

Digital library surpasses initial goal of 1 million books

PITTSBURGH, Pennsylvania (AP) — Nearly a decade ago, computer scientists at Carnegie Mellon University embarked on a project with an astonishingly lofty goal: Digitize the published works of humankind and make them freely available online.

The architects of the project said this week they have surpassed their latest target, having scanned more than 1.5 million books — many of them in Chinese — and are continuing to scan thousands more daily.

"Anyone who can get on the Internet now has access to a collection of books the size of a large university library," said Raj Reddy, a computer science and robotics professor at the university who spearheaded the project.

The latest phase in the development of the so-called Universal Library, called the Million Book Project, began in 2002 after Reddy's team successfully scanned 1,000 books.

Much of the recent work has been carried out by workers at scanning centers in India and China, helped by \$3.5 million in seed funding from the National Science Foundation and in-kind contributions from computer hardware and software makers.

The United States, China and India each have contributed \$10 million in cash and contributions to the project, undertaken with partners at China's Zhejiang University, India's Indian Institute of Science and Egypt's Library at Alexandria.

At least half the books are out of copyright or scanned with the permission of copyright holders. Excerpts of copyright-protected works are available, though organizers expect complete texts to become available eventually.

The project is not the first of its kind. Online search engine operator Google Inc. and software giant Microsoft Corp. have begun similar endeavors, though Carnegie Mellon representatives say theirs is the largest university-based digital library of free books and that its purpose is noncommercial.

It's a step toward the creation of an online library that would make traditionally published books available to all, said Reddy. "The economic barriers to the distribution of knowledge are falling," he said in a statement.

Michael Shamos, a Carnegie Mellon computer science professor and copyright lawyer working on the project, said the library's mission included making vast amounts of information freely available and preserving rare and decaying texts, among other things.

Books have been borrowed for scanning from various institutions and individuals worldwide, though institutions in Europe declined to participate, he said.

The library so far has books published in 20 languages, including 970,000 in Chinese, 360,000 in English, 50,000 in the southern Indian language of Telugu and 40,000 in Arabic. ■
