Large-scale Image Digitization: the Milwaukee Polonia Project

Michael Doylen, UW-Milwaukee
Ann Hanlon, UW-Milwaukee

WiLSWorld July 2013
Large-scale digitization

2010 OCLC survey of special collections and archives defines large-scale digitization for special collections this way: “…a systematic effort to consider complete collections – rather than be selective at the item level, as has been the norm for many projects – and using production methods that are as streamlined as possible while also accounting for the needs of special materials.”
Reception area of Kwasniewski’s Park Studio, 1917
Young newly wed couple, family, and friends, 1921
First Holy Communion, 1923

First Holy Communion photographs were a large part of Kwaśniewski’s business. He photographed these young ladies from the Josef Wozniak family in 1923.
Mitchell Street, 1920s

Mitchell Street was the major commercial district for the Polish community on Milwaukee’s South Side.
Kosciuszko Reds, circa 1912

The Kosciuszko Reds played on Milwaukee’s South Side. This photograph was taken on a day when the Reds took on the Chicago American Giants, a Negro League Team.
Children of St. Joseph’s Orphanage, 1924

Orphanages helped poor families through difficult times by temporarily caring for children. Most of these children had at least one parent and likely returned home after a few years.
Polish dance group, circa 1928

Milwaukee’s Poles, like other ethnic groups, passed their customs and traditions on to the younger generation.
Collection by Formats
Kwasniewski Access Database

- Job number
- Customer name
- Customer address
- Date
- Extent
- Format
- Dimensions
- Physical condition
- Subject headings
Digitization Quotes

- **Outsourcing**
  - Vendor 1: $4.61/item or $121,695 total
  - Vendor 2: $8.16/item or $215,408 total

- **In-sourcing**
  - Vendor 2: $13.78/item or $363,764 total
Digitization Quotes

• Outsourcing
  – Vendor 1: $4.61/item or $121,695 total
  – Vendor 2: $8.16/item or $215,408 total

• Insourcing
  – Vendor 2: $13.78/item or $363,764

• Actual cost, DIY
  – $0.81/item or $32,786
Digitization Quotes

Cost per item

Vendor 2 in-sourced: 13.78
Vendor 2: 8.16
Vendor 1: 4.61
UWM (actual): 0.81
Digitization Method

- Digital camera
  - 3-person team for digitization
  - Rail system for queuing, capturing, re-housing
April 2012 pilot

- Epson 10000XL flatbed scanner
  - 1 image / 2 ¾ minutes (excluding cropping multiples)
  - 2 ½ years

- Digital camera
  - 1 image / 40 seconds
  - 2 months
Digitization Method

Estimated time to digitize 27,000 glass plates

- Flatbed scanner: 2 1/2 years
- Digital camera: 2 months
The Kwasniewski Photographs at the UW-Milwaukee Libraries
Digitization Method

- Digital camera
  - Set up structure to support plates
  - Staging area for unwrapping and re-wrapping plates
The rail system prototype and pilot
Rail system final product

Updated rail system with plastic high-density rails; transparent rail over light-box; additional monitor for quality-control
Rail system final product

- Rails demo
Challenges: maintaining workflow

Rail digitization in action, December 2012

Note three-person “assembly-line”
Challenges

• Success and its consequences:
  – Maintaining sufficient storage for file production and processing
Existing database description re-purposed to create basic, individual item descriptions for each image.
Metadata and Description

- Religion—altar boys—family/group
- Urban Scenes—transportation—insurance photos of automobile wrecks and injuries
- Sports—baseball
# Metadata and Description

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
</tr>
<tr>
<td><strong>Subject TGM</strong></td>
</tr>
<tr>
<td><strong>Subject Kwasniewski</strong></td>
</tr>
<tr>
<td><strong>Topic Code</strong></td>
</tr>
<tr>
<td><strong>Photographer</strong></td>
</tr>
<tr>
<td><strong>Photographer's Note</strong></td>
</tr>
<tr>
<td><strong>Digital Publisher</strong></td>
</tr>
<tr>
<td><strong>Date of Photograph</strong></td>
</tr>
<tr>
<td><strong>Address</strong></td>
</tr>
<tr>
<td><strong>Original Item Medium</strong></td>
</tr>
<tr>
<td><strong>Original Item Size</strong></td>
</tr>
<tr>
<td><strong>Original Item Location</strong></td>
</tr>
<tr>
<td><strong>Original Collection</strong></td>
</tr>
<tr>
<td><strong>Repository</strong></td>
</tr>
<tr>
<td><strong>Digital ID</strong></td>
</tr>
<tr>
<td><strong>Digital Collection</strong></td>
</tr>
<tr>
<td><strong>Serial Number</strong></td>
</tr>
<tr>
<td><strong>Job Number</strong></td>
</tr>
</tbody>
</table>
**Metadata and Description**

- 35,000 entries / records
- Digital ID added to existing spreadsheet
- Separate batch upload spreadsheets prepared
- Most updates done at batch level
Enhanced metadata

80% of the collection will have basic image processing and metadata.

20% will be designated for enhanced image processing and metadata.
Enhanced metadata

80% of the collection will have basic image processing and metadata

20% “enhanced” categories

Activities and Organizations (1,104 images)
Business and Community (210)
Education (1,990)
Kwasniewski family (734)
Military (273)
Residential (210)
Sports (190)
Urban Scenes (1,356)
Timeline

April 2012: Rails and Metadata pilots

July 2012: Hire half-time project manager

August-September 2012: Project manager hires and trains students for rail digitization

September 2012-March 2013: Digitization of 5x7s

March 2013-September 2013: Digitization of remaining formats

December 2013: Online collection launches
Promotion:

• www.historypin.com

• Social-media-oriented, historic photograph sharing site

• “Pin” to map and Google street views

• Allows community interaction – comments and feedback, addition of community images, etc.
Promotion:
Promotion: historypin

[Image of a website page with a map and photos of buildings, showcasing historical content.]
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>15546</td>
<td>011305</td>
<td>15556</td>
<td>137</td>
<td>15281</td>
<td>Newman, Mrs. A.</td>
<td>03/21</td>
<td>1823</td>
<td>1923</td>
<td>1886</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>15546</td>
<td>011305</td>
<td>15556</td>
<td>137</td>
<td>15281</td>
<td>Newman, Mrs. A.</td>
<td>03/21</td>
<td>1823</td>
<td>1923</td>
<td>1886</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>15546</td>
<td>011305</td>
<td>15556</td>
<td>137</td>
<td>15281</td>
<td>Newman, Mrs. A.</td>
<td>03/21</td>
<td>1823</td>
<td>1923</td>
<td>1886</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28613</td>
<td>020501</td>
<td>28613</td>
<td>137</td>
<td>15460</td>
<td>Schütz, J.</td>
<td>10/27</td>
<td>1893</td>
<td>1923</td>
<td>1851</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>28613</td>
<td>020501</td>
<td>28613</td>
<td>137</td>
<td>15460</td>
<td>Schütz, J.</td>
<td>10/27</td>
<td>1893</td>
<td>1923</td>
<td>1851</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>28613</td>
<td>020501</td>
<td>28613</td>
<td>137</td>
<td>15460</td>
<td>Schütz, J.</td>
<td>10/27</td>
<td>1893</td>
<td>1923</td>
<td>1851</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28613</td>
<td>020501</td>
<td>28613</td>
<td>137</td>
<td>15460</td>
<td>Schütz, J.</td>
<td>10/27</td>
<td>1893</td>
<td>1923</td>
<td>1851</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>28613</td>
<td>020501</td>
<td>28613</td>
<td>137</td>
<td>15460</td>
<td>Schütz, J.</td>
<td>10/27</td>
<td>1893</td>
<td>1923</td>
<td>1851</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>28613</td>
<td>020501</td>
<td>28613</td>
<td>137</td>
<td>15460</td>
<td>Schütz, J.</td>
<td>10/27</td>
<td>1893</td>
<td>1923</td>
<td>1851</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28613</td>
<td>020501</td>
<td>28613</td>
<td>137</td>
<td>15460</td>
<td>Schütz, J.</td>
<td>10/27</td>
<td>1893</td>
<td>1923</td>
<td>1851</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>28613</td>
<td>020501</td>
<td>28613</td>
<td>137</td>
<td>15460</td>
<td>Schütz, J.</td>
<td>10/27</td>
<td>1893</td>
<td>1923</td>
<td>1851</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>28613</td>
<td>020501</td>
<td>28613</td>
<td>137</td>
<td>15460</td>
<td>Schütz, J.</td>
<td>10/27</td>
<td>1893</td>
<td>1923</td>
<td>1851</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Next: Using the collection

Now what?
THANK YOU!

- Ann Hanlon: hanlon@uwm.edu
- Michael Doylen: doylenm@uwm.edu