

Data Analysis



Jaime Haire

jaime.haire@mnhs.org

Great Rivers Cultural Heritage Network
South Dakota Partner Meeting
August 6, 2007



Introduction

- Metadata
- PastPerfect/ContentDM
- Example: Nicollet County
- Collections Survey



Data Standards

- Metadata Standards
 - Structure, Value, Format
- Existing Standards
 - MN Digital Library
 - Collaborative Digitization Program



Data Standards

- Perfect metadata vs. “reasonable” metadata
- Data enhancement – globally
- User-created content will add value



PastPerfect

- Wide spread implementation
- Common data structure
- Embedded vocabularies, authorities
- Learn one, learn all?
- Consistent usage?



PastPerfect Collections Survey

Nicollet County

- Learn more about collections in PastPerfect
 - Software version? Installation date?
 - Imported Legacy Data?
 - Types of Material?
 - How many records? % of collection in PP?
 - Guidelines for data input?



PastPerfect Sample Dataset

- Export to Excel
 - Photos
 - ~10,000 records, 141 fields
 - Objects
 - ~ 1,600 records, 219 fields



PastPerfect Data Fields

Photos – 140 fields!

- Accessions, Acquisitions
- Value, Insurance
- Loans
- Exhibits
- Location, Inventory
- Copyright, Permissions
- Condition, Maintenance
- Use of 22 custom fields?






PastPerfect Data Fields

Identify fields for GRCHN search

- Subjects
- Photographer/Studio
- People
- Place
- Notes
- Description



Data Mapping

- Fields relevant to GRCHN search?
 - People 
 - Photographer/Studio
 - People
 - Description? Subject?
 - Places 
 - Place
 - Description? Subject?
 - Keyword 
 - All of the above + more



Preliminary Data Analysis

- Review content for consistency
- Evaluate clean-up potential
- Address now with global updates?
- Address going forward with standards?



Data Evaluation

Spelling

- Minnneapolis
- Mineapolis
- Minneapulos
- Minneapollis
- Minneapolis



Data Evaluation

Abbreviations

- Saint Paul (40 hits)
- St. Paul (100 hits)
- St Paul (5 hits)
- St.Paul (1 hit)



Data Evaluation

- Name: Joseph Nicollet
- Name: Nicollet, Joseph
- First name: Joseph
- Last name: Nicollet



Mixed Content

- Description field:
 - B/W digital image of “skit depicting the early years of Nicollet, Jerry Benson, left, sells a peddler’s license to Elmer Boeke before he can peddle his wares;” for the 1970 Nicollet Centennial, July 16-19, 1970.
- People Search results?



Data Evaluation

Control

- automobile vs. car
- street vs. road
- home vs. house



Levels of Descriptions

- celebrations vs. weddings
- livestock vs. cows
- kitchen vs. stove



Data Evaluation

Context

- Format type: Photograph
- Collection Name
- Missing objects



Collections Survey

- Learn more about collections
 - PastPerfect, ContentDM
 - Indexes
 - Vital Records
 - Census
 - Cemetery
 - Other digital content