Into the Future: CLiMB-2

Marilyn Domas White
University of Maryland

CNI Meeting, April 2005
Washington, DC
Status as of April 2005

• Assembled CLiMB-2 team at UMD.
• Met with internal partners at UMD.
• Met with potential external partners.
• Established goals and objectives.

• Selecting and negotiating with partners.
• Writing prospectus for Mellon.
Goals of CLiMB-2

• Improving the cataloging toolkit, especially to extract terms from unstructured text.

• Testing its applicability in different kinds of situations: images, texts.

• Testing its use in a normal cataloging situation.
Goals of CLiMB-2

• Evaluating its impact on image cataloging.

• Evaluating its impact on actual searching by image searchers.
Research in CLiMB-2

• Effectiveness of CLiMB toolkit for extracting metadata.

• Impact on image cataloging practice.

• Interaction with thesauri for art.

• Impact on retrieval.
Research Questions

• How can the CLiMB toolkit mine structured and unstructured texts more effectively for metadata terms? Issues include:
  Reference (different types of anaphora, both explicit and implicit)
  Topic tracking
  Comparative/constructive constructions
  Narrative structure
  Machine-learning techniques
Research Questions

• In terms of rhetorical and discourse structure, how do we characterize text that is most effective for yielding metadata terms?

• What procedures can enhance mapping extracted terms to controlled vocabularies?
Research Questions

• How does the CLiMB toolkit affect existing work practices for image catalogers?

• How can CLiMB generated terms be channeled to revise or update existing thesauri?
Research Questions

• How do CLiMB generated terms affect searching for images by:
  Image professionals, including art historians

Others wanting to use images:
  Anthropologists
  Historians
  Urban planners
Building CLiMB Partnerships

- Image collections
- Catalogers and expertise
- Users
- Texts
Tasks for Partners

• Provide metadata for image collection.

• Provide access to supporting texts, preferably in electronic form.

• Use the toolkit to generate metadata for the collection.
Tasks for Partners

• Evaluate toolkit for usability and other aspects.

• Facilitate access to clients for search studies.

• Provide advice and guidance on topics relevant to CLiMB-2.
Scenario 1

- Partner has images, catalogers, image system, some electronic texts.
- Partner creates electronic versions of other texts.
- CLiMB provides toolkit and trains partner’s catalogers in its use.
- CLiMB observes use in normal cataloging situation.
- CLiMB and partner conduct search studies with partner’s normal client group.
- CLiMB conducts search studies with other user groups accessing partner’s system.
Scenario 2

- Partner provides images and existing metadata for use at CLiMB.
- CLiMB transfers text into electronic format.
- CLiMB uses toolkit and adds CLiMB terms to records.
- Collection is installed at UMD.
- CLiMB tests with UMD faculty and students, possibly other groups.
Scenario 3

- Partner provides images and metadata.
- CLiMB transfers text into electronic format.
- CLiMB trains partner’s CLiMB-funded employee and supervisor in use of toolkit.
- Partner’s employee uses toolkit and adds CLiMB terms to records, under partner’s supervision.
- Partner retains collection at his site.
- CLiMB uses partner’s clients in studies of searching effectiveness.
Benefits to Partners

• Metadata for subject access.
• Perhaps some funds to support use of toolkit.
• Knowledge that partner is involved in research to improve access to images – good for PR.
• Acknowledgment in all CLiMB presentations and publications.
• Participation in CLiMB’s annual meetings with External Board.
• Gratitude of CLiMB team.
CLiMB-2 Project Team

Judith Klavans, co-PI
Marilyn Domas White, co-PI
Jimmy Lin, UMD
Eileen G. Abels, UMD
Dagobert Soergel, UMD
Brooke Rosenblatt, UMD
Angela Giral, CU (ret)
Rebecca Passonneau, CU
CLiMB Websites

CLiMB-1: www.columbia.edu/cu/cria/climb

CLiMB-2: www.umiacs.umd.edu/~climb