

Frequently Asked Questions

THE PROJECT

What is the Cornell/Microsoft project?

Cornell University Library has entered into a partnership with Microsoft to digitize a significant number of books and to put the volumes online using the Microsoft Live Book Search service. The initiative focuses on works already in the public domain and will allow students, researchers, and scholars around the world online access to books from Cornell's outstanding collections. The project supports the library's long-standing commitment to make its collections broadly available and Cornell President David Skorton's goal to increase the impact of the university beyond campus boundaries.

Why is Cornell collaborating with Microsoft?

The Cornell-Microsoft partnership is a win-win situation for Cornell. By working with Microsoft, we participate in content selection decisions to ensure that scholarly materials that support Cornell's academic programs are available in digital form. Microsoft will give the Library high-quality digital images of all the materials, allowing the Library to provide worldwide access through its own digital library and to share the content with non-commercial academic initiatives and non-profit organizations. Lastly, the agreement with Microsoft is non-exclusive and the Library is free to join other partnerships with the same or different content in digitizing its collection.

Who is doing the digitizing? Where is it being done?

Cornell and Microsoft will be teaming with Kirtas Technologies in Victor, N.Y. to digitize the materials for Live Book Search. Kirtas is a recognized pioneer of revolutionary scanning solutions that enable high-quality, non-destructive bound document digitization at up to 2,400 pages per hour. The majority of the material will be transported to Victor for digitization, and a small portion of rare and fragile material will be scanned onsite at Cornell using a Kirtas scanning machine.

Will books be damaged or destroyed during this process?

No. The scanning vendor for this project uses a digitization process that does not damage the book. A bound book is placed in a book cradle, which allows the book to be digitized without damaging the spine. Pages are digitized by overhead digital cameras, leaving the books intact. A mechanical arm then turns the pages as they are digitized (this process is probably more gentle than turning pages by human hand). We expect all our books to be returned to the shelves in the same condition they were in before scanning.

Who is funding this project?

Microsoft is funding the digitization of the material — this includes all costs associated with scanning. Cornell University Library is contributing staff and infrastructure to support the long-term preservation of the digitized files.

How long will the project take?

Initially, we have entered into a one-year, non-binding agreement with Microsoft with the option to renew. The books slated for the first year of digitization will take approximately 18 months to digitize and make available online.

Are there any other libraries involved?

The University of Toronto, the British Library, the New York Public Library, the University of California, Yale University, Princeton Theological Seminary, and the American Museum of Veterinary Medicine are also collaborating with Microsoft.

How does this help faculty and students?

When surveyed about their needs, the Library's users rate access to full-text online as one of their highest priorities, so this partnership will enable us to respond to student and faculty expectations. Although Cornell has a long history of digitizing books for scholarly access, this project will make substantially more books available for these purposes much sooner than would otherwise be possible.

SELECTION

How many books are being digitized?

For the first year of the project, we are targeting to digitize around 100,000 monographs - English language and in public domain

How is the selection process made?

We continue to work on refining our selection policy. The first batch of materials sent were public domain monographs (pre-1923) from Olin Library and Engineering Library that are in English. Some of the subject strengths in the CUL that could theoretically be digitized include agriculture, American history, English literature, astronomy, food and wine, general engineering, the history of science, home economics, hospitality and travel, human sexuality, labor relations, Native American materials, ornithology, veterinary medicine, and women's studies.

What is not being selected?

Currently no copyrighted material, non-English language books, or serials are under consideration for digitization.

Why spend time and money to digitize old, "low-use" books?

Our experience digitizing library material suggests that digitizing and making "low-use" material available online will boost their use and increase the long-term returns on our investment in the purchase, storage, and maintenance of Cornell's extensive print collections.

Are books under copyright being scanned?

Currently we will not be scanning any copyrighted material.

USER IMPACT

How can I tell if a book is charged out to digitization?

The general public will be able to tell that items are charged out, but it will not be apparent from the public catalog that these items are charged out for digitization. Public service staff can assist users by accessing the MARC record in the online catalog. The 903 field in the MARC record will reflect that the book has been sent to be digitized. It will have "Kirtas" and the barcode in the field. For example: 903 __ |a Kirtas |d 20060901 |p 31924029885658

How long will books be unavailable?

Books will be checked out for the project for up to 3 months. We hope as the work flow processes become smoother that the actual time the books will be out of the library will be shorter. If you request (hold/recall) a title that is checked out for this project, your request will automatically be transferred to Interlibrary Loan or Borrow Direct ensuring that you receive the title you need as quickly as possible.

There is a 2006 book I need to place on course reserve. Is that likely to be charged out to digitization?

No. During the first year of the project we are only digitizing books in the public domain. We are currently focusing on books published prior to 1923.

ACCESS

How will I be able to access the digitized books?

There will be two mechanisms to access the books, a Microsoft delivery platform and a Cornell Library delivery mechanism (yet to be determined). By the end of 2007, users should be able to access material at the Microsoft Live Book site (now in beta version):

<http://search.live.com/results.aspx?q=&scope=books&FORM=BBIR>

When will digitized images be available using Cornell's platform?

We expect to start receiving the digitized images in August 2007. The LSDI Access group is in the process of assessing different delivery options. One of the scenarios reviewed involves linking the PDF versions of the books to the Voyager records as they start arriving, however we don't anticipate implementation of such a system until 2008.

Will there be a cost for using the digitized images?

No. Access to the books both from Microsoft and Cornell site will be free of any charges.

Will the books be available to people outside the University?

All books will be made freely available to the world.

FURTHER INFORMATION

Who can I contact for more info?

Contact Oya Rieger (oyr1@cornell.edu) for further information.