Digitization of the First World War
Canadian Expeditionary Force (CEF) files

Presented by: Sophie Dazé
A/ Director, Digitization Services Division

October 28th 2017
Who we are

Library and Archives Canada Digitization Services Division

- Gatineau Preservation Centre
- Division in operation since 1999
- Grown from a team of 2 to over 80
- 2016-2017 over 12 million images were digitized

Gatineau Preservation Centre located in Gatineau, Québec
Canadian Expeditionary Forces (CEF) Digitization Project

- CEF Collection is comprised of approximately 640,000 Personnel files
- Goal is to digitize and make them all available online by November 2018
- To facilitate the effort, we repurposed our high-speed auto-fed BancTec scanners.
Year One

- Pre-digitization project / Fastener and adhesive removal
- Outsource digitization trial
- Testing to establish in-house workflow and productivity to bring the project back in-house
- Operationalizing
  - Testing / Concept (3 months)
  - Network / Efficiencies
  - Hiring
- 6 months of production during year one
Conservation concerns

- Conservators had many concerns about using the BancTec for 100 year old files
- Worked with the company to mitigate these issues
The paper jogger was a major concern to conservators that its extended use would damage the material.
Step 1: Sorting

First and most important step is sorting the files to go on the various equipment. We use barcodes to track the material.
Step 1: Sorting (Cont’d)
Step 2: Scanning – BancTec

We can scan up to 30,000 images a day with the auto-feed function.

Two people scan up to 6,000 images a day with the hand-fed process.
Step 2: Other scanners

- **Fujitsu scanners** were initially used to digitize the more delicate material, which we now do on the BancTec (hand-fed process).
- It continues to be used in overflow situations.
- It also is used to digitize material that is in very damaged condition.

- **Contex scanners** are mostly used to scan oversized documents.
Step 3: Quality Control

- Technicians assigned to perform Quality Control functions look out for: cropping, colour, density, page orientation, page order contents and file naming. If any issues are found the technician is required to rescan the document.
All the documents are scanned at 300 PPI uncompressed TIFF 24 bit sRGB colour

PDF files are saved using Adobe 1.7 spec w/ JPEG2000 compression.

BancTec scanners and the PDF compression workstations have a 10 Gigabit fibre connection

Archiving is performed by an IBM tape library using Comm Vault software to LTO6 tapes
Successes & Challenges

**Successes**
First archival institution using the BancTec scanners for archival digitization
National and international institutions looking for your guidance
2016-2017 7.2 millions images - 200000 over our projected goal
Currently outputting 600,000+ images a month
Creating the BancTec hand feed technique which has doubled our fragile document scanning capabilities
As of today, Over 80% completed with approx. 8M images remaining to digitize within the next year

**Challenges**
Sorting techniques as it is one of the most important step for the project it drives our production
Properly qualified people for sorting Attention to detail is a key asset, finding the right employees can be challenging
IT infrastructure… storage… network speed…
Celebrity soldiers

- **Lieutenant Lester Bowles Pearson**
  14th Prime Minister of Canada (1963-1968)

- **Lieutenant William Avery Bishop (AKA Billy Bishop)**
  Flying Ace
  Multiple award recipient

- **Lieutenant Colonel John McCrae**
  Canadian poet, physician, author, artist and soldier during First World War
  Author of the *In Flanders Field* poem
In Flanders fields the poppies grow
Between the crosses, row on row,
That mark our place; and in the sky
The larks, still bravely singing, fly
Scarce heard amid the guns below.

We are the Dead. Short days ago
We lived, felt dawn, saw sunset glow,
Loved and were loved, and now we lie
In Flanders fields.

Take up our quarrel with the foe:
To you from failing hands we throw
The torch; be yours to hold it high.
If ye break faith with us who die
We shall not sleep, though poppies grow
In Flanders fields.