

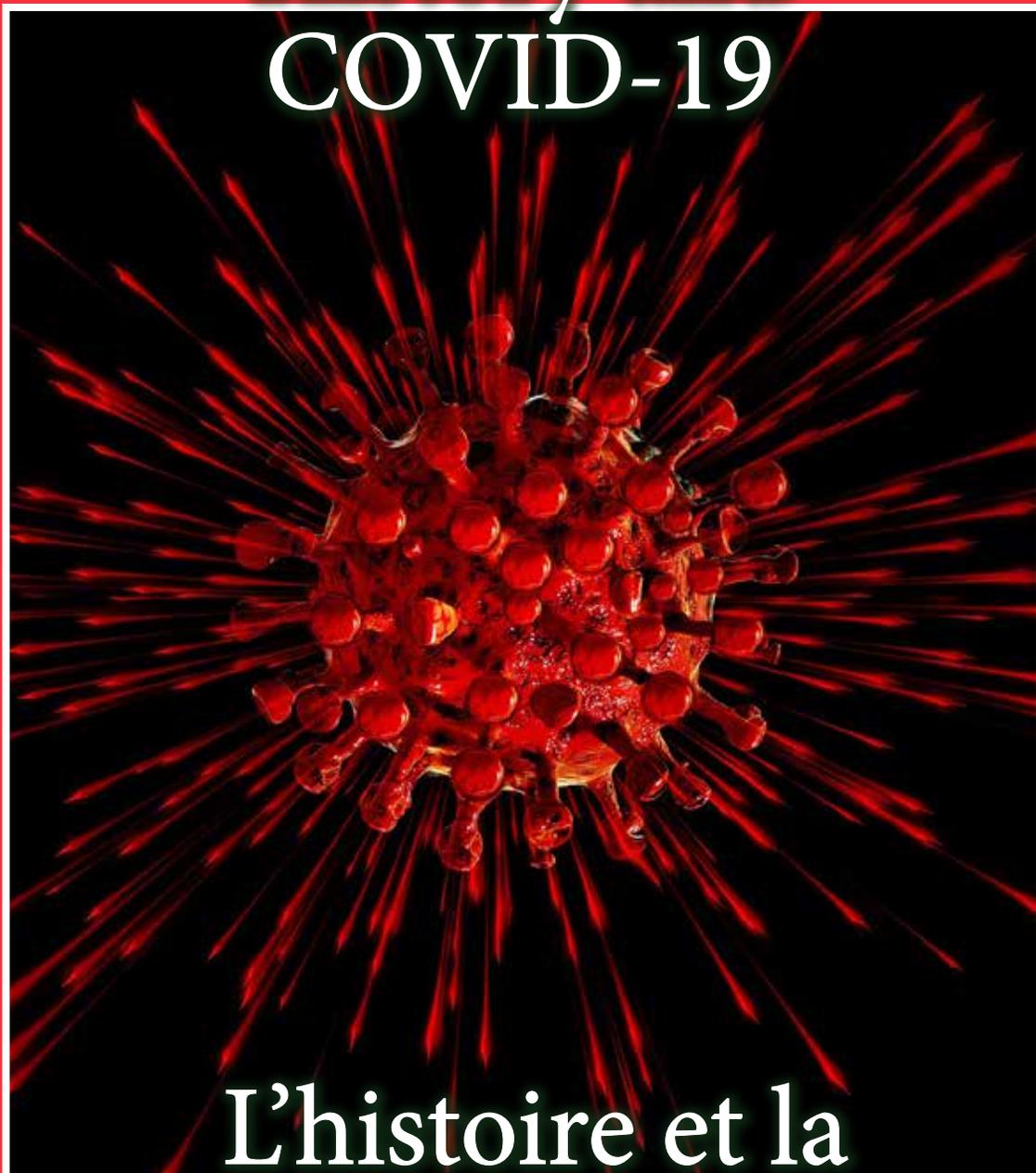
INSECTIONS

Canadian Historical Association  Société Historique du Canada

3.2 2020

where people meet history and history meets people | au carrefour de l'histoire et des collectivités

History and COVID-19



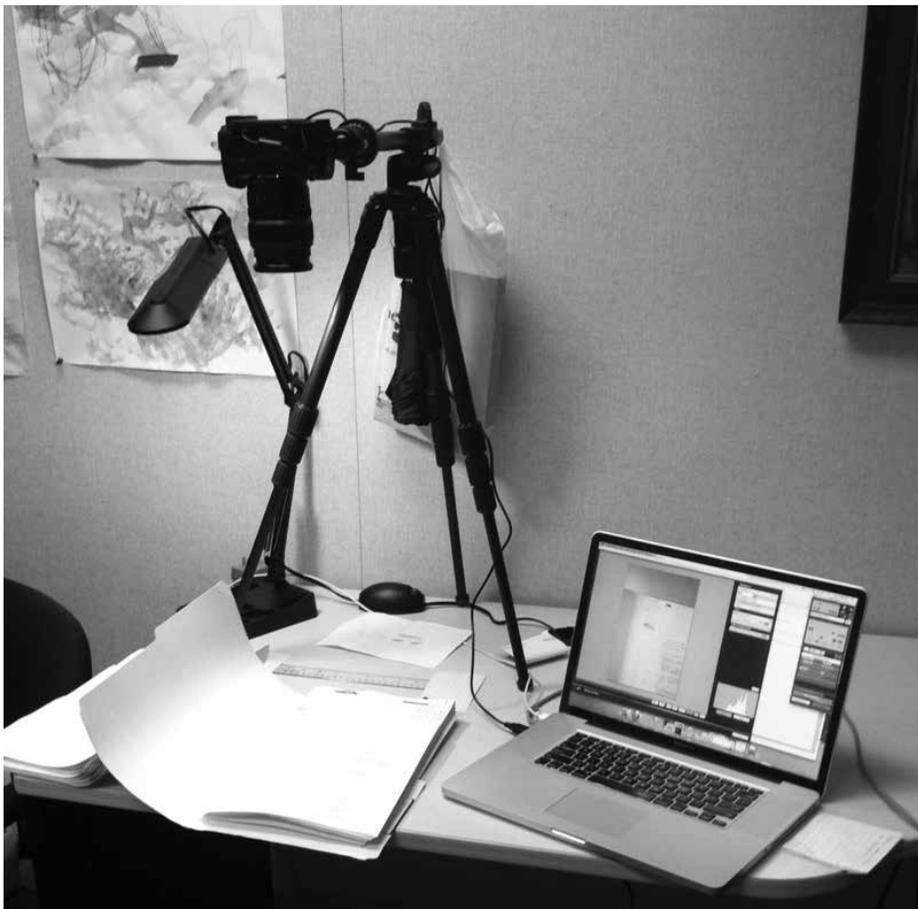
L'histoire et la COVID-19

Technology for Archival Research

In a recent paper for the American Historical Association, Ian Milligan, who studies digital research methodologies, discussed a recent survey of historians in Canada. He found that, while 92 per cent used a digital camera, 90 per cent had no formal training. Over 40 per cent took over 2000 images for their last project. But 70 per cent simply used their own device rather than a specialized piece of equipment.

Despite the pervasiveness of this technology, there are surprisingly few resources to assist scholars in choosing the proper technology for their research. A great deal depends, of course, on the nature of your work. In particular, for those historians who are working with typed text on paper, there is an entire ecosystem that you can employ to scan, process, organize, and file archival documents. The most effective scanning system for archival research includes the following (when used with a laptop, including Mac or PC):

- Tripod with adjustable head and a horizontal arm [Induro CLT204 Classic CF Tripod and BHDO Ballhead]



- Canon or Nikon DSLR Camera with Live-View Shooting [Canon EOS 80D]
- Rechargeable spare batteries [Canon Battery Pack LP-E6N]
- Media Card (64GB to 128GB)
- Software: AbbyFine Reader (corporate), Adobe Acrobat

This camera|tripod system is ideal for archival research. The 4-section tripod with ballhead enables you to attach a camera with the lens facing down towards the work table (see image). In other words, while you are sitting at the table and turning pages in a book or file folder, the camera is hovering about a foot above your head and six inches in front of you. Some people use a remote control to control the shutter. But these devices are unreliable. You might have to rescan hundreds of fuzzy images because the auto focus failed or the document was not properly centred. This is why the Canon|Nikon DSLRs with Live-View function are a useful tool for scanning at archives. These cameras enable you to connect a laptop to the camera while it is suspended above your workstation. You receive a live-view on the laptop of the document as it appears from the lens of the camera. And, more importantly, you can control the shutter and camera settings with a laptop. With this system, you never have to get up, bend over, or manually check the camera. You can adjust the position of your document or zoom the camera without leaving your seat. There's also no fear of having to rescan poorly focussed or off-centre documents because you have a live image view of the document before each scan.

There is one significant drawback to this system: cost. The Induro tripods and ballheads will cost approximately \$500. The camera will be \$1000 or more. And accessories could cost an additional \$300 (camera case, portable hard drive, etc). Once you understand the system, it is possible to experiment with workarounds, such as used equipment or alternative technologies. Still, to my knowledge, only Canon and Nikon DSLRs offer the live-view function, which I have found to be indispensable in scanning large volumes of documents.

There are, to be sure, numerous alternatives to this system. As Milligan discovered with his survey, many people simply use their smartphone. There are also point-and-shoot cameras with WiFi capability that enable

you to control them from your smartphone or tablet. Point-and-shoot cameras have become so advanced in recent years that, much like the DSLRs, they can be attached to tripods and produce high resolution images. In fact, there are several new point-and-shoot cameras that produce higher resolution images than many DSLRs [e.g., Canon Powershot G1 X Mark III]. These cameras are not only less expensive but, through WiFi, enable you to view the document before each scan. However, there are drawbacks. You cannot view the document on a larger screen or easily adjust focus and position. Not all archives provide reliable WiFi (or the WiFi is secured and cannot be used for this purpose). A WiFi connection is also slower and vulnerable to shutting down. In other words, it will take much longer to scan documents because point-and-shoot cameras rely on WiFi and do not have the option to tether the camera to a laptop.

There are a few other tips to consider when scanning documents at the archives. Purchase a camera with rechargeable batteries and at least one spare battery. The spare battery allows you to continue scanning throughout the day while simultaneously charging the other battery. If you are using the DSLR system, you can configure the camera to save each image to the media card in the camera as well as a folder on your laptop after each scan. If the folder on your computer is in a cloud-based folder,

It is not necessary to be a photography expert. At the very least, if you are working with thousands of documents displayed on a large computer monitor from home or in an office, it is worth producing quality images that are easy to read and skim.

such as Google Drive or Dropbox, you will have multiple simultaneous backups of each image, which you can access at any time or place. Finally, a minor detail that can be easily overlooked is carrying a counterweight for the horizontal arm on the tripod. In most cases, you will need to stretch the horizontal arm to its limit in order to completely hover over your documents. If the camera is heavy, such as a DSLR, you'll need a counterweight to avoid tipping over. The Induro tripod has a hook on the arm. A simple solution is to carry a bag to the archives: place a book inside the bag and hang it off the hook at the other end of the tripod.

It is not necessary to be a photography expert. It helps, however, to take into consideration a couple of factors. At the very least, if you are working with thousands of documents displayed on a large computer monitor from home or in an office, it is worth producing quality images that are easy to read and skim. The most important issue when working in archives is the ISO setting (i.e., shutter speed). In a low-light setting, it helps to adjust

your camera settings to a low ISO and slow down the shutter speed because archives will not permit the use of a camera flash. It takes an extra couple of seconds to take the picture, but the camera will be able to capture more ambient light and produce a more readable image. The disadvantage with this technique is that you might produce a slightly grainy image, which could impact readability and text-recognition. And you must be using a tripod – any shift in motion while taking the scan can result in a fuzzy image. In other words, use a high ISO if there is plenty of lighting in the archives and reduce the ISO if the archives has poor lighting. Because Library and Archives Canada and (most) provincial archives provide sufficient lighting, it is easiest to use the auto focus setting on their camera (a few cameras, such as Olympus and Nikon, have document settings for scanning text).

Finally, there is a suite of software tools to facilitate processing and organizing your files if you are scanning typed text. With the DSLR-laptop system, you can specify file names for each document or file folder as you scan. Alternatively, you can manually review your scans at home and use programs such as Automator (Mac) to quickly re-name hundreds of files at the same time. In either case, divide the files into folders on your computer. That will facilitate processing the files using OCR (Optical Character Recognition) software. Abbyy Fine Reader is among the best OCR software tools. Unfortunately, while AFR provides a Mac-based version of the software, it is vastly inferior to its Windows system. The Corporate version of the software is expensive, but it has an incredibly helpful tool called Hot Folders. It will automatically identify every folder and merge all the image files in that folder into a single PDF file with searchable text. For instance, I recently scanned the entire collection of newsletters for a local women's organization in McBride (British Columbia). As I was scanning, I adjusted the camera settings for each month/year of the publication (eg. McBride_march78_001). I then transferred the image files to a series of folders divided by month and year. AFR automatically generated a PDF file for every single issue of the newsletter. Using Spotlight in Mac, I can keyword-search thousands of documents to identify individuals or key themes and issues. Moreover, the software will remove blurriness, sharpen edges, and arrange angled scans to improve the readability of each document.

Digital technologies have transformed archival research. It was not long ago when most historians were using pencil and paper for archival research. I routinely scan 5000 to 10,000 images during a two to three-week research trip. In fact, the system is so efficient that it is sometimes more productive to scan everything and read at home. It might take 20 to 30 seconds to skim a document to determine if it is relevant to your research whereas a camera scan takes a few seconds. Moreover, this technology has enabled the creation of several online digital archives, such as statefunding.ca and HistoryOfRights.ca. The use of these technologies could, over time, provide alternative sources of research than traditional archives.

Dominique Clément, University of Alberta