

## **WANTED: Partner Institutions for CLiMB-2**

### **(Computational Linguistics for Metadata Building)**

---

CLiMB-2 is focusing on the following tasks:

- Improving the image cataloging toolkit, especially to cover extraction of terms from unstructured text.
- Testing its applicability to different types of art from a range of different collections
- Testing its use in a normal cataloging situation.
- Evaluating its impact on actual searching by users.

To accomplish these tasks, CLiMB-2 needs partner institutions. *Note: For background information on CLiMB-1, see the last paragraph.*

#### **What is a partner institution?**

A partner institution can participate in many ways in the CLiMB-2 project, e.g., by providing access to collections, supporting textual materials, staff, and clients. We are seeking several partners for CLiMB-2. We need at least one partner who would be willing to catalog a subset of its collection using the CLiMB image cataloging toolkit and then allow access to it for evaluating the impact of CLiMB-generated terms on actual searching, perhaps by its own clientele. Other partners can do one or preferably more of the following tasks (since some are related).

#### **What kinds of tasks does a partner institution have to do?**

1. Provide electronic cataloging data for an image collection. Note: A collection may be a part of your collection, such as all images of art by Van Gogh or all photographs from the WPA project.
2. Provide legal access to supporting textual materials (if your institution is the publisher) or facilitate/support CLiMB-2's obtaining this permission from publishers. Ideally the text would be in electronic form.
3. Have your image catalogers use the toolkit to index, or augment subject indexing for, your collection.
4. Have your catalogers evaluate the toolkit for usability and other aspects as needed.
5. Allow CLiMB-2 researchers to solicit your clients as participants in studies of searching a collection cataloged with CLiMB-2 subject terms. Supporting this endeavor through institutional access, if necessary.
6. Provide advice and perspectives on significant problems and decisions facing the CLiMB-2 team.

#### **What does a partner institution gain?**

1. Subject cataloging (either by itself if you do not already have that for an image collection) or augmented subject cataloging for a collection. CLiMB 2-generated subject terms can either be added to a central catalog in formats such as MARC, VRA, etc., or can be tailored for your collection to be added to your local catalog.

2. Likely – some funding from the CLiMB-2 project to support subject cataloging if your cataloger(s) use the toolkit for a particular collection.
3. Knowledge that you and your institution are involved in research designed to improve access to images. You can publicize this to your client group and in funding campaigns.
4. Publicity and acknowledgment of your role in professional meetings and other press. CLiMB-2 team members will be reporting on different facets of this project as it proceeds, and your role will always be acknowledged. For example, we will be presenting the CLiMB 2 project at the semi-annual meeting of the Coalition for Networked Information (CNI) in about 3 weeks.
5. Membership on CLiMB-2's Board of External Advisors, which will meet semi-annually in meetings similar to this to discuss progress and other aspects of the project.
6. The gratitude of the CLiMB-2 team.

### **What is CLiMB-1?**

Addressing the pervasive problem of inadequate subject indexing for electronic images, CLiMB-1 suggested the novel solution of mining scholarly texts for metadata terms for images. To accomplish this task, CLiMB-1 developed a toolkit to be used by image catalogers. Based on computational linguistics, the CLiMB Toolkit processed electronic texts to identify potential subject terms, matched the terms with controlled vocabulary in standard thesauri, and allowed catalogers to select terms, either natural language phrases or controlled vocabulary terms, to be added to image cataloging records in several different formats. CLiMB-1 and its toolkit proved the feasibility of this approach for extracting metadata, tested it in several collections and with several texts, and evaluated the usefulness of the CLiMB toolkit for image professionals. CLiMB-1, based at Columbia University, under Judith Klavans as the Principal Investigator, was funded by the Mellon Foundation from 2002 to 2004.

### ***Special note:***

*In all research involving human subjects, CLiMB-2 researchers are subject to UMD's guidelines on human subjects. Regulations usually call for anonymity in reporting results, an individual participant's (not just the institution's) deciding about participating, and his ability to withdraw at any time.*