below) are difficult to apply to the Irish slaves meme archive. Fiorella Foscarini has employed the four core archival paradigms identified by Terry Cook (i.e., evidence, memory, identity, and community) to investigate the different stages of thinking about appraisal throughout history.⁷ A useful exercise in determining the most effective method of appraising the Irish slaves meme would be to apply it to these paradigms. Cook's paradigms reflect the evolution of appraisal theory, from Hilary Jenkinson's rejection of the role of the archivist in the appraisal process⁸ to the community- and people-centred approach advocated by Ixchel M. Faniel and Elizabeth Yakel.⁹ Investigating the paradigms in terms of their applicability to the Irish slaves meme also shows how the meme fits (or does not fit) into traditional ways of thinking about appraisal.

In the evidence paradigm (of which Jenkinson was a prominent advocate), the archivist was a 'guardian' of the records of the state and did not have a part to play in shaping the archive through appraisal.¹⁰ Archival records were viewed as evidence of the functions of the state.¹¹ This positivist notion of evidence makes the paradigm unsuitable for the Irish slaves meme. Furthermore, the evidence paradigm is still closely linked with organisations and their concern for accountability and authenticity.¹² Since the Irish slaves meme does not come from only one creator or institution, it is not useful to appraise it on its value as an instrument of ensuring organisational accountability.

⁷ Terry Cook, 'Evidence, Memory, Identity, and Community: Four Shifting Archival Paradigms,' *Archival Science* 13, no.2 (2013): 95-120. Cited in Fiorella Foscarini, 'Archival Appraisal in Four Paradigms,' in *Currents of Archival Thinking*, eds. Heather MacNeil and Terry Eastwood. Second Edition, 107-133.

⁸ Foscarini, 'Archival Appraisal in Four Paradigms,' 109.

⁹ *Ibid.*, 126.

¹⁰ *Ibid.*, 109.

¹¹ *Ibid.*, 108.

¹² *Ibid.*, 109.

The memory paradigm was in vogue from the 1930s to the 1970s and its most famous proponent was Theodore R. Schellenberg. In this paradigm, the archivist was given greater agency to decide which records should be kept that best demonstrated the functions and actions of a government or organisation: i.e., its 'memory.'¹³ During appraisal, archivists would assess the 'secondary value' of an organisation's records to determine which records would provide sources of public memory for historians.¹⁴ Although Schellenberg limited his scope to public records, the 'secondary value' approach can also be applied to non-governmental records. University libraries often build their collections around certain pre-determined themes. For instance, the University of North Texas (UNT) Digital Library has created an archive of defunct government websites which focuses on websites that fall within the scope of the university's curriculum.¹⁵ However, there is a risk that this approach can lead archivists to make important appraisal decisions based on shifting trends in historiography, a method that has been criticised by Cook elsewhere.¹⁶ If a university library was to conduct appraisal based on the memory paradigm, the Irish slaves meme may be overlooked unless it corresponded to popular research topics among the university's researchers.

The identity paradigm (or 'society paradigm') is a more useful framework for appraising the Irish slaves meme.¹⁷ In this paradigm, archives are tools that should serve society and not the institutions governing it, therefore archivists should aim to create a documentary heritage that reflects all aspects of society.¹⁸ In relation to this paradigm, Foscarini cites the documentary heritage approach advocated by German scholar Hans Booms.¹⁹ Of particular relevance to the Irish slaves meme is Booms' argument that the significance of a record depends on its relationship to other contemporary societal phenomena.²⁰ Thus, the meme increases in archival

¹³ Ibid., 113.

¹⁴ Ibid., 110.

¹⁵ 'Collection Development Policy for the UNT Libraries' Digital Collections,' University of North Texas, last modified 12 January, 2017. <https://library.unt.edu/policies/collection-development-digital-collections/>

¹⁶ Cook, 'Remembering the Future,' 242.

¹⁷ Ibid., 114.

¹⁸ Ibid.

¹⁹ Ibid., 118.

²⁰ Ibid., 103.

value when viewed in the context of other issues, such as the Black Lives Matter movement. Foscarini also cites American archival scholar Helen Samuels' 'documentation strategy' approach as pivotal to the identity paradigm.²¹ This methodology involves the collecting institution conducting an analysis of its core functions and existing collections in order to determine which records would supplement or fill the gaps in its archives.²² This approach allows for collaboration with other institutions with the goal of creating the most complete record of society possible.²³ It is preferable to the memory paradigm, which relies too heavily on arbitrarily assigning value on records based on historiographical trends. The identity strategy will form a key part of the methodology for appraising the Irish slaves meme archive, which will be described more fully later in the chapter.

The fourth and final paradigm, community, is the most recent stage in the evolution of appraisal theory. This approach allows communities greater agency in the creation of their own archives, avoiding the undesirable instance of archival institutions speaking for communities and thus further marginalising them.²⁴ However, this paper has asserted that the primary goal of the Irish slaves meme archive would be to provide material for scholarly research and not to create or promote a community around the archive. In fact, such an endeavour would be quite dangerous given the false and inflammatory information contained therein. In this way, the Irish slaves meme archive goes against the current trend towards community empowerment through archives. Thus, it can be concluded that of the four archival paradigms, the identity paradigm is the most suitable one for appraising the Irish slaves meme archive.

The Irish Slaves Meme in a Collection Development Strategy

²¹ Ibid., 115.

²² Terry Cook, 'What is Past is Prologue: A History of Archival Ideas Since 1898, and the Future Paradigm Shift,' *Archivaria* 43 (Spring 1997), 33.

²³ Cook, 'What is Past is Prologue,' 32.

²⁴ Foscarini, 'Archival Appraisal in Four Paradigms,' 124.

As previously argued, a large institution such as a university, national archives, or other cultural body would be the most suitable depository for the Irish slaves meme archive. For the purposes of this paper, it will be assumed that a university library will undertake to collect and preserve an archive of the Irish slaves meme. One of the main reasons for this choice is the relatively wide variety of information available related to the collection policies of university libraries, and more particularly, their special collections. Librarian and scholar Alison Cullingford considers the library's audience and mission as some of the key issues which the librarian or archivist should keep in mind when acquiring materials to develop their special collections.²⁵ Of course, institutions differ in their approach to collecting and appraising archives. For instance, Leeds University Library has developed a strategy whereby it assesses potential acquisitions based on their worth as 'heritage,' 'legacy,' or 'self-renewing' resources.²⁶

In general, university libraries and archives give priority to the acquisition of material which would add to their existing collections, rather than creating entirely new ones. For instance, the acquisition policy of the James Hardiman Library Archives at National University of Ireland, Galway states that its focus is on acquiring documents to supplement its existing Irish Language and Literature collection, among other collections.²⁷ North Carolina State University (NCSU) has developed a collecting strategy to identify and harvest hashtags on social media which relate to the collecting areas of its research centre.²⁸ In this way, NCSU has combined techniques of appraising social media content according to the collection development policies of the university. The UNT Digital Library has created the UNT Data Repository, an archive of data used by its researchers, which includes some social media content.²⁹ Its collection development

²⁵ Alison Cullingford, *The Special Collections Handbook*, Second. (London: Facet Publishing, 2016), 85.

²⁶ 'Collections Strategy for Leeds University Library,' Leeds University Library, accessed 13 June 2019. https://library.leeds.ac.uk/downloads/download/144/collections_strategy

²⁷ 'QA819 Acquisition Policy for the James Hardiman Library Archives,' James Hardiman Library, last modified 27 March 2011. <http://library.nuigalway.ie/media/jameshardimanlibrary/content/documents/collections/Archives-Acquisition-and-Evaluation-Policy.pdf>

²⁸ 'Identifying Hashtags Related to Special Collections Collecting Areas,' NC State University Libraries, accessed 13 June 2019. <https://www.lib.ncsu.edu/social-media-archives-toolkit/collecting/identifying-hashtags-related-to-collecting-areas>

²⁹ 'UNT Data Repository,' University of North Texas, accessed 13 June 2019. <https://digital.library.unt.edu/explore/collections/UNTDRD/>

policy states its goal of preserving and making accessible the ‘research data outputs’ of UNT researchers, thereby basing its appraisal decisions on the research interests of the university community.³⁰

The National Archives in the United Kingdom (TNA) has published guidelines for organisations to devise their own collection development policy.³¹ These guidelines will provide the basis for the approach to acquisition and appraisal that this paper proposes should be taken in the creation of the Irish slaves meme archive. Cullingford advises libraries to use the toolkit provided by TNA when formulating their collection development policy.³² This toolkit consists of several questions which the archivist must answer when acquiring and appraising new material. Asking these questions of the Irish slaves meme archive would demonstrate how the collection could potentially be appraised by a university library. The TNA collection development toolkit is appropriate for the purposes of the Irish slaves meme as it encourages the adoption of Samuels’ documentation strategy.³³ Additionally, TNA argues that archivists need to be active collectors rather than passive receivers of records, especially in the digital age when there is greater danger that records will not survive in their original format and context.³⁴ This emphasis on preserving digital material is further evidence of the suitability of TNA’s collection policy guidelines for the Irish slaves meme archive.

The first section of the guide, ‘About the collections’, pertains to the collections already held by the institution. For the purposes of this exercise, let us assume that the university library in question already possesses a collection of social media data related to the discussion of race and the Irish slaves meme is a potential addition to this collection. In this section, the archivist

³⁰ ‘Collection Development Policy for the UNT Libraries’ Digital Collections,’ University of North Texas.

³¹ The National Archives, *Collection Development Tools and Guidance*, 2011.

<http://www.nationalarchives.gov.uk/documents/archives/collection-development-tools-and-guidance.pdf>

³² Cullingford, *The Special Collections Handbook*, 85.

³³ The National Archives, *Collection Development*, 11.

³⁴ *Ibid.*, 11-12: ‘[Collections development] is about the survival of evidence of aspects of modern life which may not be sustainable or recoverable within the lifespan of current adults. Collecting which is mainly passive does not entirely fulfil the cultural and evidential roles of archives.’

is asked to reflect on the archival value of the existing collection and to identify gaps in the collection that need to be filled.³⁵ Here, the archival value of the Irish slaves meme should also be considered. The meme archive would potentially enrich the existing collection by providing evidence of such phenomena as the backlash to the Black Lives Matter Movement, white victimhood, or the spread of racially-motivated false information on social media. The guidelines advise readers to consult the National Archives Collection Strategy for assistance in determining the collection aims of their institution. The principle of inclusion forms a part of this strategy and it urges archivists to 'encourage the contribution of a broad range of interests, views and issues from all sectors of society, including marginalised and minority groups.'³⁶ The Irish slaves meme would certainly qualify as a valuable addition to the collection if the university library followed these criteria.

Question 4 pertains to the existing or potential audience for the collection.³⁷ The existing users of the collection could be researchers with an interest in media studies, race relations, civil rights activism, social justice, et cetera. With the addition of the Irish slaves meme to the collection, this already broad range of possibilities of research could be further widened. Potential new users could include those researching the evolution and spread of fake news, contemporary manifestations of racism, or Irish-American identity. Therefore, adding the Irish slaves meme to the library's collection would create a new audience for the library.

Question 5 advises archivists to be aware of potential overlap between their collection and those of other archives or libraries.³⁸ This demonstrates the importance of conducting institutional functional analysis (as proposed by Samuels). The archivist needs to be fully informed of the aims and functions of the institution in order to have a clear idea of the themes

³⁵ Ibid.,6.

³⁶ The National Archives, *The National Collections Strategy*, 2009.

<https://www.nationalarchives.gov.uk/documents/ncs-vision-strategy-and-principles.pdf>

³⁷ The National Archives, *Collection Development*, 7.

³⁸ Ibid.

to be prioritised in the collection development policy. In this instance, the university library has already collected material on the theme of race on social media and the Irish slaves meme would be appraised based on its relation to this theme. The library would need to compare its collection to those of other institutions to ensure that there is no duplication of effort and of financial expense on the same theme. This issue feeds into Question 6: 'Will a partnership be useful to safeguard and provide access to this collection?'³⁹ The Irish slaves meme could be a useful resource for multiple areas of research. Thus, the university library could enter a partnership with another such institution with different collection development aims in order to lessen the financial burden and to converge resources. For example, another university library aiming to collect material on the theme of meme culture could undertake to archive the Irish slaves meme in a joint effort with our library. Documenting Ferguson is an example of collaboration between institutions in the creation of digital archives.⁴⁰

The subsequent questions deal with more practical issues, such as the storage and custody of the archive and whether any financial or technical limitations would be an obstacle to its long-term preservation.⁴¹ The guidelines also recommend that the archivist reflect on issues surrounding access restrictions and confidentiality.⁴² These matters will be discussed in greater detail in later chapters of this paper. However, it is important to recognise that these issues are also a key part of the appraisal process.

It is also worth noting some elements that are not mentioned in the TNA toolkit which are relevant when appraising the Irish slaves meme for inclusion in a collection. Colin Post, a student of Information and Library Science, has described the myriad approaches to appraisal

³⁹ Ibid., 8.

⁴⁰ Washington University Libraries has worked in conjunction with other universities and cultural institutions to create a locally-sourced archive of the events surrounding the shooting of Michael Brown by police in Ferguson, Missouri in August 2014. 'Project Explanation and Purpose,' Documenting Ferguson Committee, September 2014, 1. <http://digital.wustl.edu/ferguson/DFP-Plan.pdf>

⁴¹ The National Archives, *Collection Development*, 9-10, 19.

⁴² Ibid., 20-22.

that have been taken in order to create 'living, breathing' digital archives. He points out that a key difficulty in appraising web-based material lies in judging how much contextual information should be archived.⁴³ For example, one website could contain embedded videos or hyperlinks to dozens of other websites. The archivist needs to decide how much of this extra information to collect as part of their appraisal strategy, as removing the website from its context diminishes its archival value.⁴⁴ The interactive nature of social media platforms makes this issue particularly relevant to the Irish slaves meme archive. For instance, one post can contain hundreds of comments from other users, and these comments may contain links to other websites. In Figure 12 below, there are 7,500 comments in total. In appraising this post for inclusion in the archive, the archivist must decide if all, some, or none of the comments should be saved. In theory, only comments that add context to the post should be kept. Those containing irrelevant information should be discarded. However, it is unlikely that an archivist would have the time to read every single comment on every post to determine which ones should be kept. Therefore, the safest course of action would be to keep all comments regardless of content, as to discard all of them would deprive the original posts of valuable context.

⁴³ Post, 'Building a Living, Breathing Archive,' 70.

⁴⁴ Ibid.



Fig 12. Facebook post about the treatment of Irish slaves with 7,500 comments.

A case can also be made for archiving critical social media reactions to the Irish slaves meme. This would be beneficial as it would show how the meme is interpreted by different audiences and would further highlight the context in which the meme has developed. It can be observed that many negative reactions to the meme also employ the same hashtags as those promoting the meme's narrative.⁴⁵ The archive collection policy could also be expanded to include hashtags such as #irishslavesmyth to provide further evidence of critical reactions to the meme. Finally, Cullingford warns against collecting archives based solely on the interests of current researchers affiliated with the university, as there is a risk that the collection will become 'orphaned' if those researchers leave or the department associated with that topic is closed. She advocates for collecting items that offer fresh and unique perspectives so as to attract

⁴⁵ Among the top results of a search of #irishslaves are several Tweets that refute the myth of Irish slavery. Accessed 13 June 2019. <https://twitter.com/hashtag/irishslaves?lang=en>

future researchers and guarantee the relevance of the archive in the future.⁴⁶ Taking this criterion of archival value into account, the Irish slaves meme would earn its place in the university library's collection due to its unique and controversial nature.

Conclusion

This chapter has discussed in detail various concerns with acquiring and appraising the Irish slaves meme archive, including where the archive belongs and within which framework it should be appraised. The chapter has argued that the meme can be effectively appraised by a university library following the guidelines set out by TNA on the basis of the documentation strategy approach. Institutions should be mindful of the increasing relevance of social media to academic research and should expand their collection development policies to reflect this. The following chapter will focus on the next stage of the archival process: harvesting and preservation.

⁴⁶ Cullingford, *The Special Collections Handbook*, 87.

Chapter 3: Harvesting and Preservation

Introduction

This chapter will continue with the assumption from the previous chapter that a university library will be responsible for the creation of the Irish slaves meme archive. As shall be shown, social media archiving presents particular difficulties that other types of digital archiving do not. This chapter will examine the challenges involved in harvesting (collecting) and preserving social media data and will discuss some methods of overcoming these challenges, with the aim of finding the best method for creating the Irish slaves meme archive.

Theoretical Background

Much of the scholarly literature around web archiving is not specifically focused on social media archiving, due in part to the relatively recent evolution of social media as a popular form of communication. As technology and policies develop, there is still much to be written about how archival institutions can engage with the process of acquiring and managing social media data. Literature written as recently as five years ago is already out of date. For these reasons, this chapter will be more practically-based than theory-based. Perhaps more can be learned from examining the strategies employed by social media archiving projects that have been initiated in recent years.

Arguably the most useful information to be found on web archiving has been produced by the Digital Preservation Coalition (DPC), a non-profit agency with the mission of ‘secur[ing] our digital legacy’.¹ The DPC is a collaborative effort of several archival and library agencies across the UK and Ireland. It has released multiple *Technology Watch Reports* and a *Digital Preservation Handbook* in order to provide up-to-date guidelines for institutions to collect and

¹ ‘About’, Digital Preservation Coalition, accessed 13 June 2019. <https://www.dpconline.org/about>

preserve digital materials. Sara Day Thomson (Research Officer at the DPC) has written a *Technology Watch Report* on the subject of social media archiving. This report focuses on the various approaches to harvesting and managing social media data through APIs (Application Programming Interfaces) and it provides much of the secondary material for this chapter.² Thomson argues that the most difficult step in the process of social media preservation is gaining access to this data as it is closely guarded by the platforms that own it.³ Thomson focuses her investigation on Twitter's complex and restrictive policy regarding access and re-use of its data, detailing the obstacles faced by institutions and describing case studies of projects in which Twitter data was successfully harvested.⁴

Harvesting Data: Problems and Possible Solutions

Naturally, the first issue to face is that of collecting the relevant data to create the Irish slaves meme archive. Thomson argues that the most effective method for harvesting social media data is through APIs rather than traditional web crawlers like Heritrix.⁵ She defines an API as an 'interface for communicating with the back end of a social media system and enabling such functions as querying or requesting copies of social media data.'⁶ APIs are more appropriate data harvesting tools for social media platforms given the high level of user interaction and the complex web of information contained on social media web pages. Simply taking a snapshot of a Facebook page at one moment in time is not a complete record of that page. Thomson asserts that the background functions and data of social media platforms have a vital role to play in

² Sara Day Thomson, *Preserving Social Media: DPC Technology Watch Report 16-01*, (Digital Preservation Council, 2016), 7.

³ Thomson, *Preserving Social Media*, 8.

⁴ *Ibid.*, 29.

⁵ *Ibid.*, 7.

⁶ *Ibid.*, 35. For a more detailed explanation of APIs, see Brügger, *The Archived Web*, 66: 'What is available through an API is not the web as it would have appeared in a web browser, but the individual elements making up a web page, such as profile information, images, status updates, and "likes" on social media, as well as information not visible in the web browser—for instance, information about geolocation. In this sense, what was archived through an API was not "the web" in the form of an HTML page, but fragments that could be knit together to form a web page.'

creating a 'more authentic and complete record.'⁷ Each social media platform has its own API, or several types of API. For example, suppose a researcher wants to find out more information about a particular Tweet containing the hashtag #irishslaves. The Twitter API can provide information in the form of raw data about the location of the user, the amount of followers the user has, or the number of likes or shares the post has accumulated.⁸ This is valuable contextual information and the collecting institution would benefit from preserving it when it creates a dataset of the Irish slaves meme.

However, gaining access to this data is not a simple task. Thomson delineates the three main options available to researchers and institutions seeking to collect social media data.⁹ The first option involves using the API provided by the platform to harvest the desired data. If the archival institution does not have the technical expertise to carry this out, another (more expensive) option would be to buy the API data from a third-party reseller or archiving service, such as Gnip or Archive-It. However, both of these strategies come with an unavoidable disadvantage. As they both employ the platform's API to gather data, they must abide by the platform's API terms of use. For example, Twitter allows researchers to access data but prohibits the public sharing of this data by third parties, such as archives or libraries.¹⁰ Thus, the long-term preservation and access to the API-harvested data is severely restricted. Worryingly, Twitter is seen as having the most lenient data use policies of all the major social media companies.¹¹

An alternative strategy involves directly negotiating a partnership with the social media platform for the use and preservation of its data. Although this is only a realistic option for the

⁷ Thomson, *Preserving Social Media*, 7.

⁸ *Ibid.*, 11.

⁹ *Ibid.*, 9.

¹⁰ 'Developer Agreement and Policy,' Twitter, last modified 25 May 2018.

<https://developer.twitter.com/en/developer-terms/agreement-and-policy>

¹¹ Wasim Ahmed, 'Using Twitter as a data source: An overview of current social media research tools,' LSE Impact Blog, 10 July 2015. <https://blogs.lse.ac.uk/impactofsocialsciences/2015/07/10/social-media-research-tools-overview/>

largest institutions, this is not without precedent. In 2010, Twitter gifted access to its full archive to the Library of Congress. However, this project failed in its initial aims to preserve and provide access to all public Tweets as the amount of data collected was simply too large for the library to process and the archive is still inaccessible to researchers.¹² Since 2017, the Library of Congress has narrowed its collection scope and will only archive selected Tweets based on themes or events.¹³ Although the Library of Congress Twitter Archive proved to be too ambitious to succeed, it does provide hope that Twitter is open to engaging with cultural institutions. Twitter also maintains a partnership with the Sunlight Foundation and the Open State Foundation, for the development of Politwoops, a tool which preserves and publishes politicians' deleted Tweets in order to ensure accountability and transparency in public life.¹⁴

The emerging partnership between Facebook and Social Science One, an independent research commission, also demonstrates how social media data can be accessed and utilised by the academic community. As part of this innovative project, Facebook will allow full access to its database to a commission made up of respected scholars from multiple disciplines which then filters the data to chosen researchers based on pre-approved research questions.¹⁵ The focus of the first Facebook/Social Science One project will be on the effect of Facebook on democracy and elections.¹⁶ Facebook further specifies some possible research strands that could be included within the scope of the project such as 'misinformation' and 'polarizing content.'¹⁷ However, this initiative is more concerned with accessing data than collecting and preserving it, as there is no mention in the project explanation paper of any aim to preserve the datasets it creates in the long-term. One can safely assume that this is due to the strict privacy and use

¹² Elisabeth Fondren and Meghan Menard McCune, 'Archiving and Preserving Social Media at the Library of Congress: Institutional and Cultural Challenges to Build a Twitter Archive,' *Preservation, Digital Technology & Culture* 47, no.2 (2018): 39.

¹³ Fondren and Menard McCune, 'Archiving and Preserving Social Media,' 39.

¹⁴ Michelle Conlin, 'Twitter to revive Politwoops, archive of politicians' deleted tweets,' Reuters, 31 December 2015. <https://www.reuters.com/article/us-usa-election-politwoops-idUSKBN0UE16520151231>

¹⁵ Gary King and Nate Persily, 'A New Model for Industry-Academic Partnerships,' Working Paper, last modified 15 January 2019, 6-7. <https://gking.harvard.edu/files/gking/files/partnerships.pdf>

¹⁶ King and Persily, 'A New Model,' 11.

¹⁷ *Ibid.*, 12.

restrictions Facebook imposes on its data.¹⁸ Thus, the Irish slaves meme could be a valid area of research as part of this project given its ‘polarizing content,’ but there is a chance that the dataset created on the meme could be lost after the project is finished. Overall, this partnership is a step in the right direction, but does not do enough to secure the long-term preservation of Facebook data.

Unless the institution is able to engage in a partnership with the social media platform, the only option available for creating the Irish slaves meme archive is to harvest the data through APIs. However, the particular case of the Irish slaves meme provides further obstacles to the already complex process of collecting data. In order to create as complete an archive as possible, the library would need to collect data from at least 2013, when the meme first began to be widely circulated on social media.¹⁹ However, third party archiving services such as Archive-It or Social Feed Manager do not have access to historical data, i.e., data that is more than 30 days old. Despite the fact that data resellers are primarily interested engaging with businesses for current market research, some resellers like Gnip (Twitter’s official data reseller) and Datasift provide access to historical data which can prove useful for heritage institutions.²⁰ Gnip uses APIs to gather data from Twitter, Facebook, and Instagram, but it only offers historical data from Twitter (going back to 2006). From 2012 to 2015, Datasift provided historical data from Facebook.²¹ Unfortunately, there is no longer any option for collecting historical data from Facebook’s API (nor consequently from Instagram’s API, since the platform is owned by Facebook). Therefore, the original scope of the Irish slaves meme archive will have to be narrowed to only include historical data from Twitter. The ability of archivists to fully capture phenomena on social media is severely hindered by social media platforms’ limitations on the periods available for capture through their APIs. If an archival institution cannot afford to purchase historical data from resellers, archivists only have a short window of time in which to

¹⁸ ‘Facebook Platform Policy,’ Facebook for Developers, last updated 25 April 2019.
<https://developers.facebook.com/policy/>

¹⁹ Amend, ‘How the Myth of the "Irish slaves" Became a Favorite Meme of Racists Online.’

²⁰ Thomson, *Preserving Social Media*, 12.

²¹ ‘Historics Archive Schema,’ Datasift Developers, accessed 13 June 2019.

<https://dev.datasift.com/docs/products/stream/features/historics/historics-access/historics-archive-schema>

predict whether certain current events or trending hashtags will be significant in the future and to decide whether or not to harvest data based on this prediction. Making appraisal decisions under these conditions is not ideal and more practical and affordable methods need to be found to prevent loss of important data.

Thus, the university library has no choice but to rely on Gnip to access historical data from Twitter. However, the library has more freedom of choice in relation to acquiring current data on the Irish slaves meme. Data resellers are expensive so the library may find third party archiving services to be a more viable option. Unlike data resellers, third party archiving services are popular providers of social media data for cultural and academic institutions. Archive-It is a subscription service operated by the Internet Archive which has over 200 partners, most of which are academic or cultural heritage organisations.²² The service allows users to create specific collections which they can then catalogue with metadata. Although the data collected is usually made publicly accessible on the Archive-It website, institutions can arrange to keep the collection private (a more desirable option for the controversial and inflammatory Irish slaves meme archive, as discussed earlier). For example, the Virginia Tech Crisis, Tragedy, and Recovery Network used Archive-It to gather its collection of Twitter feeds gathered during the aftermath of the Boston Marathon Bombing in April 2013.²³ However, the Archive-It tool is less effective at harvesting social media data than capturing websites. Its method of capturing websites at particular points in time is more useful if one wishes to access entire Twitter profiles or Twitter feeds at a certain moment. It cannot harvest and preserve a collection of individual Tweets based on keyword or hashtag searches.

Social Feed Manager (SFM) is an open source software package which was developed by the George Washington University (GWU) Libraries. It can be used to harvest data from Twitter,

²² Maureen Pennock, *Web-Archiving: DPC Technology Watch Report 13-01*, (Digital Preservation Coalition, 2013), 21.

²³ 'Boston Marathon Bombing: Twitter and RSS feeds,' Archive-It, accessed 13 June 2019. <https://archive-it.org/collections/3645>

Tumblr, Flickr, and Sina Weibo. Unlike Archive-It, SFM was designed specifically for social media data harvesting. It combines API harvesting with Heritrix, a tool which incorporates the information that is embedded or linked in the tweet, such as images or links to other websites. This results in a more complete set of metadata.²⁴ The Schlesinger Library at Harvard University is currently compiling a collection of digital resources related to the #MeToo movement (including social media posts, news articles, and online discussions) for release in late 2019.²⁵ The Library is employing SFM to conduct weekly harvests of hashtags and plans to purchase eighteen million historical Tweets from the first year of the hashtag's existence.²⁶ The project initially aimed to harvest data from Facebook and Instagram pages using Webrecorder, a tool which captures individual web pages, but the wide scope of the project made this option untenable.²⁷ This is further proof that the Irish slaves meme archive cannot include Facebook and Instagram in its scope until Facebook's API policies change. It remains to be seen how the Schlesinger Library will be able to negotiate the strict terms of Twitter's API policy when it comes to providing access for researchers to the collection.

Taking into account the advantages and disadvantages of each harvesting tool, it would appear that the best course of action for acquiring data for the Irish slaves meme archive is to buy historical data from Gnip and to employ SFM to harvest current data. In the short term, the archive will have to be restricted solely to Twitter data, with the possibility that Facebook and Instagram may be added in the future if their policies change.

Long-Term Preservation

²⁴ Christopher J. Prom, 'Social Feed Manager: Guide for Building Social Media Archives,' 7 June 2017 <https://gwu-libraries.github.io/sfm-ui/resources/SFMReportProm2017.pdf>

²⁵ 'The Collection,' #MeToo Project Schlesinger Library, accessed 13 June 2019. <https://www.schlesinger-metoo-project-radcliffe.org/collection>

²⁶ Nora Caplan-Bricker, 'The Challenge of Preserving the Historical Record of #MeToo,' *The New Yorker*, 11 March 2019. <https://www.newyorker.com/tech/annals-of-technology/the-challenge-of-preserving-the-historical-record-of-metoo>

²⁷ Caplan-Bricker, 'The Challenge of Preserving the Historical Record of #MeToo.'

As we have seen, there are several methods of acquiring historical and current data from Twitter. However, the institution is then faced with another obstacle, that of long-term preservation and dissemination of the data. The developers of SFM state clearly that it cannot function as a preservation tool. The archival institution itself is responsible for ensuring the long-term preservation of the data harvested through SFM's API tool.²⁸ As mentioned previously, Twitter prohibits the public distribution of data harvested from its API. However, it does allow the distribution of Tweet IDs (the unique number given to a single Tweet), user IDs (the number given to Twitter account holders), and up to 50,000 Tweets per day in spreadsheet or PDF format.²⁹

Preserving datasets of Tweet IDs has emerged as a useful method of circumventing the restrictions of Twitter's developer policy. However, harvesting and preserving Twitter data through Tweet IDs is a complicated process. The data first needs to be harvested in its original JSON format.³⁰ To this end, Ed Summers, Lead Developer at the Maryland Institute for Technology in the Humanities, has developed a tool named twarc (a portmanteau of Twitter Archive) which can be used to harvest and archive Twitter data in JSON format using Twitter's API.³¹ The data is presented to the public in the form of Tweet IDs, which can then be hydrated (i.e., reconverted into a readable JSON format) through twarc.³² It is a long and tedious process, but it is the only option currently available to ensure that an institution's Twitter archive can be publicly accessible in the long term. Summers developed the twarc tool to archive Twitter data as part of the DocNow initiative and a catalogue of twarc-harvested Tweet ID datasets is

²⁸ Justin Littman, Daniel Chudnov, Daniel Kerchner, Christie Peterson, Yecheng Tan, Rachel Trent, Rajat Vij and Laura Wrubel, 'API-based social media collecting as a form of web archiving,' *International Journal on Digital Libraries* 19, no.1 (March 2018): 28.

²⁹ Ed Summers, 'Tweets and Deletes: Silences in the Social Media Archive,' On Archivy, Medium, 14 April 2015. <https://medium.com/on-archivy/tweets-and-deletes-727ed74f84ed>

³⁰ JSON displays Tweets in the form of 'attributes,' with fields for the username, the location of the user, the text of the Tweet, etc. For an example of a Tweet converted into JSON format, see:

<https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/intro-to-tweet-json.html>.

³¹ 'Twarc,' GitHub, accessed 13 June 2019. <https://github.com/DocNow/twarc>

³² For a visualisation of how Tweet IDs can be hydrated back into JSON format, see: <https://developer.twitter.com/en/docs/basics/twitter-ids.html>

available on the DocNow website.³³ SFM also incorporates twarc into its harvesting mechanism. The major drawback of this method of preservation is that the Tweet IDs of Tweets that were deleted after the point of harvesting cannot be hydrated at a later date; thus, that material is lost and inaccessible to researchers.

This strategy was tested during the Social Repository of Ireland project.³⁴ The Social Repository filtered data through Twitter's API using hashtags and keyword searches, similar to the approach proposed in this paper for collecting the Irish slaves meme. In order to work within Twitter's strict data policy, the Social Repository proposes that its archive can only be accessed via log-in by verified academics affiliated with an institution and would impose limits on the number of spreadsheet downloads per user.³⁵ There have been no further updates on the future of the Social Repository of Ireland since the study's end in 2015. Nevertheless, this does appear to be a promising model for collecting and preserving social media data.

However, it is debatable whether or not the Irish slaves meme should be publicly accessible. There are certain ethical concerns, such as privacy and the sensitivity of the material, that would discourage an institution from allowing the public to freely access the data. These issues will be discussed in much greater detail in Chapter 5. There are some examples of institutions limiting access to their Twitter archives which can provide some inspiration for the Irish slaves meme archive. During May and June 2014, the UNT Digital Library compiled a Twitter dataset of the #YesAllWomen debate around misogyny and violence against women.³⁶ As is the norm, the library offers unlimited access to its collection of Tweet IDs. However, it restricts use of its collection of Tweets in JSON format, requiring users to log in via an institutional account. This

³³ 'Tweet ID Datasets,' Documenting the Now, accessed 13 June 2019. <https://www.docnow.io/catalog/>

³⁴ This was a one-year feasibility study conducted over 2014/2015 by the Insight Centre for Data Analytics at NUI Galway and the Digital Repository of Ireland. The study tested the possibility of creating 'a richly annotated archive of social media responses to major events and topics of interest in Ireland, and to preserve this archive in a trusted digital repository.' Thomson, *Preserving Social Media*, 29.

³⁵ Ibid.

³⁶ Mark Edward Phillips, "Yes All Women" Twitter Dataset, UNT Digital Library, June 2014. <https://digital.library.unt.edu/ark:/67531/metadc304853/>

method ensures the long-term preservation of the harvested data in JSON format and privileges researchers affiliated with the university as they can access Tweets that may have been deleted and thus cannot be hydrated through the publicly-available Tweet IDs. The University of Miami does not allow the public to access its Twitter Archive of the online reaction to the normalisation of diplomatic relations between the USA and Cuba in 2014, restricting access solely for the University of Miami community.³⁷ This method is perhaps the most efficient as the institution does not have to devote time and resources to generating a collection of Tweet IDs for public use; the archive need only exist in its original JSON format since it is not being publicly disseminated. If access to the Irish slaves meme archive was restricted in this way, it would also limit the archive's potential use as a propaganda tool for white supremacists.

Conclusion

Current restrictions on the use of social media data greatly hinder cultural and academic institutions' efforts to preserve internet culture. In the absence of an overall strategy that incorporates the harvesting and preservation of both historical and current data, this chapter recommends mixing and matching different tools and methods to create the Irish slaves meme archive. Going forward, archives and other cultural institutions should campaign for social media companies to amend their data use policies in order to provide greater opportunities for the preservation of social media data, as it is a vital record of twenty-first century life. The next chapter will consider how arrangement and description methods can contribute to enriching this record with meaning and context.

³⁷ 'Finding Aid Collection CHC5404 - 2014 US-Cuba Policy Change Twitter Archive,' University of Miami Cuban Heritage Collection, University of Miami Libraries, last updated January 2019.
<https://atom.library.miami.edu/chc5404>

Chapter 4: Arrangement and Description

Introduction

As stated in the previous chapter, Twitter data will have to form the basis of the Irish slaves meme archive due to Facebook and Instagram's restrictions on use of its data through APIs. Thus, the aim of this chapter is to suggest the best method for arranging and describing the Irish slaves meme Twitter archive. The topics of arrangement and description of social media archives have received little to no consideration by archival scholars. Hopefully, this paper will be one of many future attempts to theorise on different possible approaches to arranging and describing social media data. In each of its two sections, this chapter will provide a theoretical background to arrangement and description in the digital age. This chapter will also evaluate the practice of arrangement and description in several existing Twitter archives in order to determine the most practical and effective approach to arranging and describing the Irish slaves meme archive.

Section I: Arrangement

i. Arrangement in Theory

The theories of provenance and original order (or *respect des fonds*), once paramount to the physical arrangement of archival material, have come under scrutiny in recent decades. In 1993, Terry Cook was one of the first archival scholars to question the assumption that the archival fonds is a product of one single provenance (or creator) and thus can only be arranged according to this provenance.¹ Cook asserts that archivists should see the fonds as a malleable concept rather than a static physical entity.² Although Cook's work primarily focuses on the

¹ Terry Cook, 'The Concept of the Archival Fonds in the Post-Custodial Era: Theory, Problems and Solutions,' *Archivaria* 35 (Spring 1993): 26.

² Cook, 'The Concept of the Archival Fonds,' 31-32. This is especially true for digital archives, where the traditional static 'record' has given way to data which can be produced and manipulated by a variety of sources and which gets its meaning from its relationships with other data and from the larger context of its creation and use.

implication of the conceptual fonds for archival description, others have applied his thinking to arrangement.

Geoffrey Yeo has gone a step further than Cook by arguing that the fonds is a conceptualisation and 'aggregations' of records are collections based on this conceptualisation.³ He denies the existence of any organic original order based on provenance, asserting that all archival collections are created and curated by humans and are thus not naturally occurring, despite what traditional archival theory has promulgated.⁴ This opens up the possibility for multiple overlapping fonds which means the records can be arranged into a variety of different orders, easily enabled by the technology of the digital age.⁵ On the arrangement of digital archives, Yeo expands on David Weinberger's theory of the 'third order of order.'⁶ In this system, users are free to make their own aggregations out of the collections provided by the archive, and no one single order is 'inevitable' nor can be made more legitimate than others, including the order imposed by the creator.⁷ The collections envisaged by Yeo are dynamic and are not restricted by boundaries or hierarchies of provenance; collections can be assembled based on users' needs and records can exist in multiple collections at one time.⁸ Yeo echoes Cook's assertion that the arrangement of records is far less important than their contextual relationships with other records.⁹ Yeo's arguments follow the turn in archival theory away from arrangement and towards description as the most valuable tool for contextualising and thus giving meaning to records in the digital era.¹⁰ However, Yeo does see the benefit in documenting the original

³ Geoffrey Yeo, 'Bringing Things Together: Aggregate Records in a Digital Age,' *Archivaria* 74 (Fall 2012): 46.

⁴ Yeo, 'Bringing Things Together,' 48.

⁵ *Ibid.*, 49.

⁶ *Ibid.*, 58. Cook asserts that Weinberger's 'third order of order' theory dictates that 'resources can be arranged into as many sequences as may be desired and users can organize their work independently of the limitations imposed by analog systems.'

⁷ *Ibid.*, 65.

⁸ *Ibid.*, 69.

⁹ *Ibid.*, 71.

¹⁰ Yeo, 'Continuing Debates about Description,' 166.

order of the collection, as it may give some key contextual information and add meaning to the records.¹¹ This view is echoed by Jane Zhang.¹²

ii. Arrangement of Social Media Archives in Practice

This section will examine how social media data is arranged in online repositories, with a view to finding the best arrangement for the Irish slaves meme archive. As this archive would most likely have to be presented to the public as a dataset of Tweet IDs due to Twitter's API terms of use, it would be prudent to compare how such datasets are arranged by other institutions, using the catalogue of Tweet ID datasets provided on the DocNow website as a framework.

Although DocNow has contributed to the development of the twarc tool, the project does not dictate how the datasets should be arranged. The datasets themselves are hosted on other websites, either those of the collecting institutions or the Internet Archive website. Therefore, each collecting institution is given the freedom to arrange its datasets according to its own wishes. Tweet ID numbers are assigned by Twitter based on the time of posting.¹³ This is the basic order in which the data is presented. However, twarc allows institutions to filter results and thus create collections based around different criteria, such as keywords, location, or even gender of the data creator.¹⁴ The Ferguson Tweet ID dataset, hosted on the Internet Archive website, contains two collections of Tweet IDs harvested during two periods using the search term 'ferguson.' This enormous aggregation of over 28 million Tweet IDs does not appear to have undergone any further filtering or rearranging after the point of capture.¹⁵ GWU Libraries compiled a collection of Tweet IDs on the subject of the 2018 referendum to repeal the Eighth Amendment of the Irish Constitution prohibiting abortion. This collection is notable in that it includes Tweets from both sides of the debate, using search terms such as #repealthe8th and

¹¹ Yeo, 'Bringing Things Together,' 77.

¹² Jane Zhang, 'Original Order in Digital Archives,' *Archivaria* 74 (Fall 2012): 186.

¹³ 'Twitter IDs (snowflake),' Basics, Twitter Developers. Accessed 13 June 2019.

<https://developer.twitter.com/en/docs/basics/twitter-ids.html>

¹⁴ 'Twarc,' GitHub, <https://github.com/DocNow/twarc>

¹⁵ 'Ferguson Tweet IDs,' Internet Archive, accessed 13 June 2019. <https://archive.org/details/ferguson-tweet-ids>

#savethe8th. However, the IDs are not separated according to these hashtags and are gathered in one collection without any apparent arrangement strategy.¹⁶ There are some advantages to keeping the entire dataset in one file. For instance, not dictating the perimeters of the collection beyond the point of harvesting offers more freedom to researchers. Once researchers hydrate the IDs, they can search the data using their own criteria. The order in which the data will appear will depend on the search terms used. Therefore, the so-called ‘original order’ of the dataset does not have much significance.

The UNT Digital Library’s ‘Yes All Women’ collection is arranged into separate files according to the hashtags used. Tweet IDs containing image URLs are also contained in separate files.¹⁷ This approach to arrangement makes sense as it reflects the various conflicting sides of the conversation. However, it might also be useful to gather the data into one file so that researchers would be able to easily examine how different Tweets relate to each other. For example, a Tweet in the #NotAllMen collection could have been a reply to a Tweet containing the hashtag #YesAllWomen. If both Tweet IDs were in the same collection, it would be easier to see the connection between them and would give more context to the conversation.

There are several methods of arranging Tweet IDs within a collection, with each method having benefits and drawbacks depending on how the data is used. There are some issues that are particular to the Irish slaves meme that need to be considered when deciding how to arrange the Tweet IDs for the archive. If the collecting institution employs both keyword searches and hashtag searches to gather data, it is possible that two separate terms can be contained within one post. For example, a Twitter user could mention the word “reparations” and also use the hashtag #irishslaves in the same Tweet. This can have consequences if the resulting dataset is separated based on different search terms, as is the case with the ‘Yes All Women’ dataset. If

¹⁶ Justin Littman, ‘Ireland 8th Tweet IDs,’ Harvard Dataverse, V1, accessed 13 June 2019.
<https://doi.org/10.7910/DVN/PYCLPE>

¹⁷ ‘Items List,’ “Yes All Women” Twitter Dataset, UNT Digital Library, accessed 13 June 2019.
<https://digital.library.unt.edu/ark:/67531/metadc304853/m1/>

one was to apply Weinberger's 'third order' theory, it is possible for the same Tweet ID to exist in multiple collections. Nevertheless, it would be useful if some connection could be made between the two instances of the ID, perhaps via metadata. A more difficult problem to overcome is related to images. As mentioned in the Introduction, the Irish slaves meme is often shared in the form of images. Separating the images from the text of the post could diminish the meaning of the post as it strips it of some of its context. The most desirable solution to this dilemma is to keep the entire dataset in one collection and to ensure that its metadata is fully searchable so that users can filter the data depending on their own criteria. Metadata is vital to ensuring that contextual links can be made between items and collections, as this is where most of the meaning is to be found. The next section will focus more thoroughly on descriptive metadata practices.

Section II: Description

i. Approaches to Metadata

This section will investigate different approaches to describing social media archives through metadata. Digital curation scholar Stacy T. Kowalczyk identifies three main types of metadata: descriptive metadata for search and discovery of items or collections; structural metadata which informs the user about the relationships between various parts of the digital object; and administrative or preservation metadata which documents the technical details of the digital object, such as its format or size.¹⁸ While all three types of metadata are necessary for providing context for the archival objects, the importance of each type of context varies depending on the needs of the user.¹⁹ Kowalczyk further argues that traditional standards of descriptive metadata do not go far enough to contextualise digital objects. Robust technical and structural metadata is essential to ensure the long-term usability of digital archives.²⁰ Zhang has found that item-level administrative and structural metadata is usually automatically-generated

¹⁸ Stacy T. Kowalczyk, *Digital Curation for Libraries and Archives* (Santa Barbara, California Denver, Colorado: Libraries Unlimited, an imprint of ABC-CLIO, LLC, 2018), 52-53.

¹⁹ Kowalczyk, *Digital Curation*, 53.

²⁰ *Ibid.*, 51.

by the creator or recordkeeping system, thus archivists tend to focus their manual descriptive efforts on the higher level of series or collections.²¹ Consequently, this chapter will focus on possible approaches to creating descriptive metadata for the Irish slaves meme archive rather than administrative or structural metadata, as this is the area which most directly involves the archivist.

Yeo and others have criticised the current widely-followed standards for description as they reinforce hierarchical structures that do not take account for the fluidity of relationships between records in the digital era.²² Scholars have put forward myriad possible approaches to description, although the question of provenance remains of paramount importance in all of these approaches.²³ The postmodern approach can be seen as the most relevant to the description of social media data. Tom Nesmith, a proponent of this approach, asserts that 'societal provenance' should form the backbone of description of records.²⁴ This perspective frees description from its hierarchical structure and reflects the broader web of stakeholders involved in the creation of a record.²⁵

The societal provenance method is particularly suitable for describing the Irish slaves meme archive. The records contained in this archive come from thousands of separate creators but their provenance can be linked by the overarching phenomenon of the myth of Irish slavery. For example, the provenance of the Irish slaves meme archive could be extended to include the Black Lives Matter movement, as many of the posts on the subject of the Irish slaves myth are in response to this movement. Thus, while Black Lives Matter has not directly created the Irish

²¹ Zhang, 'Original Order,' 190.

²² Yeo, 'Continuing Debates about Description,' 172.

²³ *Ibid.*, 164.

²⁴ Nesmith's 'societal provenance' method expands the traditional notion of provenance to include 'all of those people, organizations, and entities whose decisions and actions account for the records' existence, characteristics and continuing history.' Tom Nesmith, 'The Concept of Societal Provenance and Records of Nineteenth-Century Aboriginal-European Relations in Western Canada: Implications for Archival Theory and Practice,' *Archivaria* 60 (2006): 352.

²⁵ Yeo, 'Continuing Debates about Description,' 175.

slaves meme, it is an important part of the societal context in which the meme has evolved and this should be indicated in the description of the archive.

Parallel to the debate around provenance is the discussion of how metadata can adapt to the demands and opportunities of digital technology. Since items in a digital repository can be easily reorganised into countless different collections by users depending on their needs, there is a greater need for relational metadata for digital items than their analogue counterparts.²⁶ Yeo postulates that relational descriptive systems are the most effective means of representing the complex relationships between items in a digital collection.²⁷ Indeed, many scholars have begun to emphasise the importance of item-level metadata for technical as well as contextual preservation. Zhang and Mauney argue that systems need to be developed to increase the granularity (level of detail) of contextual information available at item level, as this will add ‘a new dimension of digital discoverability’ and contextualisation to digital records.²⁸ However, resources are often not available to give in-depth item-level metadata and institutions frequently rely solely on automatically-generated item-level technical metadata provided by harvesting tools such as SFM or ArchiveSocial.²⁹

The OCLC Research Library Partnership Web Archiving Metadata Working Group (WAM) was formed to review existing standards of description of web-based materials in libraries and archives and to recommend best practices for composing finding aids with descriptive metadata for web archives. The WAM’s review of SFM is particularly relevant for this paper as this tool was highlighted in the previous chapter as the most suitable one for creating the Irish

²⁶ Ibid., 180.

²⁷ Ibid., 181.

²⁸ Jane Zhang and Dayne Mauney, ‘When Archival Description Meets Digital Object Metadata: A Typological Study of Digital Archival Representation,’ *The American Archivist* Vol 76 No. 1 (Spring/Summer 2013): 191.

To clarify, an ‘item’ in the Irish slaves meme archive would be one single Tweet. A collection could be made up of all Tweets containing a certain hashtag. If sufficient granular metadata is given to each Tweet in each collection (such as geolocation, time of posting, number of followers of the creator of the Tweet), a researcher could make connections between different Tweets in different collections based on particular criteria. Thus, granular and relational descriptive metadata can work together to make the data more useful to researchers.

²⁹ Yeo, ‘Continuing Debates about Description,’ 183

slaves meme archive. The WAM describes the descriptive strengths and weakness of SFM as an API data-harvesting tool. As SFM is still in its development stage, its focus so far has been on automatically generating provenance metadata as this is the most important metadata needed at the collection stage. Therefore, descriptive metadata at item level is not supported by its system.³⁰ However, SFM does support descriptive metadata in other ways by generating a change log which records when collections are made or changed and comments can be added to this log by the archivist.³¹ This element of the tool promotes transparency of archival practices and allows archivists to document their actions and acknowledge their potential biases when creating and arranging collections. The documentation of the archivist's intervention in the archiving process has been identified by postmodernist archival scholars as a vital function of description.³²

The WAM includes a 'Description' section as part of its guidelines for descriptive metadata as this is where 'the content and context of the site or collection can be most clearly articulated.'³³ The most essential part of the description of the Irish slaves meme archive would be the explanation of the institution's rationale for creating the collection and the contextual information about the theme of the archive. As with other controversial archival collections, there are sure to be those who would object to the very existence of the meme archive.³⁴ It is therefore important for the institution to justify the creation of the archive and to provide in-depth background information about the origins, evolution, and effects of the myth of Irish slavery. If technology does not yet allow for detailed descriptive metadata for each item in the

³⁰ Mary Samouelian and Jackie Dooley, *Descriptive Metadata for Web Archiving: Review of Harvesting Tools*, (Dublin, OH: OCLC Research, 2018): 17.

³¹ Samouelian and Dooley, *Descriptive Metadata for Web Archiving*, 17.

³² Yeo, 'Continuing Debates about Description,' 174.

³³ In this section, the WAM strongly suggests that the institution provides the following information: its rationale for creating the collection, background information about the theme or creator(s), and a statement of scope and content (i.e., an abstract). Jackie Dooley and Kate Bowers, *Descriptive Metadata for Web Archiving: Recommendations of the OCLC Research Library Partnership Web Archiving Metadata Working Group* (Dublin, OH: OCLC Research, 2018): 22.

³⁴ Boles, "'Just a Bunch of Bigots,'" 56.

archive, archivists should ensure that as much contextualisation as possible is provided in the description of the collection as a whole.

It can be argued that widespread standards for descriptive metadata are unnecessary and flexibility is key to being able to adapt description to the changing nature of records and users' interactions with them.³⁵ Thus, in the following comparison of metadata supplied by social media archives, the only standard that will be used as a criterion of judgement will be the extent of the metadata, i.e., the more information provided, the better. Context is key, especially for the Irish slaves meme archive.

ii. Evaluation of Descriptive Metadata for Social Media Archives

This section will compare the depth and effectiveness of descriptive metadata of several social media archives to identify a suitable example for the Irish slaves meme archive. The metadata requirements for inclusion in the DocNow Twitter ID Dataset catalogue are quite basic.³⁶ The level of metadata available to users varies depending on what is provided by the website which hosts the dataset. The dataset of #UniteTheRight Tweet IDs is housed on the Internet Archive website.³⁷ The metadata contains links to other collections based on its subject tags and describes the scope and content of the collection. However, the description is lacking some contextualisation of the background of the Unite the Right rally in Charlottesville, Virginia in 2017 as this is limited to one paragraph. The rally was organised by white supremacists and a significant proportion of Twitter users using the hashtag #unitetheright during that period expressed racist views.³⁸ A collection of such incendiary material produced during a highly

³⁵ Yeo, 'Continuing Debates about Description,' 173.

³⁶ One is only obliged to add the collection's title, creator, a URL to the website hosting the dataset, the publishing date, the number of Tweet IDs included, the dates of capture, tags, and a description of the background and/or contents. 'Tweet ID Datasets,' Documenting the Now. <https://www.docnow.io/catalog/>

³⁷ Ed Summers, 'UniteTheRight Tweet IDs,' 2017, Internet Archive. <https://archive.org/details/unitetheright-ids>

³⁸ Julia Ebner and Jacob Davey, 'The far right has learned to mobilise and radicalise. Charlottesville's a wake-up call,' *The Guardian*, 14 August 2017. <https://www.theguardian.com/commentisfree/2017/aug/14/far-right-charlottesville-mobilise-radicalise-white-supremacists-coalition>

contentious period in American history would merit a much more detailed description of its background and context.

The Tweet ID datasets harvested through SFM and preserved in the GWU Libraries/Harvard Dataverse all contain README.txt files which provide additional information on the decisions made when collecting the data.³⁹ This lends transparency to the archival process that is lacking in the metadata offered by the Internet Archive website. These datasets also direct users on how to retrieve the data using hydration software and indicate the restrictions on re-use of the data; the Unite the Right metadata does not provide this information. The WAM asserts that users desire more information about the provenance of archives, including documenting the institution's decision-making process relating to acquisition and appraisal, the exact dates of capture and any changes that occurred to the collection over time.⁴⁰ Additionally, the University of California Guidelines for Born-Digital Archival Description state that processing information is the most important component of a finding aid for a born-digital collection.⁴¹

Social media datasets hosted on university library websites appear to offer the most extensive metadata of the examples studied for this paper. For example, the University of Nevada, Las Vegas (UNLV) Libraries provides a detailed collection guide for its archive of Tweet IDs on the subject of the mass shooting in Las Vegas in October 2017.⁴² This guide describes the historical background of the archive, its scope and contents, processing information, and an inventory of the archive's contents. Hyperlinks to other UNLV collections on the same subject are also

³⁹ For example, see Laura Wrubel, Justin Littman and Dan Kerchner, '2018 U.S. Congressional Election Tweet Ids', Harvard Dataverse, V1, accessed 13 June 2019. <https://doi.org/10.7910/DVN/AEZPLU>; Littman, 'Ireland 8th Tweet Ids.'

⁴⁰ Dooley and Bowers, *Descriptive Metadata for Web Archiving*, 12.

⁴¹ The authors argue that 'decisions made during processing can greatly affect who, what, where, when, why and how researchers access and understand the digital material within a given collection.' Annalise Berdini, Charles Macquarie, Shira Peltzman, and Kate Tasker, 'Describing Digital: The Design and Creation of a Born-Digital Archival Description Standard at the University of California Libraries,' *Journal of Western Archives* 9, no.1 (2018): 8. <https://digitalcommons.usu.edu/westernarchives/vol9/iss1/10>.

⁴² 'Guide to the Web Archive on the October 1, 2017 Shooting in Las Vegas, Nevada,' University of Nevada, Las Vegas. University Libraries. Special Collections and Archives. (2018). <https://www.library.unlv.edu/speccol/finding-aids/MS-00866.pdf>

embedded in the metadata, providing relational context for the collection.⁴³ In a similar fashion, the University of Miami offers an extensive finding aid for its 2014 US-Cuba Policy Change Twitter Archive.⁴⁴ Although the collection is not accessible to the general public, the finding aid contains detailed information about the historical background, scope, and limitations of the archive. This gives researchers a window into the process of assembling the collection as well as the rationale for creating the collection in the first instance.

The collections that have been examined are all described at collection level. Due to the relatively complex process involved in converting the Tweet IDs into a readable format and the large amount of data contained in each collection, it has not yet been possible for collecting entities to add descriptive metadata at item level. Given the limited technical capabilities and legal restrictions on releasing Twitter data to the public, it appears that limiting extensive descriptive metadata to collection level and allowing harvesting tools to automatically generate technical metadata at item level is currently the best strategy for describing the Irish slaves meme archive.

Conclusion

Arrangement and description can contribute in vital ways to providing meaning and context to records in a collection. As has been shown, there are certain strengths and weaknesses in the capacity of web archiving tools for ordering and describing data. Ideally, the Irish slaves meme archive would be arranged and described in such a way that the context and relationships between data could be easily found by researchers. Of course, there are some ethical concerns relating to providing extensive descriptive metadata about the sensitive data contained in the Irish slaves meme archive. This and other ethical issues will be discussed further in the next chapter.

⁴³ 'Tweet Identifiers 2017.' UNLV University Libraries Digital Collections, accessed 13 June 2019. <http://d.library.unlv.edu/digital/collection/p17304coll5/id/20087/rec/1>

⁴⁴ 'Finding Aid Collection CHC5404 - 2014 US-Cuba Policy Change Twitter Archive,' University of Miami Libraries.

Chapter 5: Ethics

Introduction

Ethics have long been a concern of the archival profession, with every national association of archivists having their own code of ethics. However, the rapid growth of Web 2.0 technologies has created challenging new ethical questions for archivists. Privacy and access present the most pressing ethical dilemmas for collectors and users of social media archives. For several reasons, ethics should be a key concern for the institution which will create the Irish slaves meme archive. As the archive will contain sensitive social media data from creators expressing controversial or hateful views, it can be argued that the privacy of these creators needs to be protected. In addition, the ethical mediation of access to the archive is of utmost importance so that the archive does not become a source of propaganda or a means to further spread the false narrative of Irish slavery. Firstly, this chapter will examine some ethical dilemmas around privacy and access. The chapter will then proceed to review some solutions to these dilemmas, evaluating their efficacy and feasibility. Throughout the chapter, the particular ethical challenges involved in the Irish slaves meme archive will be given special consideration.

Ethical Issues

Much of the academic literature discussed in this chapter focuses on ethical use of social media data from the perspective of researchers as opposed to archivists. However, many of the issues and solutions presented in these works can also be applied to an institution engaged in collecting and preserving social media data. Social science has been at the forefront of the ethical debate around the privacy given to social media users, partly due to the discipline's reliance on data taken from social media for research. Social scientists have put forward several sets of guidelines on how to use and present social media data in an ethical fashion.¹

Unfortunately, the field of archival studies has been slower to tackle the ethical challenges involved in social media preservation. This paper will be one of the first to specifically focus on the ethics of archiving problematic social media data.

¹ Two examples are: Leanne Townsend and Claire Wallace, 'Social Media Research: A Guide to Ethics,' University of Aberdeen (2016). Available at: https://www.gla.ac.uk/media/media_487729_en.pdf. Annette Markham and Elizabeth Buchanan, 'Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0)' (2012). Available at: <http://aoir.org/reports/ethics2.pdf>

Social media users' privacy has been a divisive area of debate among academics. Some have argued that if a user's profile is public, then their data is also public and researchers should not have any ethical qualms about freely using this data without attempting to protect the privacy of individuals.² However, others have rejected this reasoning, arguing that although researchers may have the legal right to access and publish individuals' data if it is public, they have an ethical imperative not to publish sensitive information relating to private citizens.³ The issue is further complicated by the social media platforms' ownership of their users' data. This means that social media users themselves are a third party in the relationship between the social media platform and the collecting institution. They are not given full control over how their data can be used and this can compromise their privacy. The code of ethics adopted by the International Council on Archives (ICA) in 1996 addresses the issue of preserving third-party privacy rights.⁴ Therefore, although archivists may legally be allowed to collect data, they have an ethical imperative to respect the creator's privacy. However, the degree to which this is possible or even desirable is debatable.

In fact, the very notions of 'public' and 'private' are contested in the social media era. Scholars argue that although social media users create posts on a public platform, they do not necessarily intend their information to be collected by researchers and published in academic papers.⁵ The 'contextual norms' of each social media platform dictate users' expectations of privacy and re-use of their data. Research suggests that despite the public nature of platforms

² Matthew L Williams, Pete Burnap and Luke Sloan, 'Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users' Views, Online Context and Algorithmic Estimation,' *Sociology* 51, no. 6, (2017): 1151.

³ Michael Zimmer, "'But the data is already public': on the ethics of research in Facebook,' *Ethics and Information Technology* 12, no.4 (December 2010): 323. <https://doi.org/10.1007/s10676-010-9227-5>

⁴ International Council on Archives, 'Code of Ethics,' 6 September 1996. https://www.ica.org/sites/default/files/ICA_1996-09-06_code%20of%20ethics_EN.pdf. Paragraph 7 states that archivists 'must respect the privacy of individuals who created or are the subjects of records, especially those who had no voice in the use or disposition of the materials.'

⁵ Heather Small, Kristine Kasianovitz, Ronald Blanford and Ina Celaya, 'What Your Tweets Tell Us About You: Identity, Ownership and Privacy of Twitter Data,' *The International Journal of Digital Curation* 7, no.1 (2012): 184. doi:10.2218/ijdc.v7i1.224

such as Twitter, many users create posts with a certain ‘imagined audience’ in mind, such as friends or people who share their political views.⁶ Rosenfeld asserts that this is more common than ever in the post-truth era. People are becoming more and more distrustful of those who do not share their opinions and thus frequently seek to limit their social media interaction to ‘insular, homogeneous intellectual communities’ of like-minded individuals.⁷ Users may be disturbed to find their views extracted from this context and used by archivists and researchers for purposes beyond their original intent. In this way, the boundaries between public and private are blurred by social media and thus the question of privacy is ethical as well as legal. Privacy and social media scholar Michael Zimmer argues that those who defend the indiscriminate harvesting of social media data based on the apparent public nature of the data create a ‘false dichotomy that information is either strictly public or private.’⁸

On the other hand, it has also been argued that if a social media user includes a hashtag in their post, they intend for that post to be a part of a larger conversation and will consequently have a much wider audience than a post that does not contain a hashtag.⁹ Therefore, a Tweet containing a hashtag can be viewed as being more public than other types of Tweets and this can help to simplify the question of public versus private. The institution responsible for archiving the Irish slaves meme could adopt this approach and limit its scope to solely hashtag-based collecting and avoid collecting based on keyword searches in order to protect creators’ privacy. As well as the expectation that their posts will only be viewed by an imagined audience, social media users also often presume that they will be approached by researchers (or collecting institutions) for permission to collect and re-use their data, and that this data will be anonymised before publication.¹⁰ In this context, scholars have emphasised the importance

⁶ Small et al., ‘What Your Tweets Tell Us About You,’ 184.

⁷ Rosenfeld, *Democracy and Truth*, Chapter 4, no page number.

⁸ Michael Zimmer, ‘The Twitter Archive at the Library of Congress: Challenges for information practice and information policy,’ *First Monday* 20, number 7 (June 2015). <https://doi.org/10.5210/fm.v20i7.5619>

⁹ Townsend and Wallace, ‘Social Media Research,’ 8.

¹⁰ Williams, Burnap and Sloan, ‘Towards an Ethical Framework for Publishing Twitter Data in Social Research,’ 1154.

of researchers and collecting institutions gaining the ‘informed consent’ of social media users before collecting and making their data publicly accessible.¹¹

The failure to protect the privacy of social media users’ data can have serious consequences. For example, if a user posts controversial or sensitive material online (such as racially-charged post on the topic of ‘Irish slavery’) and this data is preserved without anonymisation, there is a risk of stigmatisation or actual bodily harm if the user’s name and location can be identified through the metadata preserved in the archive.¹² The DocNow project white paper highlights the potential risk to social activists of being targeted by law enforcement if their social media data is preserved without protection of privacy.¹³ On the other hand, archival student Antoinette E. Baker postulates that concerns for privacy can be overstated and that sometimes an invasion of privacy is justified when it is done in defence of society’s values.¹⁴

Linked to the issue of privacy is that of access. Collecting institutions can further uphold the privacy of users’ data by restricting access to its collections. Restrictions on access can also be motivated by other ethical concerns. Pamela Graham asserts that access to collections of ‘morally abhorrent’ material such as hate speech should be heavily mediated so as not to ‘amplify the reach and impact’ of these messages.¹⁵ In relation to the proposed digitisation of Ku Klux Klan newspapers, librarians Chelcie Juliet Rowell and Taryn Cooksey warn that making the material openly available to the public runs the risk of turning such collections into ‘tools of

¹¹ Ibid., 1153.

¹² Ibid., 1152.

¹³ Bergis Jules, Ed Summers and Vernon Mitchell, Jr., ‘Documenting The Now White Paper. Ethical Considerations for Archiving Social Media Content Generated by Contemporary Social Movements: Challenges, Opportunities, and Recommendations,’ (April 2018): 3.

¹⁴ Antoinette E. Baker, ‘Ethical Considerations in Web 2.0 Archives,’ *SLIS Student Research Journal*, 1, no. 1 (2011): 3. <https://scholarworks.sjsu.edu/ischoolsrj/vol1/iss1/4/>

¹⁵ Pamela M. Graham, ‘Guest Editorial: Reflections on the Ethics of Web Archiving,’ *Journal of Archival Organization* 14 nos.3-4 (2017): 107. <https://doi.org/10.1080/15332748.2018.1517589>

contemporary communities of white supremacy.’¹⁶ The Data & Society research institute has preserved a collection of white supremacist online materials that it has used to inform its reports on media manipulation. The institute is currently grappling with the ethical dilemma of how to provide this material as a resource to researchers while avoiding the ‘amplification’ of these groups or unwittingly providing them with ‘new archival methods.’¹⁷ This is also a principal ethical issue for the creation of the Irish slaves meme archive; access to the material within that collection will need to be carefully considered to avoid further promoting its false narrative. Possible solutions to these issues will be proposed in the next section.

Possible Solutions

One of the most commonly-cited ways to bridge the gap between users’ expectations of the re-use of their data with the goals of the archive is to seek the informed consent of creators before releasing their data to researchers. There are different methods of attaining informed consent. Baker suggests that the social media platform make users aware that their data could be collected by an archival institution and it should give them an opportunity to opt-in or opt-out of being a part of a collection by checking a box.¹⁸ Zimmer criticises the Library of Congress Twitter Archive initiative for failing to afford Twitter users the opportunity to decline to be a part of the archive.¹⁹ Bergis Jules, co-founder of DocNow, suggests that archivists take inspiration from the theory behind archival deeds of gift and impose restrictions on the use of data in the interest of protecting users who may not be able to give their informed consent. For instance, the archival institution could decide to only keep the data for a limited number of years before it is destroyed.²⁰ Heather Small et al. suggest that social media archives should

¹⁶ Chelcie Juliet Rowell and Taryn Cooksey, ‘Archive of Hate: Ethics of Care in the Preservation of Ugly Histories,’ *Lady Science*, 2019. Accessed 13 June 2019. <https://www.ladyscience.com/blog/archive-of-hate-ethics-of-care-in-the-preservation-of-ugly-histories>

¹⁷ ‘Ethics and Archiving the Web: Documenting Hate,’ Vimeo, 27 June 2018. 27:25.

¹⁸ Baker, ‘Ethical Considerations,’ 11.

¹⁹ Zimmer, ‘The Twitter Archive at the Library of Congress.’

²⁰ Bergis Jules, ‘Some Thoughts on Ethics and DocNow,’ *Documenting DocNow*, 3 June 2016. <https://news.docnow.io/some-thoughts-on-ethics-and-docnow-d19cfec427f2>

provide an option for users to request that their data be removed from the collection.²¹ This is a way of gaining consent after the collection has already been made available to the public.

The issue of obtaining informed consent is made more complicated by the nature of the content being collected in the Irish slaves meme archive. Drawing from the 'imagined audience' argument, it can be debated whether any users would actively consent to having their controversial opinions preserved in an archive for research use. The SFM guidelines for building social media archives state that 'in many cases it is not desirable or feasible to obtain explicit consent from records creators,' such as members of hate groups.²² Social scientists Leanne Townsend and Claire Wallace argue that it can even be dangerous for researchers (or, in this case, archivists) to contact members of hate groups in order to obtain their consent for inclusion in a data collection. They further assert that too much emphasis on social media research ethics in this type of situation could lead to 'censorship of critical research' as these users are unlikely to want their data critically analysed by researchers.²³ Graham raises the valid question of whether the moral and cultural imperative to document hate speech and the like is more important than ensuring the privacy of those who propagate these views.²⁴ Taking into account the fact that the Irish slaves meme archive may be solely compiled of hashtags and are thus more public in nature than ordinary posts, and the fact that engaging with creators is not practical in this instance, it can be argued that gaining informed consent should not be a priority for the creation of the archive.

There are alternative ways through which the privacy of social media users can be protected by archivists and researchers, such as anonymisation. Baker posits that archivists should anonymise sensitive personal information of social media users; however, anonymisation

²¹ Small et al., 'What Your Tweets Tell Us About You,' 190.

²² 'Building Social Media Archives: Collection Development Guidelines,' Social Feed Manager, <https://gwu-libraries.github.io/sfm-ui/resources/guidelines>

²³ Townsend and Wallace, 'Social Media Research,' 12.

²⁴ Graham, 'Reflections on the Ethics of Web Archiving,' 107.

should not come at the expense of preserving important contextual information.²⁵ Thus, she argues that the extent and nature of anonymisation should reflect the mission of the archive.²⁶ Zimmer has criticised the current standard of anonymisation of social media data by research groups, arguing that new strategies need to be developed as individuals are too easily re-identified through gaps in anonymisation protection of their data.²⁷ Some scholars have a pessimistic outlook on the potential of anonymisation to protect individuals' privacy. Small et al. cite several studies which show that anonymised data can easily be compromised due to the connected nature of social media data.²⁸ Furthermore, some social media companies' APIs do not allow for the anonymisation of harvested data.²⁹ Sara Day Thomson and William Kilbride cite the unreliability of anonymisation techniques as a key motivation for researchers to follow specific ethical standards around the use of social media in order to protect creators' privacy.³⁰

DocNow balances its goals of allowing researchers to share their collections and of affording social media users the right to delete their content by presenting the data in the form of Tweet IDs.³¹ In theory, if the Irish slaves meme archive was presented to the public in Tweet ID form, the identities of the data creators would be protected. However, this method does not do much to anonymise the data that has not already been deleted by the creators. A preliminary scan of hydrated Tweet IDs from four datasets contained in the DocNow Tweet ID catalogue was undertaken for this paper in order to ascertain the level of anonymisation afforded to social media users whose data has been collected.³² All of the datasets contain usernames,

²⁵ Baker, 'Ethical Considerations,' 5.

²⁶ Ibid., 11.

²⁷ Zimmer, "'But the data is already public,'" 323.

²⁸ Small et al., 'What Your Tweets Tell Us About You,' 185.

²⁹ Williams, Burnap, and Sloan, 'Towards an Ethical Framework for Publishing Twitter Data in Social Research,' 1164.

³⁰ Sara Day Thomson and William Kilbride, 'Preserving Social Media: The Problem of Access,' *New Review of Information Networking* 20, nos.1-2 (2015): 269. <https://doi.org/10.1080/13614576.2015.1114842>

³¹ Baker, 'Ethical Considerations,' 11.

³² Justin Littman 'Ireland 8th Tweet Ids,' 2018, Harvard Dataverse, V1, <https://doi.org/10.7910/DVN/PYCLPE>. Ed Summers, 'UniteTheRight Tweet IDs,' 2017, Internet Archive, <https://archive.org/details/unitetheright-ids>. Mark Edward Phillips, "Yes All Women" Twitter Dataset. UNT Digital Library, June 2014, <https://digital.library.unt.edu/ark:/67531/metadc304853/>;

geolocations, and URLs (web addresses) through which one could access the profile pictures of creators. The Ireland 8th and Unite the Right datasets contain URLs for the Tweets themselves and thereby the profiles of the creators. While these URLs are not visible in the Las Vegas Shooting and 'Yes All Women' datasets, the usernames and profile descriptions of users is included. Thus profiles are easily searchable through Google and individuals can quickly be identified. This investigation shows the fundamental flaw in providing open access to this data without protection of anonymity or having gained informed consent from content creators. Researchers can rely on strategies of depersonalising the data and paraphrasing content when they present their work, but the task of anonymising large datasets is far more complex.³³ Anonymisation technology for archives needs further development before it can be relied upon to protect the privacy of creators' data. In the meantime, perhaps imposing restrictions on access to the Irish slaves meme archive would be the most effective method of limiting breaches of creators' privacy.

Restricting access to social media archives is not unprecedented. For instance, the initial deed of gift between Twitter and the Library of Congress dictated that Tweets would only be accessible six months after they were posted and even then, they would only be made available to 'bona fide researchers.' However, Zimmer has questioned how the Library will define which researchers are 'bona fide' and warns that such a practice goes against archival principles of fairness and neutrality.³⁴ Baker raises similar concerns about limiting access to collections solely to scholars, arguing that this may cause a 'research monopoly' to occur that would be avoided through provision of democratic access.³⁵ The University of Miami's US-Cuba Twitter archive is restricted to those with the university's log-in.³⁶ It can be argued that this gives an unfair advantage to researchers affiliated with that particular university. Although this argument is

'Web Archive on the October 1, 2017 Shooting in Las Vegas, Nevada, 2017-2018.' MS-00866. Special Collections and Archives, University Libraries, University of Nevada, Las Vegas. Las Vegas, Nevada. Accessed 13 June 2019. <http://d.library.unlv.edu/digital/collection/p17304coll5/id/20087/rec/1>

³³ Townsend and Wallace, 'Social Media Research,' 6 & 11.

³⁴ Zimmer, 'The Twitter Archive at the Library of Congress.'

³⁵ Baker, 'Ethical Considerations,' 5.

³⁶ 'Finding Aid Collection CHC5404 - 2014 US-Cuba Policy Change Twitter Archive,' University of Miami Libraries.

convincing, there is something to be said for restricting access to researchers affiliated with an institution, similar to the restrictions around access to online academic journals. In the case of the Irish slaves meme archive, it would be important to ensure that those accessing the material have some such affiliation. Otherwise, there is a risk that the material could be used to further spread the false narrative of Irish slavery or for other inappropriate ends. Therefore, requiring an institutional log-in is a desirable strategy, along with signing a terms of use agreement. This method is promoted by Rowell and Cooksey in relation to the digital collection of KKK newspapers, in order to balance ethical policies of access and care.³⁷ However, Zimmer argues that in order for terms of use agreements to mean anything, they must be enforced by the collecting institution and researchers must be monitored as studies show that researchers often are not aware that these agreements are legally binding.³⁸ Small et al. advocate for a tiered access approach; i.e., allowing open access to the text of Tweets while placing restrictions around access to users' profile information or geolocation.³⁹ This method could be a practical compromise between open access and strict restrictions.

Although institutions and projects may differ in their handling of the ethical conundrums involved in archiving social media, many scholars emphasise the importance of transparency and accountability in relation to ethical policies. Graham argues that archivists need to be open and transparent about the various decisions that are made at different points in the archival process in order to guarantee that collections are built ethically.⁴⁰ She cites DocNow and SFM as two promising examples of ethical transparency as these initiatives have developed tools to document actions taken during the collection and preservation process, such as SFM's action log.⁴¹

³⁷ Rowell and Cooksey, 'Archive of Hate.'

³⁸ Zimmer, "'But the data is already public,'" 320.

³⁹ Small et al., 'What Your Tweets Tell Us About You,' 190.

⁴⁰ Graham, 'Reflections on the Ethics of Web Archiving,' 108: 'Librarians and archivists will need to continue evolving how we document collecting practices to meet the inevitable need to explain ourselves and our work to those who will use the fragments of the past web that we might manage to save.'

⁴¹ Ibid., 107.

As well as being transparent about the intervention of the archivist, projects should also have clear policies that give prominence to ethical concerns. DocNow's white paper gives a detailed account of the project's ethics policies by presenting various ethical problems and how the project plans to address these problems.⁴² Ethics are given prominent placement on the homepage of the Schlesinger Library's #MeToo Digital Media Collection website.⁴³ The website provides an ethics statement and a bibliography of works that shaped the project's ethical principles.⁴⁴ Publishing this bibliography on the website gives the public a glimpse into the process behind the formulation of the project's ethics policy. The website also invites members of the public to contact the project team if they have any questions about the work, further demonstrating a commitment to transparency and user engagement.⁴⁵ However, the #MeToo Collection's ethics statement is quite brief and does not contain as much detail as the DocNow white paper. In the interest of transparency, it would be beneficial if a detailed report was written for the Irish slaves meme archive which clearly states the approach taken to tackle ethical concerns that may arise during the archival process. In addition, providing a bibliography and contact form would also be useful, as the Schlesinger Library has done.

There has been some debate over who should be responsible for archiving social media and thus who should be held accountable for the ethical issues involved – researchers or archivists and librarians.⁴⁶ As has been shown in this chapter, much more has been written about ethical use of social media data from the perspective of researchers than of archivists. However, work needs to be done to bridge this gap as archival institutions have an ethical imperative to actively collect and preserve records of historical value.⁴⁷ Therefore, archivists need to engage

⁴² Jules, Summers, and Mitchell, 'Documenting the Now White Paper.'

⁴³ '#metoo Digital Media Collection,' Schlesinger Library, accessed 13 June 2019. <https://www.schlesinger-metoproject-radcliffe.org/>

⁴⁴ 'Ethics,' Schlesinger Library, accessed 13 June 2019. <https://www.schlesinger-metoproject-radcliffe.org/ethics-statement>

'Ethics Bibliography,' Schlesinger Library, accessed 13 June 2019. <https://www.schlesinger-metoproject-radcliffe.org/ethics-bibliography>

⁴⁵ 'Ethics,' Schlesinger Library.

⁴⁶ Small et al., 'What Your Tweets Tell Us About You,' 180.

⁴⁷ International Council on Archives, 'Code of Ethics,' para. 10.

more fully with the ethical side of social media archiving. Small et al. assert that collaboration between archivists, librarians, and researchers provides opportunities for members of each field to learn from each other with the ultimate goal of achieving ethical collection, preservation, and use of social media archives.⁴⁸ Collaboration is needed to consolidate archival ethics with those of researchers. For example, different disciplines have different notions of privacy.⁴⁹ Workable ethical practices for archiving social media can be found through dialogue and the sharing of ideas between disciplines. For example, the 'Documenting Hate' panel discussion included a researcher/artist, a Master of Information student, and two media scholars.⁵⁰ Multi-disciplinary dialogues such as this one bring ethical issues of web archiving into focus and are a step in the right direction towards formulating effective ethical guidelines in this area.

Conclusion

This chapter has carefully considered the ethical issues that arise in the creation of social media archives in general and the Irish slaves meme archive in particular. The Irish slaves meme archive, as a collection of both social media data and misinformation in the form of a meme, raises some unique and pertinent questions about ethics. This chapter has addressed some of these questions in relation to creators' privacy and the provision of access to the archive. Further research and guidelines are needed to ensure that problematic social media archives can be created in a way that is ethical for all stakeholders: creators, users, and society at large.

⁴⁸ Small et al., 'What Your Tweets Tell Us About You,' 181.

⁴⁹ Ibid. For example, social scientists tend to be more cognisant of the dangers of breaching the privacy of their 'human subjects' by publishing their data, whereas humanist scholars favour a textual interpretation of social media data which renders concerns over creators' privacy less important.

⁵⁰ 'Ethics and Archiving the Web,' Rhizome.

Conclusion

As stated in Chapter 1, archives have a key role to play in helping us to confront uncomfortable truths about ourselves and the society we live in. This paper has argued that an Irish slaves meme archive could fulfil the same potential; the 'untruths' contained in this archive could act

as a tool for the exploration of issues that have a profound and very real effect on society. Social media platforms are havens in which misinformation and disinformation thrive, with consequences for the real world.¹ The increasing relevance of social media culture to all aspects of society, including politics, makes the question of how it can be archived all the more pressing. It is vital, then, for us not to dismiss false narratives such as the Irish slaves meme simply for being false. A more helpful approach to combatting the spread of misinformation would be to interrogate how and why these narratives resonate with people on social media.² Due to the ephemeral nature of social media content, present and future researchers will not have the ability to do this unless the data is carefully collected and preserved by archivists. This paper has argued in favour of archiving such material and has examined how this process can be undertaken, using the hypothetical creation of an archive of the Irish slaves meme as a case study.

Two recurring themes have emerged throughout this paper. The most obvious one is the need for archival institutions to embrace change and evolve archival theories to meet the challenges and opportunities of archiving new types of records. This theme pertains to social media data in general and specifically the Irish slaves meme archive. In terms of appraisal and acquisition, archives need to expand collection policies to include social media data as it provides a valuable resource for documenting society. For the same reason, misinformation spread on social media should also be viewed as having archival value. The technology used by archivists to harvest, preserve, arrange, and describe social media data needs further development as current standards do not go far enough to fulfill the usability potential of these archives. Collaboration with digital specialists and a loosening of traditional standards of order and description will help this process. In addition, the archival field is badly in need of ethical guidelines for handling social media data, especially sensitive or controversial material like the Irish slaves meme. Current codes of ethics are outdated and need to take into account the emerging challenges

¹ Alicia Parlapiano and Jasmine C. Lee, 'The Propaganda Tools Used by Russians to Influence the 2016 Election,' *The New York Times*, 16 February 2018. <https://www.nytimes.com/interactive/2018/02/16/us/politics/russia-propaganda-election-2016.html>

² Mina, *Memes to Movements*, 129.

faced by archivists engaged in this work. As this paper proves, new types of records and forms of expression require new archival methodologies to be developed. At its core, whatever methodology is adopted to create the Irish slaves meme archive should be fluid enough to ensure that multiple perspectives can be represented at every stage of the archival process. This is based on the postmodern theory that an infinite possible number of subjective 'truths' or meanings can be found in the archive, depending on the user's frame of reference.

Another theme that has dominated this paper is the importance of contextualisation in relation to the Irish slaves meme archive. This issue features in every chapter, which demonstrates its vital role in creating a useful, reflective, and meaningful archive. The processes of appraisal and harvesting need to take into account the amount and types of data necessary to provide maximum context for the material, such as comments, likes, or replies. The hashtags employed to gather the data (as should be laid out in the institution's collection policy) should be broad enough to capture the meme in a variety of contexts. For instance, harvesting data from #irishlivesmatter would show how the meme is presented in the context of the Black Lives Matter movement. Collecting Tweets with #irishslavesmyth would demonstrate negative reactions to the meme and the attempts to debunk the myth of Irish slavery. Of course, descriptive metadata and finding aids must ensure that extensive contextualisation is given to the collection. The importance of contextualisation cannot be overstated regarding the ethics of creating the Irish slaves meme archive. It would be unethical to create a collection of such problematic, false, and at times hateful material without placing it in its societal context and justifying why such a collection should exist in an archive. Social media posts often lose their meaning when removed from their context. The archivist has a responsibility, as a guardian of a 'repository of meaning,' to ensure that the various meanings of its records are preserved.³ Another key component of contextualisation is transparency. As we have seen, this is a major concern for archivists. Transparency in archival practice can come in the form of publishing collection development policies and ethics statements as well as documenting the actions taken

³ Greene, 'The Power of Meaning,' 55.

by archivists regarding the harvesting, preservation, arrangement, and description of data collections.

The case study of the Irish slaves meme has provided insights into the challenges of creating social media archives and creating archival collections that are controversial or contain false information. Framing this paper around the creation of a hypothetical social media archive has offered the opportunity to compare and contrast the effectiveness of various existing tools and projects in their application to a specific case study. In this way, this paper has demonstrated the great strides that have been made to preserve social media data while also pointing out the shortcomings of existent methods.

This paper offers a detailed survey of the current state of social media archiving and can function as a starting point for future research into best practices for creating and preserving social media archives, especially those of a controversial sort. Hopefully more archival institutions will undertake to create such collections in the future. There is certainly a lot of material to choose from. For instance, a collection of social media data surrounding the Pizzagate scandal would provide researchers with rich resources for investigating the origins, spread, and effect of this conspiracy theory.⁴ Further research is needed by archival scholars to investigate how the field can respond to the evolving nature of how people communicate and record their lives in the social media era. Archives need to act now to be more inclusive of different forms of information and expression in order to meet users' needs. If institutions embrace the challenges that this entails, the archival mission to 'hold up a mirror for mankind' can be realised.⁵

⁴ The 'Pizzagate' scandal was a conspiracy theory (since debunked) that emerged in 2016 which claimed that the Democratic Party in the USA was running a child sex ring out of a pizzeria in Washington, D.C. The theory was used as a 'political weapon' against Hillary Clinton during the presidential campaign. Mike Wendling, 'The saga of 'Pizzagate': The fake story that shows how conspiracy theories spread,' BBC News, 2 December 2016. <https://www.bbc.com/news/blogs-trending-38156985>

⁵ Ham, 'The Archival Edge,' 13

References

Ahmed, Wasim. 'Using Twitter as a data source: An overview of current social media research tools.' LSE Impact Blog, 10 July 2015.

<https://blogs.lse.ac.uk/impactofsocialsciences/2015/07/10/social-media-research-tools-overview/>

Amend, Alex. 'How the Myth of the "Irish slaves" Became a Favorite Meme of Racists Online.' Southern Poverty Law Center, 19 April, 2016.

<https://www.splcenter.org/hatewatch/2016/04/19/how-myth-irish-slaves-became-favorite-meme-racists-online>.

Baker, Antoinette E. 'Ethical Considerations in Web 2.0 Archives.' *SLIS Student Research Journal*, 1, no. 1 (2011): Article 4. <https://scholarworks.sjsu.edu/ischoolsrj/vol1/iss1/4/>

Berdini, Annalise, Charles Macquarie, Shira Peltzman, and Kate Tasker. 'Describing Digital: The Design and Creation of a Born-Digital Archival Description Standard at the University of California Libraries.' *Journal of Western Archives* 9, no.1 (2018): Article 10.

<https://digitalcommons.usu.edu/westernarchives/vol9/iss1/10>

Boles, Frank. "'Just a Bunch of Bigots" A Case Study in the Acquisition of Controversial Material.' *Archival Issues* 19, no. 1, (1994): 53-65. <https://www.jstor.org/stable/41101872>.

Brügger, Niels. *The Archived Web: Doing History in the Digital Age*. Cambridge, MA : The MIT Press, 2018.

Bruns, Axel. 'The Library of Congress Twitter Archive: A Failure of Historic Proportions.' Medium, 2 January 2018. <https://medium.com/dmrc-at-large/the-library-of-congress-twitter-archive-a-failure-of-historic-proportions-6dc1c3bc9e2c>.

Caplan-Bricker, Nora. 'The Challenge of Preserving the Historical Record of #MeToo.' *The New Yorker*, 11 March 2019. <https://www.newyorker.com/tech/annals-of-technology/the-challenge-of-preserving-the-historical-record-of-metoo>

Commisso, Corrie. 'The Post-Truth Archive: Considerations for Archiving Context in Fake News Repositories.' *Preservation, Digital Technology & Culture* 46, no.3 (2017): 99-102.

<https://doi.org/10.1515/pdte-2017-0010>.

Conlin, Michelle. 'Twitter to revive Politwoops, archive of politicians' deleted tweets.' Reuters, 31 December 2015. <https://www.reuters.com/article/us-usa-election-politwoops-idUSKBN0UE16520151231>

Cook, Terry. 'Evidence, Memory, Identity, and Community: Four Shifting Archival Paradigms.' *Archival Science* 13, no.2 (2013): 95-120.

Cook, Terry. 'Remembering the Future: Appraisal of Records and the Role of Archives in Constructing Social Memory.' In *Archives, Documentation, and Institutions of Social Memory: Essays from the Sawyer Seminar*, edited by Francis X. Blouin, 232-246. Ann Arbor, Michigan: University of Michigan Press, 2006.

Cook, Terry. 'The Concept of the Archival Fonds in the Post-Custodial Era: Theory, Problems and Solutions.' *Archivaria* 35 (Spring 1993): 24-37.

Cook, Terry 'What is Past is Prologue: A History of Archival Ideas Since 1898, and the Future Paradigm Shift,' *Archivaria* 43 (Spring 1997): 17-63.

Cox, Richard J. *Archival Anxiety and the Vocational Calling*. Duluth, Minn: Litwin Books, 2011.

Cullingford, Alison. *The Special Collections Handbook*, Second. London: Facet Publishing, 2016.

Datasift Developers. 'Historics Archive Schema.' Accessed 13 June 2019.

<https://dev.datasift.com/docs/products/stream/features/historics/historics-access/historics-archive-schema>

Digital Preservation Coalition. 'About.' Accessed 13 June 2019.

<https://www.dpconline.org/about>

Documenting Ferguson Committee. 'Project Explanation and Purpose,' September 2014.

<http://digital.wustl.edu/ferguson/DFP-Plan.pdf>

Documenting the Now. 'Tweet ID Datasets.' Accessed 13 June 2019.

<https://www.docnow.io/catalog/>

Dooley, Jackie and Kate Bowers, *Descriptive Metadata for Web Archiving: Recommendations of the OCLC Research Library Partnership Web Archiving Metadata Working Group*. Dublin, OH: OCLC Research, 2018.

Ebner, Julia and Jacob Davey. 'The far right has learned to mobilise and radicalise.

Charlottesville's a wake-up call.' *The Guardian*, 14 August 2017.

<https://www.theguardian.com/commentisfree/2017/aug/14/far-right-charlottesville-mobilise-radicalise-white-supremacists-coalition>

Editor-in Chief. 'Preserving Fake News.' *Preservation, Digital Technology & Culture* 46, no.3 (2017): 89-90. <https://doi.org/10.1515/pdtc-2017-9001>

'Ethics and Archiving the Web: Documenting Hate.' Vimeo. 27 June 2018.

<https://vimeo.com/277336010>

Facebook for Developers. 'Facebook Platform Policy.' Last updated 25 April 2019.

<https://developers.facebook.com/policy/>

Fanning, Bryan. 'Slaves to a Myth.' *Dublin Review of Books*, 1 November, 2017.

<https://www.drb.ie/essays/slaves-to-a-myth>.

Fondren, Elisabeth and Meghan Menard McCune. 'Archiving and Preserving Social Media at the Library of Congress: Institutional and Cultural Challenges to Build a Twitter Archive.'

Preservation, Digital Technology & Culture 47, no.2 (2018): 33-44.

<https://doi.org/10.1515/pdtc-2018-0011>

Foscarini, Fiorella. 'Archival Appraisal in Four Paradigms.' In *Currents of Archival Thinking*, edited by Heather MacNeil and Terry Eastwood, Second Edition, 107-133. Santa Barbara, California Denver, Colorado: Libraries Unlimited, an imprint of ABC-CLIO, LLC, 2017.

GitHub. 'Twarc.' Accessed 13 June 2019. <https://github.com/DocNow/twarc>.

Graham, Pamela M. 'Guest Editorial: Reflections on the Ethics of Web Archiving.' *Journal of Archival Organization* 14 nos.3-4 (2017), 103-110.

<https://doi.org/10.1080/15332748.2018.1517589>

Greene, Mark. 'The Power of Meaning: The Archival Mission in the Postmodern Age.' *The American Archivist* 65 (Spring/Summer 2002): 42-55.

<https://doi.org/10.17723/aarc.65.1.l914668v881wv19n>

Hale, Julie. 'Documenting the Now: Archiving Social Media for Generations to Come.' University Libraries, Washington University in Saint Louis, 20 January 2016.

<https://library.wustl.edu/8497-2/>.

Ham, F. Gerald. 'The Archival Edge.' *The American Archivist* 38, no.1 (January 1975): 5-13.

<https://www.jstor.org/stable/40291712>.

Ham, F. Gerald. *Selecting and Appraising Archives and Manuscripts*, Chicago: Society of American Archivists, 1993.

Handler, Jerome S. and Matthew C. Reilly. 'Contesting "White Slavery" in the Caribbean: Enslaved Africans and European Indentured Servants in Seventeenth-Century Barbados'. *New West Indian Guide* 91 (2017): 30-55. <https://doi.org/10.1163/22134360-09101056>.

Harmony Labs. 'VizLab.' Accessed 13 June 2019. <https://harmonylabs.org/vizlab>

Harris, Verne. 'The Archival Sliver: Power, Memory, and Archives in South Africa' *Archival Science* 2, nos. 1-2 (2002): 63-86.

Hogan, Liam. 'All of my work on the "Irish slaves" meme (2015-'18).' Medium, 12 March, 2017.

<https://medium.com/@Limerick1914/all-of-my-work-on-the-irish-slaves-meme-2015-16-4965e445802a>

Hogan, Liam. 'Open letter to Irish Central, Irish Examiner and Scientific American about their "Irish slaves" disinformation.' Medium, 8 March 2016.

<https://medium.com/@Limerick1914/open-letter-to-irish-central-irish-examiner-and-scientific-american-about-their-irish-slaves-3f6cf23b8d7f>.

Hogan, Liam. 'Two years of the 'Irish slaves' myth: racism, reductionism and the tradition of diminishing the transatlantic slave trade.' openDemocracy, 7 November 2016.

<https://www.opendemocracy.net/en/beyond-trafficking-and-slavery/two-years-of-irish-slaves-myth-racism-reductionism-and-tradition-of-diminis/>

Hogan, Liam, Laura McAtackney, and Matthew C. Reilly. 'The Irish in the Anglo-Caribbean: Servants or Slaves?' *History Ireland* 24, no. 2 (April 2016): 18-22.

Householder, Mike. 'New Michigan museum showcases racist artifacts.' About the Jim Crow Museum, Ferris State University. 19 April, 2012.

<https://www.ferris.edu/HTMLS/news/jimcrow/more.htm#>

Hunt, Elle. 'What is fake news? How to spot it and what you can do to stop it.' *The Guardian*, December 17, 2016. <https://www.theguardian.com/media/2016/dec/18/what-is-fake-news-pizzagate>

International Council on Archives. 'Code of Ethics.' 6 September 1996.

https://www.ica.org/sites/default/files/ICA_1996-09-06_code%20of%20ethics_EN.pdf

Internet Archive. 'Ferguson Tweet IDs.' Accessed 13 June 2019.

<https://archive.org/details/ferguson-tweet-ids>

James Hardiman Library. 'QA819 Acquisition Policy for the James Hardiman Library Archives.' Last modified 27 March, 2011.

<http://library.nuigalway.ie/media/jameshardimanlibrary/content/documents/collections/Archives-Acquisition-and-Evaluation-Policy.pdf>

Jimerson, Randall. 'Archives for All: Professional Responsibility and Social Justice.' *The American Archivist* 70, no. 2 (September 2007): 252-281.

<https://doi.org/10.17723/aarc.70.2.5n20760751v643m7>.

Jordan, Don and Michael Walsh. *White Cargo: The Forgotten History of Britain's White Slaves in America*. New York: New York University Press, 2008.

Juhasz, Alexandra. 'A Preservationist's Guide to #100hardtruths- #fakenews: One Fake News Preserve.' *Preservation, Digital Technology & Culture* 46, no.3 (2017): 103-108.

<https://doi.org/10.1515/pdte-2017-0011>

Jules, Bergis. 'Some Thoughts on Ethics and DocNow.' Documenting DocNow, 3 June 2016.

<https://news.docnow.io/some-thoughts-on-ethics-and-docnow-d19cfec427f2>

Jules, Bergis, Ed Summers and Vernon Mitchell, Jr. 'Documenting The Now White Paper. Ethical Considerations for Archiving Social Media Content Generated by Contemporary Social Movements: Challenges, Opportunities, and Recommendations.' (April 2018).

<https://www.docnow.io/docs/docnow-whitepaper-2018.pdf>

King, Gary and Nate Persily, 'A New Model for Industry-Academic Partnerships.' Working Paper.

Last modified 15 January 2019. <https://gking.harvard.edu/files/gking/files/partnerships.pdf>

Kowalczyk, Stacy T. *Digital Curation for Libraries and Archives*. Santa Barbara, California Denver, Colorado: Libraries Unlimited, an imprint of ABC-CLIO, LLC, 2018.

Lamoree, Karen M. 'Documenting the Difficult or Collecting the Controversial.' *Archival Issues* 20, no. 2 (1995): 149-153. <https://www.jstor.org/stable/41101913>.

Lau, Joyce. 'Master and Servant.' *New York Times*, 27 April, 2008.

<https://www.nytimes.com/2008/04/27/books/review/Lau-t.html>

Laughland, Oliver. 'Reparations: Democrats renew debate over how to heal the legacy of slavery.' *The Guardian*, 21 March, 2019. <https://www.theguardian.com/us-news/2019/mar/21/reparations-democrats-2020-slavery>

Leeds University Library. 'Collections Strategy for Leeds University Library.' Accessed 13 June 2019. https://library.leeds.ac.uk/downloads/download/144/collections_strategy

Littman, Justin. 'Ireland 8th Tweet IDs.' Harvard Dataverse, V1, 2018.

<https://doi.org/10.7910/DVN/PYCLPE>

Littman, Justin, Daniel Chudnov, Daniel Kerchner, Christie Peterson, Yecheng Tan, Rachel Trent, Rajat Vij and Laura Wrubel. 'API-based social media collecting as a form of web archiving.' *International Journal on Digital Libraries* 19, no.1 (March 2018): 21-38.

Markham, Annette and Elizabeth Buchanan. 'Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee (Version 2.0).' (2012).
<http://aoir.org/reports/ethics2.pdf>

Mina, An Xiao. *Memes to Movements: How the World's Most Viral Media is Changing Social Protest and Power*. Boston: Beacon Press, 2019.

Moffitt, Jim. 'Choosing a Historical API.' Gnip, 20 June 2017.
<http://support.gnip.com/articles/choosing-historical-api.html>

Myers, Cayce and James F. Hamilton. 'Social Media as Primary Source.' *Media History* 20, no. 4 (2014): 431-444.

Nesmith, Tom. 'The Concept of Societal Provenance and Records of Nineteenth-Century Aboriginal-European Relations in Western Canada: Implications for Archival Theory and Practice.' *Archivaria* 60 (2006): 351-360.

NYU Press. 'History.' Accessed 13 June 2019. <https://nyupress.org/about-nyupress/history/>

O'Callaghan, Sean. *To Hell or Barbados: The ethnic cleansing of Ireland*. Dingle: Brandon, 2000.

Panofsky, Aaron and Joan Donovan. 'Genetic Ancestry Testing Among White Nationalists.' *SocArXiv*, 17 August, 2013. doi:10.31235/osf.io/7f9bc

Parlapiano, Alicia and Jasmine C. Lee. 'The Propaganda Tools Used by Russians to Influence the 2016 Election.' *The New York Times*, 16 February 2018.
<https://www.nytimes.com/interactive/2018/02/16/us/politics/russia-propaganda-election-2016.html>

Pearce-Moses, Richard. *A Glossary of Archival and Records Terminology*. Chicago: Society of American Archivists, 2005.

Pennock, Maureen. *Web-Archiving: DPC Technology Watch Report 13-01*. Digital Preservation Coalition, 2013.

Phillips, Mark Edward. "Yes All Women" Twitter Dataset. UNT Digital Library, June 2014.

<https://digital.library.unt.edu/ark:/67531/metadc304853/>

Post, Colin. 'Building a Living, Breathing Archive: A Review of Appraisal Theories and Approaches for Web Archives.' *Preservation, Digital Technology & Culture* 46, no.2. (2017): 69-77. <https://doi.org/10.1515/pdtc-2016-0031>

Prom, Christopher J. 'Social Feed Manager: Guide for Building Social Media Archives.' 7 June 2017. <https://gwu-libraries.github.io/sfm-ui/resources/SFMReportProm2017.pdf>

Rhizome. 'Ethics and Archiving the Web.' Accessed 13 June 2019. <https://eaw.rhizome.org/>

Rosenfeld, Sophia. *Democracy and Truth: A Short History*. Philadelphia: University of Pennsylvania Press, 2019. Ebook edition. No page numbers.

Rowell, Chelcie Juliet and Taryn Cooksey. 'Archive of Hate: Ethics of Care in the Preservation of Ugly Histories.' *Lady Science*, 2019. Accessed 13 June 2019.

<https://www.ladyscience.com/blog/archive-of-hate-ethics-of-care-in-the-preservation-of-ugly-histories>

Samouelian, Mary and Jackie Dooley, *Descriptive Metadata for Web Archiving: Review of Harvesting Tools*. Dublin, OH: OCLC Research, 2018.

Schlesinger Library. Ethics.' #metoo Project Schlesinger Library. Accessed 13 June 2019.

<https://www.schlesinger-metoo-project-radcliffe.org/ethics-statement>

Schlesinger Library. 'Ethics Bibliography.' #metoo Project Schlesinger Library. Accessed 13 June 2019. <https://www.schlesinger-metoo-project-radcliffe.org/ethics-bibliography>

Schlesinger Library. 'The Collection.' #metoo Project Schlesinger Library. Accessed 13 June 2019.

<https://www.schlesinger-metoo-project-radcliffe.org/collection>

Sheffield, Rebecka. 'Community Archives.' In *Currents of Archival Thinking*, edited by Heather MacNeil and Terry Eastwood, Second Edition, 351-376. Santa Barbara, California Denver, Colorado: Libraries Unlimited, an imprint of ABC-CLIO, LLC, 2017.

Small, Heather, Kristine Kasianovitz, Ronald Blanford and Ina Celaya, 'What Your Tweets Tell Us About You: Identity, Ownership and Privacy of Twitter Data,' *The International Journal of Digital Curation* 7, no.1 (2012), 174-197. doi:10.2218/ijdc.v7i1.224

Social Feed Manager. 'Building Social Media Archives: Collection Development Guidelines.' Last updated 13 March 2017. <https://gwu-libraries.github.io/sfm-ui/resources/guidelines>

Stack, Liam. 'Debunking a Myth: The Irish Were Not Slaves, Too.' *New York Times*, 17 March, 2017. <https://www.nytimes.com/2017/03/17/us/irish-slaves-myth.html>

Stearns, Josh. '20 Projects Receive Funding to Combat Misinformation and Build a More Trustworthy Public Square.' Democracy Fund, 22 June 2017.

<https://www.democracyfund.org/blog/entry/20-projects-receive-funding-to-combat-misinformation-and-build-a-more-trust>

Summers, Ed. 'Tweets and Deletes: Silences in the Social Media Archive.' On Archivy, Medium, 14 April 2015. <https://medium.com/on-archivy/tweets-and-deletes-727ed74f84ed>

Summers, Ed. 'UniteTheRight Tweet IDs.' Internet Archive, 2017.

<https://archive.org/details/unitetheright-ids>

The National Archives. *Collection Development Tools and Guidance*, 2011.

<http://www.nationalarchives.gov.uk/documents/archives/collection-development-tools-and-guidance.pdf>

The National Archives. *The National Collections Strategy*, 2009.

<https://www.nationalarchives.gov.uk/documents/ncs-vision-strategy-and-principles.pdf>

Thomson, Sara Day. *Preserving Social Media: DPC Technology Watch Report 16-01*. Digital Preservation Council, 2016.

Thomson, Sara Day and William Kilbride. 'Preserving Social Media: The Problem of Access.' *New Review of Information Networking* 20, nos.1-2 (2015), 261-275.

<https://doi.org/10.1080/13614576.2015.1114842>

Towson, Anna. 'Collection Development Policy for Archives, Museum Objects and Ephemera.' LSE Library, June 2015. <http://www.lse.ac.uk/library/assets/documents/Archives-Collection-Development-Policy-PDF.pdf>

Townsend, Leanne and Claire Wallace. 'Social Media Research: A Guide to Ethics.' University of Aberdeen (2016). https://www.gla.ac.uk/media/media_487729_en.pdf.

Tucker, Neely. 'In Mississippi, defenders of state's Confederate-themed flag dig in.' *The Washington Post*, 18 August, 2015.

https://www.washingtonpost.com/lifestyle/style/mississippi-resists-calls-to-change-its-confederate-themed-flag/2015/08/18/cd590924-41d3-11e5-846d-02792f854297_story.html?utm_term=.586bed5e530b

Tumblr. "'But the Irish were slaves too.'" Accessed 13 June 2019.

<https://theishwereslavestoo.tumblr.com/>

Twitter. 'Developer Agreement and Policy.' Last modified 25 May, 2018.

<https://developer.twitter.com/en/developer-terms/agreement-and-policy>

Twitter Developers. 'Twitter IDs (snowflake).' Accessed 13 June 2019.

<https://developer.twitter.com/en/docs/basics/twitter-ids.html>

University of Miami Libraries. 'Finding Aid Collection CHC5404 - 2014 US-Cuba Policy Change Twitter Archive.' University of Miami Cuban Heritage Collection. Last updated January 2019.

<https://atom.library.miami.edu/chc5404>

University of Nevada, Las Vegas. 'Guide to the Web Archive on the October 1, 2017 Shooting in Las Vegas, Nevada.' University Libraries. Special Collections and Archives. (2018).

<https://www.library.unlv.edu/speccol/finding-aids/MS-00866.pdf>

University of Nevada, Las Vegas. UNLV University Libraries Digital Collections. 'Tweet Identifiers 2017.' Accessed 13 June 2019.

<http://d.library.unlv.edu/digital/collection/p17304coll5/id/20087/rec/1>

University of North Texas. 'Collection Development Policy for the UNT Libraries' Digital Collections.' Last modified 12 January, 2017. <https://library.unt.edu/policies/collection-development-digital-collections/>

Virginia Tech: Crisis, Tragedy, and Recovery Network. 'Boston Marathon Bombing: Twitter and RSS feeds.' Accessed 13 June 2019. <https://archive-it.org/collections/3645>

Wallace, David A. 'Archives and Social Justice.' In *Currents of Archival Thinking*, edited by Heather MacNeil and Terry Eastwood, Second Edition, 271-297. Santa Barbara, California Denver, Colorado: Libraries Unlimited, an imprint of ABC-CLIO, LLC, 2017.

Wendling, Mike. 'The saga of 'Pizzagate': The fake story that shows how conspiracy theories spread.' *BBC News*, 2 December 2016. <https://www.bbc.com/news/blogs-trending-38156985>

Williams, Matthew L., Pete Burnap and Luke Sloan, 'Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users' Views, Online Context and Algorithmic Estimation,' *Sociology* 51, no. 6 (2017): 1149-1168.

<https://doi.org/10.1177/0038038517708140>

Wrubel, Laura, Justin Littman and Dan Kerchner. '2018 U.S. Congressional Election Tweet Ids.' Harvard Dataverse, V1, 2019. <https://doi.org/10.7910/DVN/AEZPLU>

Yeo, Geoffrey. 'Bringing Things Together: Aggregate Records in a Digital Age.' *Archivaria* 74 (Fall 2012): 43-91.

Yeo, Geoffrey. 'Continuing Debates about Description.' In *Currents of Archival Thinking*, edited by Heather MacNeil and Terry Eastwood, Second Edition, 163-192. Santa Barbara, California Denver, Colorado: Libraries Unlimited, an imprint of ABC-CLIO, LLC, 2017.

Zhang, Jane. 'Original Order in Digital Archives,' *Archivaria* 74 (Fall 2012): 167-193.

Zhang, Jane and Dayne Mauney. 'When Archival Description Meets Digital Object Metadata: A Typological Study of Digital Archival Representation.' *The American Archivist* 76, no. 1 (Spring/Summer 2013): 174-195.

Zimmer, Michael. "'But the data is already public'": on the ethics of research in Facebook.' *Ethics and Information Technology* 12, no.4 (December 2010): 313-325.
<https://doi.org/10.1007/s10676-010-9227-5>

Zimmer, Michael. 'The Twitter Archive at the Library of Congress: Challenges for information practice and information policy.' *First Monday* 20, no. 7 (June 2015).
<https://doi.org/10.5210/fm.v20i7.5619>

Zinn, Howard. 'Secrecy, Archives, and the Public Interest'. *The Midwestern Archivist* II, no. 2 (1977): 14-26.