

2023

Archiving “sensitive” social media data: ‘In Her Shoes’, a case study

LORRAINE GRIMES Dr

Maynooth University, lorraine.grimes@mu.ie

Kathryn Cassidy Dr

Trinity College Dublin, k.cassidy@tchpc.tcd.ie

Murilo Dias

Trinity College Dublin, mdias@tcd.ie

Clare Lanigan

Royal Irish Academy, c.lanigan@ria.ie

Aileen O'Carroll Dr

Maynooth University, aileen.ocarroll@mu.ie

See next page for additional authors

Follow this and additional works at: <https://elischolar.library.yale.edu/jcas>



Part of the [Archival Science Commons](#), [Digital Humanities Commons](#), and the [Women's Studies Commons](#)

Recommended Citation

GRIMES, LORRAINE Dr; Cassidy, Kathryn Dr; Dias, Murilo; Lanigan, Clare; O'Carroll, Aileen Dr; and Singhvi, Preetam (2023) "Archiving “sensitive” social media data: ‘In Her Shoes’, a case study," *Journal of Contemporary Archival Studies*: Vol. 10, Article 19.

Available at: <https://elischolar.library.yale.edu/jcas/vol10/iss1/19>

This Article is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in *Journal of Contemporary Archival Studies* by an authorized editor of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

Archiving “sensitive” social media data: ‘In Her Shoes’, a case study

Cover Page Footnote

Thank you to The Wellcome Trust who funded this research.

Authors

LORRAINE GRIMES Dr, Kathryn Cassidy Dr, Murilo Dias, Clare Lanigan, Aileen O’Carroll Dr, and Preetam Singhvi

ARCHIVING “SENSITIVE” SOCIAL MEDIA DATA: “IN HER SHOES,” A CASE STUDY

In recent years, social media has played an increasingly significant role in activist and social movements. As archivists and researchers, we recognize the imperative of capturing social change by preserving this data. Archiving social media is a relatively new phenomenon and an area that needs greater clarity, understanding, and uniformity. Recent archiving of social movements such as #MeToo, #BlackLivesMatter, and others have advanced progress in this field, but digital archiving and social media are rapidly expanding areas that need best-practice clarification and better regulation.¹ Activist archiving has emerged from the shift in the profession toward an “archival practice which, rejecting professional advocacy of neutrality and passivity, acknowledges the role of the record keeper in ‘actively’ participating in the creation, management and pluralization of archives and seeks to understand and guide the impact of that active role.”² Michelle Caswell argues that the acts of collecting and preserving records affirm the existence of communities that have historically been silenced, erased, or marginalized as a political act.³

Our project used the “In Her Shoes: Women of the Eighth” Facebook dataset as a case study. This page was set up during the 2018 referendum to remove the constitutional ban on abortion in Ireland. On the page, women shared their experiences of being denied access to abortion care in Ireland. This article explores the ethical issues involved with copyright, consent to archive, appropriate language and metadata, and anonymization. This article addresses the challenges faced and provides recommendations for others working with social media collections that include sensitive data. We argue that more research on how we archive social media data, particularly sensitive data, and clearer guidance and information-sharing is needed.

The Digital Preservation of Reproductive Health: Archiving the Eighth project is run by the Digital Repository of Ireland (DRI). The project intends to provide long-term preservation and access to the many at-risk archives generated by grassroots women’s reproductive health movements during the campaign to repeal the Eighth Amendment (the constitutional ban on abortion). The project collects, catalogs, and preserves much of the born-digital content generated by grassroots organizations. One of the project’s core aims is to archive and preserve the “In Her Shoes: Women of the Eighth” Facebook page. The page was founded by pro-choice activist Erin Darcy in Ireland in January 2018. Prior to 2019, abortion was extremely restrictive in Ireland, only legal in cases where the woman’s life was at risk. The page was a political response to a national debate, motivated by an upcoming referendum on abortion, and created a space where women could share their experiences of abortion. Women posted their personal experiences of being refused healthcare, barriers to access, traveling for healthcare abroad, illegally ordering pills online, and the emotional impact of this. The page was instrumental in changing voters’ minds about abortion and had 100,000 followers by the time of the referendum.⁴ The stories shared capture an insight into the realities surrounding the illegality of abortion.

Why Archive?

¹ Rosenbaum, “Archiving #MeToo”; Uimonen, “#MeToo in Sweden”; Carter, “Of Things Said and Unsaid.”

² Cifor et al., “What We Do Crosses Over into Activism,” 73.

³ Caswell, “Towards a Survivor-Centered Approach to Records Documenting Human Rights Abuses,” 314.

⁴ McDonnell, “In Her Shoes.”

In recent decades, there has been a growing interest in community-based and participatory archives, due in large part to an increasing recognition that people and communities need to be empowered to make decisions about how their histories are documented and preserved.⁵ This is particularly valuable for groups that have been underrepresented in the archives or previously overlooked. Grassroots and community-based archives can be a space in which the archive becomes a significant tool for discovery, education, and empowerment. Archivists recognize that power lies with those who make appraisal decisions in regards to the following questions: What documents are considered significant? Whose lives are worth preserving/remembering? How does one access the archive?

The literature on the formation of the archives of women's right organizations references archiving to inspire contemporary and future activism.⁶ A collection of archival groups using the hashtag #ArchivesForBlackLives advocates questioning traditional archival authority, disrupting the status quo of memory, and decolonizing conventional and mainstream approaches to information, knowledge, records, and archives. Community and activist archiving play an important role both in promoting a more diverse society and in fostering community identity among groups often marginalized by more powerful elites.⁷ Archiving social movements is rooted in a diverse public attempt to bring attention to marginalized aspects of societal history, often for purposes of transforming them.⁸ R. C. Jimerson rejects the idea that archiving is impartial or "objective," and argues that all archives are politically engaged and should be willing to give voice to marginalized groups and address social issues in the interest of democracy.⁹ Women's experiences and those most marginalized in society have traditionally been excluded in the archives. Archivists have a responsibility to ensure that women's lived experiences are not lost.

Today, according to Elisabeth Fondren and Meghan Menard McCune, "social media is supplementing and in some cases supplanting letters, journals, serial publications and other sources routinely collected by research libraries."¹⁰ Therefore, it is important for society and social history to collect and preserve these experiences. Historians often face the frustration that many sources do not survive; women's experiences have not been preserved because they were not deemed of historical importance. One of the lessons of the #MeToo archive is the reminder that the sources are there; if we ask different questions, they emerge from the shadows.¹¹ An ethically informed archivist looks for real stories outside of "safe" (e.g., academic, governmental) spaces. Tonia Sutherland urges us to think about where we collect material and to be sensitive to the nature of the source.¹²

The Archiving Reproductive Health project archives campaign materials from groups that sought the expansion of abortion care in Ireland. We believe that traditionally women's organizations have largely been excluded from archives, and the focus of the project is to collect and preserve women's experiences. The In Her Shoes Facebook page is particularly important to archive as it captures personal "hidden histories," the silent suffering of women

⁵ Sutherland, "Archival Amnesty," 15.

⁶ Sadler and Cox, "'Civil Disobedience' in the Archive."

⁷ Jimerson, "Archives for All," 266.

⁸ Morrone, *Informed Agitation*, 211.

⁹ See Jimerson, *Archives Power*.

¹⁰ Fondren and McCune, "Archiving and Preserving Social Media at the Library of Congress," 33.

¹¹ Rosenbaum, "Archiving #MeToo," 260.

¹² See Sutherland, "Archival Amnesty."

and the shame and stigma associated with abortion, and captures a time of emerging social change.

Social Media Archival Practices

Lisa Lindström has found that there are significant differences in archiving practices for social media data such as opting or not for explicit consent, limiting access to collections, and setting the terms of anonymization. She notes that some archives choose complete anonymization of their social media material while others offer limited anonymization along with consent from those non-anonymized. Yet others did not carry out any anonymization.¹³ Ally McCrow-Young’s research involved taking information data from Instagram, which led to ambiguity as to whether the data was “personal data,” even if the data was not personally identifiable.¹⁴ Where there is uncertainty, archivists tend to opt for the most conservative method in order to avoid legal challenges. Current copyright legislation in Ireland fails to directly address social media.¹⁵ As Chloé S. Georas argues, copyright law is “anachronistic and inhibitory to archiving projects,” while the legal privacy framework “is not sufficiently nuanced to accommodate the complex scales of visibility deployed by people on social networks.”¹⁶ The lack of consistency in archival practices makes it difficult to define what best practices are. As more and more researchers are opting to use social media as data, further clarification and guidance around social media archiving and using social media for research is needed.

In Her Shoes: Consent to Archive

The In Her Shoes Facebook dataset was provided by the administrators of the page to the DRI. As is the case of all digital objects donated there, “The DRI is not the owner of the copyright to the digital object.”¹⁷ Although the DRI had permission from the administrators, copyright belongs to the individual authors of the posts. Contributors to the page privately sent their stories to the administrators. Every post was published on the page by the administrator, so no private Facebook account was linked to any one story. The posts remained in the storyteller’s own words, only edited by administrators for obvious spelling errors and formatted for ease of reading. Getting explicit consent to archive from each of the 1,394 authors was not possible. When the contributor emailed or privately messaged the administrators with their story, the administrator then deleted these messages in case the page or email was hacked. This was a real fear as the page was under threat by pro-life groups and individuals who repeatedly reported the page on Facebook in an attempt to have it taken down.¹⁸ In addition, many women sent their stories through fake/proxy accounts and were no longer contactable. The administrator of the page sought consent from each contributor to post on the Facebook page, therefore contributors were made aware that the post would be public.

Under General Data Protection Regulation (GDPR) legislation, article 6 (1) states that data processors may retain and process personal data under five additional exemptions including for archival reasons.¹⁹ This exception is also outlined in the Irish Data Protection Act 2018, 42 (1),

¹³ See Lindström, “Archiving in the Era of Online Activism.”

¹⁴ See McCrow-Young, “Approaching Instagram Data.”

¹⁵ Breed, “Capturing a Moment.”

¹⁶ Georas, “Networked Memory Project,” 472.

¹⁷ “DRI Notice and Action Policy,” Digital Repository of Ireland, 11 April 2017, <https://repository.dri.ie/catalog/-vh5499702>.

¹⁸ Darcy, *Women of the Eighth*.

¹⁹ “Article 6: Lawfulness of Processing,” European Union GDPR, 2016, <https://gdpr.eu>.

which states, “Subject to suitable and specific measures being taken to safeguard the fundamental rights and freedoms of data subjects, personal data may be processed, in accordance with Article 89, for (a) archiving purposes in the public interest; (b) scientific or historical research purposes.”²⁰ Although the contributor’s own personal Facebook account was not linked to the page, archivists need to be aware of the wider impact of archiving. As Katharina Kinder-Kurlanda and colleagues argue, “Even if a user of an online platform knowingly shares a piece of information by posting it on the platform, Big Data analysis can publicize and amplify it in a way the user never intended or agreed to.”²¹ For our project, we recognized that the In Her Shoes stories were written at a time when abortion was illegal in Ireland, when there was an upcoming contentious referendum, and when sharing these stories publicly carried great personal risk. In sharing these stories at a time of high risk, we believe the women who contributed the stories were both aware of and willing to share their life experiences knowing there were potential dangers in doing so. Nevertheless, we decided to make the collection restricted access, only available to those engaged in teaching and research. There are a number of reasons for why we came to this decision including cataloging and anonymization issues.

Minimized Anonymization Process

In most cases the authors of the posts did not provide their names or personal identifiable information about others. Nevertheless, there were nine instances where women left their names at the bottom of their story. As a team we discussed the possibility that some women may have wanted ownership of their stories and what happened to them, and therefore wanted to have their names publicly associated. By removing their names from the stories, were we imposing anonymity? According to Eveline Vlassenroot and colleagues, non-pseudonymized data cannot be disseminated unless first, there is explicit consent; second, the personal data has been made public by the data subject themselves; and third, the personal data is closely related to the subject’s public/historical character.²² Given that we were unable to get direct consent from the contributors, we decided that all names, initials, and pseudonyms would be redacted from each post.

We are conscious that many women who contributed to the In Her Shoes Facebook page wanted to publicly share their stories in order to counter public stigma against abortion and to share their lived experiences. The question was, what level of anonymization did we want to impose on the stories? Initially, we started with a heavily anonymized process removing any direct or indirect identifiers. Examples of direct identifiers are a person’s full name, address, date of birth, and so on. Indirect identifiers are information that on their own are not enough to identify someone but, when linked with other information, could be used to deduce the identity of a person. For example, these might include age, gender, education, and employment status. After some time we realized that by removing both direct and indirect identifiers we were heavily redacting information, and we felt this was damaging the quality of the data. For example, age could be an important factor in the story if the woman was particularly young at the time of pregnancy. Below is an example of a story that we had heavily redacted, with square brackets indicating the material redacted.

Post 1: 13 May 2018 at 07:32

²⁰ “Data Protection Act 2018,” Irish Statute Book, 2018, <https://www.irishstatutebook.ie/eli/2018/act/7/section/-42/enacted/en/html#sec42>.

²¹ Kinder-Kurlanda et al., “Archiving Information from Geotagged Tweets,” 1.

²² Vlassenroot et al., “Web-Archiving and Social Media,” 121.

I am [in my mid 30s] & I have [medical condition]. I was diagnosed at approx. 2 weeks old. My health was gradually deteriorating as expected for most people with [medical condition] with age. In [early 2010s] there was medical breakthrough for people with my [medical condition] & the new wonder drug [name of drug] was born.

This level of redaction makes the stories difficult to follow and almost impossible for the researcher to use. In the case above, we decided to keep the information but redact only the name of the hospital she attended. Decisions on all redactions were given explanatory notes that the wider team discussed and agreed upon. We acknowledge this is a very specific case that some medical practitioners might remember, and thus there is a slim but possible chance that the author could be identified.

Nonetheless, we chose a “minimized anonymization process” for the project. This process recognizes that some details of the story may be recognizable to people who are close to the author such as close family, relations, or friends. To protect the identity of the author, the collection is not open to the public but restricted to those for teaching and research purposes. Decisions about the minimized anonymization process were made on the basis that an individual cannot be identified with reasonable effort by the general public based on the data provided, or by combining any public data with the data provided. It is our intention to protect the author from being identified by the general public, rather than from people who know and are familiar with the story. Below are the guidelines we developed for the minimized anonymization process.

Table 1. Minimized Anonymization Process: General Guidelines

Information Redacted	Local areas in Dublin and other cities and townlands in Ireland are redacted. Birthdays or dates of birth are redacted. The author’s occupation and partner’s occupation are redacted.
Information Kept	Large cities in England and Ireland are not normally redacted. Age is not redacted. Year or dates are not redacted. Medical conditions are not redacted. Indirect mention of people (e.g., “my sister”) are not redacted. (Names are redacted.) British Pregnancy Advisory Service (BPAS) and Marie Stopes (agencies providing reproductive healthcare) are not redacted as many Irish women use these services abroad. Hospital names are not normally redacted. In some cases, accounts of treatment by the hospital toward the patient might expose the patient to accusations of defamation by hospital staff. If this could potentially be the case, we anonymize the hospital.

Below are examples of stories that were anonymized using the minimized process.

Post 2: 09 May2018 at 09:40

I am [name redacted]. I am 30 years of age from the west of Ireland, mother to a beautiful boy, 2 years of age, partner to a strong minded, [occupation redacted] man

and I have been aching to scream my story from the rooftops. . . . I got my letter, my appointment was for the [date redacted] just after my 27th birthday. . . . On Tuesday the 24th of February 2015 at 6:30am we flew to Manchester from Dublin. . . . It turned out that we were all on the same flight home that night at 21:40, the loved ones holding us as I at 27 and her at 14 just sat there numb, bleeding, abandoned by our country, lost to go back to an island that didn't care or want us, an island that if they found out would through [*sic*] us in jail for 14 years.

In this case the location is not specific enough to be traceable. The author's husband's occupation has been redacted as it might make him potentially identifiable. We have redacted the date they found out she was pregnant because it was just after her twenty-seventh birthday, which might allow someone to work out her date of birth. We decided to include the exact date and time of appointments and flights. Some close family and friends might be familiar with these dates but they would not be recognizable to others, and therefore we feel are not a risk to their identity.

Post 3: 08 May 2018 06:45

He came from the new family who had moved into our small town and for some reason I was drawn to him. I was in 6th year at school, top of all my classes. Nice little weekend job at the local café. . . . My dying grandfather told me with his last words to me "you're a terrible daughter you broke your mother's heart." . . . I had a baby girl. She was a beautiful baby. She was also a very ill baby. She had to be resuscitated a couple times at birth, she had heart & lung problems. After a while in intensive care in Dublin they found out she had a disability. My family are carriers. We didn't know until we had tests. My sisters have decided to never have children because of it and I have a suspicion that one sister had an abortion because of it. Since then my daughter has needed full time care. She is now an adult. I have never had another relationship or any further children. My life is spent looking after my daughter's needs. I work during the few hours she is at day services.

From this case we know that the father was in sixth year in school at the time of the pregnancy. We know that the author was also in sixth year and working in the local café at the weekend. Someone local could potentially know the individuals if they were familiar with the story, but we felt that the fact that they were both in school was an important factor to the story. We decided not to redact the details on schooling and work. Furthermore, the author tells a very specific story about her grandfather's last words to her. This could potentially identify the author to others in her family who are quite likely to know this information. However, the likelihood of anyone outside of close friends and family knowing this information is small, and we felt it was of value to the story. The author identifies her baby as having heart and lung problems. She describes medical information about not only her daughter but her family. Her family did not consent to have this information shared. Knowing that the author had a child in the 1990s while she was in primary school, and that the child was born with complications, would not identify her to the general public but she could be identifiable to family members and possibly other close contacts. We noted that the person posted the story to argue in favor of abortion, despite having a disabled child. The disability is a key identifying part of the story but also a key part of the story to the author. For this reason we decided to retain the material.

Cataloging the Dataset

We developed a cataloging process combining research coding with cataloging. A postdoctoral researcher read each individual post and sociologically coded each post using Anselm Strauss and Juliet Corbin’s grounded theory approach.²³ The grounded theory approach involves identifying common topics that emerged from the posts, for example, fatal fetal anomaly, postnatal depression, and so on. Using MaxQDA software, a Computer Assisted Qualitative Data Analysis Software (CAQDAS), the researcher developed a “codebook” based on analysis of research material. When the codebook was finalized, the researcher and digital archivist on the team decided on appropriate catalog terms using the Library of Congress vocabulary and Hasset vocabulary.²⁴ Catalog terms were decided based on the sociological terms used. See table 2 for the breakdown of grounded theory codebooks and approved catalog terms.

Table 2. Codebooks and Catalog Terms

Codebook	Catalog Term	Library
Abusive childhood Support from parents/family	Family environment	Hasset
Abusive relationship	Intimate partner violence	Library of Congress (LOC)
Adoption	Adoption	Hasset
Alcohol or drug abuse	Substance abuse	LOC
Anti-abortion protests	Pressure groups	Hasset
Attempted miscarriage through self-harm	Self-harm	Hasset
Bias/judgement from healthcare workers Kindness from healthcare workers	Health professionals	Hasset
Catholic Church/religious influence	Religious beliefs	Hasset
Child sexual abuse	Child sex abuse	LOC
Contraception	Contraception	Hasset
Difficulties getting pregnant	Miscarriage	Hasset
Fatal fetal abnormality (FFA)	Antenatal diagnostic tests	Hasset
Financial struggles	Finances	Hasset
Homelessness	Homelessness	Hasset
Medical complications before/after abortion/medical condition	Pregnancy complications	Hasset
Mental health	Mental health	Hasset
Lack of consent to medical procedure	Informed consent (medical law)	LOC
Migrant	Migrants	Hasset
Miscarriage	Miscarriage	Hasset
Post-natal depression	Post-natal depression	Hasset
Psychiatric hospital admission	Psychiatry	Hasset
Protests	Protest	Hasset
Rape, sexual assault, nonconsensual sex	Sexual consent	Hasset
Sex education	Sex education	Hasset
Single mother	Single mother	LOC

²³ Strauss and Corbin, *Grounded Theory in Practice*.

²⁴ “Controlled Vocabularies”, Library of Congress, <https://www.loc.gov/librarians/controlled-vocabularies/#Library%20of%20Congress%20Subject%20Headings>; “Hasset Thesaurus,” UK Data Service, 2023, <https://hasset.ukdataservice.ac.uk/thesaurus-search.aspx>.

Social services	Social services	LOC
Sterilization	Sterilization	LOC
Stop eating as response	Eating disorders	Hasset
Suicidal thoughts/attempts	Suicide	Hasset

Many of the stories within the collection include sensitive issues such as intimate partner violence, suicide because of unplanned pregnancy, and attempts to bring on miscarriage by harmful and dangerous means. The question is how we adequately categorize the collection using appropriate language. This was especially difficult when it came to issues of rape, sexual assault, and sexual violence. The author may have used language such as “he forced himself on me” but not explicitly used the terms “rape,” “sexual assault,” or “sexual coercion” to describe the experience. Should the archivist have the authority to define the author’s experience? In cases such as this, we decided to use the subject term “sexual consent,” implying that there was a question of consent (see table 2). Due to distressing material within the posts, a trigger warning/sensitivity alert was added in the metadata of each post.

Additional Precautions

In order to deal with some of the ethical issues arising throughout the project, we formed a Stakeholder Advisory Forum, bringing together a group of experts on ethics, archiving, research, and reproductive healthcare to ensure more equitable decision-making.²⁵ The forum also includes representatives of the grassroots volunteer groups whose collections we are archiving. As we made decisions on cataloging and anonymization, the forum was consulted for their input and approval.

Ethical approval for the collection was sought and granted by Maynooth University’s Ethical Committee. The project team also developed an “Ethical Protocol,” spelling out the ethical challenges and clearly explaining our decisions to address each of these challenges. The purpose of the protocol is “to ensure the quality and integrity of our research and archiving processes; to outline the ethical issues which arise from this project and provide solutions to these challenges; and to ensure transparency in the project.”²⁶ The protocol was published alongside the initial launch of the collection on 8 March 2022.

In order to inform those who had posted to the Facebook page, we developed along with the protocol an FAQ document explaining the archiving procedure and what we were planning to do. This information was posted on the “In Her Shoes” Facebook page and our website. It outlined that the posts on the page were being archived and answered questions that someone who posted their story on the Facebook page might have, such as, What will this mean? Why is the project doing this? And what do I need to do? It stated, “Access to the majority of individual abortion stories will be restricted to people engaged in teaching and research, with access provided on a case by case basis following requests. These posts can only be used for educational purposes. . . . The In Her Shoes Facebook page does not contain any links to personal Facebook accounts. We are not archiving the comments/likes on each of the posts.”²⁷

²⁵ “Stakeholder Advisory Forum,” Archiving Reproductive Health, Digital Repository of Ireland, 2021, <https://dri.ie/archiving-reproductive-health/#stakeholder-advisory-forum>.

²⁶ “Ethics Protocol: Archiving Reproductive Health,” Digital Repository of Ireland, 8 March 2022, <https://doi.org/10.7486/DRI.6t05h703d-1>.

²⁷ Erin Darcy, “In Her Shoes: Women of the Eighth,” Facebook, 2018, <https://www.facebook.com/InHerIrishShoes/>.

We have a contact email for anyone who has questions or concerns about the project but we have not received any emails to date.

We have developed a Notice and Action policy particularly for queries in relation to the In Her Shoes collection. The Notice and Action policy states, “On receiving a request, we will try to ease any concerns of anonymity, traceability or otherwise by altering the post if possible in a mutually agreed way. Where this occurs, the post will indicate that it has been altered or redacted. . . . If the complaint is in relation to consent to archiving, then the item and its metadata will be removed, and will be permanently deleted from the repository, if that is the wish of the petitioner.” There have been no queries or requests for the removal of any story to date. Those who request access to any collection on the DRI must agree to the DRI End User Agreement, which states that they must “agree to use the digital objects only in accordance with this End User Agreement” and “ensure that full compliance with all data protection laws is observed when the digital objects, or derived material, are published or presented in a public forum.”²⁸

We developed a self-care protocol for archivists and researchers intending to work with the In Her Shoes dataset. For the project, the postdoctoral researcher was required to conduct a detailed reading of all 1,394 posts to the In Her Shoes page. The stories contain multiple accounts of rape, sexual assault, intimate partner violence, attempted suicide, and other forms of hurt. Each one of these posts was read and sociologically coded and then read for a second time and anonymized. We were conscious that vicarious trauma can occur following exposure to stories of traumatization, and to the intense emotions exhibited by those who have experienced trauma.²⁹ In order to deal with the potential trauma of working with sensitive and traumatic data, the researcher developed a Researcher and Archivist Self-Care Protocol.³⁰ We published this protocol alongside the In Her Shoes collection and have made it available to all researchers working with traumatic data. It offers suggestions for trauma mitigation: ways researchers/archivists can practice self-care when working with traumatic and distressing data. Those wishing to work with the In Her Shoes collection are encouraged to review the self-care protocol and develop their own.

Conclusion

This article explores the ethical challenges surrounding archiving social media, particularly in relation to sensitive data. We discuss issues such as individual consent and copyright, finding the correct level of anonymization, cataloging traumatic experiences, and ensuring the balance between protecting identity and sharing important lived experience. Anonymization can undermine the purpose of the story to the point where it no longer bears the author’s original intentions, therefore in our project we opted for a minimized level of anonymization along with restricted access to the collection. Cataloging sensitive data can be challenging, and appropriate language is important. We understand that language also changes with time and that the terms chosen may need to be reviewed in future. We hope this article will offer insight and guidance for those working to archive similar datasets to the In Her Shoes collection. We understand that as social media archiving develops and with possible forthcoming legislative changes, these decisions may need to be reviewed. More clarity around the practice of social media archiving

²⁸ Digital Repository of Ireland End User Agreement, DRI, 2023, <https://repository.dri.ie/pages/terms>.

²⁹ See Lambert, “Vicarious Traumas,” and Markowitz, “The Better to Break and Bleed With.”

³⁰ Lorraine Grimes, “Self-Care Protocol for Archivists and Researchers: Archiving Reproductive Health,” Digital Repository of Ireland, 8 March 2022, <https://doi.org/10.7486/DRI.7079k0713-1>.

is needed, and archivists would greatly benefit from greater information-sharing on social media archiving, particularly in relation to archiving sensitive data.

Overall, we argue that the benefit to archiving the In Her Shoes Facebook page outweighed the risks. With the recent overturning of *Roe v Wade* and the rollback of provisions of abortion services in the United States and other places around the world, the personal accounts within the In Her Shoes collection provide an important insight into the realities surrounding illegal abortion. We hope that preservation of women's narratives will contribute toward destigmatization of abortion by encouraging further research in the area. Archiving this collection may lead to important publications within the fields of reproductive healthcare and human rights, and possibly influence legislative and policy change, not only nationally but internationally. As Marika Cifor and colleagues argue, knowing the histories and legacies of activist movements enables current activists to "know that they're part of a larger movement. It is empowering to learn that you are part of a larger and ongoing struggle for justice."³¹ Furthermore, the effort of archiving the collection acknowledges the work of the administrator in the collection of these stories and, most importantly, ensures that the women who shared their experiences are remembered and their stories not forgotten.

References

- Breed, Miana. "Capturing a Moment: The Practices and Ethics of Social Media Archiving." M.S. degree, University of North Carolina, 2019. https://cdr.lib.unc.edu/concern/-masters_papers/6395wc44v.
- Carter, Rodney. "Of Things Said and Unsaid: Power, Archival Silences, and Power in Silence." *Archivaria* 61 (2006): 215–33. <https://archivaria.ca/index.php/archivaria/article/view/12541>.
- Caswell, Michelle. *Archiving the Unspeakable: Silence, Memory, and the Photographic Record in Cambodia*. Madison: University of Wisconsin Press, 2014.
- Caswell, Michelle. "Toward a Survivor-Centered Approach to Records Documenting Human Rights Abuse: Lessons from Community Archives." *Archival Science* 14 (2014): 307–22. <https://doi.org/10.1007/s10502-014-9220-6>.
- Cifor, Marika, et al. "What We Do Crosses Over into Activism: The Politics and Practice of Community Archives." *Public Historian* 40, no. 2 (May 2018): 69–95.
- Darcy, Erin. *Women of the Eighth*. Dublin: New Island Press, 2020.
- 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 and Culture* 47, no. 2 (2018): 33–44. <https://doi.org/10.1515/pdte-2018-0011>.
- Georas, Chloé S. "Networked Memory Project: A Policy Thought Experiment for the Archiving of Social Networks by the Library of Congress of the United States." *Laws* 4, no. 3 (2014): 469–508. doi: [10.3390/laws3030469](https://doi.org/10.3390/laws3030469).

³¹ Cifor et al., "What We Do Crosses Over into Activism," 91.

- Jimerson, R. C. “Archives for All: Professional Responsibility and Social Justice.” *American Archivist* 70, no. 2 (2007): 252–81.
- Jimerson, R. C. *Archives Power: Memory, Accountability and Social Justice*. Chicago: American Library Association, 2009.
- Kinder-Kurlanda, Katharina, et al. “Archiving Information from Geotagged Tweets to Promote Reproducibility and Comparability in Social Media Research.” *Big Data and Society* 4, no. 2 (July–December 2017): 1–14. doi: [10.1177/2053951717736336](https://doi.org/10.1177/2053951717736336).
- Lambert, Sharon. “Vicarious Trauma: The Impact of Working with Survivors of Trauma School of Applied Psychology.” In *Toward a Trauma-Responsive Criminal Justice System: Why, How and What Next?* 5th Annual Irish Criminal Justice Agencies (ICJA) Conference, 2018, 30–36. [https://www.acjrd.ie/images/PDFs/icja/ICJAC REPORT 2018 - Toward a - Trauma-Responsive Criminal Justice System Why, How and What Next.pdf](https://www.acjrd.ie/images/PDFs/icja/ICJAC_REPORT_2018_-_Toward_a_Trauma-Responsive_Criminal_Justice_System_Why,_How_and_What_Next.pdf).
- Lindström, Lisa. “Archiving in the Era of Online Activism: Challenges and Practices of Collecting and Providing Access to Activist Social Media Archives.” MLIS thesis, Lund University, 2019. <https://lup.lub.lu.se/student-papers/search/publication/8980793>.
- Markowitz, Ariana. “The Better to Break and Bleed With: Research, Violence, and Trauma.” *Geopolitics* 26, no. 1 (2019): 94–117. doi: [10.1080/14650045.2019.1612880](https://doi.org/10.1080/14650045.2019.1612880).
- McCrow-Young, Ally. “Approaching Instagram Data: Reflections on Accessing, Archiving and Anonymising Visual Social Media.” *Communication Research and Practice* 7, no. 1 (2021): 21–34. <https://doi.org/10.1080/22041451.2020.1847820>.
- McDonnell, Ciara. “In Her Shoes: How a Facebook Page Became the Voice of the Repeal the Eighth Movement.” *Irish Examiner*, 8 October 2020. <https://www.irishexaminer.com/-lifestyle/healthandwellbeing/arid-40060805.html>.
- Morrone, Melissa, ed. *Informed Agitation: Library and Information Skills in Social Justice Movements and Beyond*. Sacramento, CA: Litwin Books and Library Juice Press, 2013.
- Rosenbaum, Judith. “Archiving #MeToo: Past, Present, and Future.” *American Jewish History* 104, nos. 2–3 (2020): 251–63. <https://doi.org/10.1353/ajh.2020.0027>.
- Sadler, Rosa, and Andrew Martin Cox. “‘Civil Disobedience’ in the Archive: Documenting Women’s Activism and Experience through the Sheffield Feminist Archive.” *Archives and Records* 39, no. 2 (2018): 158–73. doi: [10.1080/23257962.2017.1387523](https://doi.org/10.1080/23257962.2017.1387523).
- Strauss, Anselm, and Juliet Corbin, eds. *Grounded Theory in Practice*. Thousand Oaks, Calif.: Sage, 1997.
- Sutherland, Tonia. “Archival Amnesty: In Search of Black American Transitional and Restorative Justice.” *Journal of Critical Library and Information Studies* 1, no. 2 (2017): 1–23. doi: [10.24242/jclis.v1i2.42](https://doi.org/10.24242/jclis.v1i2.42).
- Uimonen, Paula. “#MeToo in Sweden: Museum Collections, Digital Archiving and Hashtag Visuality.” *Ethnos* 85, no. 5 (2020): 920–37. doi: [10.1080/00141844.2019.1640264](https://doi.org/10.1080/00141844.2019.1640264).

Vlassenroot, Eveline, et al. "Web-Archiving and Social Media: An Exploratory Analysis." *International Journal of Digital Humanities* 2 (2021): 107–28. <https://doi.org/10.1007/s42803-021-00036-1>.