An introduction to computational access

Leontien Talboom

Digital repositories on the Web

INTERNET ARCHIVE
WayBackMachine
http://www.kb.nl/kb/manuscripts/introduction/index.html
Go
MAY JUN AUG
45 captures
24 Jun 2001 - 19 Aug 2016
2000 2001 2002

Introduction

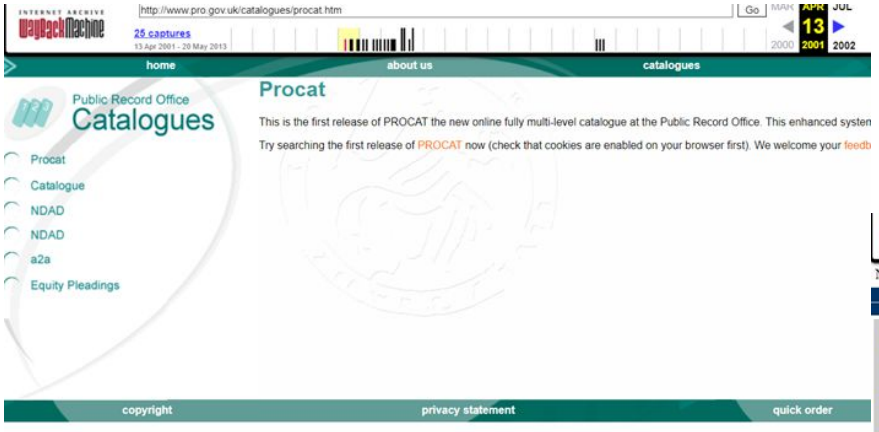
The scope and purpose of the web catalogue

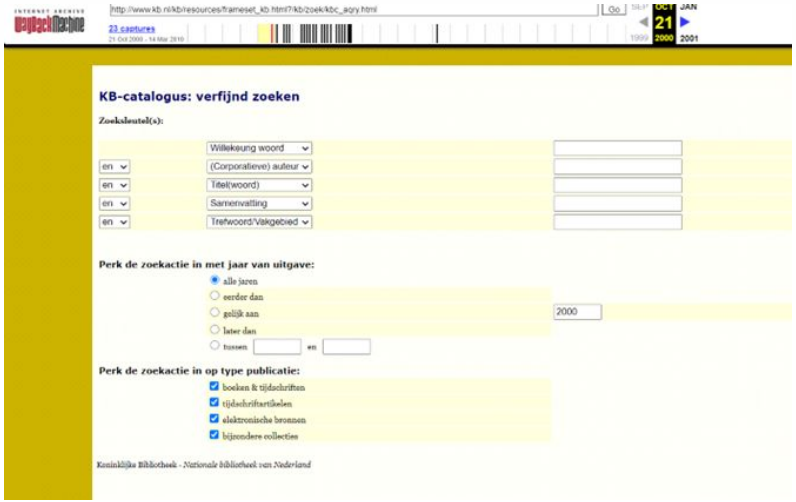
A web catalogue of a major collection of medieval illuminated manuscripts must aim to serve many different masters. Just like a published catalogue, it should provide information to anyone with a serious scholarly interest. Like a public exhibition, however, it should also be interesting and pleasurable to the occasional visitor, merely keen on finding pretty pictures.

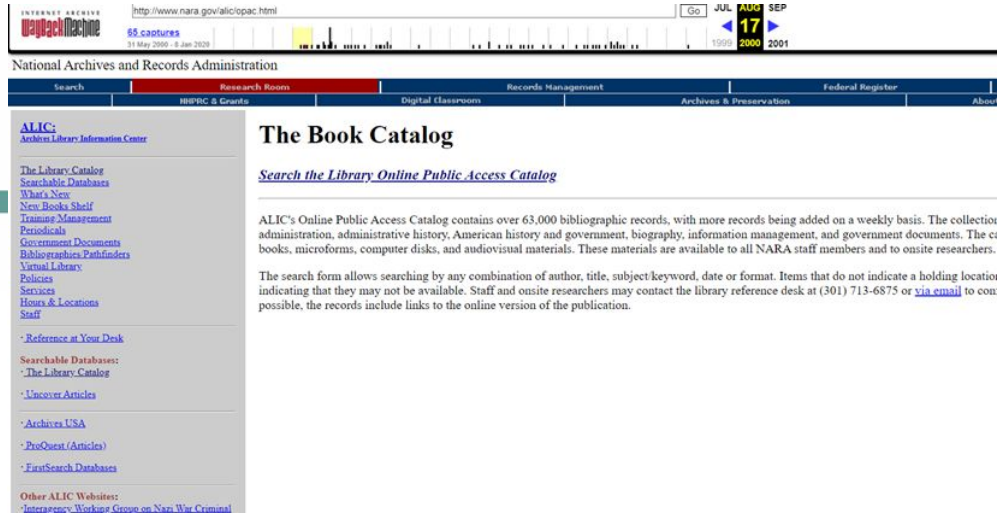
Of course, we have tried to make the access to the underlying database of medieval illumination as transparant as possible. Therefore, the search screen will hold few surprises for those who are used to consulting catalogues on medieval manuscripts. Even so, it may still be useful for them to read at least part of this prologue. For those who have no experience with manuscript catalogues or with medieval art in general, these introductory paragraphs will certainly be useful, because we shall show something of the richess of the material. We hope that the type of information we describe here will stimulate you to explore the online catalogue which you will reach when you click search. For the moment the web catalogue is first of all committed to offering a search facility for the collection. This emphasis on the database and its access should not prevent you from browsing around. As a starting point for browsing through the collection you could, for example, simply select a *Shelfmark* from the first index on the search screen. It is our intention to supplement the search facility of the catalogue with a *browse* option which will give access to 'temporary exhibitions' that focus on particular aspects of the collection.

Subject matters

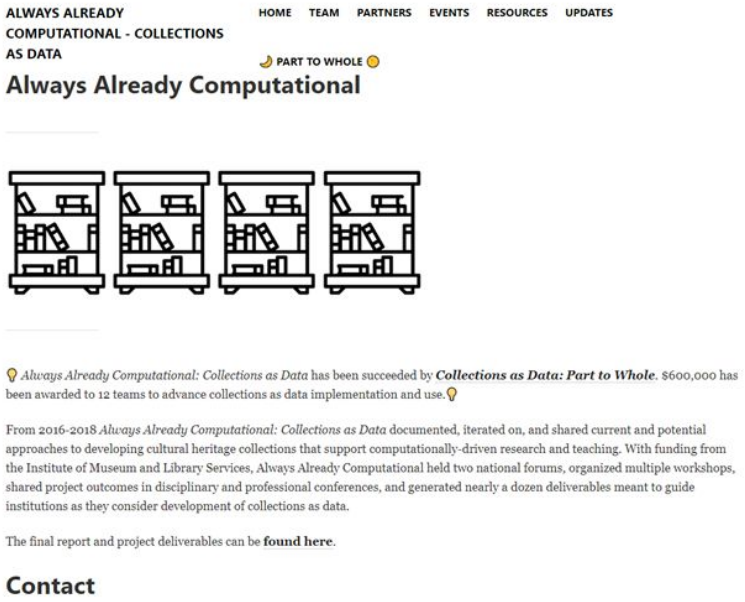
We begin with a few notes on the iconographic retrieval facilities of the catalogue, because the illumination is a focal point of the collection, and because a great deal of time and effort has been invested in the subject access to the miniatures, initials and border decorations. First of all, captions have been written for every picture and an alphabetically arranged list of these captions can be called forth by clicking on the index button at the box *Iconographic descriptions*.At a later stage of the project a complete list of the words used in the captions will be added to the retrieval facilities.


http://www.pro.gov.uk/catalogues/procat.htm
26 captures
13 Apr 2001 - 20 May 2013
APR 13 2001
JUL 2002
home
about us
catalogues
Public Record Office
Catalogues
Procat
Catalogue
NDAD
NDAD
a2a
Equity Pleadings
Procat
This is the first release of PROCAT the new online fully multi-level catalogue at the Public Record Office. This enhanced systen
Try searching the first release of PROCAT now (check that cookies are enabled on your browser first). We welcome your feedb
copyright
privacy statement
quick order


http://www.kb.nl/kb/resources/frameset_kb.html?kb/zoek/kbc_agry.html
23 captures
21 Oct 2000 - 14 Mar 2010
JAN 21 2001
KB-catalogus: verfijnd zoeken
Zoeksleutel(s):
Willekeurig woord
(Corporatieve) auteur
Titel(woord)
Samenvatting
Trefwoord/Vakgebied
Perk de zoekactie in met jaar van uitgave:
alle jaren
eerder dan
gelijk aan
later dan
tussen
en
2000
Perk de zoekactie in op type publicatie:
boeken & tijdschriften
tijdschriftartikelen
elektronische bronnen
bijzondere collecties
Koninklijke Bibliotheek - Nationale bibliotheek van Nederland


http://www.nara.gov/alic/opac.html
65 captures
31 May 2000 - 8 Jan 2020
JUL AUG SEP
17
1999 2000 2001
National Archives and Records Administration
Search
Research Room
Records Management
Federal Register
NHPRC & Grants
Digital Classroom
Archives & Preservation
ALIC:
Archives Library Information Center
The Library Catalog
Searchable Databases
What's New
New Books Shelf
Training Management
Periodicals
Government Documents
Bibliographies/Pathfinders
Virtual Library
Policies
Services
Hours & Locations
Staff
· Reference at Your Desk
Searchable Databases:
· The Library Catalog
· Uncover Articles
· Archives USA
· ProQuest (Articles)
· FirstSearch Databases
Other ALIC Websites:
·Interagency Working Group on Nazi War Criminal
The Book Catalog
Search the Library Online Public Access Catalog
ALIC's Online Public Access Catalog contains over 63,000 bibliographic records, with more records being added on a weekly basis. The collectio
administration, administrative history, American history and government, biography, information management, and government documents. The ca
books, microforms, computer disks, and audiovisual materials. These materials are available to all NARA staff members and to onsite researchers.
The search form allows searching by any combination of author, title, subject/keyword, date or format. Items that do not indicate a holding locatio
indicating that they may not be available. Staff and onsite researchers may contact the library reference desk at (301) 713-6875 or via email to con
possible, the records include links to the online version of the publication.

The Digital Shift



- Digitised material
- Born-digital material
- Computing capacity
- Tools and examples

What is computational access?

Any type of access that requires computational methods:

-This is not limited to state-of-the-art projects

-Is available in a digital environment and scalable

-If granted by copyright laws or other regulations, accessible to all

Four different approaches:

- Terms of Use

-Bulk dataset

-API

-Platform

Terms of Use

Internet Archive's Terms of Use, Privacy Policy, and Copyright Policy

Terms of Use
31 Dec 2014

This terms of use agreement (the "Agreement") governs your use of the collection of Web pages and other digital content (the "Collections") available through the Internet Archive (the "Archive"). When accessing an archived page, you will be presented with the terms of use agreement. If you do not agree to these terms, please do not use the Archive's Collections or its Web site (the "Site").

Access to the Archive's Collections is provided at no cost to you and is granted for scholarship and research purposes only. The Archive, at its sole discretion, may provide you with a password to access certain Collections, provided that you complete any required application process and provide accurate information in your application. You may use your password only to access the Collections in ways consistent with this Agreement — no other access to or use of the Site, the Collections, or the Archive's services is authorized. You agree not to interfere with the work of other users or Archive personnel, servers, or resources. Further, you agree not to recirculate your password to other people or organizations. Please report any unauthorized use of your password promptly to info@archive.org. You acknowledge that you have read and understood the Archive's Privacy Policy and agree that the Archive may collect, use, and distribute information pursuant to that policy. If you provide any content to the Archive, you grant the Archive a nonexclusive, royalty-free right to use that content.

Some of the content available through the Archive may be governed by local, national, and/or international laws and regulations, and your use of such content is solely at your own risk. You agree to abide by all applicable laws and regulations, including intellectual property laws, in connection with your use of the Archive. In particular, you certify that your use of any part of the Archive's Collections will be limited to noninfringing or fair use under copyright law. If a Creative Commons or other license has been declared for particular material on the Archive, to the extent you trust the declaration and declarer (which is rarely the Internet Archive), you may use the content according to the terms and conditions of the applicable license. In using the Archive's site, Collections, and/or services, you further agree (a) not to violate anyone's rights of privacy, (b) not to act in any way that might give rise to civil or criminal liability, (c) not to use or attempt to use another person's password, (d) not to collect or store personal data about anyone, (e) not to infringe any copyright, trademark, patent, or other proprietary rights of any person, (f) not to transmit or facilitate the transmission of unsolicited email ("spam"), (g) not to harass, threaten, or otherwise annoy anyone, and (h) not to act in any way that might be harmful to minors, including, without limitation, transmitting or facilitating the transmission of child pornography, which is prohibited by federal law and may be reported to the authorities should it be discovered by the Archive.

You agree that we may contact you from time to time with surveys or other questions regarding your opinions about and uses of the Archive, as well as with information we believe may be of interest to you. We encourage you to respond to these surveys because we value your input, which will assist us in improving the Archive. In addition, we request that, according to standard academic practice, if you use the Archive's Collections for any research that results in an article, a book, or other publication, you list the Archive as a resource in your bibliography.

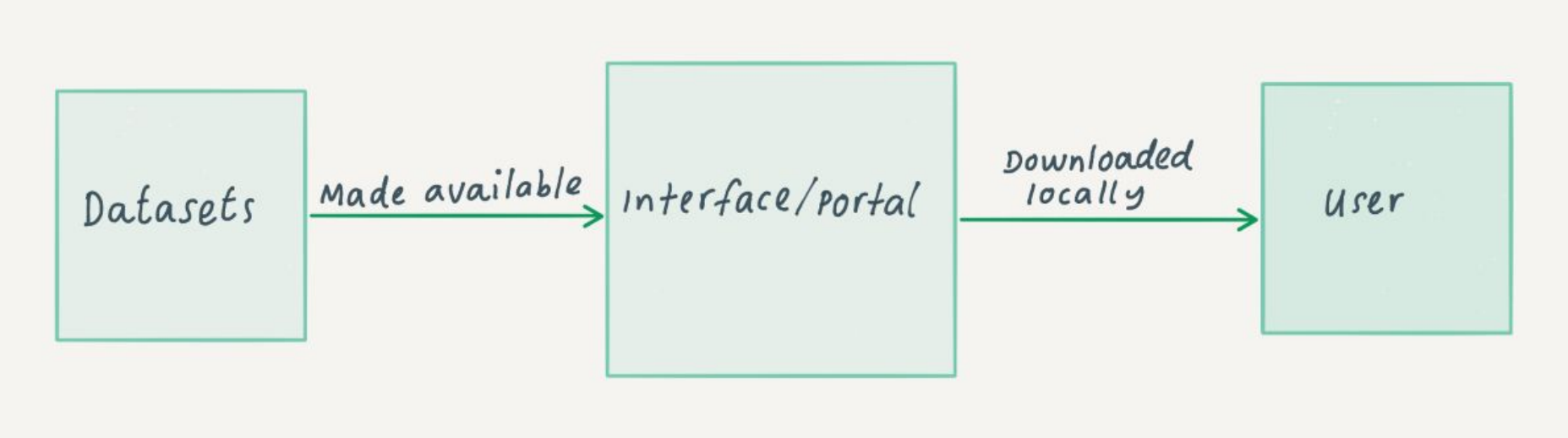
While we collect publicly available Internet documents, sometimes authors and publishers express a desire for their documents not to be included in the Collections (by tagging a file for robot exclusion or by contacting us or the original crawler group). If the author or publisher of some part of the Archive does not want his or her work in our Collections, then we may remove that portion of the Collections without notice.

The Archive may immediately terminate this Agreement at its sole discretion at any time upon written notice (including via email) to you. Upon termination, you agree that the Archive may immediately deactivate any password it has issued to you and bar you from accessing the Collections or the Site.

The Archive may modify this Agreement from time to time, and your continued use of the Collections and/or the Site constitutes your acceptance of any and all modifications. The Archive will attempt to notify you of substantial modifications via the email address that you have registered with us, if any.

Because the content of the Collections comes from around the world and from many different sectors, the Collections may contain information that might be deemed offensive, disturbing, pornographic, racist, sexist, bizarre, misleading, fraudulent, or otherwise objectionable. The Archive does not endorse or sponsor any content in the Collections, nor does it guarantee or warrant that the content available in the Collections is accurate, complete, noninfringing, or legally accessible in your jurisdiction, and you agree that you are solely responsible for abiding by all laws and regulations that may be applicable to the viewing of the content. In addition, the Collections are provided to you on an as-is and as-available basis. You agree that your use of the Site and the Collections is at your sole risk. You understand and agree that the Archive makes no warranty or representation regarding the

Bulk dataset


Datasets
Made available
Interface/Portal
Downloaded locally
User

Datasets

Downloadable, non-consumptive book data.

HTRC releases research datasets to facilitate text analysis using the HathiTrust Digital Library. While copyright-protected texts are not available for download from HathiTrust, fruitful research can still be performed on the basis of non-consumptive analysis of features extracted from full text. These features include volume-level metadata, page-level metadata, part-of-speech-tagged tokens, and token counts. Additionally, HTRC has partnered with advanced researchers to release a derived dataset, Word Frequencies in English-Language Literature, 1700-1922.

HTRC Extracted Features Dataset

Page-level features from 17.1 million volumes [v.2.0]

Description

The HTRC Extracted Features Dataset v.2.0 is composed of page-level features for 17.1 million volumes in the HathiTrust Digital Library. This version contains non-consumptive features for both public-domain and in-copyright books.

Features include part-of-speech tagged term token counts, header/footer identification, marginal character counts, and much more.

A full explanation of the dataset's features, motivation, and creation is available at the EF Dataset documentation page

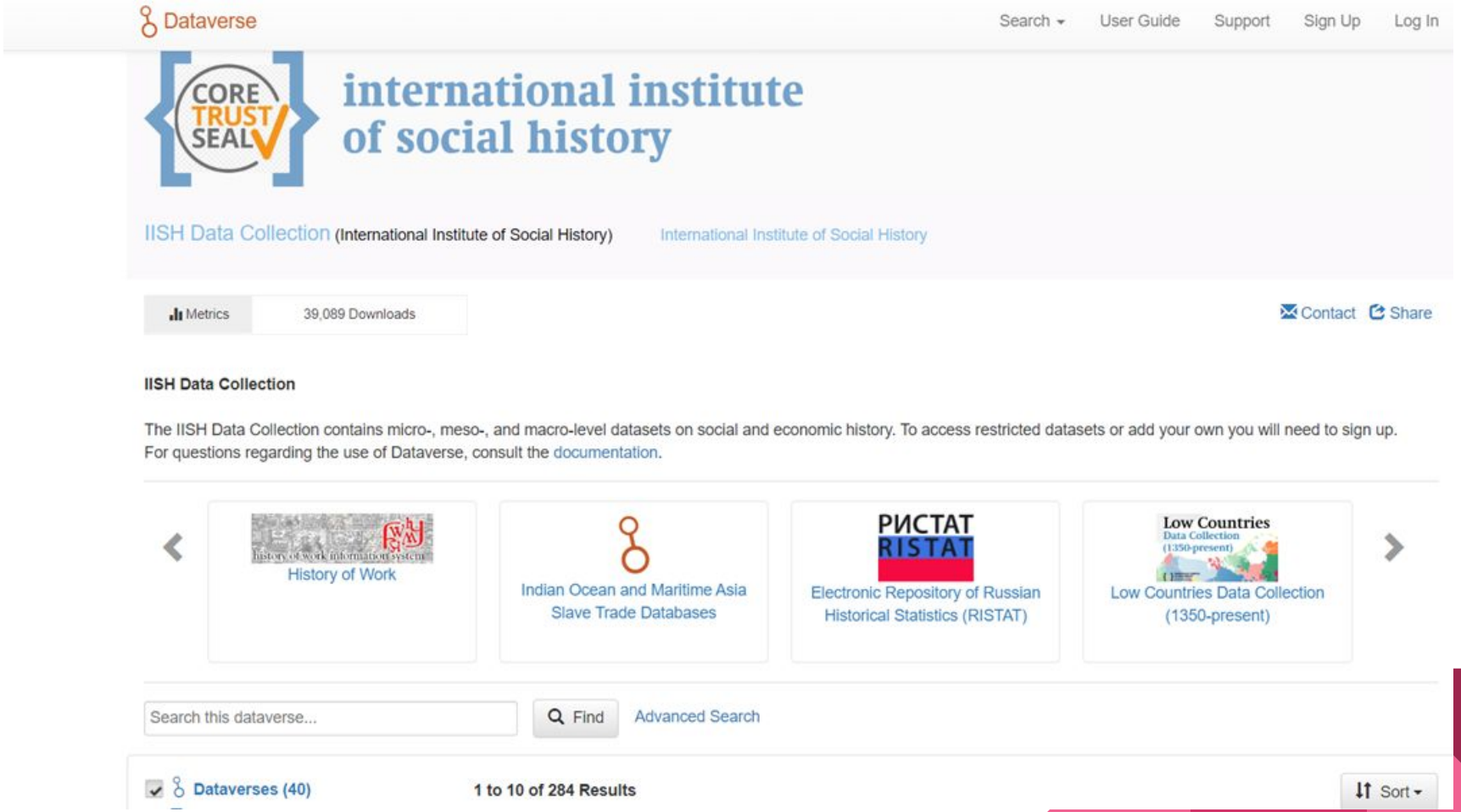
Download the data

All 17.1 million files as well as custom subsets of the EF data are accessible using `rsync`, as

Contents

# of volumes	17,123,746
In-copyright	10,550,952
Public domain and Creative Commons	6,572,794
# of pages	6,221,631,336
# of tokens	2,906,819,723,689

Resources


Dataverse
Search
User Guide
Support
Sign Up
Log In
CORE TRUST SEAL
international institute of social history
IISH Data Collection (International Institute of Social History)
International Institute of Social History
Metrics
39,089 Downloads
Contact
Share
IISH Data Collection
The IISH Data Collection contains micro-, meso-, and macro-level datasets on social and economic history. To access restricted datasets or add your own you will need to sign up.
For questions regarding the use of Dataverse, consult the documentation.
History of Work
Indian Ocean and Maritime Asia Slave Trade Databases
РИСТАТ RISTAT
Electronic Repository of Russian Historical Statistics (RISTAT)
Low Countries Data Collection (1350-present)
Low Countries Data Collection (1350-present)
Search this dataverse...
Find
Advanced Search
Dataverses (40)
1 to 10 of 284 Results
Sort

master · 3 branches · 58 tags · Go to file · Code

john-halderman Automatic monthly update. cf5a0b3 26 days ago 73 commits

.gitattributes	May 16th update; moving datasets to large file storage; added domain ...	4 years ago
Artists.csv	Automatic monthly update.	26 days ago
Artists.json	Automatic monthly update.	26 days ago
Artworks.csv	Automatic monthly update.	26 days ago
Artworks.json	Automatic monthly update.	26 days ago
README.md	Automatic monthly update.	26 days ago

README.md

The Museum of Modern Art (MoMA) Collection

The Museum of Modern Art (MoMA) acquired its first artworks in 1929, the year it was established. Today, the Museum's evolving collection contains almost 200,000 works from around the world spanning the last 150 years. The collection includes an ever-expanding range of visual expression, including painting, sculpture, printmaking, drawing, photography, architecture, design, film, and media and performance art.

MoMA is committed to helping everyone understand, enjoy, and use our collection. The Museum's website features 88,338 artworks from 26,443 artists. This research dataset contains 138,210 records, representing all of the works that have been accessioned into MoMA's collection and cataloged in our database. It includes basic metadata for each work, including title, artist, date made, medium, dimensions, and date acquired by the Museum. Some of these records

About

The Museum of Modern Art (MoMA) collection data

Readme

Releases 58

v1.56 Latest
26 days ago

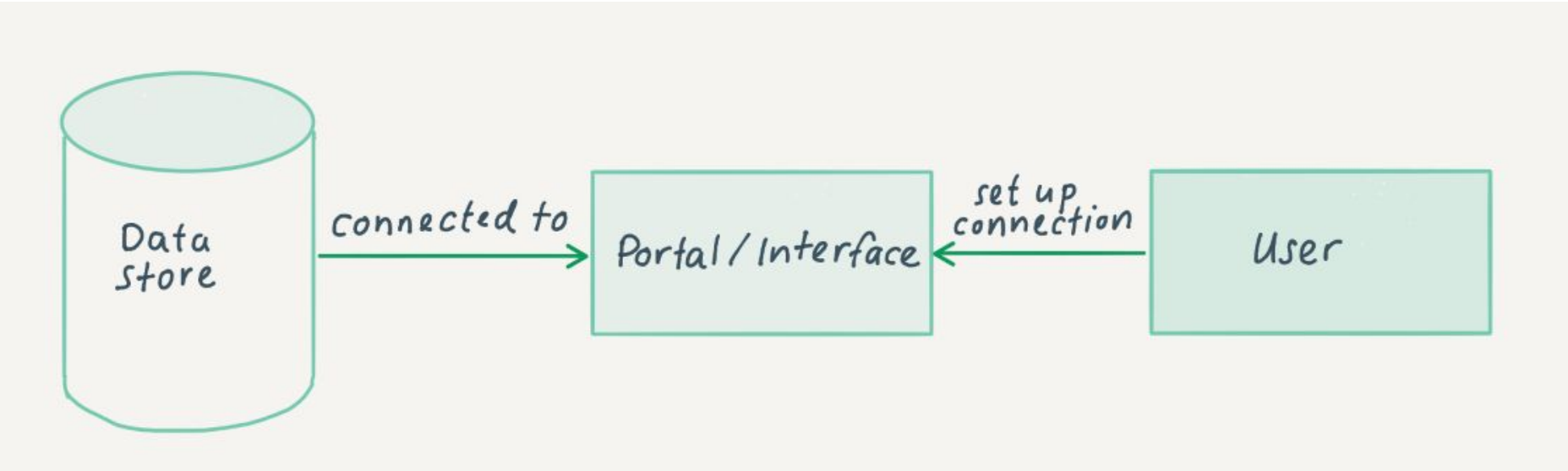
+ 57 releases

Contributors 2

john-halderman john-halderman

momadm momadm

API


Data store
connected to
Portal / Interface
set up connection
User

Wellcome Collection Developers

Catalogue Images Datasets Search...

Make something

Use open APIs and datasets to make something new with our collections

Catalogue

Create apps that search our museum and library collections, using the same API that we do.

Images

Access our images using standard International Image Interoperability Framework (IIIF) APIs.

Datasets

Download openly available data about our collections, for use in your research.

Follow along

We work completely in the open, find out what we're up to

Stacks

Read about developments at Wellcome Collection.

GitHub

View code from the Wellcome Trust on GitHub.

Cardigan

Explore our UX design system and UI components.

nullhandle added missing quotation mark `4fab816` on 26 Nov 2018 **81** commits

ait-implementation	Publish Archive-It's next implementation	3 years ago
ait-specification	Fix link syntax	3 years ago
general-specification	Fix some long-standing typos	3 years ago
lockss-implementation	Update and rename README to README.md	4 years ago
utilities	Create README.md	3 years ago
README.md	added missing quotation mark	2 years ago

README.md

WASAPI Data Transfer APIs

This is the public repository for work on the Web Archiving Systems API (WASAPI) data transfer APIs. The intention for these APIs is to provide a standardized mechanism for export and import of web archive data (and perhaps, ultimately, derivative data and capture metadata) between diverse systems for preservation, replication, research use, data delivery, and other purposes. Design and development is being carried out by project partners on the Institute of Museum and Library Services-funded National Leadership Grant, LG-71-15-0174, "Systems Interoperability and Collaborative Development for Web Archiving" (PDF) in consultation with a technical working group and based on feedback from the web archiving community.

Clients & Utilities

WASAPI data transfer APIs

Readme

Releases

No releases published

Contributors 3

ursafoot ursafoot

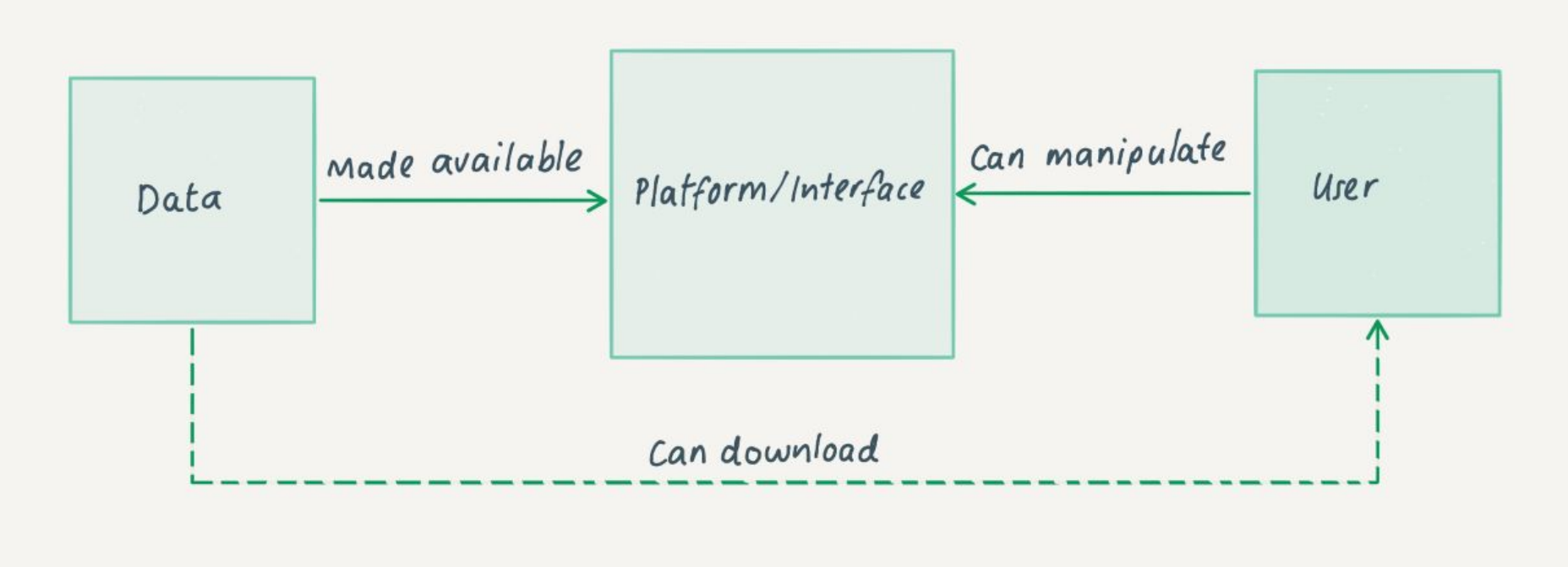
jeffersonbailey jeffersonbailey

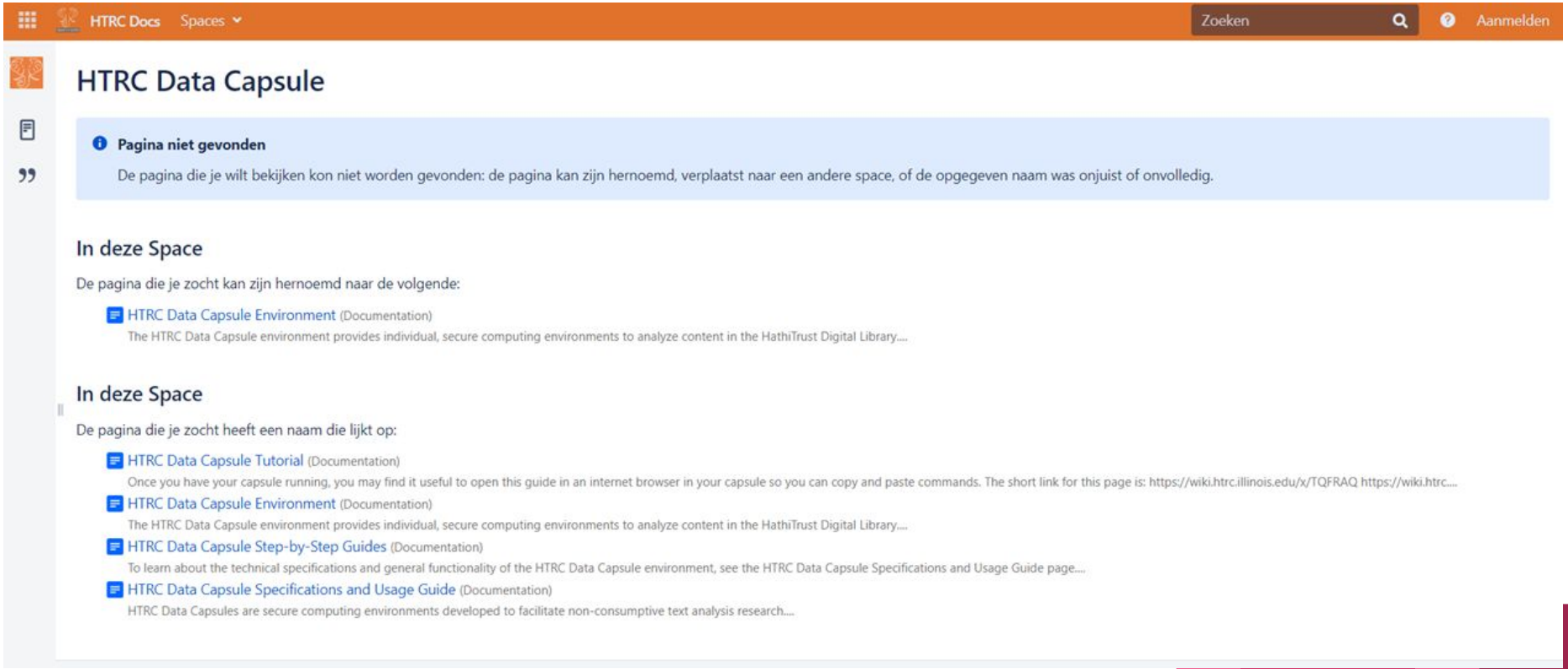
nullhandle nullhandle

Languages

Python 100.0%

Platform

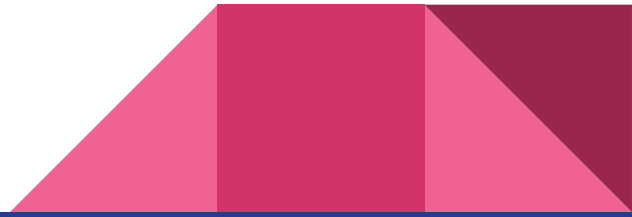

Data
made available
Platform/Interface
Can manipulate
User
Can download


HTRC Docs
Spaces
Zoeken
Aanmelden
HTRC Data Capsule
Pagina niet gevonden
De pagina die je wilt bekijken kon niet worden gevonden: de pagina kan zijn hernoemd, verplaatst naar een andere space, of de opgegeven naam was onjuist of onvolledig.
In deze Space
De pagina die je zocht kan zijn hernoemd naar de volgende:
HTRC Data Capsule Environment (Documentation)
The HTRC Data Capsule environment provides individual, secure computing environments to analyze content in the HathiTrust Digital Library....
In deze Space
De pagina die je zocht heeft een naam die lijkt op:
HTRC Data Capsule Tutorial (Documentation)
Once you have your capsule running, you may find it useful to open this guide in an internet browser in your capsule so you can copy and paste commands. The short link for this page is: https://wiki.htrc.illinois.edu/x/TQFRAQ https://wiki.htrc....
HTRC Data Capsule Environment (Documentation)
The HTRC Data Capsule environment provides individual, secure computing environments to analyze content in the HathiTrust Digital Library....
HTRC Data Capsule Step-by-Step Guides (Documentation)
To learn about the technical specifications and general functionality of the HTRC Data Capsule environment, see the HTRC Data Capsule Specifications and Usage Guide page....
HTRC Data Capsule Specifications and Usage Guide (Documentation)
HTRC Data Capsules are secure computing environments developed to facilitate non-consumptive text analysis research....

CLARIAH
MEDIA SUITE
Data
Tools
About
Learn
Contact
Login
Important notice for active users:
As of today, we've switched to a new login system. This requires you to import your old workspace. Please visit the import page before you begin
IMPORT YOUR WORKSPACE

Computational Access Guide

- Loads of resources, but not for our community
- First steps
- Basic explanation of a number of definitions



Questions?

Email: lkt39@cam.ac.uk

Twitter: @makethecatwise