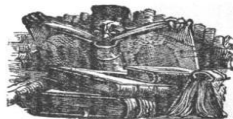


Digitized Collections Online

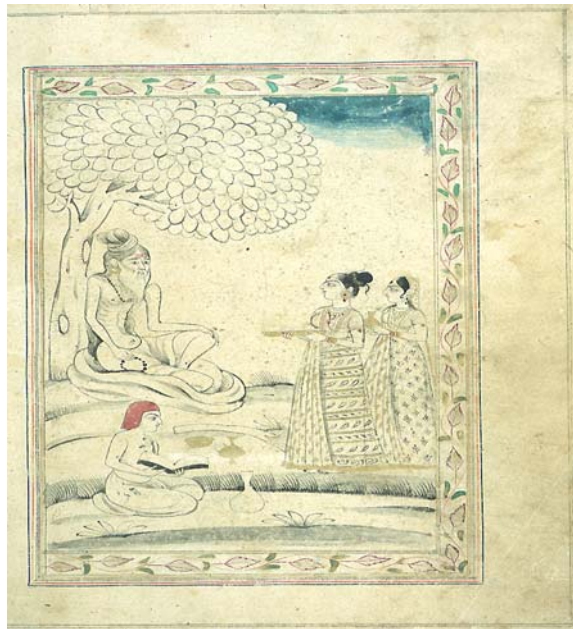
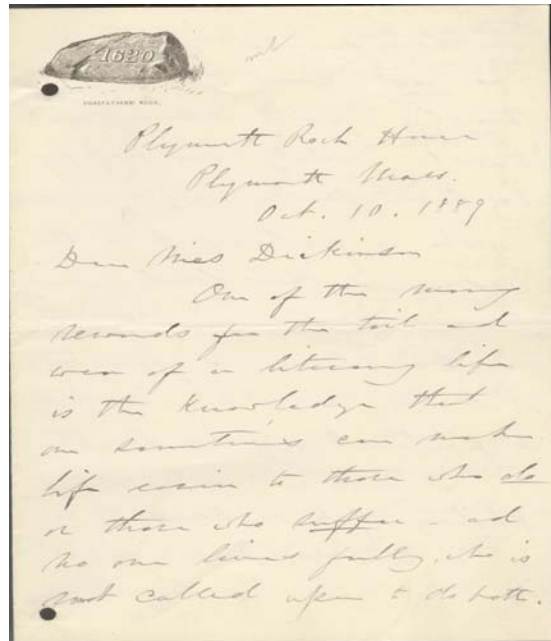
New York Society Library
November 3, 2006

THE
COMPLETE POEMS
OF
CHARLES DICKENS



NEW YORK
WHITE, STOKES, & ALLEN
MDCCCLXXXV

Digitized by Google



Keren Fleshler, Cataloger

kfleshler@nysoclib.org

1. What is digitizing or the process of digitization?

Digitizing is the process of converting data to digital format for processing by a computer.

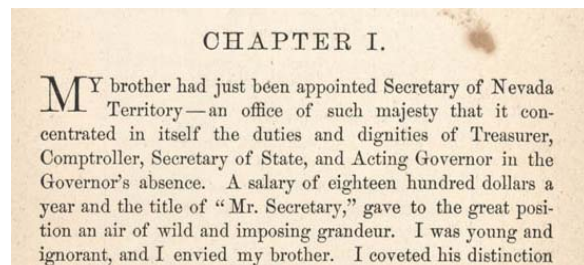
Digitized texts and images are scanned and then uploaded to the web in their original structure or design, or they may undergo a process involving **OCR**, or Optical Character Recognition.

OCR uses computer software designed to translate images of handwritten or typewritten text, usually captured by a scanner, into machine-editable text¹ (such as Notepad, Microsoft Word, etc.). OCR is helpful because it allows users to search for words in the text, but is problematic because often times words are translated incorrectly in the process.

OCR version

My brother had just been appointed Secretary of Nevada Territory -- an office of such majesty that it concentrated in itself the duties and dignities of Treasurer, Comptroller, Secretary of State, and Acting Governor in the Governor's absence. A salary of eighteen hundred dollars a year and the title of "Mr. Secretary," gave to the great position an air of wild and imposing grandeur. I was young and ignorant, and I envied my brother.

Scanned-only version



Some materials are "born digital" and do not to be digitized. This refers to such resources as Word documents, e-journals, photos taken with a digital camera, or websites.

Why digitize?

Institutions around the world are focusing on digitization projects in order to share their resources around the globe, promote their collections, and to preserve older or rare material for posterity. The digitized item often becomes the surrogate for material that is often too brittle to be used.

Digitization reduces the distance between repositories to a keystroke, eliminates barriers erected between different types of research materials, and reunites separated collections.²

¹ http://en.wikipedia.org/wiki/Optical_character_recognition

² <http://www.ncecho.org/guide/intro.asp>

What is a digitization project?³

A digitization project is, mostly simply, the process by which a repository, or group of repositories, selects materials of high historical or informational value in order to digitize them and make them available to the public on the internet. The project usually involves scanning, indexing material by author, title, or subject keywords, and uploading them to the web. Libraries, archives, museums, historical societies, and universities, and other related institutions are also focusing on digitization projects as way to preserve older or rare materials, and to make them accessible beyond library walls.

Repositories often collaborate in digitization projects in order to strengthen online collections by filling in gaps, and to share the burden of high costs. A large number of projects are paid for by outside grants.

Often times, the scanning of materials can be done on-campus at a number of universities, or they may be shipped off to service companies specializing in conservation and digitization. Scanning of items such as books may be done directly from the pages or from microfilm copies.

A digitization project may focus on a specific time period, geographic location, person or family, or subject. It may also respond to the need to preserve particularly delicate collection.

2. What is a digital collection?

Digital collections are the products of digitization projects. A collection of materials converted to machine-readable format for preservation or to provide electronic access, usually by libraries, museums, or archives. A digital collection consists of digital objects that are selected and organized to facilitate their access and use.

Digital collections often include books, manuscripts, letters, photographs, journals, magazines, newspapers, documents, and audio and visual clips. They also frequently complement ongoing or past exhibitions.

³ Helpful resources for digitization projects:
HANDBOOK FOR DIGITAL PROJECTS: A Management Tool for Preservation and Access published by the Northeast Document Conservation Center.
<http://www.nedcc.org/digital/dighome.htm>
Planning Digital Projects for Historical Collections, NYPL
<http://digital.nypl.org/brochure/index.html>
Digital Imaging resources and links
<http://palimpsest.stanford.edu/bytopic/imaging/>

Copyrights and permissions for digital material

The majority of books digitized by libraries and through projects such as Google Books and Project Gutenberg are **public domain** books. These are books that are out of copyright, and therefore can be distributed to the public and copied freely.

Books published in the United States prior to 1923 are public domain books, and may be digitized without obtaining any permissions and put online in full-text.⁴ Works of the United States Government and various other governments are also excluded from copyright law and may therefore be considered to be in the public domain.

In some cases, permissions for public use of digitized material must be obtained from the holding institution. The terms of use for a particular institution/collection are normally found on the website for the collection.

System requirements

Java, download from: <http://java.com>
Adobe Reader (PDF viewer), Adobe Flash Player, Adobe Shockwave Player:
Download all three from www.adobe.com

For video players: Windows Media Player, Real Player, and Quicktime Player are most commonly used.

A few sites may require you to download specific reader software. Some may also allow you to download the e-resources to your PDA or iPod.

If you do not have the appropriate software to view images of text, there is usually an icon on the webpage you are viewing that will link you to the website where you can download the software.

3. How would I search for digitized resources in my area of interest?

1. If you are aware of an institution that holds or collects in your area of interest, then there is a good chance it may have a digital collection on the subject. To find digital collections at a university, go to its library homepage and look for links for "Digital xxx" or something similar. Sometimes, a library

⁴ Copyright timelines on books published after 1923 inside and outside the United States are accessible online at <http://www.copyright.cornell.edu/> (click on link for Public domain chart)

Digitized Collections Online
New York Society Library, November 3, 2006

collection may have its own page and the digital projects may be found there. Other times checking the rare books and special collections department page is also helpful.

2. If your research topic is geographically based, there is a good chance a library, university, or historical society in the area of interest may have a digital collection on your topic. The same hold true for research on people influential in certain geographical areas.

3. A Google search may pull up digital collections by a combination of name, subjects, or institutions.

Example searches:

Women's history digital collections- pulls up sites with women's history digital collections at various institutions.

UCLA digitization projects – pulls up link all of UCLA's online collections

These searches may also help you pull up other digital collections that may include your subject in some of the documents.

4. You may use the one of the comprehensive Digital Collection Guides at the beginning of the following set of collection links.

Remember: Even if an institutional website does not offer digital collections, often times a helpful "links" page can point you in the right direction. Example: a small historical society in Vermont may have a links page that will point you to a major Vermont history digitization collaboration being administered by a network of Vermont libraries and universities, even if the historical society itself is not taking part.

As with all research, online research takes time and patience. Every collection database has a unique design and its own set of quirks that may take time to learn, but the rewards are great.

A Selection of Digital Collections

There are two ways to view materials in digital collections:

Browse: Search a listing of names, titles, and keywords

Search: Enter a search word(s) to locate a specific document

Guides to Digital Collections

British Library, links to European digital collections

<http://www.bl.uk/collections/epcdigital.html>

Institute of Museum and Library Services, Digital Collections and Content

<http://imlsdcc.grainger.uiuc.edu/>

Library of Congress directory to digital collections, by institution

http://www.loc.gov/rr/print/resource/223_piccat.html

E-Book Collections

Bartleby: includes fiction and non-fiction, but also full-text dictionaries, encyclopedias, quotation books, and other reference materials (OCR)

<http://www.bartleby.com/>

Internet Archive- searches e-books from American and Canadian libraries, as well as from the Million Books Project, Project Gutenberg, Open Source Books, and classic children's books.

<http://www.archive.org>

Google Books

<http://books.google.com/>

Listing of public domain books available on the web with links to corresponding Google Book pages

<http://zuhouse.org/dp/gfound1.html>

ManyBooks.Net (can download books to PDA or iPod)

<http://manybooks.net/>

Online Books Page, Univ. of Pennsylvania (over 25,000 English works)

<http://digital.library.upenn.edu/books/>

Project Gutenberg
<http://www.gutenberg.org/>

Project Gutenberg – Australia
<http://gutenberg.net.au/>

Early Books

British Library: Turning the Pages
<http://www.bl.uk/onlinegallery/ttp/ttpbooks.html>

British Library: Treasures in Full
<http://www.bl.uk/treasures/treasuresinfull.html>

Gutenberg Digital
<http://www.gutenbergdigital.de/gudi/start.htm>

Renaissance in Print: 16th century French Books in the Douglas Gordon Collection
<http://www.lib.virginia.edu/rmds/collections/gordon/>

Literature E-book Projects

Early American Fiction, University of Virginia (books and manuscripts)
<http://etext.lib.virginia.edu/eaf/pubindex.html>

American Verse Project (OCR only)
<http://www.hti.umich.edu/a/amverse/>

Periodicals

Library of Congress, 19th century periodicals (American Memory)
<http://memory.loc.gov/ammem/ndlpcoop/mahtml/snchome.html>

Cornell University's Making of America Journal Collection (linked to LOC project)
<http://library5.library.cornell.edu/moa/>

Internet Library of Early Journals

20-year consecutive runs of: Gentleman's Magazine, The Annual Register, Philosophical Transactions of the Royal Society, Notes and Queries, The Builder, Blackwood's Edinburgh Magazine

<http://www.bodley.ox.ac.uk/ilej/>

Vanity Fair Caricatures

http://www.lib.virginia.edu/rmds/collections/vanity_fair/index.html

Utah Digital Newspapers

<http://www.lib.utah.edu/digital/unews/>

Washington State Historical Newspapers

<http://www.secstate.wa.gov/history/newspapers.aspx>

British Library Online Newspaper Archive

<http://www.uk.olivesoftware.com/>

Freedom Journal, Wisconsin

First African-American Periodical, 1827-1829

<http://www.wisconsinhistory.org/libraryarchives/aanp/freedom/>

Northern New York Historical Newspapers

<http://news.nnyln.net/>

Brooklyn Daily Eagle, 1841-1902

<http://www.brooklynpubliclibrary.org/eagle/>

National and local histories and periodicals

American Memory, Library of Congress

<http://memory.loc.gov/>

British History Online (offers citation feature)

Includes books, maps

<http://www.british-history.ac.uk/>

Colorado History (including collection of Colorado paper, example of previously microfilmed material put online)

<http://www.cdpheritage.org>, then click on Collections from menu

Digital Library of Wales: Digital Mirror

Manuscripts, maps, images, sound and video

http://www.llgc.org.uk/drych/index_s.htm

Historic Pittsburgh (Books, images, censuses, maps)
<http://digital.library.pitt.edu/pittsburgh/>

Making of America: E-Books (scanned and OCR versions)
<http://www.hti.umich.edu/m/moagrp/>

Salem Witch Trials, from the University of Virginia
<http://www.salemwitchtrials.org/home.html>

Women's history

American Women's History: a research guide (links to several online collections)
<http://www.mtsu.edu/~kmiddlet/history/women/wh-digcoll.html>

Women Working: 1800-1930, Harvard University
<http://ocp.hul.harvard.edu/ww/faqs.html>

Maps

David Rumsey Map collection
<http://www.davidrumsey.com/>

Religion and Medicine

American Theological Library Association (example of different phases in which new libraries joined the initiative)
<http://www.atla.com/digitalresources/>

Turning the Pages Online, National Library of Medicine
<http://archive.nlm.nih.gov/proj/ttp/intro.htm>

Islamic Medical Manuscripts
<http://www.nlm.nih.gov/hmd/arabic/arabichome.html>

Other Collections:

Clearinghouse of Government Documents Digital Projects

"The ALA Government Documents Round Table is collecting this information to facilitate cooperation and partnerships between libraries for digitization projects."

<http://www.gl.iit.edu/services/ref/SearchDatabase.htm>

Columbia Papyrus Project

<http://www.columbia.edu/cu/lweb/projects/digital/apis/>

Complete Work of Charles Darwin Online (in process until 2009)

<http://darwin-online.org.uk/>

Image, Sound, and Visual collections

Archive.org: extensive collection of moving images and audio

<http://www.archive.org>

Indianapolis Museum of Art, African Masks (example of zoom+drag)

<http://www.ima-art.org/cycles/>

Bodleian Ballads (images and sound)

<http://www.bodley.ox.ac.uk/ballads/>

Steven Spielberg Jewish Film Archive, in association with the Hebrew University of Jerusalem : access to over 300 Jewish history and Zionist films in English and Hebrew

<http://www.spielbergfilmarchive.org.il/>

Victoria Museum, Zoological illustrations, Caught and Colored

<http://www.museum.vic.gov.au/caughtandcoloured/Collection.aspx>

Collections by Institution:

Columbia University

<http://www.columbia.edu/cu/lweb/digital/collections/>

University of Virginia (OCR)

<http://etext.lib.virginia.edu/>

University of Pittsburgh

<http://digital.library.pitt.edu/>

National Library of Medicine

Digitized Collections Online
New York Society Library, November 3, 2006

<http://www.nlm.nih.gov/onlineexhibitions.html>

Cornell University

<http://cdl.library.cornell.edu/>

UC, Berkeley

<http://sunsite.berkeley.edu/Collections/>