CLiMB

Computational Linguistics for Metadata Building

Using Computational Linguistic Techniques to Harvest Image Descriptors
CLiMB Briefing

• Overview of CLiMB Goals and Past Progress
  Judith L. Klavans
• Moving from CLiMB-1 at Columbia University to
  CLiMB-2 at University of Maryland: Research
  Issues and Partnerships
  Marilyn Domas White
• New Collections and Directions
  Angela Giral
The CLiMB Project

• What are the problems that CLiMB addresses?
• What are the potential CLiMB solutions?
• How did CLiMB come about?
• Who have been CLiMB participants?
• What are the next steps and goals?
Problems in Image Access

- Cataloging digital images

- Traditional approach:
  
  Manual expertise
  
  - labor intensive
  
  - expensive

- Can automated techniques help?
Can we harvest image descriptors?
CLiMB Technical Contribution

CLiMB will identify and extract
• proper nouns
• terms and phrases

from text related to an image:

By September 14, 1908, the basis of the Greene's final design had been worked out. It featured a radically informal, **V-shaped plan** (that maintained the original **angled porch**) and interior volumes of various heights, all under a constantly changing **roofline** that echoed the rise and fall of the mountains behind it. The **chimneys** and **foundation** would be constructed of the **sandstone boulders** that comprised the local geology, and the **exterior** of the house would be sheathed in **stained split-redwood shakes**.

CLiMB Overall Goals

The essence of CLiMB:
- Use scholars themselves as “catalogers” by utilizing scholarly publications
- Enhance existing descriptive metadata
- Help cataloger with subject access
- Improve retrieval of images

Outcomes:
- **Research**: Impact on retrieval through increased numbers of descriptors
- **Practice**: Suite of CLiMB tools
- **Resources**: Vocabulary list(s) that can be used by other visual resource professionals
CLiMB History

• CLiMB-1  2002-2004 Columbia University
  – How did it start?
  – What were goals?
  – What were achievements?
  – Next steps

• CLiMB-2 – 2005 University of Maryland
  – Pursuing next steps
  – CLiMB-2 goals and partners - Marilyn White
  – New collections - Angela Giral
How Did CLiMB Start?

- Within the context of interdisciplinary digital library research at Columbia
- Center for Research on Information Access
  - Mission – Link Libraries with DL Technologies
- Search for a libraries-centric project
- Took 2 years to conceptualize
CLiMB Then and Now

• Evolution over time
• Shift from use of
  – diffuse text to
  – more tightly coupled text and image and back to
  – more diffuse relationships between text and image
• Better understanding of criteria
• Clearer idea of possible techniques
CLiMB-1 Project Team

Judith Klavans, PI
Stephen Davis
Angela Giral
Patricia Renfro
Bob Wolven
Robertta Blitz
Rebecca Passonneau
David Elson
CLiMB-1 Achievements

• Created collection selection criteria

• Built and tested prototype CLiMB image cataloging toolkit

• Evaluated use of cataloging toolkit
Key Open Issues

- Processing texts with more diffuse relation between text and image
- Refining CLiMB toolkit
- Utilizing existing vocabularies
- Incorporating into image search platform
- Testing with users, both catalogers and image searchers
Building CLiMB Partnerships

Columbia University

CLiMB Advisory Board

University of Maryland

New Partners
Building Partners

- Collections
  - Museums
  - Libraries
- Users and user groups
  - Catalogers and end users of many types
- Tool builders for digital libraries
- Evaluation, evaluation, evaluation
Next Steps

• Initial project – Ended June 2004

• Planning grant at UMD – January 2005
  – College of Information Studies (CLIS)

• Schedule for 2005:
  – March – Partner meeting
  – April – Coalition for Networked Information (CNI)
  – Submit prospectus May 2005
CLiMB Briefing

• Overview of CLiMB Goals and Past Progress  
  Judith L. Klavans

• Moving from CLiMB-1 at Columbia University to CLiMB-2 at University of Maryland: Research Issues and Partnerships  
  Marilyn White

• New Collections and Directions  
  Angela Giral
Thank you
Please email your comments
jklavans@umd.edu
whitemd@umd.edu
giral@columbia.edu