



Canadian Research
Knowledge Network

Réseau canadien
de documentation
pour la recherche

Searchable Handwriting: An AI-Powered ICR Pilot Project with Transkribus Software

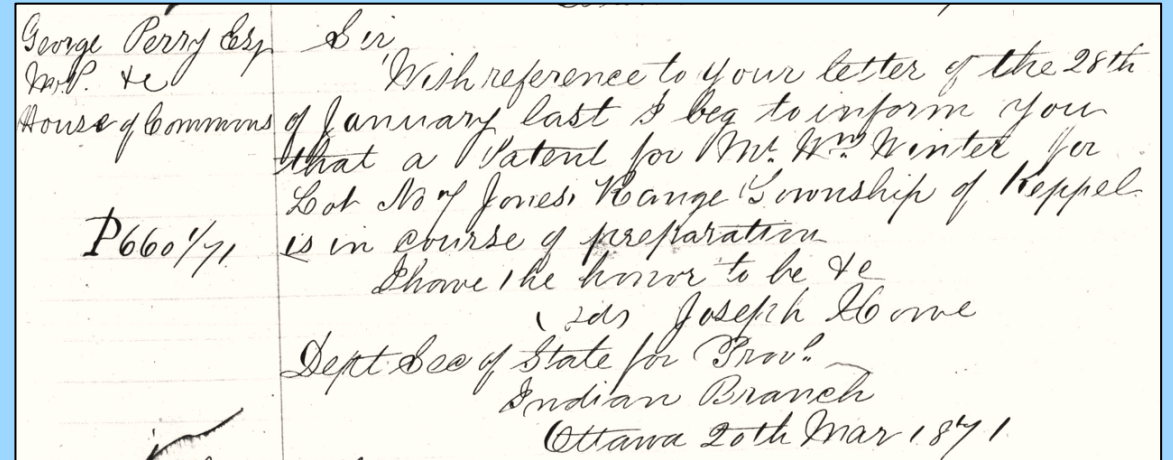
Jason Friedman, Senior Manager, Heritage Services

crkn-rcdr.ca

Discovery Challenges

Digitized microfilm

- Limited finding aids/indexes
- Limited metadata
- Limited search capabilities
- Lots of handwriting means no OCR, no full text search

A handwritten letter on a piece of paper, likely a microfilm frame. The text is written in cursive and is somewhat difficult to read due to the handwriting. The letter is addressed to George Perry Esq. and discusses a patent for Mr. Wm. Winter for Lot No. of Jones Range Township of Keppel. The letter is dated Ottawa 20th Mar 1871 and is signed by Joseph Howe, Esq., Indian Branch, Dept. Sec of State for Prov.

George Perry Esq. Sir
Wm. W. Winter
House of Commons
P660/71
With reference to your letter of the 28th
of January last I beg to inform you
that a Patent for Mr. Wm. Winter for
Lot No. of Jones Range Township of Keppel
is in course of preparation
I have the honor to be &c
(sd) Joseph Howe
Dept. Sec of State for Prov.
Indian Branch
Ottawa 20th Mar 1871

Solution: Intelligent Character Recognition (ICR)



Uses AI to perform character recognition



Can be used on handwritten materials



Environmental scan conducted in Fall 2022

Solution: Transkribus



Produced by READ-COOP



Used by researchers/historians working with parts of the Héritage collection



Initial sample sent for testing

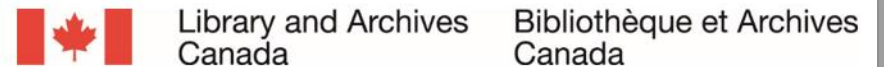
1-2% error rate for typewritten materials
5-7% error rate for handwritten materials



Next step: pilot project

Implementing: ICR Pilot Project

- Partnership with LAC
- Full scope: **2 million** images
- Began in March 2023
- Initial batch of **600K+**
- Additional **500K** identified
- **~1 million** images to be identified



First Batch: Departmental Letterbooks

All are handwritten with relatively consistent writing and format



Cover a breadth of content relevant to a variety of research topics



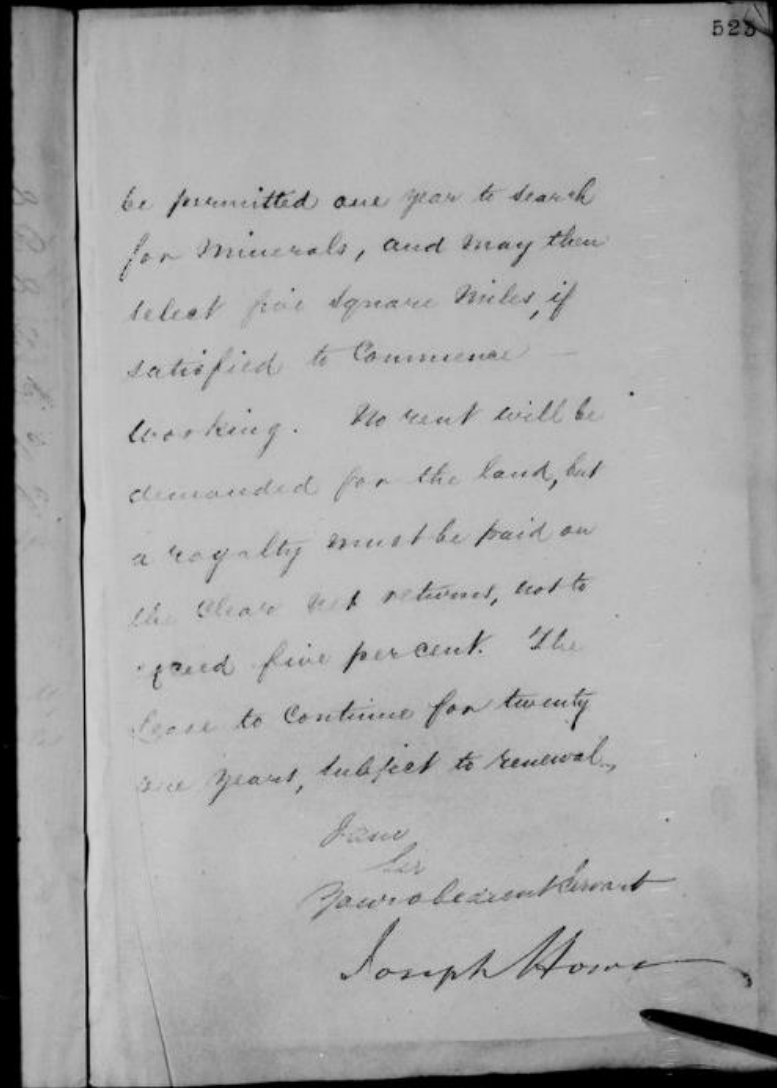
Could benefit from ICR as they are poorly described



Time period of 1871-1929 (low risk of sensitive content)



Initial results: out of the box



Dept. of Indian & Northern Affairs, Letterbook, 21
November 1871 - 6 March 1872, (R.G. 30, Volume 4397)

PUBLIC ARCHIVES
ARCHIVES PUBLIQUES
CANADA



be permitted one year to search
for minerals, and may then
select five Square miles, if
satisfied to commence
working. No rent will be
demanded for the land, but
a royalty must be paid on
the clear net returns, not to
exceed five percent. The
Lease to continue for twenty
one years subject to renewal.

I am
Sir
Your obedient Servant
Joseph Howe
PUBLIC ARCHIVES
ARCHIVES DUBLIQUES
CANADA

Pook LOPY
a 523

Initial results:
comparison of
two models

Poor Copy

98

beyond clearing 4 acres on
the part of Section No. 8 in
the Township of Penefather
purchased for Settlement
by Mr R. Bannon in the
employ of the Messrs Green,
and the erection of the
bodies of two houses and
timber prepared for building
a Dock on wharf, the
renewal will be made
upon the express under-
standing that no transfer



beyond clearing 4 acres on
the part of Section No. 8 in
the Township of Penefather
purchased for settlement
by Mr. R. Bannon in the
employ of the Messrs Green
and the erection of the
bodies of two houses and
timber prepared for building
a Dock on wharf, the
renewal will be made
upon the express under-
standing that no transfer

November 10.
PUBLIC ARCHIVES
ARCHIVES DUBLIQUES
CANADA

Poor COPY
94

beyond cleared to acres on
the part of Section of in
the Township of Pether
purchased for Settlement
by Mr R. Banon in the
employ of the Messrs Creen
and the creation of the
bodies of two houses and
timber prepared for building
a Dock or Wharf the
renewal will be made
upon the express under-
standing that no transfer

I.
PUBLIC ARCHIVES
ARCHIVES PUBLIQUES
CANADA

POOR COPY
8

Dept. of Indian & Northern Affairs, Letterbook, 27
November 1871 - 6 March 1872, (R.G. 10, Volume 4377)

PUBLIC ARCHIVES
ARCHIVES PUBLIQUES
CANADA



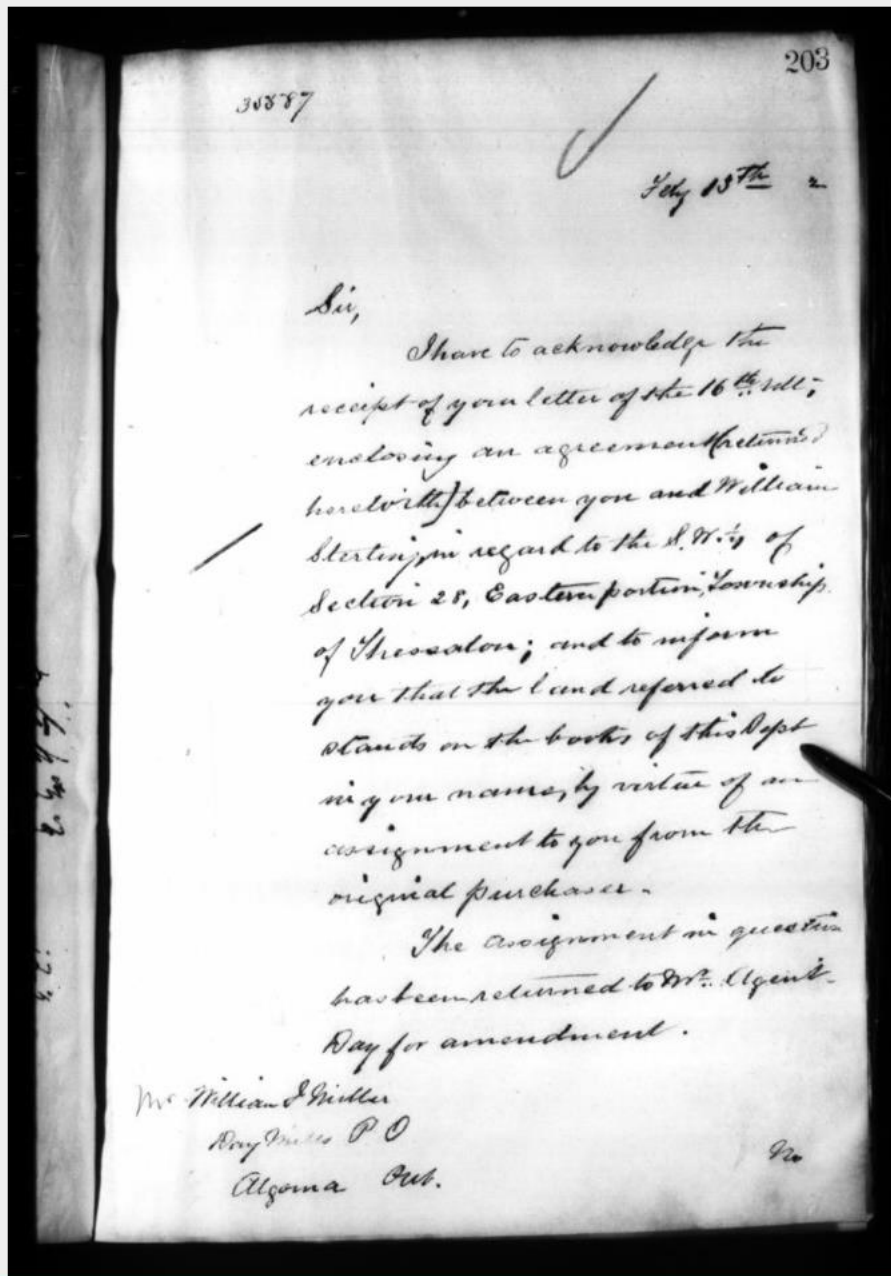
Search within this reel:

Algoma

Search

Matching images: **Image 599**

(3.8G)



29703

14th March

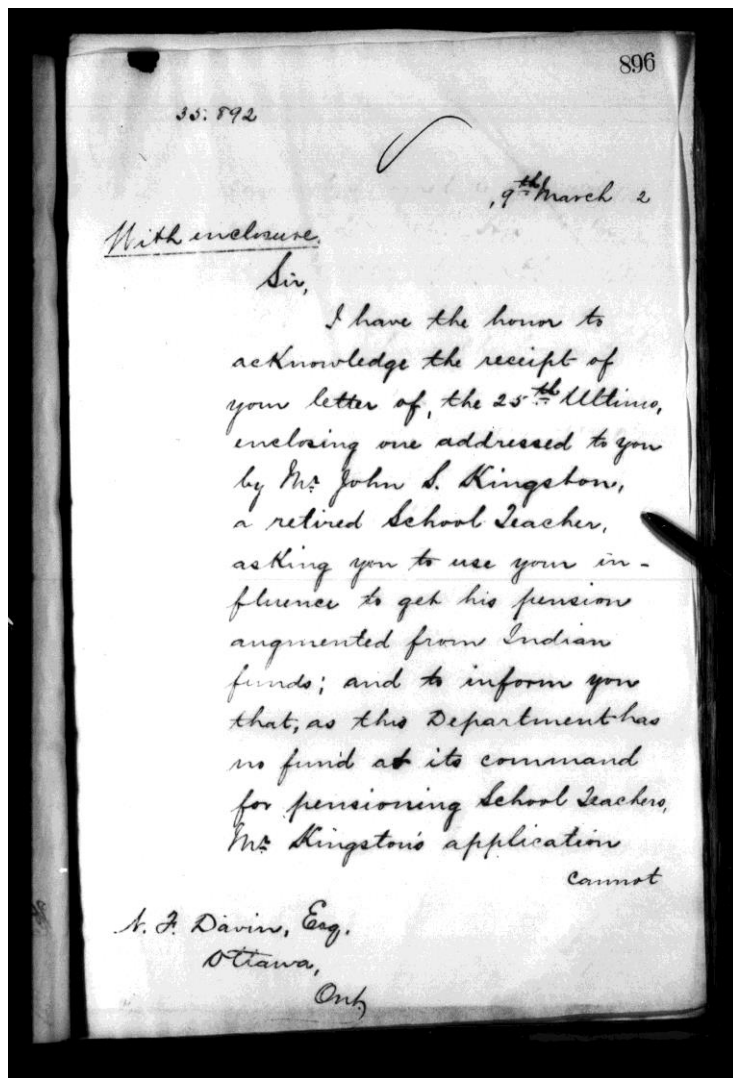
Sir,

In reply to your letter of the 22nd ultimo, I have to inform you that, should the name of Mr. Duchesnois be added to the Pay list of the Lake Nipissing Band as proposed by you, the Department fears an inconvenient precedent would be established.

I am,

Sir,

Copy/Paste

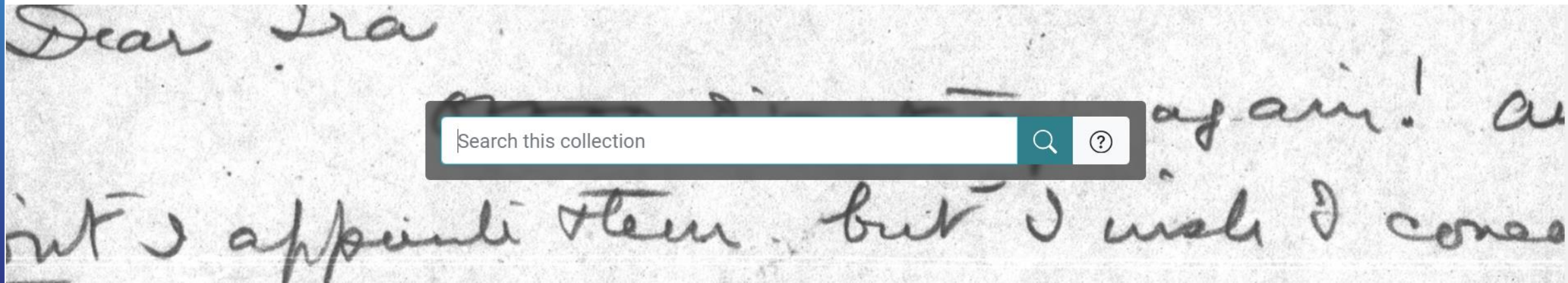


9th March 2 With enclosure. Sir, I have the honor to acknowledge the receipt of your letter of the 25th Ultimo, enclosing one addressed to you by Mr John S. Kingston, a retired School Teacher, asking you to use your influence to get his pension augmented from Indian funds; and to inform you that, as this Department has no fund at its command for pensioning School Teachers, Mr Kingston's application cannot A. F. Davin, Esqr. Ottawa Ont

Early Results

“In the last few weeks, I’ve been noticing more keyword hits for Letterbook volumes in recent Héritage searches, including a couple new finds that were super helpful for developing specific claims for First Nations.”

– *Mark Werner, First Peoples Law*



Current Status/Next Steps

- Over **300K** transcribed images online
- **400K** with READ-COOP
- **500K** in process of being exported
- **1 million** additional to be identified

Department of Indian Affairs : Departmental letterbooks : C-8600
① 1326 images, 1325 with full-text search
Matching images: [Image 1051](#) [Image 1052](#)

Indian and Inuit Affairs Program : Modified duplex numeric system : C-7992
① 2559 images, 2557 with full-text search
Matching images: [Image 1799](#) [Image 2440](#) [Image 2502](#) [Image 2513](#)

Department of Indian Affairs : Departmental letterbooks : C-7515
① 1457 images with full-text search
Matching images: [Image 558](#) [Image 615](#) [Image 780](#) [Image 781](#) [Image 1219](#)

Department of Indian Affairs : Departmental letterbooks : C-7507
① 1180 images with full-text search
Matching images: [Image 277](#) [Image 999](#)

Department of Indian Affairs : Departmental letterbooks : C-8601
① 1310 images, 1308 with full-text search
Matching images: [Image 228](#)

Indian and Inuit Affairs Program : Modified duplex numeric system : C-8000
① 2618 images, 2613 with full-text search
Matching images: [Image 1662](#)

1 2 3 ... 6 Next

Let's talk ICR!

- AIMLComm Mattermost
- Intelligent Character Recognition channel



A decorative vertical element on the left side of the slide, consisting of several parallel white lines of varying thicknesses and lengths, some with rounded ends, creating a modern, abstract border.

Thank you!!!

jfriedman@crkn.ca