

Loren Fantin  
CRKN Digital Heritage  
October 18, 2019

OurDigitalWorld.org



# COLLABORATING. CREATING. SHARING.



OurDigitalWorld is a non-profit organization that supports digital stewardship of community cultural heritage.

We collaborate with libraries, archives, museums, historical societies, and community groups to create sustainable digital collections and bring together for discovery the extraordinary stories that define and shape our communities.

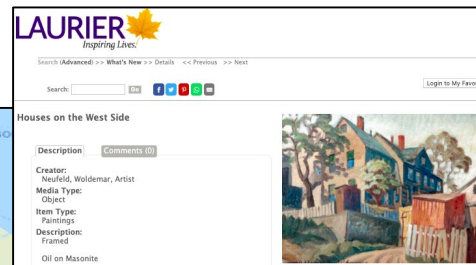


# Digitization Strategy

Create opportunities and build capacity, particularly at the grassroots/local level

- Foster equitable participation
- Create reusable cultural heritage data
- Build up the capabilities of GLAM sector

# Digital Toolkit (VITA Partners)



*Federated Women's Institutes of Ontario*  
DIGITAL COLLECTIONS



# “Community” collections

```
<record xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:georss="http://www.georss.org/georss">
  <dc:identifier>
    http://vitacollections.ca/brightonarchives/3446621/data
  </dc:identifier>
  <dc:title> Goodfellow Road # 655 </dc:title>
  <dc:creator> Ralph de Jonge </dc:creator>
  <dc:publisher> Brighton Digital Archives </dc:publisher>
  <dcterms:temporal> 2017 </dcterms:temporal>
  <dc:description>
    Owner: The McCreedy Family Address: 655 Goodfellow Road Structural History: Records indicate the following agricultural
    structures were erected at this address, although some may no longer remain standing. This chronological list provides a
    built heritage narrative to accompany the image of the remaining barn(s) on the property. 1900 Single Family Detached, 1.75
    Storey, 2180 sq. ft. 1900 Shed, 360 sq. ft. 1900 Type I Barn, 2.0 Storey, 972 sq. ft. 1910 Type I Barn, 2.0 Storey, 2172 sq.
    ft. 1960 Type I Barn, 2.0 Storey, 1182 sq. ft. Owner: The McCreedy Family Address: 655 Goodfellow Road Structural History:
    Records indicate the following agricultural structures were erected at this address, although some may no longer remain
    standing. This chronological list provides a built heritage narrative to accompany the image of the remaining barn(s) on the
    property. 1900 Single Family Detached, 1.75 Storey, 2180 sq. ft. 1900 Shed, 360 sq. ft. 1900 Type I Barn, 2.0 Storey, 972
    sq. ft. 1910 Type I Barn, 2.0 Storey, 2172 sq. ft. 1960 Type I Barn, 2.0 Storey, 1182 sq. ft.
  </dc:description>
  <dc:subject>Farms</dc:subject>
  <dc:subject>Farming</dc:subject>
  <dc:subject>Farmers</dc:subject>
  <dc:subject>Farm life</dc:subject>
  <dc:subject>Barns</dc:subject>
  <dc:subject>Architecture</dc:subject>
  <dc:subject>Agriculture</dc:subject>
  <dc:subject>Agricultural facilities</dc:subject>
  <dcterms:issued>2017-12-04T08:43:53Z</dcterms:issued>
  <dcterms:modified>2017-12-12T15:17:31Z</dcterms:modified>
  <dc:type> Image </dc:type>
  <dcterms:spatial> Brighton Codrington </dcterms:spatial>
  <dc:language> eng </dc:language>
</record>
```

See Alternate Formats:

[Dublin Core](#) | [RSS](#) | [Atom](#) | [Solr](#) | [Mods](#) | [RDF](#)

# “Local” history = national heritage



Search (Advanced) >> What's New >> Details << Previous >> Next

Search:  [Go](#) [f](#) [t](#) [p](#) [g](#) [m](#) [Login to My Favourite](#)

## Japanese Internment Camp

Description

Comments (0)

**Media Type:**  
Image

**Item Type:**  
Photographs

**Description:**  
Rear L – R: 1) ? 2) Fred Vogami 3) ? 4) ? 5) ? 6) ? 7) Jack Sadoio Shikitani 8) ? 9) ? 10) ? 11) Willie Kimitoshi Utsunomiya 12) Swede Sawada 13) George Suzuki, Middle L – R: 1) ? 2) Butch Watanabe 3) Sam Hagino 4) Fred Akira Shititani 5) Koichi Nishikaze, Front L – R: 1) Mickey Hogara 2) George Mochizuki 3) ? 4) John Kikuo Shikatani 5) ? 6) Art Tateishi 7) Sandy Ono



[Comment on this item](#)

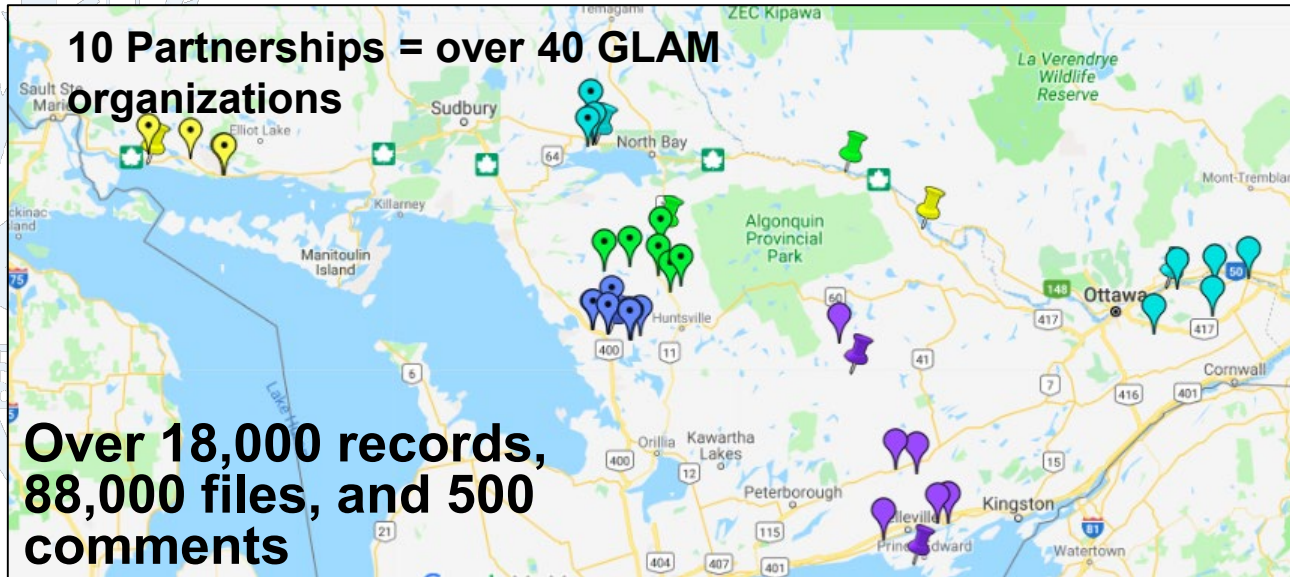
[Create electronic postcard](#)

[Full Image](#)

# Community Digitization Project (2009-2011)

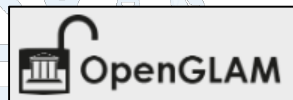
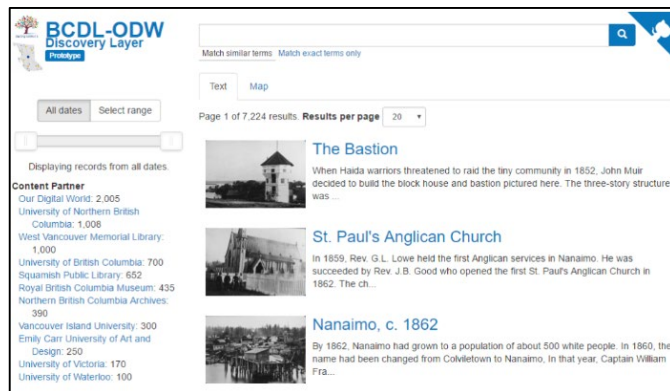
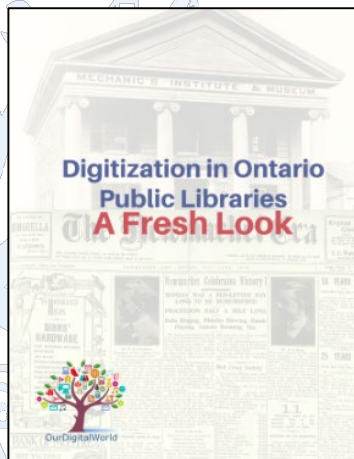
**10 Partnerships = over 40 GLAM  
organizations**

**Over 18,000 records,  
88,000 files, and 500  
comments**





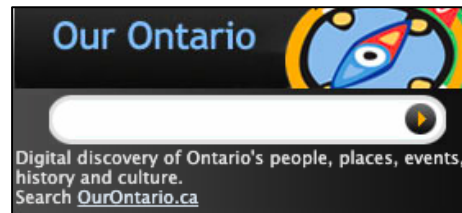
# Fostering an ecosystem for digitization





# Discovery as part of digitization

- Discovery is local/thematic/regional/provincial
- Discoverable, online access to millions of digital assets and community stories
- Largest online public archive of Ontario newspapers, from 1810-2018:
  - [Ink.ourontario.ca](http://Ink.ourontario.ca)
  - [news.ourontario.ca](http://news.ourontario.ca)



# Newspaper digitization is a priority

- a footprint for nearly every sports team, community event, church social, business opening/closing, birth, marriage, death, celebration, criminal action, expansion, decline, opinions, gatherings, elections, trends, weather patterns and more...

## National Heritage Digitization Strategy Survey

“What published  
material most  
urgently needs  
digitization in your  
community?”

NEWSPAPERS



LAST COPIES



RARE BOOKS



GOVT PUBS







# Newspaper Projects

- Bird's eye perspective
- Expertise
- Copyright
- Advocacy

For each project:

- Identify best copies
- Shepherd the digitization process
- Migrate & clean index records
- Process the digital files: OCR, issue recognition, etc.
- Ingest to toolkit, map data (index or article to newspaper, etc.)
- Custom display site



# Clarington Digital Newspaper Collection

Clarington Public Library • Clarington Museums and Archives



Search (Advanced) >> What's New >> Details >> 1 << Previous Record Next Record >>

Search Site:

Canadian Statesman (Bowmanville, ON), 29 Mar 2006, p. 1

Pages:   Search within this title:   PDF version of the page

<< Previous Issue Current Issue Next Issue >> [Browse Issues](#) [More about this Newspaper](#)

[small \(640 x 480\)](#) [default \(1000 x 800\)](#) [large \(1600 x 900\)](#) [extra large \(2400 x 1200\)](#)

# The Canadian Statesman

durhamregion.com ♦ Pressrun 24,150 ♦ Founded 1854 ♦ MARCH 29, 2006 ♦ 60 Pages ♦ Optional 3 Week Delivery \$5/\$1 Newsstand



## Farmers feeling snubbed by province





# OCR

```

1 look user?
2 'delia-15110101'IT.E. House or MD? 'Coun?1110101'
3 '11111' 'delia-15110101'
4 'Mr. Zimmerman, can I find out? 2 - 11111, Mr.
5 Zimmerman? From the report, I can find out?
6 the use of the word 'delia-15110101' as a
7 'delia-15110101' as a
8 'delia-15110101' as a
9 'delia-15110101' as a
10 'delia-15110101' as a
11 'delia-15110101' as a
12 'delia-15110101' as a
13 'delia-15110101' as a
14 'delia-15110101' as a
15 'delia-15110101' as a
16 'delia-15110101' as a
17 'delia-15110101' as a
18 'delia-15110101' as a
19 'delia-15110101' as a
20 'delia-15110101' as a
21 'delia-15110101' as a
22 'delia-15110101' as a
23 'delia-15110101' as a
24 'delia-15110101' as a
25 'delia-15110101' as a
26 'delia-15110101' as a
27 'delia-15110101' as a
28 'delia-15110101' as a
29 'delia-15110101' as a
30 'delia-15110101' as a
31 'delia-15110101' as a
32 'delia-15110101' as a
33 'delia-15110101' as a
34 'delia-15110101' as a
35 'delia-15110101' as a
36 'delia-15110101' as a
37 'delia-15110101' as a
38 'delia-15110101' as a
39 'delia-15110101' as a
40 'delia-15110101' as a
41 'delia-15110101' as a
42 'delia-15110101' as a
43 'delia-15110101' as a
44 'delia-15110101' as a
45 'delia-15110101' as a
46 'delia-15110101' as a
47 'delia-15110101' as a
48 'delia-15110101' as a
49 'delia-15110101' as a
50 'delia-15110101' as a
51 'delia-15110101' as a
52 'delia-15110101' as a
53 'delia-15110101' as a
54 'delia-15110101' as a
55 'delia-15110101' as a
56 'delia-15110101' as a
57 'delia-15110101' as a
58 'delia-15110101' as a
59 'delia-15110101' as a
60 'delia-15110101' as a
61 'delia-15110101' as a
62 'delia-15110101' as a
63 'delia-15110101' as a
64 'delia-15110101' as a
65 'delia-15110101' as a
66 'delia-15110101' as a
67 'delia-15110101' as a
68 'delia-15110101' as a
69 'delia-15110101' as a
70 'delia-15110101' as a
71 'delia-15110101' as a
72 'delia-15110101' as a
73 'delia-15110101' as a
74 'delia-15110101' as a
75 'delia-15110101' as a
76 'delia-15110101' as a
77 'delia-15110101' as a
78 'delia-15110101' as a
79 'delia-15110101' as a
80 'delia-15110101' as a
81 'delia-15110101' as a
82 'delia-15110101' as a
83 'delia-15110101' as a
84 'delia-15110101' as a
85 'delia-15110101' as a
86 'delia-15110101' as a
87 'delia-15110101' as a
88 'delia-15110101' as a
89 'delia-15110101' as a
90 'delia-15110101' as a
91 'delia-15110101' as a
92 'delia-15110101' as a
93 'delia-15110101' as a
94 'delia-15110101' as a
95 'delia-15110101' as a
96 'delia-15110101' as a
97 'delia-15110101' as a
98 'delia-15110101' as a
99 'delia-15110101' as a
100 'delia-15110101' as a


```



# OCR and full-text searching



## Canadian Statesman (Bowmanville, ON), 28 Jan 1926

 Canadian Statesman (Bowmanville, ON), 28 Jan 1926,  
Ance E. Quibell, Toronto

lice Everett Quibell, daugh-  
late W. A. Quibell and Mrs  
passed away Thursday, Jan-  
t., at her home, 49 Robina

Special P

Pages/Parts:

4 5



... Quibell, daugh-  
Sto ter of the late W. A. Quibell and Mrs  
<em>Quibell</em>, passed away ...  
ville, ON), June 17, 1943

sister Mrs.  
art. Deceas-  
url Bowdway

COMING AUCTIONS  
Hoover's Furniture Sale, June 19;  
Quibell's furniture sale, June 26.

Pages/Parts:

1 3 8

# Leveraging

- Part of our ph (and necessit leverage fund work with par (grants prepa
- MHSO projec OCR'ed 15 m languages, in Inuktitut

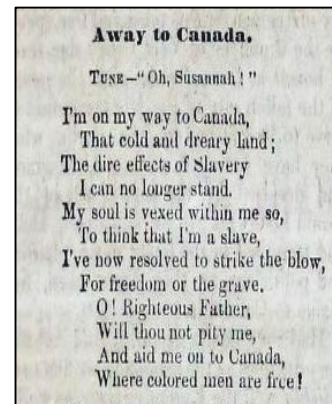
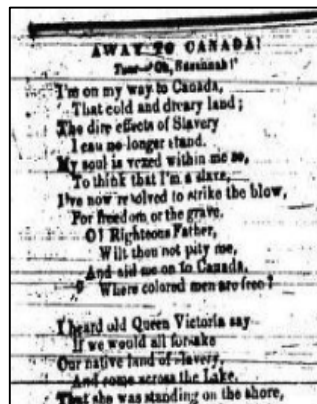
- # Leveraging
- Part of our ph (and necessit leverage fund work with par (grants prepa
  - MHSO projec OCR'ed 15 m languages, in Inuktitut

[illegible]

```
Branch: master ▾ odw-font-training / inuktitut /  
artunit addition of inuktitut files  
..  
font_properties addition of inuktitut files  
iku.inttemp addition of inuktitut files  
iku.itoday.exp0.box addition of inuktitut files  
iku.itoday.exp0.tif addition of inuktitut files
```

# Digitization is evolving (1)

- "Best" original = better OCR quality
- Improvements in OCR (machine learning) have been phenomenal



<http://news.ourontario.ca/abolition/search>

# DIY Digitization

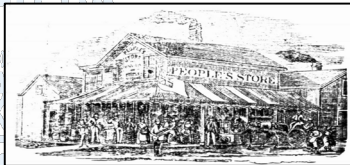
- Most expensive part of “digitizing” newspapers is digitization because of the sheer volume
- Experimenting with DIY macro-photography scanner
- Up to 2 reels a day (~2000 newspaper pages)

***Newspapers are the richest source  
of social science data***



# Re-digitization: Scanning service vs. macro-photography

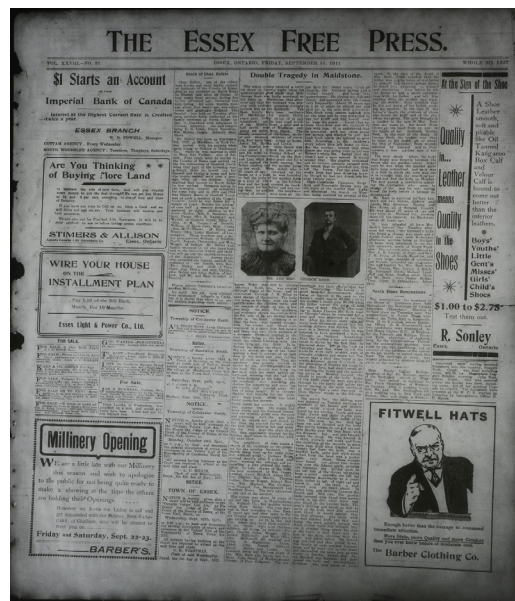
FALSE RUMOR.—Telegrams last week informed us that Old London was in excitement over a startling rumor to the effect that Queen Victoria had died suddenly at Balmoral. A later report relieved the public anxiety by announcing that the news was false. The excitement occasioned during the short period that elapsed between the publication of the first report and the contradiction showed how deeply the news of Her Majestys' death would affect her subjects in all parts of the world, and the feeling of relief exhibited on receipt of the contradiction of the painful rumor testified the respect and admiration felt for the Sovereign by all classes of her subjects at home and abroad. It is the sincere wish of her people that many years may elapse before they will be called upon to mourn her decease.



FALSE RUMOR.—Telegrams last week informed us that Old London was in excitement over a startling rumor to the effect that Queen Victoria had died suddenly at Balmoral. A later report relieved the public anxiety by announcing that the news was false. The excitement occasioned during the short period that elapsed between the publication of the first report and the contradiction showed how deeply the news of Her Majestys' death would affect her subjects in all parts of the world, and the feeling of relief exhibited on receipt of the contradiction of the painful rumor testified the respect and admiration felt for the Sovereign by all classes of her subjects at home and abroad. It is the sincere wish of her people that many years may elapse before they will be called upon to mourn her decease.



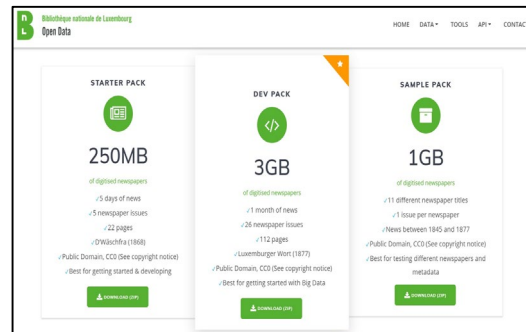
# Re-digitization





# Challenges and opportunities ahead

- Interoperable data sets for newspapers - (we are working on this)
- Machine learning and AI for newspapers and digital collections
- Privacy and best practices more critical than ever
- Open, trusted datasets for researchers
- Metrics around use, social and economic benefits
- **Climate change!**



# Collaboration and Partnerships



OurDigitalWorld is proud to be part of a generous community. We thank our sponsors for their ongoing support and advocacy.



University  
of Windsor



.. ontario library association



Hamilton  
Public  
Library




ORION



OLRC





Digital collections are *public* and *social assets* -  
connecting people, connecting research and fostering  
relationships around the digital assets.

**Merci!**

**Loren Fantin**

**ourdigitalworld.net**

**lfantin@ourdigitalworld.org**

---