A Critical Reflection on Social Media Research Using an Autoethnographic Approach

Jordan Eschler
University of Washington
Information School
jeschler@uw.edu

Abstract
The landscape of social research is changing rapidly due to the wealth of data that researchers can use and analyze from the social media engagements of real people. As social media researchers, we use a spectrum of methods to study use and behavior, and the extent to which we engage users in our research affects the knowledge we hope to gain. In this paper, I use an autoethnographic approach to analyze the interplay between my publicly-available social media content, my private social data, and my timeline of lived experience in diagnosis and treatment of Stage IV cancer to demonstrate that the knowledge researchers gain solely from social media data is partial ("almost" knowing). Using the case of my own experience, I argue that research findings through social media, and the knowledge contributed through such research, would be improved by engaging more frequently in qualitative research that includes social media users.

1. Introduction

With the explosion of engagement with social media among people of many socioeconomic and cultural backgrounds [36], researchers may now access copious public data on various social media platforms that can give glimpses into everyday life experiences of individuals. Access to publicly-available social media artifacts (short status posts, blogs, and photo or video data) creates an opportunity to gather material from some populations and communities where no substantial data collection opportunities existed before. This access to public user data expands researchers’ ability to ask new questions, or old research questions in new contexts.

The spectrum of research methods used in the social media realm ranges from “big data” and computational methods to close readings through content analysis or ethnographic work (for the purpose of this paper, I will refer to the latter close reading type research as “small data” research). In big data approaches, larger sets of data are engaged to generate patterns of use [7], information about social networks (e.g., [9]), or frame observations in terms of concepts such as virality (e.g., [34]) or sentiment (e.g., [21]). In this paper, I am focusing on research in the small data realm, in which qualitative approaches are used to closely examine use and interactions via social media. The analysis in this paper is presented to argue for directly engaging users more frequently in small data social media research; in particular, I believe eliciting users’ reflections and interpretations of their own social media use and content would lead to better questions and new research insights.

The inclusion of users in small data research has been a particular concern of mine in the past year for both personal and professional reasons. One, because I was diagnosed with and treated for Stage IV Hodgkin’s lymphoma, and two, because I use small data research methods to explore the experience of cancer among other young adults in my own work. The autoethnography work described here has been a test to see what insights I can gain by constructing meaning out of my social media use patterns and content during cancer diagnosis, treatment, and remission phases.

Ultimately, the contributions of this paper are threefold. First, the use of autoethnography lends personal perspective from a member-researcher [1] of a vulnerable population to the critical discussion about social media research. Second, I point to good work done in other disciplines which can help researchers to think about including social media users in research designs which typically focus solely on social media data (i.e., artifacts). Finally, I discuss how this autoethnographic exercise has informed my subsequent work with young adult cancer patients and survivors, and outline practical approaches researchers can use to include users in social media research.
2. Background

I have framed the background to my approach in this paper in three parts. First, I explain how autoethnography is used in general, as well as in this paper. Then, I cite literature from the disciplines of museum studies and Indigenous Knowledge for the purpose of articulating concepts of interpretation, curation, and power in artifact research (e.g., social media data). Finally, I point to recent human-computer interaction research that both engages users in social media research, which are good models of the type of research I urge here, and elicits important user perspectives and reflections about social media artifacts.

2.1. The use of autoethnography

Broadly, online ethnography methods incorporate engagement with users in studying online spaces as cultures and cultural artifacts [23]. As one type of ethnographic method, autoethnography calls for the researcher to become the informant. Specifically, I use an autoethnographic exercise that reconstructs the timeline of my cancer diagnosis, treatment, and remission to examine what can be known about my cancer experience by combining my social media artifacts and critical reflection on the experience. Autoethnography studies are the “production of highly personalized accounts where authors draw on their own experiences to extend understanding of a particular culture” ([38]; p. 3). Autoethnography is never generalizable; however, the insights it produces are intended to enlighten the researcher through reflection (and benefit other researchers, ideally, through publication). Through this method, I add one more critical viewpoint in the ongoing dialogue regarding the use of social media data in research (e.g., [16]).

Cunningham and Jones [13] invoke the benefits of autoethnography for its utility in self-reflection on technology tool use. The use of social media artifacts, I found through this process, is particularly powerful: the artifacts that I recovered in obtaining my social media data confronted me with patterns of use, and specific acts, that I sometimes could not remember or rationalize in hindsight. This reflection on my own social media data has helped to form a personal, critical viewpoint of social media research that underlines the importance of engaging with social media users in small data methods. In this vein, autoethnography has been successfully used, particularly in health information systems [37], to help researchers understand and empathize with participants in research. To articulate issues associated with artifact research, I turn in the next section to museum studies and Indigenous Knowledge literature, which are helpful for understanding concepts such as power and ownership in research where artifacts are used to inform research about people.

2.2. Artifact research in museum studies and Indigenous Knowledge systems

Museum studies and Indigenous Knowledge scholars grapple with the presentation of knowledge that profits from separating physical artifacts from production and context. If we view social media researchers as using methods analogous to those of a researcher studying cultural, physical artifacts (e.g., collection, curation, interpretation, and display of artifacts), then it is useful to learn from these disciplines. I have found the criticisms and reflections in these disciplines to be helpful for conceptualizing problematic aspects of collecting and interpreting artifacts such as social media data.

Particularly among museum studies theorists of late (post-1990), the literature calls for considering responsibilities of collecting, using, and displaying artifacts as a process of constant decision-making and value-weighing [17]. In the “curatorial” role that we as researchers fill, we not only collect, but select data to investigate our research questions. We also ultimately share specific example artifacts that illustrate our findings (even, as with a museum of physical objects, we digitally store vastly more artifacts from data collection than we could ever hope to illuminate individually). The act of characterizing meanings attached to artifacts chosen for display is a vital research act: such artifacts “provide vehicles for people to think with and more importantly to feel with” ([25]; p.54).

“Diplaying” social media artifacts in our research – to illustrate our findings – should not prevent us from involving social media users in populations of interest. To this end, it is helpful to learn from Indigenous Knowledge scholars, who strive to include populations of interest in relevant research and artifact stewardship. Agrawal [2] discusses researchers’ act of separating artifacts from creators in terms of power differentials in research. Issues of artifact ownership also arise [4], related to the intricacies of power and control between researchers and those researched. I note here that, at an individual level, the state of “ownership” of social media property reflects a similarly unsettled matter of “rights” (see also [31]). For this paper, the act of performing autoethnography eliminated the problems of power and ownership described above. Insofar as I could recover the data I produced on social media (and
Finally, Indigenous Knowledge scholars have long examined and discussed the ramifications of extracting knowledge from its cultural and “embodied” context [6]. The resulting extraction ostensibly results in “research ideology that puts the accumulation of knowledge ahead of the interests of the people studied” ([32]; p. 20). The solution to regaining a balance between researcher and research participant interests, Menzies [32] argues, is for researchers to better incorporate “the interests of people studied.” In the same way, small data social media researchers should position their methods in terms of social media user involvement. For example, if researchers make the choice to exclude users from the formulation of research questions, but collect and analyze users’ social media data to answer those questions, what are the implications and assumptions about the researcher’s interests versus those of the participants (who may or may not know they are participating in research)?

The following section gives an overview of previous research that has employed user reflection on social media artifacts and use. These studies are good examples of the type of inclusive user research I argue for in this paper, and demonstrate insights that become accessible when including users in social media research.

2.3. User-assigned meanings to digital artifacts

Gulotta et al. [22] found that digital information systems are “rapidly changing” the way users perceive what constitutes a “meaningful possession.” Lindley et al. [30] echo this finding, noting that participants interviewed in their study were not altogether clear on which photos from their social media archives were even “worth keeping” (p. 755). Additionally, in a conceptually related study, Zhao and Lindley [42] found that giving users the ability to design “keepsakes” from social media data led to various reasons for sentiment associated with digital artifacts, often associated with meanings of offline relationships. Given these findings, researchers may find themselves in the position of imparting weight and meaning, and inferring values, using artifacts that are themselves changing in terms of weight, meaning, and values for their creators.

The studies above are examples of user-inclusive social media research, and demonstrate how users can reflect upon social media artifacts to yield new insights to artifact meaning and significance. Researchers should acknowledge that social media artifacts were perhaps never expected by their creators to live beyond the fleeting moment of a strong emotion or lived experience expressed in a social media post. By incorporating social media users in the research process we can employ methods that elicit reflection and interpretation of social media artifacts. As in the above studies, this practice enables researchers to triangulate input from users and public social media artifacts, thereby increasing the nuance of research findings from small data inquiry. In the following section, I describe the collection and analysis of my own batch of “small data.”

3. Methods and analysis

The methods involved in reconstructing the “timeline” of my cancer experience consisted of merging two separate streams of data over the time period between January 2014 (initial diagnosis) and October 2014 (first clear scan that marked the official “No Evidence of Disease,” or NED, medical status). The first stream of data, which reflects an “objective” medical timeline, consists of paperwork such as permission forms for procedures, outpatient instructions, and medical bills. The purpose for constructing the timeline of events from medical artifacts was threefold: first, I suffered memory loss from chemotherapy and was unable to recall the order of events, so this timeline gave me reliable information about where and when certain events occurred. Second, creating a timeline allowed me to see clusters of social media activity, or prolonged periods of social media silence, over standard units – weeks and months. Third, triangulating the medical timeline versus my own production of content gave me valuable information about the context of social media data production.

The second stream of data was constructed from downloads of digital data from Twitter, Facebook, and Gmail. These three platforms represent the bulk of my social media use in producing content; other platforms I used for information included cancer forums, but the interaction there was limited to reading posts, and a handful of private messages exchanged with other users. For Facebook, I pulled down my own zip file, which anyone can request, and looked over the segregated files of photos and statuses that the download file provides (through https://www.facebook.com/help/). I toggled between the raw data in the download and reviewing older posts on my timeline, in order, to see how the timeline “read.” Twitter data was provided in a .csv file download (https://twitter.com/settings/your_twitter_data) in reverse chronological order. I read through these posts but did not view these posts on Twitter, as I had with
Finally, with email, I did not read all messages kept over the time period of diagnosis to remission, but instead consulted the email archive during periods of silence on social media, to fill in gaps in the timeline.

The two streams of data were combined by overlaying social media posts on the medical timeline. That way, I could track the correspondence of social media use (or not) with significant events from my medical history. “Significant events” here invariably corresponded with spikes in medical bills, e.g., periods of hospitalization. For example, I know which dates I was an inpatient at the hospital, and therefore could identify related social media posts, if applicable. Particularly in periods of low social media use, I found emails to be valuable archives of thoughts and feelings that I shared privately, one-to-one, with people closest to me. The audiences for each type of artifact can differ greatly; these differences are outlined in Table 1.

Much as with any approach I would take to getting comfortable with understanding trends in use, I quantified trends in Twitter and Facebook engagement types first (e.g., original Tweets and Retweets; text posts, and photos). My characteristic social media use varied by platform. I posted regularly on Facebook – 162 posts during the period, or about every other day, on average. However, my Twitter posting (1,039 posts January to October 2014) was much more uneven, as I show in the analysis below. Once I understood patterns of use, I dove into the content of social media use to determine when I was producing “cancer content” versus “non-cancer content.”

For the purpose of this exercise, “cancer content” was anything I posted that mentioned my experiences related to treatment, tests, side effects, etc. “Non-cancer content” was everything else, which I imagine I would have posted regardless of medical status. The purpose of identifying “cancer content” was primarily to figure out how prevalent it was in my online presence, through this crude counting method. From there, I was able to examine the format of content posted (pictures, hyperlinks to other content, or text), the nature of the content posted (e.g., tone) and the reception of the content, if applicable (comments on Facebook or responses on Twitter).

### 4. Results

I used two framings to make sense of the medical timeline overlaid with social media use. First, I looked at significant events on the medical timeline to assess use of social media around such events. Significant events included medical milestones in treatment, such as first and final chemotherapy sessions, hospitalizations, and changes in my status, such as achieving NED status. Second, I looked at trends in social media use – where was I posting frequently, or not as frequently? In addition, where are there gaps in my apparent social media use, and what was happening at that time, based on the email and the medical timeline? By using these two approaches, I am able to connect my social media use with the timeline of the medical cancer experience, but also include and examine moments of personal importance that may not be significant medically. Findings from the approaches are detailed below.

#### 4.1. Significant events

The types of events I identified on the medical timeline included, in order: diagnosis; chemotherapy
treatments (with the first and final treatments being particularly important to me personally); two hospitalizations; medical clearance to resume travel; and the clear body scan demonstrating NED status. The choices I made around social media use at the time of each of these events differ both in terms of how I shared information about individual episodes, and how the choices I made changed over time during my experience.

First, news of my diagnosis was shared almost exclusively over email (and phone calls, though that data are not available for review). In the emails, I informed those closest to me with a personal note. Facebook use during the period of diagnosis – which lasted about a month from the first bad scan to node biopsy and staging – includes posts about everyday life: status updates about laundry, watching a movie, etc. In a sense, the cancer experience was truncated from a social media point of view – I didn’t mention cancer for a wide “audience” at all until treatment started. Without access to more private correspondence (e.g., e-mail), my emotional experience of diagnosis would be invisible to a researcher using public social media data.

The occasion of my first chemotherapy treatment was when I decided to make a broader “announcement” to friends on Facebook in the form of a photo post. I mentioned only that it was my first chemotherapy treatment, but not the specific diagnosis. The comments on that post were overwhelmingly supportive, which is to be expected; the audience was limited to people I knew fairly well.

The “first chemo” post had been my first post to Facebook in some time, probably due to the fact that I was asleep most days, or spending hours at the cancer clinic getting blood transfusions. I recall checking over my Facebook post security settings very carefully before posting the “first chemo” picture to make sure only friends could see it. For this reason, if a researcher were to pull my public Facebook posts to study the “appearance” of cancer on social media, this post would ostensibly not be visible. At this point, the experience of cancer was still very new, and the loss of control in my day to day routine forced me to be cautious about appearing emotional or scared in a public way (I recall Goffman and the concept of self-presentation [18]).

I think it is important to note that I rejected using a specialized platform to coordinate updates (specifically, CaringBridge, which offers private or public updates for people who require caregiving for chronic or acute medical conditions; www.caringbridge.org). When a friend suggested that I update friends and family using Caring Bridge, I was initially taken aback. This was because I had recently followed two acquaintances on Caring Bridge, who ultimately died from cancer. I associated Caring Bridge as being used by people who were “sicker than I was.” Facebook seemed like a more normal and natural platform to share treatment progress.

Neither instance of hospitalization was clearly documented on social media. I used my phone to take frequent “selfies” simply to see what I looked like, since there were no mirrors in my hospital rooms. These selfies never made it to social media, and are therefore another aspect of the experience a researcher might miss, particularly if she were interested broadly in technology use during medical experiences. However, I did sometimes allude on social media to other outpatient procedures if I could make a joke. Such jokes often belied apprehension in my case (although the use of humor in information sharing is not necessarily uncommon in the case of young adult cancer; see [24]).

My last chemotherapy treatment was celebrated with a photo of me holding my “final chemo” poster, which had names of people I wanted to thank written on it (see Figure 1). I knew final chemo or radiation treatment pictures, in which the patient holds a celebratory sign or decorated cake, were a tradition in the cancer community from seeing posts on Facebook groups I followed. In fact, I re-posted the picture of me to one of my favorite Facebook groups. The picture I posted to this group is public, meaning if a researcher were to scrape data from the group, she could see my picture. For this reason, I carefully weighed whether I wanted to post it, but in the end I felt the positive encouragement afforded from posting the picture meant more to me than privacy in this instance.

Figure 1. “Last chemotherapy” poster shared on Facebook
4.2. Trends in social media use

I posted more than I realized about cancer during the course of treatment. This has to do with using Facebook as a platform to update friends and family; these posts added up over time. Specifically, posting pictures of myself taken during infusion to Facebook became routine on treatment days every two weeks. My caregivers also reminded me to update my Facebook picture posts when they took me to treatment, which probably reduced the work for them in updating others about how I was doing. These Facebook pictures were always set to update only “friends,” though I did “friend” some distant family members to give them access to the updates.

Facebook use diverged from that of Twitter in posting pictures. It was unusual for me to post a picture of myself using my public Twitter account; I counted only two cancer-related photo posts on Twitter, and both were “selfies,” poking fun at my baldness at the time. I made almost zero Twitter posts during the period between diagnosis and the first few chemotherapy treatments. This is not unusual, however - I post sporadically on Twitter, with spikes in activity during academic conferences. The rest of the time, I generally use the platform to consume information, such as news updates. I use Twitter primarily as a public academic profile, and it took me some time to link my private experiences with cancer with my public persona.

My Twitter data also reveals a mini-pattern that involves side effects from steroid therapy. I took ultra-high doses of Prednisone for six months. During that time, my Twitter usage jumped significantly (see Figure 2). I struggled with mania during this period and slept four or five hours per night (a common side effect; see [28]), often filling time with reading and posting on social media. Many of the Tweets are incoherent, content-wise. It is unclear if a researcher without contextual information would notice or be able to interpret this pattern. This observation leads me to believe that incorporating qualitative data from users in combination with small data could be a powerful method to understand micro-patterns like this.

The greatest number of posts to Facebook occurred when I was medically cleared to travel at the end of treatment, despite the fact that this was not an official “medical” moment. This was another “celebration” moment that I shared with Facebook friends only. It was probably the most personally significant milestone. By sharing this update as a friends-only post, I continued to “hide” my personal cancer experience from public view on Facebook. In contrast, achieving NED status was the ultimate medical milestone, but I made a simple Facebook status update, again for friends only.

Although many of my posts regarding my personal cancer experience were visible only to friends, I should point out that I did make regular public posts about lymphoma and young adult cancer survivorship awareness on Facebook, and I continue to do so now. This is the most “public” thing about cancer that I share on Facebook, although I have also transitioned to making public posts about having lymphoma on Twitter (my account is set to public). In that way, I would be identifiable as “active” in the young adult cancer movement to the public, or to a researcher using public data.

5. Discussion

I am struck by the clarity I gained from triangulating data from several sources (medical, social media content, and experiential) to make sense of my
episodic social media use. The above results indicate some advantages to triangulating user reflections and viewpoints with social media data. However, it can be daunting to design research that directly involves social media users (particularly if research questions focus on social media use of potentially vulnerable individuals [20, 29]). As researchers, we potentially receive limited expert assistance from Institutional Review Boards (IRBs); as Cutcliffe and Ramcharan [14] point out, IRB committees “draw largely on a medical and quantitative research tradition” (p. 1000). In the sections below, I synthesize themes from the results and indicate implications for research designs that can include social media users, and not just their data.

5.1. What we lose when we exclude users

One issue that complicated the analysis of my data, but no doubt helped to make the examination of my digital footprint during the cancer experience more authentic to research processes involving use of third-party participant data, was the memory loss caused by chemotherapy treatment. I am unable to recall my state of mind in posting much of my social media content, much as a researcher is unable to divine user intent by merely examining online user data. However, given the opportunity to reflect on the artifacts yielded insights that enriched my observations about use patterns and social media content. This aspect of triangulating user viewpoints and data is one that seems extensible to other qualitative and/or small data research designs.

For a researcher, one problem would be accessing sufficient cancer-related content by studying only my public “small data” set. Most of my cancer-related posts just would not show up in a collection of public social media data. As stated above, I do make public advocacy posts, particularly about lymphoma symptom awareness, but a researcher would miss out on most of my photo posts, and artifacts that indicated a change in illness phase (e.g., starting treatment or achieving NED status). The absence of public data reduces researcher access to artifacts, but including information from the user – such as through interviews or a walk-through of the user’s Facebook timeline – could illuminate “missing” artifacts and prompt dialogue about sharing habits.

Even if a researcher accessed my private Facebook posts, she would see a pattern in the data that I shared more good news than bad online. This is one example of how a cancer patient may “skew” her online presence in various ways, as would any individual managing her self-presentation online (e.g., Goffman’s theories of self-presentation [18]; in this case, Goffman’s work regarding social stigma is also relevant [19]). Acts of prospective or retrospective self-censorship may consequently limit the value of knowledge gained by examining solely social media data, perhaps particularly when related to significant life events, such as cancer treatment. The act of “confronting” artifacts and eliciting user perspectives and reflections, as described in the following section, can compensate for missing context around social media data, and add depth to research inquiry.

5.2. Users can confront artifacts to add depth

The ability to elicit user perceptions and reflections on social media behavior (what users think they do) is limited without triangulating content and patterns of use. An example from my case would be that I made a very firm decision early in treatment that I would not share on social media about being in the hospital. However, when I examined the data, I did make one Facebook post about being in the hospital. It was a one-panel comic about being in the hospital that I found on one of my cancer forums that same day. I remember feeling camaraderie with another cancer patient also in the hospital, and the comic was funny, so I shared it. This is the kind of exception, or override, to a personal social media “rule” for behavior that requires incorporating the content creator/sharer to understand, to reconcile differences in memory or perception of use versus actual behavior.

Related to the above issue, in working within the young adult cancer survivor community, I have observed instances of working-age cancer patients being encouraged to “scrub” their social media accounts for mentions of cancer to increase future employability and dateability, for single cancer survivors ([11], [12]). This is a “culture,” or “community” level reason to investigate actual social media use and content in conjunction with user reflection and perception. Because of this messaging in the young adult cancer community, as well as the basis of existing theories on self-presentation ([18], [19]), it is important to incorporate the user to reflect on self-censorship subsequent to the creation of social media content.

Leveraging the memory and reflections of social media users can also better incorporate how the users’ judgment about social media use shifts over time. Social media content is not created in a cultural vacuum. Specific to the cancer experience, the medical patient in Western culture in general (see [10]) and the cancer patient in particular (see [33], [40]) has shifted over time in the cultural consciousness, from a stigmatized individual to one of a “fighter.” The social media artifacts produced by users and studied by researchers are specific products of current societal values. For these reasons, researchers would gain
valuable insight by including users in small data research to elicit reflection on creation, revision, or deletion of content over time. To this end, the following section reviews practical implications for including users in different phases of research, informed by the autoethnographic work I have presented above.

5.3. Practicalities of involving users in research

Having gone through the process of examining my own social media “trail” throughout the cancer experience, I have incorporated some of my own observations from this work into my subsequent research designs. Below, I outline how these observations have helped me to design research that includes users directly, employing ethical inclusion of users in different phases of the research process (explained in helpful detail by Sharkey et al. [39]).

First, the needs and customs of the population of interest – in my case, young adult cancer survivors – should be incorporated in identifying a research problem area and formulating questions (as suggested in the Indigenous Knowledge work of Menzies [32]). Since each research study exposes a population to some level of risk, we must also ensure that research studies are worthwhile to justify the risk [5]. Allowing the population of interest to discuss the questions posed to inform a research study is one way of testing such a risk/benefit balance.

Involving the population of interest to inform the research question is a particularly hands-on approach, but it can be useful from a social-constructionist stance and to dismantle power differences in multiple phases of research [3]. This practice has the added benefit of extended engagement with the population, which can build trust between the researcher and the community of research interest. Even where the researcher does not engage directly with individuals in the population, Bruckman [8] has advised that the researcher should, at a minimum, become familiar with conventions of online behavior in the community of interest. For example, in the case of my own experience, it would benefit a researcher studying cancer patients online to understand the custom of sharing a “last chemotherapy” picture.

In the subsequent research phase, data collection, understanding the structure of the code that facilitates shapes, and prohibits online behaviors [27] is an essential starting point. Many platforms facilitate the sharing of photos, for example. Having examined how I shared my own photo data, and the way my feelings about my appearance changed over time, I have decided not to use or publish photographs of cancer patients or survivors shared to social media [e.g. 15].

Another advantage to consulting users when studying social media data is to understand how the production of social media content is facilitated by affordances that can become invisible when data is disconnected from the social media platform [41]. An example of such affordances would be Twitter’s privacy settings (binary; public/private), versus Facebook’s more nuanced settings for lists, etc. These affordances provide different abilities to expressing private acts [35]. When one platform allows a different level of dexterity in managing online “self-presentation” than another, the observations of phenomena under study might vary greatly, and therefore impact research findings.

Related to data collection, one proxy social media researchers can use is the Terms of Service agreement associated with each social media platform [8]. Having analyzed how I chose to share publicly (or not) via Facebook and Twitter, I see that this approach obviously calls for an assumption that individuals understand privacy settings, or else it is meaningless. However, a reasonable balance to the risk of collecting data from unwitting participants is an application of extreme caution in the subsequent research publication phase. As a research participant for this autoethnography, for instance, I was comfortable analyzing many pictures that I would not share here. I see this as another reason to approach users directly for engagement in social media research. Building rapport with users can facilitate sharing and talking about their social media data, which can increase research access and enriches the data collected at the same time.

The designs of data analysis and publication of any illustrative data should balance the level of risk users are exposed to by the data collection strategy. As stated in King’s [26] early commentary on the coming changes to research using social media data, the reporting of findings from public data should go to reasonable lengths to avoid identifying individual users. In the case that users are included in data collection – such as in interviews that elicit reflections on social media data – it is possible to construct informed consent for users about how their data will be shared. I also believe that publishing or sharing social media research should extend to the community studied, if possible. In my research designs, I have worked to include communication back to research participants, and with the larger community of research interest.

In sum, efforts to include social media users in research processes, and to inform the broader population of research interest about social media research findings, have numerous potential benefits. For the researcher, discussion of findings with the community of interest may improve research validity.
through post-hoc member checking. Discussion with community members may also lead to better future research questions. These benefits are supported by insights from museum studies and Indigenous Knowledge literature, the self-reflection that occurred through the autoethnography presented in this paper, and discussions I have generated with members of the young adult cancer community in my own work.

6. Limitations

Autoethnography research is not generalizable. The analysis I performed here is intended as a critical reflection on what I have termed small data research in social media, e.g., content analysis or ethnographic work. In addition, the linear (timeline) approach I have taken to analyze and interpret the data determines the scope of insights I have been able to glean from the process, and is by no means the only way to interpret this “small data” set.

7. Conclusion

Access to social media data introduces new considerations for all phases of research design, and researchers should recognize responsibilities to the populations we study via their social media content. Researchers must consider underlying issues of creating meaning from artifacts divorced from creators and context. With regard to these responsibilities, social media researchers effectively grapple with epistemological and ethical quandaries similar to those discussed in literature such as museum studies and Indigenous Knowledge systems. These disciplines have increasingly worked toward including artifact creators and stewards in scholarship and research; in employing the same approach, social media researchers can work to involve users in research.

While it is not always possible to painstakingly remarry social media data with creator and context, as I have done in my autoethnographic analysis, researchers should continue to work toward incorporating users in our “small data” social media research. By doing so, researchers can improve their research questions, enhance the validity of findings over time, and potentially extend benefits of research findings to the broader research population of interest. My contribution to the argument for this approach is offering some insight to the bounds of what can be known from the digital artifacts I “left behind” during my experience as a cancer survivor, and providing critical reflection by inhabiting the roles of both informant and researcher.

Acknowledgements

This work is supported by the Paul Evan Peters Fellowship (2014-16). I am indebted to the reviewers, who gave invaluable input for revising and strengthening this paper, and to Shawn Walker, Dr. Andrew Miller, and Dr. Lisa M. Vizer for their critical feedback.

8. References