Building community at distance: a datathon during COVID-19

Samantha Fritz and Ian Milligan
Department of History, University of Waterloo, Waterloo, Canada

Nick Ruest
York University Libraries, York University, Toronto, Canada, and

Jimmy Lin
David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, Canada

Abstract

Purpose – This paper aims to use the experience of an in-person event that was forced to go virtual in the wake of COVID-19 as an entryway into a discussion on the broader implications around transitioning events online. It gives both practical recommendation to event organizers as well as broader reflections on the role of digital libraries during the COVID-19 pandemic and beyond.

Design/methodology/approach – The authors draw on their personal experiences with the datathon, as well as a comprehensive review of literature. The authors provide a candid assessment of what approaches worked and which ones did not.

Findings – A series of best practices are provided, including factors for assessing whether an event can be run online; the mixture of synchronous versus asynchronous content; and important technical questions around delivery. Focusing on a detailed case study of the shift of the physical team-building exercise, the authors note how cloud-based platforms were able to successfully assemble teams and jumpstart online collaboration. The existing decision to use cloud-based infrastructure facilitated the event's transition as well. The authors use these examples to provide some broader insights on meaningful content delivery during the COVID-19 pandemic.

Originality/value – Moving an event online during a novel pandemic is part of a broader shift within the digital libraries' community. This paper thus provides a useful professional resource for others exploring this shift, as well as those exploring new program delivery in the post-pandemic period (both due to an emphasis on climate reduction as well as reduced travel budgets in a potential period of financial austerity).

Keywords Interdisciplinary, Team formation, COVID-19, Datathon, Online events, Web archives

Paper type Research paper

Introduction

It happened very quickly. Over the last week of February, our Canadian-based team was closely monitoring the ever-growing COVID-19 outbreak in the USA – the discovery of

community spread in Washington State, followed by the first community case in New York City. As scholarly events began to cancel, mostly in Europe, the last-minute cancellation of the American Physical Society's March meeting in Denver, CO made clear that North America was no longer immune from academic disruption (Cho, 2020).

Our team quickly realized the inevitable impact the virus would have on our in-person datathon, which was scheduled to be hosted at Columbia University's Butler Library in New York City between 26 and 27 March 2020. This marked the fourth datathon in a series of events supported by The Andrew W. Mellon Foundation, and for some of us, the eighth datathon organized under the broader umbrella of the "Archives Unleashed" moniker (Milligan *et al.*, 2019, pp. 265–268; Ruest *et al.*, 2020, p. 2). Our generous hosts at Columbia University Library and the Ivy Plus Libraries Confederation had provided us access to datasets, and we were looking forward to exploring their collections with our interdisciplinary attendees. Our project team had considerable experience bringing together anywhere between 15 and 50 individuals to work collectively on problems involving Web archive analysis. These interdisciplinary gatherings of humanists, social scientists, librarians, archivists and computer scientists, provided an opportunity to gather and closely collaborate on problems, share meals and help both tackle technical challenges as well as build community. Yet suddenly our emphasis on close in-person events to physically gather attendees was reason to pause.

Throughout 1 and 2 March 2020, our team concluded that an in-person datathon was irresponsible given rising concerns around the virus and the formation of emergency management committees at universities across North America. This was not an easy decision to make, as while some events had cancelled, the majority were still continuing as planned [1]. As important as our digital library event was to our project and (hopefully) to our community, the benefits did not outweigh the risk. We still strongly believed in our mission, however, and decided to proceed by transforming our in-person datathon into a new, all-virtual format.

In this article, we explore the ways in which COVID-19 impacted our datathon, using it as a window into the broader implications around transitioning on-campus and in-person events into remote events in a way that continues to serve the broader digital library community and other stakeholders. As one of the first events to make the rapid online transition, our experiences are broadly generalizable to other digital library events that are confronting the "new normal" of COVID-19. Recent pandemics, such as H1N1 influenza in 2009, led to localized school closures in the USA. However, the H1N1 pandemic did not see widespread use of nonpharmaceutical interventions such as mandated physical distancing. Indeed, one of the fears around school closures during the 2009 pandemic was that students would congregate in public places such as libraries (Navarro *et al.*, 2016, p. 12). This argument supported school reopenings. Indeed, a comprehensive overview of pandemic responses shows that we need to look back a century to the 1919 Spanish Flu pandemic to see similar widespread adoption of draconian (if necessary) nonpharmaceutical interventions such as school and other public space closures (Saunders-Hastings and Krewski, 2016).

Our article begins with an overview of the Archives Unleashed project and its situation within the digital libraries field, notably how the project operates at the intersection of tool builders, researchers and digital content access providers. We believe that this makes our experience broadly relevant. We then explore how we navigated the process of reshaping our in-person datathon to an online one, all while highlighting specific lessons learned and applying them to the broader shifting landscape confronting digital libraries. In particular, we were fortunate that our decision to host technical infrastructure in the cloud facilitated

our transition. The article then concludes with recommendations of best practices for others exploring similar implications of shifting services and events online in the COVID-19 era and beyond, before making some brief final insights on the changing nature of remote library services and events.

Digital libraries, the Archives Unleashed project and our datathons

Digital libraries have been with us for a long time, and their provision is part of a long continuous tradition. While the mission of the library has evolved, the core aims of collecting, organizing, safeguarding and making information accessible have remained foundational (Rumsey, 2016). Trace foundations of digital libraries can be seen through the "social science data archives," of the 1950s, while subsequent decades transformed the landscape into what we now recognize as the digital library. This was firstly seen in the steady adoption of technology in libraries in the 1970s and 1980s, and then with a surge and "accelerated growth of digital libraries" in the 1990s, influenced by the rapid development of the Web (Altman, 2008, p. 153). Digital libraries represent a relationship between information, collections and technological advances, both in terms of developing infrastructure to collect, manage, search, retrieve and access digital content, as well as understanding the shift in the types of collections that libraries are working with. Much of the emphasis within digital libraries has shifted towards acquiring and delivering information while ensuring its integrity and is increasingly focused on growing scales of information (Xie and Matusiak, 2016). Digital libraries do not exist in a vacuum, of course. They are interconnected with external organizations, institutions, projects, groups and beyond, forming communities that benefit from and that are influenced by the work done in the digital libraries space.

It is very much within this digital library landscape that the Archives Unleashed project exists. The project is guided by an impetus to make large-scale Web archival data accessible to researchers and curators working with born-digital cultural heritage. Founded in 2017 with the support of the Andrew W. Mellon Foundation, the Archives Unleashed project aims to make petabytes of historical internet content accessible to scholars and others interested in researching the recent past. In this context, internet content refers to material saved in the form of Web archives, which can be understood as any form of deliberate and purposeful collecting and preserving of web-based material (Brügger, 2010, p. 25). Web archival data is largely preserved in the ISO-standardized WARC file format, which due to its standardization has facilitated a vibrant ecosystem of Web archival tools and platforms. Yet the WARC file is difficult to use, necessitating projects that can take the data found in them and translate them for use in other platforms and by researchers with varying levels of technical expertise.

The Archives Unleashed project grows out of an understanding that Web archives, which have been collected by the internet Archive and other libraries and memory institutions around the world since roughly 1996, are increasingly central to an understanding of the world around us. Historians, for example, tackling any topic from the mid-1990s onwards would benefit from using Web archives to answer scholarly questions; a historian of the 9/11 attacks, for instance, or the dot.com boom, childhood in the early 2000s, Tamagotchi pocket pets and beyond can all benefit from the millions of pages that have been collected (Milligan, 2019, pp. 19–20). The problem is that while all of this data has been collected – the Internet Archive, for example, as of writing has over 40 PB of unique Web data captured from over 635 *billion* webpages – access has lagged and tools to analyze this data at scale are critical. While the examples above skew historical, Web archives are naturally useful to other disciplines as well: political scientists, sociologists, digital

humanists and any others who seek to explore contemporary culture (See the 40 chapters in Brügger and Milligan, 2018 for an overview).

Archives Unleashed thus lives at the intersection between tool builders, researchers and digital content access providers. Reflecting this, the three co-principal investigators of the Archives Unleashed project come from disparate disciplines: a historian, a computer scientist and a librarian. Providers are those who archive the Web, largely the Internet Archive and memory institutions (notably public, academic and national libraries). In other words, Archives Unleashed is a project that brings together librarians and archivists who want to deploy tools for their own communities of researchers and scholars, digital humanities scholars who want to work directly with data at scale, as well as collecting institutions who want to make sure that their Web archives can be fruitfully used. In this, the project finds itself at a nexus of the digital libraries field.

To achieve the goal of lowering barriers to access, the project has three main pillars: the Archives Unleashed Toolkit, an Apache Spark command line-based tool to work with WARC files directly; the Archives Unleashed Cloud, a cloud-hosted GUI front-end to interact with the Toolkit; and the Archives Unleashed Datathons, which are events that bring together tool builders, researchers and content providers to foster community and use of our infrastructure. Recognizing the importance of in-person gatherings, the Andrew W. Mellon Foundation has supported travel grants (usually C$1,000–$1,300 per attendee) as well as event and team support, to ensure broad participation in an intensive atmosphere of interdisciplinary collaboration. Returning to the impact of COVID-19, as of early March 2020, everything was in place for our final datathon. We were looking forward to showcasing new and emerging tools. Yet the best made plans must be forced to pivot in a time of change and upheaval.

At first glance, digital libraries and the communities around them appear to be well positioned for our sudden "online surge." Yet, this assumption downplays the disruption that this sudden shift means for all of us within the community: dealing with personal fears around contracting or spreading the COVID-19 virus; worry for families; anxieties around the tensions facing marginalized and essential workers; the closure of child care facilities compounding domestic tensions; and dire news around rising death tolls, case counts and lockdown or "shelter-in-place" orders. Life is not "normal, but online." The mandated shift to working from home had implications for staff, programming and users, as staff and faculty quickly shifted to virtual delivery of services. As with all digital libraries, the pandemic forced our project to shift platforms to accommodate how we would deliver our instructional experiences, forge new collaborations and connect with our users and the broader community.

Let's consider the shape of our in-person datathons before reflecting on the adaptation process. The datathon, or hackathon, is a well-developed model within fields as varied as computer science, digital libraries, civic engagement and beyond. Briscoe and Mulligan provide a good definition and overview:

> A hackathon is an event in which computer programmers and others involved in software development collaborate intensively over a short period of time on software projects. These hackathons are encouraging of experimentation and creativity, and can be challenge orientated. From holding large numbers of these events, the hackathon phenomenon has emerged as an effective approach to encouraging innovation with digital technologies in a large range of different spaces (music, open data, fashion, academia, and more). (Briscoe and Mulligan, 2014, p. 2).

Hackathons or datathons have been used in many diverse fields. This includes academia, where they have seen a wide use in the digital humanities and elsewhere. Most events are

in-person, but some are hybrid events, where projects can proceed asynchronously after an introductory session.

We can also understand a datathon within the library field as part of a broader transformation of the profession owing both to new technologies as well as corresponding new approaches to building community relationships. Within the digital libraries field, the rise of Web 2.0 has led to arguments that this new platform can be used to develop "participatory culture." As coined by Jenkins *et al.*, a "participatory culture is also one in which members believe their contributions matter, and feel some degree of social connection with one another." (Jenkins *et al.*, 2009, p. 3). The rise of participatory culture has had implications for libraries as well, transforming a traditional provider/receiver relationship to one that can "create spaces for conversation, critical inquiry, and the expression of alternative viewpoints"; in other words, to provide users with the "tools to participate more fully in the construction of knowledge." (Deodato, 2014, p. 751). Later in this article, we note the impact that the study of participatory culture has had on aspects of our datathon.

All of the above form the context of the Archives Unleashed event. The goal was to explicitly provide a collaborative, hands-on environment to work with Web archive analysis tools and to build community between attendees. Care was taken to assemble an intentionally balanced group that is about one-third librarians/archivists, one-third researchers from the social sciences and humanities and one-third tool builders from computer science and other technical fields (Milligan *et al.*, 2019, p. 268). This is the context behind the event that we realized we needed to move online in a short span of two or three weeks.

Moving online: Should your event go online, and if so, what needs to change?

How do you plan and then implement a shift from a tried-and-true in-person model to an online one, especially on a relatively short timeline, when all of the team's planning, funding framework and experience is predicated around an in-person event? Teamwork, planning and flexibility emerged as key considerations. Several of the specific lessons learned during our pivoted-to-online datathon are broadly applicable to the shifting landscape confronting digital library events more broadly. While the lessons draw from the evidence of our specific example, we make them as generalizable as possible.

For context, our project team was fortunate in two key respects. First, like many but not all in the digital libraries field, we have a flexible working environment. Our team already had some members on 100% work-from-home and others who worked in a hybrid work-from-home/office working arrangement. This meant that all the organizers had good home internet connectivity, a familiar place to work at home, and – crucially – familiarity with systems that enabled a daily mixture of synchronous and asynchronous project work. In other words, all of our systems worked entirely the same at home as they did when in our normal campus settings, thanks to our experience as well as employers who foster flexibility.

Second, the organizers had an established environment and workflow to run all processing in the cloud. Nobody was going to be using the power of their own laptop to crunch data at the actual event. We had learned the hard way in previous events that even if everybody is together in one room, we do not want them to all use their own laptops (Milligan *et al.*, 2019, p. 268). Reasons for this included data transfer speeds, unique computing environments and administrative restrictions. We had thus already done months of work to prepare collections, load them onto cloud infrastructure and allocate resources attendees would access. In essence, we had ideal technical circumstances to support the migration of our event online. All that was needed was to replicate the social interactions. This, of course, is the most difficult component.

Indeed, these social dynamics were what underpinned our initial intention to host a physical, in-person event. We would not have budgeted tens of thousands of dollars on these events if we thought virtual was an ideal substitute, and all of us are aware of the climate impact of travel and still considered the benefits to outweigh the negatives. Remote work misses a lot of the social dynamics that we take for granted when everybody is in one room, which is especially critical when people do not have existing social connections as in the case of our event. Our goal was to forge new connections, and this would be challenging in a virtual environment. When we decided to transition our event online, we had several key considerations to reflect on.

The first key consideration was to reflect on whether our event could be run at all. This required an understanding that this virtual event would not simply be replicating our in-person event "online." For the broader digital library community, there are three broad categories into which events fall: those that can be easily moved online with minor changes, those that require significant adjustments to fit the online environment but that can still reasonably do so, and those unsuitable for the online environment. While obvious, this categorization is a key consideration. It is rarely straightforward to determine which category your event might fall into.

From March onwards, many institutions were forced to confront this question. This necessitated both an assessment of staff bandwidth, as well as re-examining the realities of what would or would not work in a virtual environment. Some events, such as the large gathering of academics (spanning the humanities, social sciences and academic librarians) in Canada, initially decided to move online and then, after a few weeks of reflection, decided to cancel (Congress 2020, 2020). Other organizers – such as the large international gathering of digital humanists – announced the cancellation of their in-person event, and then took time to assess the feasibility of a virtual event (DH 2020, 2020). If one has the benefit of time, a period of reflection is ideal: if not, as in the case of our datathon, a rapid assessment needs to be made.

We resisted the impulse to go online simply because other events were. We wanted to go online only if we could achieve our goals meaningfully. Recall that our overarching goal was to serve the Web archiving community in a way that allowed them to collaborate with individuals from different disciplines and to explore unfamiliar content. If at the end of the day, we could still forge those connections between people and expose them to new content, our overall goal could be met even if the event itself would undergo dramatic change.

The second key consideration was to then determine the shape that our online event would take. Would it be an "online" event in that it leveraged the affordance of technologically-mediated platforms? Or would it be "remote" in that it tried to replicate an in-person event at a distance? As with the move to widespread remote teaching in the post-secondary sector, the most pressing question was whether our event would be synchronous, asynchronous or a hybrid of the two. It became clear that we could not do a fully synchronous event for two days.

Even without a global pandemic reshaping the world, a long conference call is draining (Supiano, 2020). We needed to become truly online, not remote. With daycares and schools shuttered, workplaces thrown into disarray, and beyond, we settled on a mixture of synchronous moments where all attendees would come together, as well as moments where teams and individuals could work as they saw fit. To accommodate these workflows, we decided to use Zoom for synchronous moments and Slack for asynchronous ones. Our schedule is provided in Table 1.

Asynchronous work requires no less of a commitment from organizers and attendees alike. In a physical room, organizers can look around a room and see what works and

what does not work. This does not always come from raised hands or directly asked questions, but rather requires organizers to be able to read body language and sense frustration if a tool was not working or a collaboration was foundering. For the virtual datathon, we required each team to create their own Slack channel, in which we participated, keeping (literal) tabs open. While we encouraged participants to notify us by mentioning our usernames, we also proactively monitored channels and connected with participants.

Even with asynchronous work, time availability and commitment from participants varies dramatically. We thus stressed the importance of flexibility. Kids, dogs, cats, not to mention dire news from the ongoing pandemic, all of these made appearances throughout the day. We wanted to make sure that health and safety, broadly defined, was the number one priority. We did not take attendance, and encouraged teams to be considerate of all of their members.

The third and final consideration was to ensure that we could test and do a dry run of all components before being able to run the datathon online. Some of this is technical: can the chosen communication platform host the volume of participants expected to attend? Have there been updates or changes to the applications you are working with? How can you seamlessly switch presenters, both amongst organizers as well as any participants that may need to give a presentation? We made no assumptions and did two full dry-runs of the event, bringing in students to help ensure our chosen systems could handle the participant surges. Apart from technical testing, we recognized that the "flow" of the event would be largely be determined by setting expectations amongst organizers and attendees: making sure to "mute" or "unmute" all users when appropriate to ensure background noise is kept to a minimum; reminding users to send messages to organizers as we cannot see them suffering in silence in their living rooms or home offices; and, crucially, being as descriptive as possible at all times. At in-person events, people learn by modelling behaviour and seeing what others are doing. They see what questions people are asking and ask similar ones, and in many cases, chat about what they want to get out of the event or problems they are facing over a quick coffee or walk to the refreshment station. By providing direct and descriptive instructions, we instilled a proactive workflow, which in turn helped us to react throughout the event.

These three considerations – considering whether an event should or should not be online, determining the new nature of an event, and then testing with a dry-run – were all instrumental behind our ultimate decision to shift online. But how could we ultimately, during the two days of the event itself, instill a sense of community and teamwork?

Thursday (26 March 2020)	
9:15 a.m.	Participants connect in with Zoom
9:30 a.m.	Opening, welcome, introductory presentations
10:15 a.m.	Team formation
10:30 a.m.	Demo of notebooks + Google Colab setup
10:45 a.m.	Let the hacking begin! (via Slack)
4:30 p.m.	Check-in with participants (Zoom)
Friday (27 March 2020)	
9:30 a.m.	Check-in with participants (Zoom)
2:00 p.m.	Check-in with participants (Zoom)
3:30 p.m.	Final presentations (Zoom)

Table 1.
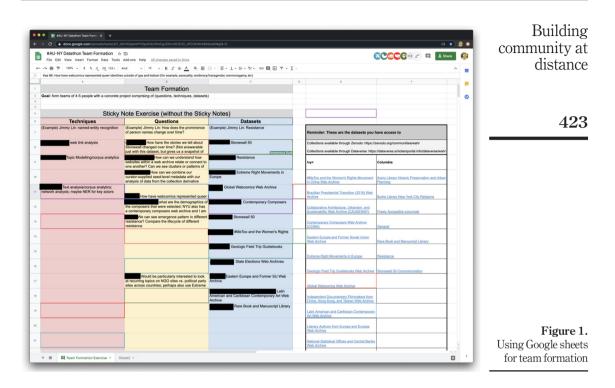Archives Unleashed
online datathon
schedule

The shift online: sticky notes and teamwork in the cloud

One part of the datathon – the team formation activity – is worth exploring in detail. For the organizers, this tends to be both the best part of the in-person event in terms of personal satisfaction, and it is also the aspect that helps us meet the objectives of the event. This would be just as true during the virtual event. To assemble the teams that will form the basis of the datathon for the remainder of the event, we use a "sticky notes exercise" to bring teams together, a method adapted from the aforementioned field of participatory design (Milligan *et al.*, 2019, p. 266; Walsh, 2011, pp. 1062–1063; Walsh *et al.*, 2013, p. 2895). Using three differently coloured sticky notes, participants are asked to write down a short description of a data set they are interested in, a research technique and a research question. For example, an attendee would write that they were interested in "named-entity recognition" on a red post-it note for their technique interest, the question of "how does the prominence of person name changes over time?" on a yellow sticky note for their research question, and "Stonewall 50th Anniversary" on a blue sticky note for their dataset preference. Throughout this group formation process, the organizers help group and organize the notes into emerging themes (i.e. clustering similar data sets, questions and techniques into physical areas of the wall surfaces). This requires active moderation. After several iterations, where we stop the room and ask various clusters of people to talk about what they are interested in, teams begin to organically emerge. This exercise is perhaps the most critical moment of the in-person event. It not only forms the teams, but crucially, begins the process of intensive networking and allows organizers to get a birds-eye view of what questions and tools the community is interested in.

The stakes are high. If teams don't properly form, or if people are drowned out, we risk a wasted opportunity. Even for people who do not end up on the same team, the conversations during this intensive period of mixing and brainstorming seed ideas and expose participants to diverse perspectives in the field. Naturally, this is also the part of the event that relies the most on physicality: people move around the room, using various designated walls or easels to adhere notes to. Organizers read and cluster notes before physically gathering people to talk to each other. These small groupings continually report back to the whole on the status of their conversations, people move between them, and finally after about 30 min of active movement, teams are established. It is great for team formation, but you can imagine, it is the sort of thing that has made us especially hesitant to run during a global pandemic. In other words, what was great to spread ideas was also sadly a perfect way to spread a virus.
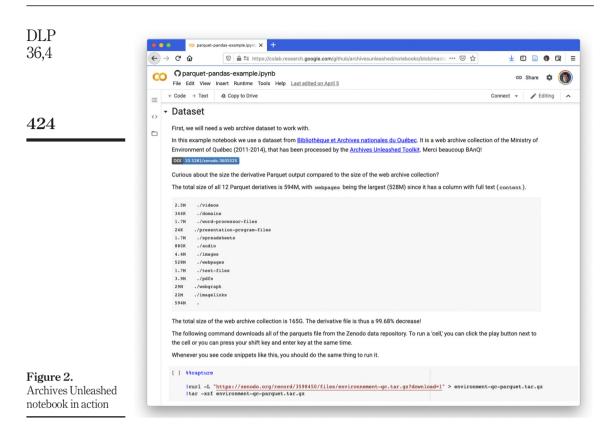
Enter the replacement "virtual sticky notes" process, seen in Figure 1. Recent scholarship has suggested that, in small-group ideation settings at least, a digital version of sticky-note exercise may encourage more interaction than physical ones; and, in any case, did not lead to significant differences in ideation *outcomes* (Jensen *et al.*, 2018b, p. 609; Jensen *et al.*, 2018a, p. 224; Chulvi *et al.*, 2017, p. 7). Other scholarship, however, has noted the difficulties of "replicat[ing] the material affordances of paper that enable such rich flexibility." (Harboe and Huang, 2015, p. 99) Given our specific needs, which was to bring people together into teams – where most of the idea generation would subsequently happen – we were curious whether our experience would bear out these earlier experiences.

Using a Google Sheet, which permitted anonymous contributions, participants used a mixture of writing in cells and using the comments function to thread discussions, which allowed for our team to offer a "running commentary" on the interests. We replicated the colours in the three columns: one for technique, one for research question, and one for data set. Indeed, in the absence of people being able to see the visual clusters, our team paid close attention to emerging teams and offered this commentary which was almost akin to live sportscasting commentary. It leveraged the platform well.

The exercise was different, but it worked on its own merits. This exemplified our desire
to be an "online" datathon, not just a remote one, by taking advantage of various
technologically mediated communication platforms. In any case, seeing all of the
anonymous cursors flying around the page was reminiscent of the physical movement
that we were fondly used to in our in-person events. It was the best of both worlds. The
teams that were ultimately formed all tackled fascinating topics: a team looking at
contemporary musical composers, Latin American and Caribbean Contemporary art Web
archives, queer webcomics and an exploration of narratives in the Stonewall 50th
anniversary collection [2].

With teams formed and collaborations begun, both needed to be sustained to allow for
ongoing collaboration. To facilitate both computing in the Cloud *and* the sharing of code
snippets, we developed a Google Colab-based workflow. Google Colab is a free service that
provides a Web-browser based notebook and runs in the Google Cloud. As we had
previously developed Python-based notebooks to facilitate datathon collaborations and the
use of our collections-as-data, this was ideal for our workflow (Deschamps *et al.*, 2019,
p. 337). Notebooks were pre-generated, and teams would be encouraged to "fill in the
blanks": swapping in collection IDs and research questions of interest and thus exploring
results. An example datathon notebook can be seen in Figure 2. Unfortunately, here is where
we ran into trouble during the first day of our event; despite extensive testing, Google had
quietly launched Colab Pro a few days before. As a result, many of the free services offered
by Colab were degraded. Unfortunately, Colab Pro was available only to American
residents, meaning that we could not pay to quickly upgrade the notebooks during the
event.

Remember the idea of flexibility? Our team was forced to quickly change course. Fortunately, we had generated all the derivative datasets (full text of collections, hyperlink diagram networks, other statistical information), and could quickly make them available to teams independent of the Colab-hosted notebook approach. We hastily gave a webinar on how to use the derivatives, directed teams, and worked late into the evening of Day One to make sure the new process was in good shape. We were glad we had been prepared to hastily adjust on the spot, and in doing so were able to rescue the event.

Lessons learned and broader insights

Our participants brought with them a wide range of backgrounds, skills and experiences when it came to work with Web archives. Despite the move to an online environment and a number of challenges, all of our participants produced final projects that demonstrated creativity, ingenuity and the truest sense of collaboration. Given our successes in migrating this event, what are the broader insights, questions, challenges and impacts that we can explore around the rapid shift online during the global pandemic?

First, it is clear that digital library programming such as our datathon can fill a void and still deliver a diverse array of content. Offering continuity in terms of access is important. Witness the response to the controversial National Emergency Library provided by the Internet Archive, with many authors and rights holders decrying their loss of revenue while the Internet Archive points to shuttered physical libraries and long wait lists for e-books

(Grady, 2020). For many researchers and others stuck at home, these collections are critical. Student and faculty researchers alike are turning to digitized primary sources to continue research agendas, especially pressing for those on the academic job market or in shorter one-year MA programs with uncertain funding extensions (MacEachern and Turkel, 2020). There are silver linings in the pivot online as well. Virtual replacements can also expand accessibility, reduce carbon footprints and force us to reconsider the costly (in all respects) of physically gathering. In the post-pandemic world, we will need to continue to ensure users see the value of these virtual events. In this, we can see opportunities to reach a broader audience and to help further expose collections and resources to larger audiences.

Second, it is clear that transitioning to online-first environments requires different skills and approaches. Losing the ability to scan a room to see varying levels of engagement and needing to pivot rapidly in ever-changing circumstances are examples of skills critical across the broader digital library community as we shift online. Under these challenging circumstances, many commercial library providers have opened resources for free for a limited amount of time (International Coalition of Library Consortia, 2020); giving an opportunity to see how a digital library might look like in a world free from VPNs and proxies. The challenge is to engage many of our new users in novel, technology-mediated forms.

Yet our online experience suggests the broader challenges we will face in cultivating vibrant virtual events, even after the pandemic subsides. Users still crave a "human" connection, a theme which came up during our datathon. Sharing a screen was not as convenient as having somebody working over your shoulders, and there were intangible benefits from chatting about projects over coffee and being located together. None of these are unique to libraries, as our society navigates the shift to widespread remote working and suggest that there will be a role for hybrid events in the future. It is critical to make sure we do not lose people along the way as we move online as well, especially as we continue to recognize inequalities of connectivity and access. Privileging asynchronous over synchronous interaction can help, as long as accessibility is kept at front of mind. Addressing the difference between social dynamics in an online environment versus the physical world, with a goal of allowing everybody to feel included and part of an event, requires active and deliberate work. We need to be aware that individuals will be accessing events from a variety of environments, some from their home offices, other from kitchen tables, and even some from campus and café parking lots (Fugett, 2020).

If we are to deliver meaningful content and services then we need to use this chaotic time to re-examine how we deliver services. COVID has increased the demand for digital library services, and thanks to planning and foresight, the supply is there – especially important as information professionals stress the need for high-quality and reliable information (Ostman, 2020). As it is clear that our reliance on digital libraries and technology is not just a matter of convenience, but of necessity, these lessons learned can help inform others thinking about making this shift in the future.

Conclusions

At our Archives Unleashed online datathon, participants brought with them a wide range of backgrounds, skills and experiences when it came to exploring Web archives. Our experience sheds light on how the shift to a virtual learning and collaborative environment for a datathon can provide valuable insights and lessons for organizations elsewhere in the digital library field. Given the depth and breadth of the effects of this pandemic caused by a novel virus, it is worth underscoring the uncharted nature of the territory we are all in. As we reimagine and reshape our day-to-day business, we need to develop and critically reflect

on new methods of interaction and collaboration. Not everything will be perfect, but with flexibility and an emphasis on empathetically engaging with the circumstances of our users and collaborators, the field can still make strides in community building and accessibility. Indeed, some of the lessons learned during this period will be useful in the post-pandemic period, due to both positive reasons – an effort to combat climate change – as well as negative ones, with travel budgets being slashed as institutions and governments attempt to shrink deficits.

Should you transition your own event online? There are no firm answers, as always, but the experience and broader discussion provided in this article can raise a number of considerations for teams facing similar situations. First, the most important component is an overall assessment of whether the event should be run online at all, by reflecting on the shift of components, technical realities, and most crucially, on whether a modified event will still provide a meaningful way of connecting with attendees. Second, a determination of the nature of an online event, understanding that how users participate and adapt will be key to making sure they feel valued and see value in attending the virtual event. And, finally, if it does run, to make sure to both test all components of the event and to prepare for a descriptive and proactive approach in leading exercises and guiding participants.

Uncharted territory is ahead; while there have been past severe pandemics, none have led to widespread social distancing in an age of technology-mediated communication. We hope that our experiences in moving our Archives Unleashed datathon can help other organizations navigate this shift and contribute to a new and growing literature on how to best quickly adapt and dynamically serve our communities even when we cannot be physically together.

Notes

1. For example, Code4Lib 2020 – which draws many of the same attendees as we did – continued to run and was held 8–11 March 2020 in Pittsburgh.

2. Team presentations and names can be found at: https://archivesunleashed.org/new-york/, shared with permission of participants.

References

Altman, M. (2008), "Digital libraries", in Garson, G.D. and Khosrow-Pour, M. (Eds), *Handbook of Research on Public Information Technology*, Information Science Reference, Hershey, New York, NY, pp. 145-162.

Briscoe, G. and Mulligan, C. (2014), "Digital innovation: the hackathon phenomenon (no. 6)", CreativeworksLondonWorkingPapers, available at: www.creativeworkslondon.org.uk/wp-content/uploads/2013/11/Digital-Innovation-The-Hackathon-Phenomenon1.pdf (accessed 21 May 2020).

Brügger, N. (2010), "Web archiving: between past, present, and future", in Burnett, R., Consalvo, M. and Ess, C. (Eds), *The Handbook of Internet Studies*, Wiley-Blackwell, Malden, pp. 24-42.

Brügger, N. and Milligan, I. (Eds) (2018) *SAGE Handbook of Web History*, Sage, London.

Cho, A. (2020), "We had to act.' how coronavirus fears forced physics society to nix giant meeting", Science, 4 March, available at: www.sciencemag.org/news/2020/03/we-had-act-how-coronavirus-fears-forced-physics-society-nix-giant-meeting (accessed 21 May 2020).

Chulvi, V., Mulet, E., Felip, F. and García-García, C. (2017), "The effect of information and communication technologies on creativity in collaborative design", *Research in Engineering Design*, Vol. 28 No. 1, pp. 7-23.

Congress 2020 (2020), "Novel coronavirus (COVID-19) and congress 2020", available at: www.congress2020.ca/covid-19 (accessed 18 April 2020).

Deodato, J. (2014), "The patron as producer: libraries, web 2.0, and participatory culture", *Journal of Documentation*, Vol. 70 No. 5, pp. 734-758.

Deschamps, R., Ruest, N., Lin, J., Fritz, S. and Milligan, I. (2019), "The archives unleashed notebook: Madlibs for jumpstarting scholarly exploration of web archives", *Proceedings of the 2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, Institute of Electrical and Electronics Engineers, pp. 337-338.

DH 2020 (2020), "Cancellation of the in-Person DH2020 conference", available at: https://dh2020.adho.org/ (accessed 18 April 2020).

Fugett, K. (2020), "I live in rural America cut off from the internet. The pandemic has made me more isolated than ever", Vox, 9 April, available at: www.vox.com/first-person/2020/4/9/21214105/coronavirus-internet-rural-america (accessed 19 April 2020).

Grady, C. (2020), "Why authors are so angry about the internet archive's emergency library", Vox, 2 April, available at: https://www.vox.com/culture/2020/4/2/21201193/emergency-library-internet-archive-controversy-coronavirus-pandemic (accessed 19 April 2020).

Harboe, G. and Huang, E.M. (2015), "Real-world affinity diagramming practices: bridging the paper-digital gap", *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, Association for Computing Machinery, pp. 95-104.

International Coalition of Library Consortia (2020), "ICOLC COVID19 complimentary expanded access specifics", available at: https://docs.google.com/spreadsheets/d/1xiINlF9P00tO-5lGKi3v4S413iujYCm5QJoKUG19a_Y/edit#gid=2027816149 (accessed 19 April 2020).

Jenkins, H., Purushotma, R., Weigel, M., Clinton, K. and Robison, A.J. (2009), *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*, MIT Press, Cambridge, MA.

Jensen, M.M., Rädle, R., Klokmose, C.N. and Bødker, S. (2018a), "Remediating a design tool: implications of digitizing sticky notes", *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18*, Association for Computing Machinery, pp. 1-12.

Jensen, M.M., Thiel, S.-K., Hoggan, E. and Bødker, S. (2018b), "Physical versus digital sticky notes in collaborative ideation", *Computer Supported Cooperative Work (Cscw)*, Vol. 27 Nos 3/6, pp. 609-645.

MacEachern, A. and Turkel, W.J. (2020), "A time for research distancing", Active History, available at: http://activehistory.ca/2020/03/a-time-for-research-distancing/ (accessed 19 April 2020).

Milligan, I. (2019), *History in the Age of Abundance? How the Web is Transforming Historical Research*, McGill-Queen's University Press, Kingston and Montreal.

Milligan, I., Casemajor, N., Fritz, S., Lin, J., Ruest, N., Weber, M.S. and Worby, N. (2019), "Building community and tools for analyzing web archives through datathons", *Proceedings of the 2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, Institute of Electrical and Electronics Engineers, pp. 265-268.

Navarro, J.A., Kohl, K.S., Cetron, M.S. and Markel, H. (2016), "A tale of many cities: a contemporary historical study of the implementation of school closures during the 2009 pA(H1N1) influenza pandemic", *Journal of Health Politics, Policy and Law*, Vol. 41 No. 3, pp. 393-421.

Ostman, S. (2020), "Fighting fake news in the pandemic", *American Libraries Magazine*, available at: https://americanlibrariesmagazine.org/blogs/the-scoop/covid-19-fighting-fake-news-pandemic/ (accessed 19 April 2020).

Ruest, N., Lin, J., Milligan, I. and Fritz, S. (2020), "The archives unleashed project: technology, process, and community to improve scholarly access to web archives", *Proceedings of the 2020 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, Association for Computing Machinery.

Rumsey, A.S. (2016), *When We Are No More: How Digital Memory is Shaping Our Future*, Bloomsbury Press, London.

Saunders-Hastings, P.R. and Krewski, D. (2016), "Reviewing the history of pandemic influenza: understanding patterns of emergence and transmission", *Pathogens*, Vol. 5 No. 4, pp. 66-19.

Supiano, B. (2020), "Why is zoom so exhausting?", Chronicle of Higher Education, 23April, available at: www.chronicle.com/article/Why-Is-Zoom-So-Exhausting-/248619 (accessed 21 May 2020).

Walsh, G. (2011), "Distributed participatory design", *Extended Abstracts on Human Factors in Computing Systems, CHI EA'11*, Association for Computing Machinery, pp. 1061-1064.

Walsh, G., Foss, E., Yip, J. and Druin, A. (2013), "FACIT PD: a framework for analysis and creation of intergenerational techniques for participatory design", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'13, Association for Computing Machinery*, pp. 2893-2902.

Xie, I. and Matusiak, K.K. (2016), "Introduction to digital libraries", in Xie, I. and Matusiak, K.K. (Eds), *Discover Digital Libraries*, Elsevier, Oxford, pp. 1-35.

Corresponding author
Ian Milligan can be contacted at: i2millig@uwaterloo.ca