User Input into Image Retrieval Design

Sam Hastings
University of North Texas
School of Library and Information Sciences
The Sepia Web Survey is a project designed to gather information from the general public about a specific image collection. Digital images from the Sepia Photo Archive are presented to web users for the purpose of collecting personal knowledge about the photograph subjects and times and possible key terms.
Why?

- To share the images with the web community
- For publicity
- To identify the images or context for the images
- To contribute key terms for indexing
Denton Public Library

Photo Scanning Station

- About the Project
- Why scan?
- How to Scan Your Photos
- Hours/Directions
- FAQ

UNIVERSITY OF NORTH TEXAS
African American Museum

Sepia Thumbnails

Browse next page of Sepia thumbnails

Page 1 2 3 4 5
Technology has placed the emphasis on automatic image indexing and content-based retrieval—but how does the retrieval functionality found in these systems correlate with image information needs of real users?

User input can provide in-depth understanding of the user's information needs and his/her cognitive abilities. The understanding can be applied to design better user system functions.
Users may find images more quickly when other people have described the images.

Describing a large amount of images is very labor intensive.

Automated description procedures are only successful for narrow, well described domains.

Thus, the descriptions of images are useful when the person who wants to find an image helps to describe the image thesaurus.
Indexing vs. Users

Intellectual Property
- Image
- Museum
- Indexed
- Technology
  - Debate & Confusion!

What Usually Happens

Online Survey Concept

Digital Image data

People
Real people exist in social contexts, typically far removed from the technology and thesaurus parameters and specific Museum Communities.

To be relevant, our indexing thesaurus must be what people actually use, not the ones we imagine or wish they would use.

There is good indexing via controlled classification vocabulary from the Museum Community, but is this as effective for users as translating user descriptions into vocabulary for optimized retrieval?
The survey consists of a series of images which are displayed on a web survey page.

Participants are asked to choose four key terms to describe the subject content of an image.

Keywords can also be used to describe what is implied or symbolized by the image, including emotions and concepts.

Additional comments can be made in the comments box.

The information will feed into the development and structuring of the thesaurus in the Sepia database and in approaches to providing access to the images and image-based resources.

Sepia_Survey.htm
Queries hoped for…

- Direct Query on Descriptions from Users
  One way to abstract images is to describe them with words. We need user supplied terms to build these abstractions or descriptions or annotations.

- Browsing
  Using themes, categories or words described by users supported by a controlled vocabulary.
THANK YOU!

- All images from the Sepia Photo Archive are property of the African American Museum Fairpark Dallas
  http://www.aam.org
- Sepia Survey -
  http://digitallab.lis.unt.edu/research/sephia_survey.htm
- Questions? Hastings@unt.edu