

A Large-Scale Digitization Plan for Public Domain Works in the Killam Library, Dalhousie University.

Geoff Brown,
Digital Scholarship Librarian,
Dalhousie University Libraries,
December 2020

1. [Introduction](#)
2. [Establishing Digitization Priorities](#)
3. [Finding Potential Works For Digitization](#)
4. [Methods](#)
5. [Results](#)
6. [Discussion](#)
7. [Conclusion](#)
8. [References](#)
9. [Appendix A – JSON script used for each Library of Congress subclass in the Killam Library collection](#)
10. [Appendix B – Detailed public domain assessments of sampled items](#)

Introduction

Student engagement with print library resources has shifted dramatically over the last 20 years to a point where post-secondary curriculums rely almost exclusively on digital tools and online content. Some disciplines continue to use print collections, but there is little doubt that most students and researchers find greater value in online access to content over the inconvenience of travelling to a physical place to consult print material. Yet our print collections are immensely valuable, both as an active tool for teaching and learning, and as a scholarly record.

While it is part of the mission of research libraries to maintain these valuable print collections, we are most definitely limiting the usefulness of these collections by keeping them locked in analog form. Especially since an unknown amount of material resides in the public domain and could be digitized and made available online. The purpose the work undertaken during my 2020 study leave at Dalhousie University is to inform future digitization decisions related to the Killam Library book collection. How much of the collection is in the public domain and does it make sense digitize this content?

In this report I will explore the possibilities for a large-scale digitization initiative at Dalhousie. I am interested in studying the merits of a program that is mindful of risk and makes reasonable efforts to identify public domain works in Canada. Canada is in an enviable position with respect to content that resides in the public domain. Along with New Zealand, we are the last remaining western English-speaking countries that (in very general terms) use life of the author +50 years to determine public domain status. Most countries have a copyright term that covers life of the author +70 years. The United States has a patchwork of copyright terms that match wild swings in copyright protection from none in some decades, to 95 years after publication for some periods, and death of the author +70 years in others. When the recent Canada-US-Mexico Trade Agreement came into effect in July 2020, it meant an extension to the copyright terms in Canada from death of the author +50 years to death +70. One

possible outcome of this deal is a 20-year freeze on works entering the public domain in Canada. Although the Government of Canada is looking at other options, such as a registration requirement for meeting the copyright obligations in the trade deal, that would be out of step with what's happening elsewhere. All this being said, extensions agreed to in the Canada-US-Mexico Trade Agreement will not be retroactive. So works that are currently in the public domain as of January 1, 2020 will remain so once the new terms are established (Geist, 2019). This puts Canadian libraries in a strong position globally to initiate public domain digitization projects now and in the immediate future. Library and Archives Canada released a National Heritage Digitization Strategy in 2016 <https://www.bac-lac.gc.ca/eng/about-us/Pages/national-heritage-digitization-strategy.aspx> that encourages Canadian memory institutions to initiate such projects. With public access to our physical collections being limited in 2020 and the foreseeable future due to the global Covid pandemic, the utility of digitizing our public domain materials becomes even more apparent.

Establishing Digitization Priorities

When thinking about how to prioritize our digitization work at Dalhousie I looked at ongoing work in New Zealand and other projects in Canada such as the National Heritage Digitization Strategy (NHDS) and Canadiana. In New Zealand, Clark and Chawner (2014) discovered that many of the larger U.S. based book repositories (including Google Books, Hathi Trust, Internet Archive and Project Gutenberg) place excessive access restrictions on readers outside the United States. As part of their work, Clark and Chawner found that only 3 books out of a random sample of 100 public domain books published in New Zealand were available to New Zealand readers without restrictions. This result really calls into question the practice of relying on repositories in other countries for reliable access to public domain content, and is surprising considering the Clark and Chawner study only included books published prior to 1890. This is a very conservative date for a country that has public domain rules that are similar to Canada's. We would expect a far greater number of works published prior to 1890 to be available without restriction to New Zealand readers. For years, Canadian libraries like libraries in many other countries, have depended on large U.S. based book repositories for access to digital copies of materials in our print collections. The assumption that many libraries have made is that we should not prioritize the digitization of items that are already available in these repositories. The work by Clark and Chawner suggests that this assumption might be ill advised. After identifying items from our own collection that are in the public domain in Canada, I checked several for availability in Hathi Trust (see [Discussion](#) section below) and found that Canadian readers are denied access. Just like New Zealanders, Canadians are lumped in with all other international users of Hathi Trust and do not always get access to works that are in the public domain here. Put simply, Canadian libraries should not set digitization priorities based on the holdings and perceived access to content in U.S. book repositories.

Fortunately, there is an excellent digital repository of older public domain works available to Canadians that, due to its Canadian focus, was not considered in the New Zealand study. Canadiana started as a microfilming project in 1978 and has since grown into an online digital repository that provides Canadians with access to over 100,000 items. The monograph collection consists of works published prior to 1921 and all materials are available to Canadians without restriction. In 2018, Canadiana.org merged with CRKN and further strengthened its ability to maintain and provide access to these collections for all Canadians. When making a final decision to digitize an item from the Dalhousie University Library collections, availability of the item in Canadiana is a valid reason not to digitize. Unlike

the many repositories based in other countries, Canadiana is a dependable access point for Canadians that is based on Canadian copyright law.

Another Canadian initiative is the National Heritage Digitization Strategy (NHDS) <https://nhds.ca/>. This strategy launched in 2016 to help focus the digitization efforts of Canadian memory institutions. The work of the NHDS has been closely aligned with Library and Archives Canada and more recently CRKN/Canadiana but as the name suggests, it is mostly concerned with overall strategy and priority setting rather than the development of a specific digital repository. The NHDS appears to recommend a risk-based approach and set objectives for digitizing 90% of works published before 1917 and 50% of all monographs published before 1940 (National Digital Heritage Strategy website <https://nhds.ca/about/> accessed September 11, 2020). These dates are intended as a best guess that balances the low risk of copyright infringement prior to these dates with the widespread benefit to society of having public access to older content. As such, the NHDS approach favours expediency when it comes to making decisions about whether or not to digitize older works. Whether this approach is advisable for Canadian libraries is another question. It seems that we should take advantage of the relevant bibliographic data that describes our collections to support a more precise identification of public domain works. That being said, the objectives of the NHDS are a strong indication that it's ok to take measured risk when it comes to digitizing older works for widespread public access. The work of the NHDS highlights the fact that there is no risk-free approach to digitizing public domain works.

Like other national digitization strategies in New Zealand (<https://digitalnz.org/>) and Sweden (<http://www.digisam.se/>), the NHDS is concerned exclusively with Canadian content. These national programs are broad in terms of format (accepting objects from libraries, museums and archives) but narrow in terms of limiting themselves to a national scope. On both counts, the formats and the content support the mission of these nationally based institutions. As a university library, our collections serve an even broader academic mission and consist of content published anywhere in the world. Just like the national programs though, we must shape our focus to align with our mission. As a result, it would not be appropriate to limit our digitization efforts to just Canadian works. That is not to say that we should not be mindful of Canadian content in our collection, but that we should not concern ourselves exclusively with Canadian content.

Based on the above, I suggest digitizing potential public domain items in the Dalhousie University Library collections and utilizing both the wealth of bibliographic information in our catalogue and the public domain rules in place in Canada as of January 1, 2020. Furthermore, I suggest that we accept a certain level of risk. No public domain assessment program will be 100% accurate. This risk is already acknowledged in the "Protocol for takedown requests regarding materials in the libraries' online collections" https://cdn.dal.ca/content/dam/dalhousie/pdf/library/services/copyright-office/Takedown%20request_v1.pdf adopted by the Dalhousie University Libraries in 2015. Our bibliographic data is immensely useful in this endeavor but there is no doubt that incomplete or inaccurate bibliographic data can result in items being incorrectly identified as public domain. This risk is inherent in many other mass digitization initiatives, particularly at the national level as noted above. Accepting some risk on the copyright side though, doesn't mean that we should set aside all risk. Final decisions to digitize would be subject to some practical considerations that can only be taken into account with an item materially in hand.

One such consideration is the physical condition of the item. No digitization should proceed that would destroy the library's last copy. Other considerations would be the physical format of an item that might make digitization impractical. Physical conditions aside, the potential benefits of a digitization program that gives the reading public widespread access to the potentially significant public domain collection that resides in our research library is something that is very much in line with the public domain provisions in Canadian copyright law. As Craig (2010) points out, Canadian courts are increasingly concerned about balancing the rights of content creators with the right of the public to have access to a vibrant public domain that supports creativity and innovation. The public domain is every bit as important as the private domain of in-copyright works and by not taking advantage of the freedoms afforded to public domain works we do a disservice to our research collection.

Finding Public Domain Works for Digitization

The source data for my work is a series of data files extracted from the Novanet catalogue in 2017 by the Dalhousie Libraries Resources Unit using the Greenglass collection analysis tool. The data captures basic bibliographic information about the Killam book collection. I believe the original intent was to determine how much of our print collection is included in the Hathi Trust public domain collection in the United States. My work uses the same Greenglass generated data to answer a slightly different question: What items in the Killam Library book collection are in the public domain in Canada? With few exceptions, many of these works could be digitized and made available to the public.

To answer the above question, I turned to the already available Greenglass dataset. The important pieces of information in the Greenglass data are author main entries (including birth and death dates when available), title, and dates of publication. Novanet regularly updates our bibliographic data with information on author death dates from the Library of Congress subject authorities but since a death date of 50 or more years ago determines public domain status, 2017 data is more than sufficient. Any deaths recorded in the last 3 years will not enter the public domain for many years. A major shortcoming of the Greenglass data is that it doesn't always identify additional authors. Information about additional authors is an important piece of information in determining the public domain status of a work. In Canada, books enter the public domain 50 years after the death of the last living author. So, while the 2017 Greenglass data is a great starting point for determining what might be in the public domain, there will be times when a trained staff person would need to review an item to make a final determination about public domain status. The methods outlined below will identify all possible public domain items with the caveat that additional work could be required to make a final decision about whether or not a work is in the public domain. The additional work is not difficult and could be undertaken by library support staff with minimal training.

Methods

Identifying Possible Public Domain Items using OpenRefine

In my public domain analysis of the Killam Library book collection, I use OpenRefine software freely available at <https://openrefine.org/> to develop a repeatable method for identifying possible public domain works. I then extract a sample set of results to evaluate the relative success of this method. Again, it must be acknowledged that there is no perfect method for identifying public domain works that

is 100% accurate and without some risk of copyright infringement. The method outlined here is intended to strike a balance between accuracy, expediency and risk.

Smith (2012) provides an interesting perspective on risk as it relates to large digitization projects of heterogeneous library materials. Smith notes that many libraries strive for zero risk of copyright infringement when it comes to digitization projects and that this is entirely unnecessary. Instead, Smith suggests that libraries simply attempt to reduce the number of risky items and the chances of potential litigation. Smith goes on to recommend that libraries only ask large or prominent rights holders for permission to digitize and to have a takedown policy in place for those who object to the online availability of a particular work on the grounds that it infringes their copyright. As mentioned above, the Dalhousie Libraries already have an established takedown policy in place. So, we have already taken a significant step to mitigate risk if we pursue a project such as this.

The method described here is intended to further reduce risk by applying the basic criteria outlined in the Canadian Copyright Act <<https://laws-lois.justice.gc.ca/eng/acts/C-42/Index.html>> to determine public domain status. The rules for published literary works in Canada are that an item enters the public domain 50 years after the death of the author or 50 years after the date of publication in the case of anonymous or pseudonymous works. Expediency is achieved by using readily available software to programmatically apply the public domain rules from the Canadian Copyright Act to the approximately 640,000 items in the Killam Library book collection in a timely manner. Risk is measured by manually reviewing a random sample of items flagged by the automated method.

To start, I used MS Excel to convert the original GreenGlass data files into comma separated .csv files with UTF-8 character encoding. I then imported the files into OpenRefine and created projects based on the Library of Congress class and subclass groupings captured by the original GreenGlass data.

OpenRefine was used because it is an excellent tool for working with messy data. Local cataloguing practices can result in data that is often incomplete and inconsistent. OpenRefine allows for easy splitting of multi-valued cells (e.g. author fields with names and dates bundled together) and quick faceting of similar items in massive datasets. Both of these features are relied upon heavily in my OpenRefine workflow. The workflow outlined below represents an individual OpenRefine project. I created separate projects in OpenRefine for each Excel file that was exported from the GreenGlass collections analysis tool by the Killam Library Resources Unit in 2017. After manually creating the workflow (as described below), I then extracted the JSON (JavaScript Object Notation) so that it could be applied to each Library of Congress class and subclass in sequence. The JSON is included in **Appendix A**.

Step 1, identify works with known authors – I identified items with a personal author and created a text facet for those that died 50 years or more ago. This workflow eliminates items with no known author death date. The result is a list of items from the catalogue where the main author death date results in the item possibly being in the public domain. Public domain status is only a possibility since data about additional authors is not available in the GreenGlass dataset.

- All items with a personal author AND a death date have a dash “-”. Split cells in the Author column into several columns by the dash separator using a Regular Expression. To do this “Edit Column > Split into several columns”. Under the heading “After Splitting”, leave the box checked to have OpenRefine guess at the data type. This ensures that the death dates are stored as numbers rather than text. Keep the original field intact. Indicate that the data should be split

into 2 columns. In the separator text box enter `\d\d\d\d-` and then check the Regular Expression box. The result is 2 columns. The second column has the death dates. Change the name to “Known Death Dates”.

- Select the “Known Death Date” column then “Edit column > Add column” based on this column. Enter the column name as “Person Author Public Domain Status” and enter the GREL `if(value < 1970, "Possible Public Domain", "Not Public Domain")`.
- Create a text facet on this column to see the Possible Public Domain items based on the available death dates.

Step 2, identify items with anonymous authors – Next, I identified works from the GleenGlass dataset with no personally named author and created a text facet for those that have been published 50 years or more ago. This workflow gives us a list of items that have no individual person who is affiliated with the work as an author in the GreenGlass dataset. Items with a blank or null author field can indicate works that have either a corporate body author or no identified author. From a public domain perspective, these items can be treated the same as long as no person is indicated anywhere on the item as an author. The lack of a named person in the GreenGlass dataset is a good indicator but it holds no guarantee that no individual author is affiliated with the work.

- To group the works with no identified author with the corporate authored works go to Author>Edit cells>Transform and enter the GREL expression `if(value==null, "Title: Main Entry", value)`. The colon is important as the punctuation that will facilitate grouping the title main entry works with the corporate authored works.
- Items with corporate or meeting names tend to have a colon “:” somewhere in the Author field as a qualifier element. Select the Author column and then Edit column > Add column based on this column. Name the new column Author Type and enter the GREL expression `if(value.contains(":"), "Anonymous Author", "Identified Author")`. This creates a new column called “Author Type” that gives everything with a colon in the “Author” field a value of “Anonymous Author” and everything that does not have a colon, a value of “Identified Author”
- On the Author Type column select Facet > Text facet to see only the items with “Anonymous Author”. Select this facet.
- Now go to the Publication year. Edit cells > Common transforms > To number
- Next Edit column>Add column based on this column. Leave the GREL expression as `value` to create a copy of the Publication Year column called “Anonymous Author Public Domain Status”.
- Now go to the Anonymous Author Public Domain Status column and Edit cells > Transform. Enter the GREL expression `if(value < 1970, "Possible Public Domain", "Not Public Domain")`.
- In the Anonymous Author Public Domain Status column choose Facet > Text Facet to create a facet with the 2 values. Choose the “Possible Public Domain” facet to see all the items that have no identified personal author in the GreenGlass dataset that have been published fifty or more years ago.

Once this method was converted to a JSON script (see **Appendix A**) using the Extract feature in OpenRefine, it was very efficient to apply the script to all call number ranges as originally generated by the Killam Library Resources Unit. The OpenRefine projects along with this report are all saved in DalSpace <http://hdl.handle.net/10222/80297> for long term public access. The project files are compressed into a single file called OpenRefineProjects that contains 65 individual projects representing

Library of Congress subclasses in the Killam Library collection. OpenRefine must be installed on any user's PC who wishes to work with the projects. The only action required from Steps 1 and 2 above to see the list of possible public domain items is the final faceting action. This is represented by the final bullet point in Steps 1 and 2 described above. Spreadsheets can easily be exported to support any staff who might be engaged to work on a large-scale digitization project that results from this work.

The method above creates a large set of possible public domain items. The most significant shortcoming of the original GleenGlass dataset was the lack of author added entry data from the Novanet database. Public domain status relies on information about all authors of a work but the method described above does a reasonable job of eliminating hundreds of thousands of items that are definitely **not in the public domain**. The Novanet Office could provide a more complete bibliographic dataset that includes additional authors (known in bibliographic terms as author added entries) but author added entries rarely contain author death dates. So, even with a more complete dataset, the resulting set of possible public domain items would likely be very similar. Based on this knowledge, I decided to sample from the items identified, to determine how successful the above method was in identifying public domain items from the Killam Library book collection.

Sampling Possible Public Domain Items

To get a sense of how well the above method works to identify items in the public domain, I randomly sampled items from the result sets of the different LC classes and subclasses to get an anecdotal sense of how accurate my method was overall and whether or not there are any issues that reflect disciplinary differences in our collection or in book publishing in general that might reveal any shortcomings in my JSON script. I did this using the RAND function in MS Excel. Each OpenRefine result set was exported back into MS Excel where I added a "Random" column to each dataset. I then used the =rand() function to assign a random number to each row, and then sorted on that number from lowest to highest. For the larger sets of possible public domain items ranging from 14,803 to 17,804 items I randomly sampled 20 items for detailed follow-up. For the smaller sets of 200 to 3,238 items, I randomly sampled 10 items for detailed follow-up. In total, I checked 80 individual random items to determine the success of the JSON script. The sample sets simply gave me reasonable numbers of items to investigate more closely. Consequently, the results can only be considered anecdotal. Ultimately, the purpose of this part of the work was simply to determine how successful the JSON script is in identifying possible public domain items in the collection, based on the known lack of comprehensive author information noted earlier. In some cases, I sampled from entire classes while in other cases, I sampled from specific subclasses that hold Canadian content and specific subjects that I wanted to investigate further. The following are the classes/subclasses that I chose to gauge the success of my original script.

- D – World History
- F – History of the Americas
- H – Social Sciences
- PR – English Literature
- QH – Biology

The method for evaluating whether or not sampled public domain items are really in the public domain involved a number of steps. First was an examination of the entire bibliographic record for each sampled item. Particular attention was paid to author added entries. Missing death dates for any author affiliated

with a work were first checked in the Library of Congress Authority files. Any authors not listed there were checked in Wikipedia, which uses an authoritative and massive open knowledgebase of human and machine-readable data. Any author that did not turn up in either location was subject to a final webscale search to locate any relevant birth/death information on any website. Anonymous works were subject to an additional search in Google Books to check cover pages and tables of contents to see if any personal names appeared on the physical items. Normally, I would have checked the physical items in the collection but this was not possible due to covid-19 restrictions that eliminated access to the Killam book collection. The detailed public domain assessments for each item along with a brief narrative explanation of how I reached my conclusion for each item are included in **Appendix B**. The success of the OpenRefine workflow is measured against this final assessment. In addition to a public domain assessment, I provide a final recommendation to digitize or not that is based on both the public domain assessment **and** the presence of the item in the Canadiana repository. This last piece of information is just included for discussion purposes as is the information about whether an item is included in another repository outside of Canada such as Hathi Trust, Project Gutenberg or Internet Archive.

Results

Results from the Open Refine workflow

Table 1 provides the results from the JSON script applied to the 639,961 books in the Killam Library collection using OpenRefine. The script flagged 121,553 items representing 19% of the book collection as being possibly in the public domain based on the criteria of an author main entry death date of 50 or more years ago or, in the case of anonymous works, a publication date of 50 or more years ago. The results when viewed according to LC Class vary significantly and reveal some disciplinary trends. General Works (A), and Music and Books on Music (M) show very high potential public domain levels of 51% and 48% respectively. Conversely, Medicine (R) and Agriculture (S) show the lowest percentages of possible public domain content at 5% and 4% respectively. In terms of total size, Language and Literature (P), the largest class in the book collection is comprised of 27% possible public domain works, followed in total size by the Social Sciences with just 7% possible public domain works and Science with 11%. So, there is some significant disciplinary variation. The results from each class are summarized below.

Table 1 – Potential public domain content in each LC Class in the Killam Library book collection

LC Class	Personal Public Domain	Anonymous Public Domain	Total Possible Public Domain	Total Items	Percentage Public Domain
A -- GENERAL WORKS	751	329	1080	2108	51%
B -- PHILOSOPHY. PSYCHOLOGY. RELIGION	9068	1189	10257	44182	23%
C -- AUXILIARY SCIENCES OF HISTORY	625	205	830	3446	24%
D -- WORLD HISTORY	12889	1914	14803	59771	25%
E -- HISTORY OF THE AMERICAS	2839	227	3066	15055	20%
F -- HISTORY OF THE AMERICAS	2697	541	3238	16191	20%
G -- GEOGRAPHY. ANTHROPOLOGY. RECREATION	1218	453	1671	22586	7%
H -- SOCIAL SCIENCES	4602	1655	6257	95830	7%

J -- POLITICAL SCIENCE	1746	521	2267	20812	11%
K -- LAW	301	91	392	4933	8%
L -- EDUCATION	677	227	904	10620	9%
M -- MUSIC AND BOOKS ON MUSIC	13885	1443	15328	32175	48%
P -- LANGUAGE AND LITERATURE	46868	3076	49944	188132	27%
Q -- SCIENCE	4733	3002	7735	72722	11%
R -- MEDICINE	253	147	400	8229	5%
S -- AGRICULTURE	234	138	372	8985	4%
T -- TECHNOLOGY	287	449	736	11527	6%
U -- MILITARY SCIENCE	211	95	306	4076	8%
V -- NAVAL SCIENCE	81	17	98	1161	8%
Z -- BIBLIOGRAPHY. LIBRARY SCIENCE. INFORMATION RESOURCES	1401	468	1869	17420	11%
Total	105366	16187	121553	639961	19%

Results of sampling from select call number ranges

As stated previously, I measured the successful identification of public domain items in the set of 121,553 possible public domain works by conducting a detailed public domain assessment of a random sample of these items. Rather than sample from the entire result set, I decided to sample a combination of entire classes and subclasses to get a sense of how the script worked at a disciplinary level, a subject level and with Canadian content. World History D was chosen as a representative subject in the social sciences, History of the Americas was chosen because it includes Canadian History, Social Sciences was chosen as a general discipline, English Literature was chosen as an important humanities subject with Canadian content and Biology was chosen as a representative subject within the sciences. Table 2 below shows the subject classes and subclasses where I sampled from a list of items that my script identified as being possibly in the public domain, along with the relative percentage these possible public domain items comprised of the overall class/subclass. In the cases where I chose an entire class, this information duplicates the information in Table 1. Where I have drilled down to the subclass (e.g. PR and QH), I have updated the values to reflect that fact. In each case, I randomly sampled items from the Total Possible Public Domain column in Table 2. Each item in the respective samples was investigated individually to see how likely it really is to be in the public domain. The detailed assessments are included in Appendix B and the overall results are summarized in Table 3.

Table 2 – LC classes and subclasses chosen for a more detailed assessment of public domain works.

LC Class / Subclass	Personal Public Domain	Anonymous Public Domain	Total Possible Public Domain	Total Items	Percentage Public Domain
D – World History	12889	1914	14803	59771	25%
F – History of the Americas	2697	541	3238	16191	20%
H -- Social Sciences	4602	1655	6257	95830	7%
PR -- English Literature	16922	885	17804	52964	34%
QH -- Natural History - Biology	588	321	909	10017	9%

Overall results for the detailed assessments were quite promising. In total, 67 of the 80 items were found to be in the public domain. This represents an 84% success rate for the script. Based on the assessments, it appears that the script was quite successful in identifying public domain works in English Literature (100%) and World History (95%). It appeared to be least successful in the Social Sciences, where 65% of items flagged by the script appeared likely to be in the public domain. Not surprisingly, classes that have a very high percentage of the total collection in the public domain as flagged by the original script in Open Refine are showing the highest success rates when the items themselves are investigated to determine public domain status.

Table 3 – Success rate of detailed public domain assessments by LC class and subclass.

LC Class / Subclass	Public Domain Items	Sample Size	Success Rate
D -- WORLD HISTORY	19	20	95%
F -- HISTORY OF THE AMERICAS	8	10	80%
H -- Social Sciences	13	20	65%
PR -- English Literature	20	20	100%
QH -- Natural History - Biology	7	10	70%
Total	67	80	83%

Both with the script and the detailed assessment of sampled works, I used basic criteria to determine public domain status. A work moves into the public domain 50 years after the death of the last surviving author, at the end of the calendar year. In the case of anonymous works, they enter the public domain 50 years after publication or 75 years after the date of creation, whichever is sooner (Government of Canada, Copyright Act: Version of section 6.1 from 2002-12-31 to 2020-06-30). Since library catalogues typically only describe published items, the GreenGlass bibliographic data does not provide creation dates, my script only uses the date of publication for anonymous works. As such, the method and result described above err on the side of less risk and may not include anonymous works that are already in the public domain.

Another group of items that is not specifically addressed in the above results are posthumous works. It would be fairly difficult to identify posthumous works from the available bibliographic data used in this analysis. Posthumous works have been subject to some different copyright terms in Canadian copyright law over the years. Many posthumous works entered the public domain in 2004 if the last surviving author died anytime from 1948 right up until 1998 (University of British Columbia Scholarly Communications and Copyright Office, 2020). In a limited number of cases we could have items enter the public domain in as little as 32 years after the death of the author that are not identified in the results described above. By necessity, my method again errs on the side of less risk of copyright infringement rather than more risk.

It should also be noted that the Government of Canada has begun making changes to the Copyright Act <<https://laws-lois.justice.gc.ca/eng/acts/C-42/index.html>> that are now reflected in the online version of the Act as of October 29, 2020. As of that date, the Term of Copyright changes have only been made to subsections 6.1 "Anonymous and pseudonymous works" and 6.2 "Anonymous and pseudonymous works of joint authorship" to reflect the longer terms of 75 years after publication and 100 years after creation, that are effective as of June 30, 2020. These terms are not reflected in my script as they only

apply to materials that would enter the public domain after June 30, 2020. Pseudonymous works are difficult to ascertain using bibliographic data and the incidence of truly pseudonymous works in the collection as a whole is likely quite low as most authors who write under a pseudonym eventually become known and then the public domain rules for known authorship come into effect.

Discussion

The public domain is difficult to pin down. It changes over time in scale and scope, and varies by country. As noted previously, a Canadian perspective is both warranted and necessary. While Canadians have access to a limited number of works that are in the public domain in U.S. based book repositories such as Hathi Trust, Project Gutenberg and the Internet Archive, the level of access is not sufficient and Canadian libraries can offer more than is possible in the United States. The public domain has historically benefitted from a broader range of dates in our country. The laws of a country often reflect the values of its citizens. In Canada, we have a history of striking an equitable balance between public good and individual gain. As such, the Canadian Copyright Act up until recently placed the tipping point for financial protection of published materials at life of the author plus 50 years for the works of individuals, and date of publication plus 50 for the works of companies or other organizations. In the United States and many other European countries, that tipping point has been significantly extended to favour individual or corporate financial interests. Unfortunately for Canadians, recent trade pressure from the United States has resulted in changes to Canadian Copyright Law that will limit future growth of our public domain at the expense of the public good. The argument in support of these extensions is essentially to protect the financial interests of corporate and estate copyright owners mostly outside of Canada. Thankfully, unlike the situation in the European Community in 1993 where copyright terms were revived in member states (Geller, 2000), Canada has not made copyright extensions that revive copyright. Bill C-4 "An Act to implement the Agreement between Canada, the United States of America and the United Mexican States" (Section 34) outlines the following transitional provision related to copyright which received Royal Assent on March 13, 2020:

"No revival of copyright

34 Sections 6.1, 6.2 and 11.1, paragraphs 23(1)(a) and (b) and subsection 23(1.1) of the Copyright Act, as enacted by sections 24, 26 and 29, respectively, do not have the effect of reviving the copyright or a right to remuneration in any work, performer's performance fixed in a sound recording or sound recording in which the copyright or the right to remuneration had expired on the coming into force of those provisions of that Act."

The effect of this provision cannot be overstated. Works that were in the public domain as of January 1, 2020 will remain in the public domain. This small consolation, provides the Dalhousie University Libraries with a rare opportunity to digitize much more recent public domain works from our collection than is possible in most other parts of the world.

As mentioned previously, there are many parallels between Canada and New Zealand. Clark and Chawner (2014) discovered that many public domain works in New Zealand are not in fact available to New Zealanders from online book repositories. In particular, they found that Hathi Trust was the most problematic in terms of not providing full access to content. In the detailed public domain assessments contained in Appendix B, I noted that a lot of our public domain material is included in Hathi Trust. While the appearance of older works in Hathi Trust may seem promising for Canadians, a quick read

through the document entitled “What are the different Copyright statuses of items in HathiTrust, and what do they mean?” on the Hathi Trust website

https://www.hathitrust.org/help_copyright#RightsCodes indicates additional limits placed on international readers including Canadians. On Hathi Trust, Canadian readers will not be granted access to works published outside the United States between 1880 and 1924. To illustrate this point, I offer two easily found examples of works that are in both the Killam print collection and Hathi Trust:

- *Robert Randall and the Le Breton flats; an account of the early legal and political controversies respecting the ownership of a large portion of the present city of Ottawa* by Hamnett P. Hill, published in 1919 in Ottawa.
- *Signs & wonders* by J.D. Beresford published in England in 1921.

Both of the above works are in the public domain in Canada but not available to Canadian readers on Hathi Trust. Even more problematic is the lack of Canadian access to the first item above, which was published prior to 1925 and is not available because it is published outside the U.S. (in Canada ironically enough). Similarly, works published in the United States after 1925 are not available to Canadian readers despite being in the public domain in Canada. Again, it is easy to find examples of works that are in both the Killam print collection and Hathi Trust that Canadian do not have access to:

- *The two thieves* by Theodore Francis Powys published in the United States in 1932.
- *The summing up* by Somerset Maugham published in the United States in 1938.

These examples clearly demonstrate that Canadian libraries cannot rely on Hathi Trust for access to public domain works in Canada. Hathi Trust is quite transparent about how it provides access to readers from outside the United States and the four examples provided above follow the stated rules on the Hathi Trust website. Canadian readers are grouped in with all other readers from outside the United States. This means that Canadian readers are granted less access to public domain works than they would have under Canadian Copyright law and even less access than U.S. readers, who by many recent measures should have more restricted access than Canadians.

If we accept that Hathi Trust is not a dependable source for Canadians to access many public domain books, then the next obvious source is Canadiana Online. With the Canadiana-CRKN merger in 2018, university libraries (including Dalhousie) have an indirect connection with the organization. One of the most significant changes to happen since 2018 is that content is now freely available to anyone online. Given the history and organizational structure, Canadiana can certainly be considered a dependable source of Canadian public domain content. When considering items published in Canada for digitization, it is advisable to check Canadiana for a copy before a final decision is made to digitize.

Of course, not all of our content is Canadian. The script identified roughly 19% of the Killam Library collection that was flagged for possible digitization. As mentioned in the results, there were some disciplinary trends that are also worth considering. Music and books on music showed high possible public domain levels that may not be straight forward to assess. Musical scores and sound recordings present some unique problems separate from books on music. The difficulty with musical scores is that the original creators / composers are identified in the bibliographic record, while it is often more recent editors and their publishers who claim renewed copyright in these works. Laroche (2014) points out that the bar is actually quite high for modern editions of public domain musical scores to be candidates for renewed copyright. While there have been no cases in Canada to test his view, the courts in England

have established that a modern editor would need to make significant original contributions that would substantially affect the sound of a work, for the work to receive a new copyright. In light of this, it appears that the underlying scores themselves, absent any cover image and original introductory text could be digitized and made available. Laroche further notes, however, that music publishers have been fairly adamant in claiming copyright regardless of the legitimacy of their claims. So further discussion of the merits of digitizing this part of the collection would be worth having. Another consideration with the music collection relates to sound recordings. Sound recordings made prior to June 23, 2015 with an individual creator main entry, actually pass into the public domain sooner than my script would indicate. As a result, there are a potentially large number of recordings that are in the public domain that could be digitized and made available. The methods used here could be revised and applied separately, and specifically to sound recordings to generate an even larger list of public domain items for that format.

On the flipside we have the sciences (Q) and related disciplines such as medicine (R), agriculture (S), and even the social sciences (H) where there are a lot of newer edited, multi-authored works that reflect publishing and collecting trends in those disciplines. As noted previously, the risk of copyright infringement increases with the number of named authors who contribute to a work. Also, using the NHDS guidelines that recommend digitizing works published prior to 1940 reveals some obvious trends with respect to the Killam book collection in these subject areas. Those call number ranges with the lowest percentages of public domain items (as identified by my script), also contain the lowest percentages of books published prior to 1940. Looking specifically at some of the low outliers; the Q's (science) have only 5%, the H's (social Science) have 4%, and the R's (Medicine) have 1%. These numbers all reflect the priority given to newer items in these subject areas and could be used to steer digitization priorities towards subjects that place a higher value on older items.

The results of the detailed public domain assessments revealed another potential consideration related to digitization. This relates to public domain works that are effectively reprinted but with introductions, forewords and other copyrighted material contributed by editors. Goldstein (1983) writing about derivative works points out that copyright does not extend to pre-existing underlying material and Braithwaite (1982) writing specifically about the Canadian context notes that even literary abridgements don't necessarily benefit from renewed copyright in Canada. The detailed assessments in Appendix B, revealed a potentially large number of items in our collection that fall into such a category. For these items we might consider a low risk digitization approach that involves OCR of underlying works with a resulting transcription for the underlying work and redaction or omission of in-copyright portions of the work. The production of such works would be slightly more involved than producing a faithful scan of an item but could still be done with minimal staff training. This is similar to the approach used by volunteers working with the long-standing digitization initiative Project Gutenberg, in the United States.

Once works have been identified, digitized and, in some instances, OCR transcribed there remains work to be done in making them accessible. As with our existing digital collections, a multi-pronged approach is advisable. The Dalhousie University Libraries currently support digital collections on many platforms. Several of these platforms could be used to support a large-scale public domain digitization program as outlined in this report. Our existing DSpace repository can be used to support web access and provide the core infrastructure to support a public domain hosting project of this scale. The Dalhousie Libraries have been using this software since 2007 and it continues to serve us well. The Canadian Association of Research Libraries "Inventory of Canadian Repository Platforms"

(<https://docs.google.com/spreadsheets/d/1LYaS1fbnYNEeWTN3ribt4j5whfOIGly->

[KXYgwfp spg/edit#gid=842684754](#)) accessed in November 2020 is showing that 31 university libraries across Canada still use DSpace to support digital collections. It seems clear that, in the medium term at least, DSpace will continue to be widely used and supported. The Dalhousie Libraries currently use the Handle System with our DSpace and it works well to provide stable links to our content. An excellent feature of DSpace that is particularly relevant for a large-scale public domain project such as this, is the ability to embargo items that might need to be removed from public view. This feature provides obvious support for any take-down requests that may emerge if a copyright holder objects to our online distribution of a particular item. A feature such as this may seem minor but it provides a tangible and concrete means to support our take-down policy for items that might still be in copyright. It is proof that we have taken steps to minimize any damages that result from an item being made available by mistake. The feature also ensures that we have not wasted effort in digitizing the item in the first place. All items will enter the public domain at some point in time and this embargo feature allows us to simply set a future release date for the item that is consistent with its copyright status without any further human intervention. In addition to DSpace, we have access to Archivematica software to support digital preservation activities and Omeka could support downstream exhibiting work for select titles. Further access points can be supported within the Novanet library catalogue if desired, and provide an opportunity to connect print readers with the newly available digital editions.

Conclusion

The public domain is not simply the absence of copyright, consisting of material not worth protecting because it lacks any value. Instead, the public domain represents an endpoint or timespan beyond which the state is no longer willing to tolerate the loss of value to society by having a work under the complete control of a copyright holder. To libraries, the intellectual value of a work is clearly more important than its economic value. Our collecting priorities have little to do with the copyright status of individual items. We physically protect all of our materials and remain committed to preserving intellectual and physical access points for items in our collections. As such, libraries place as much value on public domain material as we do on in-copyright materials.

This report provides a methodology and a path forward for adding value to the public domain material owned by Dalhousie University for minimal cost and effort. I used a combination of open source and readily available desktop software to facet our collection and I propose using existing software tools for online hosting and delivery of scanned items. I did not specifically address digitization workflows and staffing since the equipment and personnel are already in place to carry out such activities. The role of the Digitization Coordinator in the university archives would be a key player in any large-scale digitization program, subject to availability and with a level of involvement that would depend on operational decision makers in the library. The initiative isn't intended to replace print items with preservation quality digital objects but the project could evolve to serve a preservation goal as well if that's considered a priority. Such a plan would require additional time and effort, especially given the potential size of the public domain component of our collection. If the public domain percentage for the other Dalhousie Libraries is close to the Killam Library figure of 19%, we are looking at possibly having 200,000 items or more that could be digitized. The scale of this effort is substantial but the outcome is unprecedented online access to our valuable print collections. This outcome is both attainable and possible to support with our current digital infrastructure.

As with any project of this scale there will be aspects that are difficult to anticipate. This report outlines a plan for digitizing public domain materials. If the plan proceeds, we could notice as we physically go through the collection, that other kinds of material in our collection need attention. As the 19% public domain figure suggests, our print collection is aging. An interesting question raised by others (notably Demas and Brogdan, 1997) is whether libraries should be more proactive in digitizing brittle materials as part of their preservation mandates, regardless of the public domain status of those materials. It is difficult to imagine a case for not doing this, especially if the originals are beyond repair. From a hosting perspective, the approach to brittle or damaged items would be largely the same as for public domain materials that are in good physical condition. Digitized copies could be deposited in DalSpace, and a public domain assessment done on the item to provide a release date for online access. As mentioned above, all print items in the collection will be in the public domain at some point, so digitization work will always be worth the effort. Even before items are publicly released online, they will be discoverable, so there is immediate value to this work.

Some commentators such as Michael Geist (2016) have been very critical of the lack of effort shown by Canadian libraries to digitize public domain works. Similar to the National Heritage Digitization Strategy, Geist recommends digitizing everything published prior to 1940. A key finding of a report by the Canadian Council of Academies (2015) is that Canadian memory institutions are lagging behind other countries in response to digital opportunities. Despite such criticisms, there is no evidence to suggest that libraries in Canada have begun to act in this area at all. In Canada, we tend to wait for a national initiative that provides a risk-free, low barrier environment for moving forward in new areas. The results of my public domain survey of the Killam Library book collection hopefully provide a low-barrier starting point for the Dalhousie Libraries to begin addressing large-scale digitization of our print collection. To date we have focused our digitization efforts almost exclusively on archival materials. This makes sense as a starting point but we can do more. The items that I have identified through the work outlined in this report represent a much greater range of materials from all subject areas that are supported by the library. These items represent a significant investment by the university and a potential showpiece for the Dalhousie University Libraries. Our leadership in this area would provide an even greater digital focus for the future and would demonstrate our ability to effectively manage and highlight much more of our collection online.

References

Bill C-4, An Act to implement the Agreement between Canada, the United States of America and the United Mexican States. 43rd Parliament, 1st Session. 2020.

<https://www.parl.ca/DocumentViewer/en/43-1/bill/C-4/royal-assent>

Braithwaite, William J. (1982). Derivative Works in Canadian Copyright law. *Osgoode Hall Law Journal*, 20(2), 191-231. <http://digitalcommons.osgoode.yorku.ca/ohlj/vol20/iss2/1>

Clark and Chawner. (2014). Enclosing the public domain: The restriction of public domain books in a digital environment. *First Monday*, 6(2). <https://doi.org/10.5210/fm.v19i6.4975>

Copyright Act. (R.S.C., 1985, c. C-42). <https://laws-lois.justice.gc.ca/eng/acts/C-42/section-6.1-20021231.html>

- Council of Canadian Academies. (2015). *Leading in the Digital World: Opportunities for Canada's Memory Institutions*. Ottawa (ON): The Expert Panel on Memory Institutions and the Digital Revolution, Council of Canadian Academies. https://cca-reports.ca/wp-content/uploads/2018/10/cofca_14-377_memoryinstitutions_web_e.pdf
- Craig, Carys, J. (2010). The Canadian Public Domain: What, Where, and to What End? *Canadian Journal of Law and Technology*, 10 (1 & 2), 221–241. <https://ojs.library.dal.ca/CJLT/article/view/5219/4715>.
- Demas, Samuel and Brogdon, Jennie L. (1997). Determining Copyright Status for Preservation and Access: Defining Reasonable Effort. *Library Resources and Technical Services*, 41(4), 323-334. <https://doi.org/10.5860/lrts.41n4.323>
- Geist, Micheal. (2016). Canada's National Digitization Plan Leaves Virtual Shelves Empty. Homepage Version. <https://www.michaelgeist.ca/2016/07/canadas-national-digitization-plan-leaves-virtual-shelves-empty/>
- Geist, Micheal. (2019). Canada Introduces USMCA Implementation Bill...Without a General Copyright Term Extension Provision. Homepage Version. <https://www.michaelgeist.ca/2019/05/canada-introduces-usmca-implementation-bill-without-a-general-copyright-term-extension-provision/>
- Geller, Paul Edward. (2000). Zombie and Once-Dead Works: Copyright Retroactivity after the E.C. Term Directive. *Entertainment and Sports Lawyer*, 18(2), 7-12. HeinOnline. <https://heinonline.org/HOL/P?h=hein.journals/entspl18&i=27>
- Goldstein, Paul. (1983). Derivative Rights and Derivative Works in Copyright. *Journal of the Copyright Society of the U.S.A.*, 30(3), 209-252. HeinOnline. <https://heinonline.org/HOL/P?h=hein.journals/jocos030&i=223>
- Laroche, Guillaume. (2014). Settling the Score: Copyright in Modern Editions of Public Domain Musical Works. *Osgoode Legal Studies Research Paper Series*, 80. <https://digitalcommons.osgoode.yorku.ca/olsrps/80>
- Smith, Kevin L. (2012). Copyright Risk Management: Principles and Strategies for LargeScale Digitization Projects in Special Collections. *Research Library Issues: A Quarterly Report from ARL, CNI, and SPARC*, no. 279, 17–23. <http://publications.arl.org/rli279/>
- University of British Columbia Scholarly Communications and Copyright Office. (2020). Copyright at UBC. <https://copyright.ubc.ca/public-domain/>

Appendix A – JSON script used for each Library of Congress subclass in the Killam Library collection

```
[
  {
    "op": "core/column-split",
    "engineConfig": {
      "facets": [],
      "mode": "row-based"
    },
    "columnName": "Author",
    "guessCellType": true,
    "removeOriginalColumn": false,
    "mode": "separator",
    "separator": "\\d\\d\\d\\d-",
    "regex": true,
    "maxColumns": 0,
    "description": "Split column Author by separator"
  },
  {
    "op": "core/column-rename",
    "oldColumnName": "Author 2",
    "newColumnName": "Known Death Date",
    "description": "Rename column Author 2 to Known Death Date"
  },
  {
    "op": "core/column-addition",
    "engineConfig": {
      "facets": [],
      "mode": "row-based"
    }
  }
]
```

```
"baseColumnName": "Known Death Date",
"expression": "grel:GREL if(value < 1970, \"Possible Public Domain\", \"Not Public Domain\")",
"onError": "set-to-blank",
"newColumnName": "Person Author Public Domain Status",
"columnInsertIndex": 12,
"description": "Create column Person Author Public Domain Status at index 12 based on column
Known Death Date using expression grel:GREL if(value < 1970, \"Possible Public Domain\", \"Not Public
Domain\")"
},
{
"op": "core/column-removal",
"columnName": "Author 1",
"description": "Remove column Author 1"
},
{
"op": "core/column-removal",
"columnName": "Person Author Public Domain Status",
"description": "Remove column Person Author Public Domain Status"
},
{
"op": "core/column-addition",
"engineConfig": {
"facets": [],
"mode": "row-based"
},
"baseColumnName": "Known Death Date",
"expression": "grel:if(value < 1970, \"Possible Public Domain\", \"Not Public Domain\")",
"onError": "set-to-blank",
"newColumnName": "Person Author Public Domain Status",
```

```
"columnInsertIndex": 11,

"description": "Create column Person Author Public Domain Status at index 11 based on column
Known Death Date using expression grel:if(value < 1970, \"Possible Public Domain\", \"Not Public
Domain\")"

},

{

"op": "core/column-addition",

"engineConfig": {

"facets": [],

"mode": "row-based"

},

"baseColumnName": "Author",

"expression": "grel:if(value.contains(\":\"),\"Corporate Author\", \"Personal Author\")",

"onError": "set-to-blank",

"newColumnName": "Author Type",

"columnInsertIndex": 10,

"description": "Create column Author Type at index 10 based on column Author using expression
grel:if(value.contains(\":\"),\"Corporate Author\", \"Personal Author\")"

},

{

"op": "core/column-addition",

"engineConfig": {

"facets": [

{

"type": "list",

"name": "Author Type",

"expression": "value",

"columnName": "Author Type",

"invert": false,

"omitBlank": false,
```

```
"omitError": false,
"selection": [
  {
    "v": {
      "v": "Corporate Author",
      "l": "Corporate Author"
    }
  }
],
"selectBlank": false,
"selectError": false
}
],
"mode": "row-based"
},
"baseColumnName": "Publication Year",
"expression": "grel:value",
"onError": "set-to-blank",
"newColumnName": "Corporate Author Public Domain Status",
"columnInsertIndex": 15,
"description": "Create column Corporate Author Public Domain Status at index 15 based on column
Publication Year using expression grel:value"
},
{
  "op": "core/text-transform",
  "engineConfig": {
    "facets": [
      {
        "type": "list",
```

```
"name": "Author Type",
"expression": "value",
"columnName": "Author Type",
"invert": false,
"omitBlank": false,
"omitError": false,
"selection": [
  {
    "v": {
      "v": "Corporate Author",
      "l": "Corporate Author"
    }
  }
],
"selectBlank": false,
"selectError": false
}
],
"mode": "row-based"
},
"columnName": "Corporate Author Public Domain Status",
"expression": "value.toNumber()",
"onError": "keep-original",
"repeat": false,
"repeatCount": 10,
"description": "Text transform on cells in column Corporate Author Public Domain Status using
expression value.toNumber()"
},
{
```

```
"op": "core/text-transform",
"engineConfig": {
  "facets": [
    {
      "type": "list",
      "name": "Author Type",
      "expression": "value",
      "columnName": "Author Type",
      "invert": false,
      "omitBlank": false,
      "omitError": false,
      "selection": [
        {
          "v": {
            "v": "Corporate Author",
            "l": "Corporate Author"
          }
        }
      ],
      "selectBlank": false,
      "selectError": false
    }
  ],
  "mode": "row-based"
},
"columnName": "Corporate Author Public Domain Status",
"expression": "grel:if(value < 1970, \"Possible Public Domain\", \"Not Public Domain\")",
"onError": "keep-original",
"repeat": false,
```

```
"repeatCount": 10,  
  "description": "Text transform on cells in column Corporate Author Public Domain Status using  
expression grel:if(value < 1970, \"Possible Public Domain\", \"Not Public Domain\")"  
}  
]
```

Appendix B – Detailed public domain assessments and digitization recommendations for sampled items

World History – D

A random sample of 15 items (0.1%) from the 14,803 possible public domain items in the World History call number range revealed 1 item that definitely isn't in the public domain. An additional item appears to be in the public domain but would need select in-copyright content redacted before being made publicly available. With 14 out of 15 sampled items being in the public domain, my OpenRefine script was effective at identifying public domain items in the World History call number ranges.

Wellington, Arthur Wellesley. Supplementary Despatches and Memoranda of Field Marshal Arthur Duke of Wellington, K. G. London, 1858. Print. This item is definitely in the public domain in Canada. It has a single author with a death date of 1852. The item is available in Hathi Trust and due to its age (published prior to 1880 and place of publication (outside the U.S.) it will likely continue to be accessible to Canadians but this is no guarantee of future access by Canadians. **Digitize.**

Trevelyan, George Otto, and Macaulay, Thomas Babington Macaulay. The Life and Letters of Lord Macaulay. New York: Harper & Brothers, 1876. This item is in the public domain in Canada. It has a single author with a death date of 1928. The item is available in Hathi Trust and due to its age (published prior to 1880 and place of publication (outside the U.S.) it will likely continue to be accessible to Canadians but this is no guarantee of future access by Canadians. **Digitize.**

Nehru, Jawaharlal, Chalapathi Rau, M., Sharada Prasad, H. Y., Nanda, B. R., Gopal, Sarvepalli, and Nanda, Bal Ram. Selected Works of Jawaharlal Nehru. New Delhi: Orient Longman, 1972. This item will need to be viewed in person to determine whether original content was contributed by the editor or advisory board. These individuals are identified as named added entries some with death dates as recently as 2008. **Digitize and OCR underlying content by Nehru. Redact content from subsequent contributors.**

Lenin, Vladimir Il'ich. Polnoe Sobranie Sochinenii. Izd. 5.. ed. Moskva: Gos. Izd-vo Polit. Lit-ry, 1960. This item looks like it is a single volume in a multi volume set with dates ranging from 1921-1988. With a single author and known death date in 1924 this particular volume is in the public domain. **Digitize.**

Arberry, A. J. The Legacy of Persia. Oxford: Clarendon, 1953. The Legacy Series. This item has a named editor who appears bibliographically as a personal name added entry whose death date would place this item in the public domain. With a title main entry though, it appears that this item is in the public domain by virtue of its publication date and having no named authors. **Digitize.**

Kinglake, Alexander William. The Invasion of the Crimea : Its Origin, and an Account of Its Progress down to the Death of Lord Raglan. Edinburgh ; London : W. Blackwood, 1901. This item is part of a multi volume set of works. The edition here is in the public domain by virtue of the death date of the only author in 1891. An earlier edition of this item is available in the Killam print collection and in Hathi Trust. Due to its age (published prior to 1880 and place of publication (outside the U.S.) that version is accessible to Canadians but this is no guarantee of future access by Canadians. **Digitize.**

Chamberlain, Neville, and Self, Robert C. The Neville Chamberlain Diary Letters. Aldershot, Hampshire, England ; Burlington, Vt., USA: Ashgate Pub., 2000. Print. It appears this item was flagged as possible

public domain because the original creator of the diaries is included in the bibliographic record as a main entry with a death date of 1940. This edited version published in 2000 by Robert C. Self is not in the public domain. **Do not digitize.**

Johnson, George William. *The Fairfax Correspondence: Memoirs of the Reign of Charles the First*. London: R. Bentley, 1848. This is volume 2 of a 2 volume set. It is in the public domain with an author death date of 1886. This item is available in Hathi Trust and due to its age will likely continue to be available to Canadian readers but there is no guarantee of future access by Canadians. **Digitize.**

Hidalgo De Cisneros, Ignacio. *Cambio De Rumbo*. Bucarest, 1970. This item is volume 2 of a 2 volume set. It is in the public domain in Canada with a single author death date of 1966 and can be digitized and made publicly available. No other digital copy is available online. **Digitize.**

Lauterbach, Richard E. *These Are the Russians*. New York: Book Find Club, 1945. This item has an author death date of 1950 with no other contributors. It is in the public domain in Canada and can be digitized and made publicly available. No other digital copy is available online. **Digitize**

Luchaire, Achille. *Innocent III*. Paris: Hachette, 1904. This item has an author death date of 1908 with no other contributors. It is in the public domain in Canada and can be digitized and made publicly available. No other digital copy is available online. **Digitize.**

Statisticheskoe Opisanie Zemli Donskikh Kazakov, Sostavlennoe v 1822-32 Godakh. *Novocherkassk: Oblastnaia Voiska Donskago Tip.*, 1891. This item has no person named author or contributor. With a publication date of 1891, it is in the public domain in Canada and can be digitized and made publicly available. No other digital copy is available online. **Digitize.**

Ruville, Albert Von, Chaytor, H. J., and Morison, Mary. *William Pitt, Earl of Chatham*. London : New York: Heinemann ; G.P. Putnam's Sons, 1907. This item is a translation of an earlier work. The death date of the original author is 1934. The death date of the first translator is 1945 and the death date of the contributor who wrote the introduction is 1927 (according to the Library of Congress Name Authorities). The death date of the second translator is unknown. Given the age of the item the uncertainty surrounding the death date of the secondary translator presents a low risk for copyright infringement. This item can be digitized and made publicly available. No other digital copy is available online. **Digitize.**

Bogdanovich, M. I. *Istoriia Tsarstvovaniia Imperatora Aleksandra I I Rossii v Ego Vremia*. S. Peterburg: Tip. F. Sushchinskago, 1869. This item has one identified author with a death date of 1882. It is in the public domain in Canada and can be digitized and made publicly available. No other digital copy is available online. **Digitize.**

Zabielin", Ivan". *Istoriia Russkoj Zhizni S" Drevniejšikh" Vremen"*. The Hague ; Paris: Mouton, 1969. This is a 1969 reprint of an earlier published item. There are no named contributors related to the 1969 reprint. The death date of the original author in 1909 puts this item in the public domain in Canada. It is not available to Canadian readers in Hathi Trust. **Digitize.**

Raine, James. *The Historians of the Church of York and Its Archbishops*. Vaduz?]: Kraus Reprint, 1965. Print. *Rerum Britannicarum Medii Aevi Scriptores; v. 71*. The author died in 1896. This 1965 reprinting

is in the public domain. The item is commercially available to Dalhousie readers on Tanner Ritchie. **Digitize.**

Adams, Henry, and Ari'fi Ta'ima'fi. Tahiti; Memoirs of Arie Taimai. Ridgewood, N.J.: Gregg, 1968. The most recent author affiliated with this reprint died in 1918. The item is in the public domain. **Digitize.**

Busch, Moritz. Bismarck; Some Secret Pages of His History; Being a Diary Kept by Dr. Moritz Busch during Twenty-five Years' Official and Private Intercourse with the Great Chancellor. London ; New York: Macmillan, 1898. The author died in 1899. This item is in the public domain and also available on Hathi Trust. **Digitize.**

McLennan, J. S. History and Present Problems. Sackville, N.B.: Argosy, 1924. Webster Lectureship in Canadian History, Mount Allison University. The author died in 1939. The item is in the public domain. **Digitize.**

Gandhi. Collected Works of Mahatma Gandhi. 2d Rev. Ed.]. ed. New Delhi: Publications Division, Ministry of Information and Broadcasting, Govt. of India, 1969. Gandhi died in 1948. This 1969 Government of India publication is in the public domain in Canada. **Digitize.**

History of the Americas – F

The History of the Americas class was chosen because it contains Canadian History. For this class, I took a random sample of 10 items. An additional consideration with Canadian content is the availability of items on the Canadiana platform. Unlike Hathi Trust, Canadiana can be considered a dependable online source for Canadians to access public domain materials online. Items from the Dalhousie collection that are in the public domain but also available on Canadiana are not recommended for digitization. Overall, 6 of 10 items in the sampled call numbers are definitely in the public domain. An additional 2 items appear to be in the public domain and present a very low risk of copyright infringement. Surprisingly only 4 items of the 10 sampled, are available in Canadiana. With an 80% success rate, the script appears to be a reasonable indicator of public domain items that should be digitized.

Canniff, William, and Great Britain. Army. King's Royal Regiment, 2d Battalion. History of the Settlement of Upper Canada (Ontario); with Special Reference to the Bay Quinté. Toronto: Dudley & Burns, 1869. This British Army publication from 1869 is in the public domain in Canada. It can be digitized but it is already available in Canadiana which is a dependable source for Canadians to access digital content online. **Public Domain but do not digitize.**

Careless, J. M. S., and Robert Craig Brown. The Canadians, 1867-1967. Toronto: Macmillan of Canada, 1967. This item has many contributors and an editor with a death date of 2009. My workflow in OpenRefine identified it as a title main entry item but closer inspection reveals many personal names associated with this work. This item is not in the public domain in Canada. **Do not digitize.**

Canada. Commission of Conservation. Committee on Waters Water-Powers. Water Powers of British Columbia, including a Review of Water Power Legislation Relating Thereto and a Discussion of Various Matters Respecting the Utilization and Conservation of Inland Waters. 1919. This item is in the public domain in Canada. It can be digitized but it is already available in Canadiana which is a dependable source for Canadians to access digital content online. **Public Domain but do not digitize.**

Robbins, John E. Encyclopedia Canadiana. 1st Deluxe ed. Ottawa: Canadiana, 1957. As an encyclopedia, the print item would need further review to determine public domain status. If the names of contributors are included in the item, it would not be in the public domain. **Do not digitize.**

Devine, P. K. Ye Olde St. John's. St. John's Newfoundland, 1939. With an author death date of 1950, this item is in the public domain in Canada. It can be digitized and made publicly available. No other digital copy is available online. **Digitize.**

Dawson, S. J., and Thunder Bay Sentinel. The Red River Expedition of 1870. Toronto: Canadiana Library Service, 1967. With an author death date of 1902, this item is in the public domain in Canada. It can be digitized but it is already available in Canadiana which is a dependable source for Canadians to access digital content online. It is also available on HathiTrust. **Public Domain but do not digitize.**

Macnaughton, John, and Vaughan, Walter. Lord Strathcona. New York, Toronto: Oxford University Press, 1926. Print. Makers of Canada Ser. Anniversary Ed Vol. 10. The primary author has a death date of 1943 and the death date of the second identified author is not indicated in the Library of Congress authority files. Given the date of publication there is a high likelihood that this item is in the public domain in Canada. There is a low risk of copyright infringement. **Digitize.**

Lajeunesse, Ernest J. and Champlain Society. The Windsor Border Region, Canada's Southernmost Frontier; a Collection of Documents. Toronto: Champlain Society for the Government of Ontario, University of Toronto. 1960. Publications of the Champlain Society. Ontario Ser.; 4. This item has a society listed as author with an individual identified as writing the introduction. In the absence of a death date for the individual, this item could be selectively digitized to exclude the introduction. **Underlying content in the public domain. Digitize with introduction redacted.**

McIlwraith, Jean N. Canada. Toronto: William Briggs, 1899. Print. Children's Study. The authors death date puts this item in the public domain but it is already available online in Canadiana. **Public domain but do not digitize.**

Glover, T. R., and Calvin, Dileno Dexter. A Corner of Empire: The Old Ontario Strand. Cambridge: U, 1937. The death dates of both authors is greater than 50 years ago. **Public domain, digitize.**

Social Sciences - H

With a success rate of 65%, (13 out of 20), the social sciences was the least successful subject area based on the detailed assessments. In many cases the author main entries are not revealing the many collaborative multi-authored works appearing in the social sciences. In some cases, the bibliographic records indicate a corporate or society author but closer inspection of the records reveals that individuals are often affiliated with the work. When reviewing items in the social sciences it appears that publication date could be a better overall indicator of public domain status. Guidance could be taken from the National Heritage Digitization strategy that encourages libraries to look for publication dates prior to 1940.

Brailsford, Henry Noel. Property or Peace. London: V. Gollancz, 1936. The author died in 1958. The item is in the public domain. **Digitize.**

Angell, Norman. Must Britain Travel the Moscow Road? London: N. Douglas, 1926. The author died in 1967. The item is in the public domain. **Digitize.**

Henry, David H. W., Clarry, J. H. C., and Canadian Institute of Chartered Accountants. The Combines Investigation Act and Management. Toronto: Canadian Institute of Chartered Accountants, 1965.

Authors are identified as added entries in the bibliographic record. No death dates are indicated. The item is likely not in the public domain. **Do not digitize.**

Habakkuk, H. J., and Postan, M. M. The Industrial Revolutions and after: Incomes, Population and Technological Change. Cambridge University Press, 1966. Cambridge Economic History of Europe; v. 6.

Many authors are indicated in the contents note in the bibliographic record. The item is likely not in the public domain. **Do not digitize.**

Chatterton, E. Keble. King's Cutters and Smugglers, 1700-1855. Philadelphia, London: J. B. Lippincott; G. Allen &, 1912. The author died in 1944, The item is in the public domain and available on Project Gutenberg. **Digitize.**

Fawcett, Millicent Garrett. Tales in Political Economy. London: Macmillan, 1874. The author died in 1929. The item is in the public domain and available on Hathi Trust. **Digitize.**

Unwalla, Darab B., Trimm, John H., Collins, June M., and Warner, W. Lloyd. The Emergent American Society. New Haven: Yale UP, 1967. The item has many additional authors. It is likely not in the public domain. **Do not digitize.**

Menger, Carl. Problems of Economics and Sociology. Urbana: U of Illinois, 1963. Copyright in this translation extends to the translator. There is no available death date for the translator and given the publication date it is unlikely that this item is in the public domain. **Do not digitize.**

Saposs, David J. Left Wing Unionism; a Study of Radical Policies and Tactics. New York: Russell & Russell, 1967. This is a reprint of the original 1926 edition. No new material is indicated in the bibliographic record. The author died in 1968. The item is in the public domain and available on Hathi Trust. **Digitize.**

Arrow, Kenneth J., Scitovsky, Tibor, and American Economic Association. Readings in Welfare Economics, Selected by a Committee of the American Economic Association. Homewood, Ill.: Published for the Association by R. D. Irwin, 1969. Ser. of Republished Articles on Economics, v. 12. The item consists of republished articles with many authors. The item is likely not in the public domain. **Do not digitize.**

Rusanov, N. S. Iz Moikh Vospominaniĭ. Berlin: Izd-vo Z.I. Grzhebina, 1923. Letopis Revoliutsii; No. 13. The author died in 1939. The item is in the public domain. **Digitize.**

Pareto, Vilfredo, and Eisermann, Gottfried. System Der Allgemeinen Soziologie. Stuttgart: F. Enke, 1962. The author died in 1923 but the work was not published until 1962. The work is in the public domain but introductory content by Gottfried Eisermann should be redacted. **Digitize.**

Marx, Karl, and Engels, Friedrich. Karl Marx, Friedrich Engels. Werke. 1. Aufl. ed. Berlin: Dietz, 1956. Vol. 8. This posthumous publication (in 41 volumes) of earlier works by Marx and Engels was published in 1956. The content is in the public domain. **Digitize.**

Marx, Karl, and Engels, Friedrich. Karl Marx, Friedrich Engels. Werke. 1. Aufl. ed. Berlin: Dietz, 1956. Vol. 17. This posthumous publication (in 41 volumes) of earlier works by Marx and Engels was published in 1956. The item is in the public domain. **Digitize.**

Heckscher, Eli F., and Soderland, Ernest Frithiof. Mercantilism. Rev. [2d] Ed. Edited by E. F. Sörderlund.. ed. London, New York: Allen & Unwin; Macmillan, 1955. This second edition of a 1935 translation is likely not in the public domain since it has been revised and translations benefit from new copyright. The death date of the translator may have been as recent as 2010. **Do not digitize.**

Dawson, Sidney Stanley, and Graves, Harry Robert. The Accounts of Trustees, Liquidators, and Receivers. London: Gee &, 1905. Accountants' Library ; 2nd Ser., Vol. 41. The author died in 1926. The item is in the public domain. **Digitize.**

Marx, Karl, and Engels, Friedrich. Karl Marx, Friedrich Engels. Werke. 1. Aufl. ed. Berlin: Dietz, 1956. Vol. 28. This posthumous publication (in 41 volumes) of earlier works by Marx and Engels was published in 1956. The item is in the public domain. **Digitize.**

Prest, A. R. The UK Economy; a Manual of Applied Economics. 2nd ed. London: Weidenfeld & Nicolson, 1968. With only a named editor this item could be in the public domain but it appears unlikely. It would need to be viewed in person to determine if individual authors are identified. **Do not digitize.**

Addams, Jane. The Second Twenty Years at Hull-house, September 1909 to September 1929, with a Record of a Growing World Consciousness. New York: Macmillan, 1930. The author died in 1935. The item is in the public domain. **Digitize.**

Hollander, Jacob H., and Evans, George Heberton. A Reprint of Economic Tracts. Baltimore: Johns Hopkins, 1900. This reprint of earlier works in economics has named editors but there's no evidence of original contributions by the editors. The item is in the public domain. **Digitize.**

English Literature - PR

This relatively large class in the Killam Library collection includes Canadian literature. I sampled 20 random items and found all of them to be in the public domain. With a relatively small number of Canadian works the category, not many items are available on Canadiana. The perfect success rate of the script is likely due to the fact that English literature doesn't appear to have a lot of collaboratively written multi authored works.

Ainsworth, William Harrison. Ballads; Romantic, Fantastical, and Humorous, to Which Is Added the Combat of the Thirty. Comprising a New Chapter of Froissart. London, New-York: G. Routledge & Sons, 1872. The author died in 1872. This item is in the public domain. **Digitize.**

Talfourd, Thomas Noon. Vacation Rambles and Thoughts : Comprising the Recollections of Three Continental Tours, in the Vacations of 1841, 1842, and 1843. London: Edward Moxon, 1845. The author died in 1854. This item is in the public domain. It is available on Hathi Trust but not in a Canadian repository. **Digitize.**

Trollope, Anthony, and Leslie, Shane. Phineas Finn, the Irish Member. London: Oxford UP, 1969. World's Classics; 447. The author died in 1882 but the introduction to this 1969 printing was written by someone who died in 1971. This printing can be digitized with the introduction redacted but there is an earlier version in the Killam collection that can be digitized and made available in its entirety. The item is

available on both Hathi Trust and Project Gutenberg but no Canadian source. The underlying content is in the public domain. **Digitize.**

Deeping, Warwick. *The Road*. Toronto: McClelland and Stewart, 1931. The author died in 1950. The item is in the public domain and not available elsewhere online. **Digitize.**

Swinburne, Algernon Charles. *The Poems of Algernon Charles Swinburne*. London: Chatto & Windus, 1912. The author died in 1909. The item is in the public domain and is available on Hathi Trust but not on a Canadian source. **Digitize.**

Leland, Thomas, and Leland, John. *Longsword, Earl of Salisbury. Flowering of the Novel*. New York: Garland Pub., 1974. The author died in 1785. This item is a reprint only. It is in the public domain. Available on Hathi Trust. **Digitize.**

Somerville, E. C., and Ross, Martin. *The Sweet Cry of Hounds*. London: Methuen, 1936. The author died in 1949. This item is in the public domain and not available elsewhere online. **Digitize.**

Blackwood, Algernon. *The Fruit Stoners; Being the Adventures of Maria among the Fruit Stoners*. London: Grayson & Grayson, 1934. The author died in 1949. This item is in the public domain and not available elsewhere online. **Digitize.**

Cowper, William, and Nicolas, Nicholas Harris. *The Poetical Works of William Cowper ...* Boston: Little, Brown, 1853. *British Poets*, Ed. by F. J. Child. The death date of both author and author added entry put this item in the public domain. The item is available on Hathi Trust. **Digitize.**

Lindsay, Norman. *Puddin' Poems : Being the Best of the Verse from the 'Magic Pudding'*. London [etc.]: Angus and Robertson, 1977. The author died in 1969 but the item was published in 1977. The bibliographic record does not indicate any author added entries or editors. The item is in the public domain but may need selective redaction. **Digitize.**

Morley, John. *Critical Miscellanies*. London: Macmillan, 1886. The author died in 1923. This item (a reprint from the Fortnightly Review) is in the public domain. There are versions available on Project Gutenberg and Hathi Trust. **Digitize.**

Corelli, Marie. *The Treasure of Heaven ; a Romance of Riches*. London: A. Constable, 1906. The author died in 1924. The item is in the public domain. It is available on Hathi Trust and Canadiana. **Do not digitize.**

Hammond, Eleanor Prescott. *Chaucer; a Bibliographical Manual*. New York: Macmillan, 1908. The author died in 1933. The item is in the public domain. A version is available in Hathi Trust. **Digitize.**

House, Humphry. *Coleridge. Clark Lectures, 1951-2*. London: R. Hart-Davis, 1953. The author died in 1955. The item is in the public domain. **Digitize.**

Harris, Frank. *The Bomb*. London: J. Long, 1908. The author died in 1931. The item is in the public domain. **Digitize.**

Moore, T. Sturge. *Mariamne*. London: Duckworth, 1911. The author died in 1944. The item is in the public domain. A version is available on Hathi Trust. **Digitize.**

Eliot, T. S. *Elizabethan Essays*. New York: Haskell House, 1964. The author died in 1965. The item is in the public domain in Canada but not in the U.S. so it is not available on Hathi Trust. **Digitize.**

Buchan, John, and Tweedsmuir, Susan. *The Clearing House*. London: Hodder and Stoughton Limited, 1946. The author died in 1940. There is a preface written by another individual who died in 1957. The item is in the public domain. **Digitize.**

Addison, Joseph, and Arnold, Thomas. *Addison*. Oxford: Clarendon, 1906. Print. Clarendon Press Ser. Addison died in 1719 and Arnold died in 1900. This item is in the public domain. **Digitize.**

Baring, Maurice. *Passing by*. London: M. Secker, 1928. The author died in 1945. This item is in the public domain. **Digitize.**

Natural History – Biology QH

A total of 7 out of 10 sampled items from this subclass appear to be in the public domain. The lower success rate could be explained by the sciences having a greater number of items with multiple authors especially with published conference proceedings being more common in the sciences.

Holt, Perry C., Parker, Bruce C., Roane, Martha K., Virginia Polytechnic Institute State University, and Association of Southeastern Biologists. *The Distributional History of the Biota of the Southern Appalachians*. Blacksburg: Virginia Polytechnic Institute and State U, 1969. Virginia Polytechnic Institute and State University. Research Division. Monograph 1. This item has multiple authors added entries and appears to be a published symposium proceedings (so there are likely many more). No death dates are provided in the bibliographic data. Due to the date of publication, it doesn't appear likely that this item is in the public domain. **Do not digitize.**

Kamm, Minnie Elizabeth Watson. *Studies on Gregarines II: Synopsis of the Polycystid Gregarines of the World, excluding those from the Myriapoda, Orthoptera, and Coleoptera*. Urbana, Ill.: U of Illinois, 1922. Illinois Biological Monographs, Vol. VII, No. 1, January, 1922. The author died in 1954 and the record indicates contributions by Henry B. Ward. The Library of Congress subject authorities indicate a death date of 1945 for this individual. The item is in the public domain. **Digitize.**

Castle, William E. *Genetic Studies of Rabbits and Rats*. Washington: Carnegie Institution of Washington, 1922. Carnegie Institution of Washington Publication; No. 320. The author died in 1962. The item is in the public domain and is available on Hathi Trust. **Digitize.**

Huxley, Thomas Henry, and Huxley, Leonard. *Life and Letters of Thomas Henry Huxley*. London: Macmillan, 1900. The death dates of both Huxleys puts this item in the public domain. It is available on Hathi Trust. **Digitize.**

Romanes, George John. *An Examination of Weismannism*. Chicago: Open Court, 1893. The author (a Canadian) died in 1894. The item is in the public domain and is available on Hathi Trust and Canadiana. **Do not digitize.**

Glass, David C., Russell Sage Foundation, Social Science Research Council, and Rockefeller University. *Genetics; Proceedings of a Conference under the Auspices of Russell Sage Foundation, the Social Science Research Council, and the Rockefeller University*. (1968). This conference proceeding from 1968 has a named editor and would have many authors by virtue of being a conference proceeding. With a publication date of 1968, it is not likely in the public domain. **Do not digitize.**

Darwin, Charles, and Keynes, R. D. Charles Darwin's Beagle Diary. Cambridge [Cambridgeshire] ; New York: Cambridge UP, 1988. The bibliographic record for this item has Charles Darwin as the author. Since Darwin died in 1882, the script flagged this as being in the public domain. The editor of these diaries likely made significant contributions to the underlying content and died in 2010. This 1988 edition is still in copyright. **Do not digitize.**

Muller, H. J., and Carlson, Elov Axel. The Modern Concept of Nature. 1st ed. Albany: State U of New York, 1973. The author died in 1967. The item is in the public domain. **Digitize.**

Lent, Herman. Atas Do Simposio Sôbre a Biota Amazonica. Rio De Janeiro: Consellio Nacional De Pesquisas, 1967. The main entry for this atlas is an organization with Herman Lent serving an editorial role rather than a content creation role. This item is likely in the public domain and presents a low risk of copyright infringement. **Digitize.**

Goldschmidt, Richard, and Dakin, William J. The Mechanism and Physiology of Sex Determination. London: Methuen &, 1923. The authors died in 1958 and 1950. The item is in the public domain. **Digitize.**