Chris Freeland

Technical Director, BHL

Director, Center for Biodiversity Informatics

Missouri Botanical Garden

São Paulo, Brasil 3 – 5 Feb 2010

BHL Architecture

Overview

Goals of BHL

- Scan public domain biodiversity literature.
- Negotiate rights to digitize copyrighted materials.
- Ingest content digitized by others.
- Provide interfaces & APIs for repository.
  - GUIs
  - Services for data mining & citation resolution

Biodiversity Heritage Library: http://biodiversitylibrary.org

BHL is purpose-driven

- BHL database was created to provide access to digitized books made available through multiple, active digitization projects
  - BHL partners
  - Botanicus from MOBOT
  - Gallica from BnF
- Initial focus on representing bound materials held on library shelves

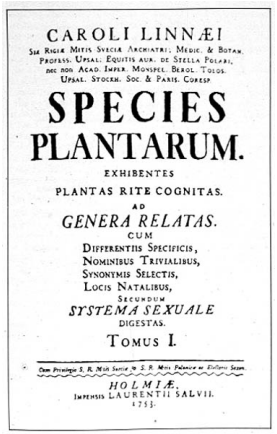
Biodiversity Heritage Library: http://biodiversitylibrary.org

Workflow

Conservation

Selection → Preparation → Digitization

Post Production → (Re)publication



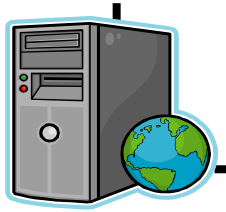
CAROLI LINNÆI
SUÆ REGIÆ MTIS SVECIÆ ARCHIATRI, MEDIC. & BOTAN. PROFESS. UPSAL. EQUITIS AUR. DE STELLA POLARI, NEC NON ACAD. IMPER. MONSPEL. BEROL. TOLOS. UPSAL. STOCKH. SOC. & PARIS. CORESP.

SPECIES PLANTARUM.

EXHIBENTES
PLANTAS RITE COGNITAS.
AD
GENERA RELATAS.
CUM
DIFFERENTIIS SPECIFICIS,
NOMINIBUS TRIVIALIBUS,
SYNONYMIS SELECTIS,
LOCIS NATALIBUS,
SECUNDUM
SYSTEMA SEXUALE
DIGESTAS.

TOMUS I.

HOLMIÆ,
IMPENSIS LAURENTII SALVII.
1753.



Prioritized Literature from TROPICOS: Top 500 titles from VAST & MOST

This list was generated by selecting all titles in TROPICOS published before 1925 and counting the number of vascular (VAST) and bryophyte (MOST) names described in each. This list is sorted by **Total taxa** and is restricted to 500 titles.

#	Tropicos ID	Title	Call number	BPH	TL2	TL2 Author	City	Total taxa	VAST	MOST
1	553	Botanische Jahrbucher fur Systematik, Pflanzengeschichte und Pflanzengeographie	QK1.B55	220.2				15,052	14,415	637
2	1063	Revisio Generum Plantarum	Rare Book QK96.K85		4021	Kuntze, C. E. O.		13,548	13,462	86
3	703	Linnaea	QK1.L56	532.04			Berlin & Halle, GERMANY	12,695	10,877	1,818
4	1051	Prodromus Systematis Naturalis Regni Vegetabilis	QK97.C2		999	A. P. de Candolle (ed.)		11,757	11,756	1
5	629	Flora Brasiliensis	FOLIO QK263.M47; reprint edil		5538	C. F. P. von Martius (ed.)		9,833	9,729	104
6	682	Journal of the Linnean Society, Botany	QK1.J823	471.16				7,599	5,886	1,713
7	786	Repertorium Specierum Novarum Regni Vegetabilis	QK1.R337	772.2				7,578	7,578	
8	586	Bulletin of the Torrey Botanical Club	QK1.B9673; EWAN	284.15			, USA	5,853	5,183	670
9	1071	Species Plantarum	INDEX REFERENCE QK91.S6		4769	Linnaeus, C.		5,736	5,675	61
10	3295	Das Pflanzenreich	QK97.P4		1713	H. G. A. Engler (ed.)..		5,455	4,955	500
11	775	Proceedings of the American Academy of Arts and Sciences	QK1.A3791	724.2				5,438	5,297	141
12	661	Hedwigia	QK1.H38	413.25			Dresden, GERMANY	5,373	393	4,980
13	648	Flora	QK1.F418	373.28				5,122	3,441	1,681
14	754	Die naturlichen Pflanzenfamilien	QK97.E47 1887-1909		1710	A. Engler & K. Prantl.		4,869	2,140	2,729
15	605	Contributions from the United States National Herbarium	QK1.U6755	331.05				4,767	4,747	20
16	676	Journal of Botany, British and Foreign	QK1.J5952	459.22				4,310	3,895	415
17	20316	Prodromus Systematis Naturalis Regni Vegetabilis	QK97.C2		999	A. L. P. P. de Candolle (ed.)		4,177	4,177	
18	576	Bulletin de la Societe Botanique de France	QK1.B91792	275.24			Paris, FRANCE	3,776	3,570	206
19	575	Bulletin of Miscellaneous Information Kew	QK1.B956	263.12				3,161	3,088	73
20	951	Flora of Tropical Africa	QK381.F56		7055	Oliver, D.		3,027	3,027	
21	922	The Flora of British India	QK349.H665.1897		2981	Hooker, J. D.		3,014	3,014	
22	1230	Repertorium Specierum Novarum Regni Vegetabilis, Beihefte	QK1.F32, QK1.F35	772.21				2,994	2,994	
23	552	Botanical Gazette	QK1.B468	220.08				2,955	2,676	279

History

- Preliminary work: MOBOT's Botanicus
  - http://www.botanicus.org
- Funded by Keck Foundation & IMLS
- Working demonstration of how nomenclators/databases (like Tropicos) can link into digitized scientific literature
- Codebase reused for BHL, then changed to fit requirements

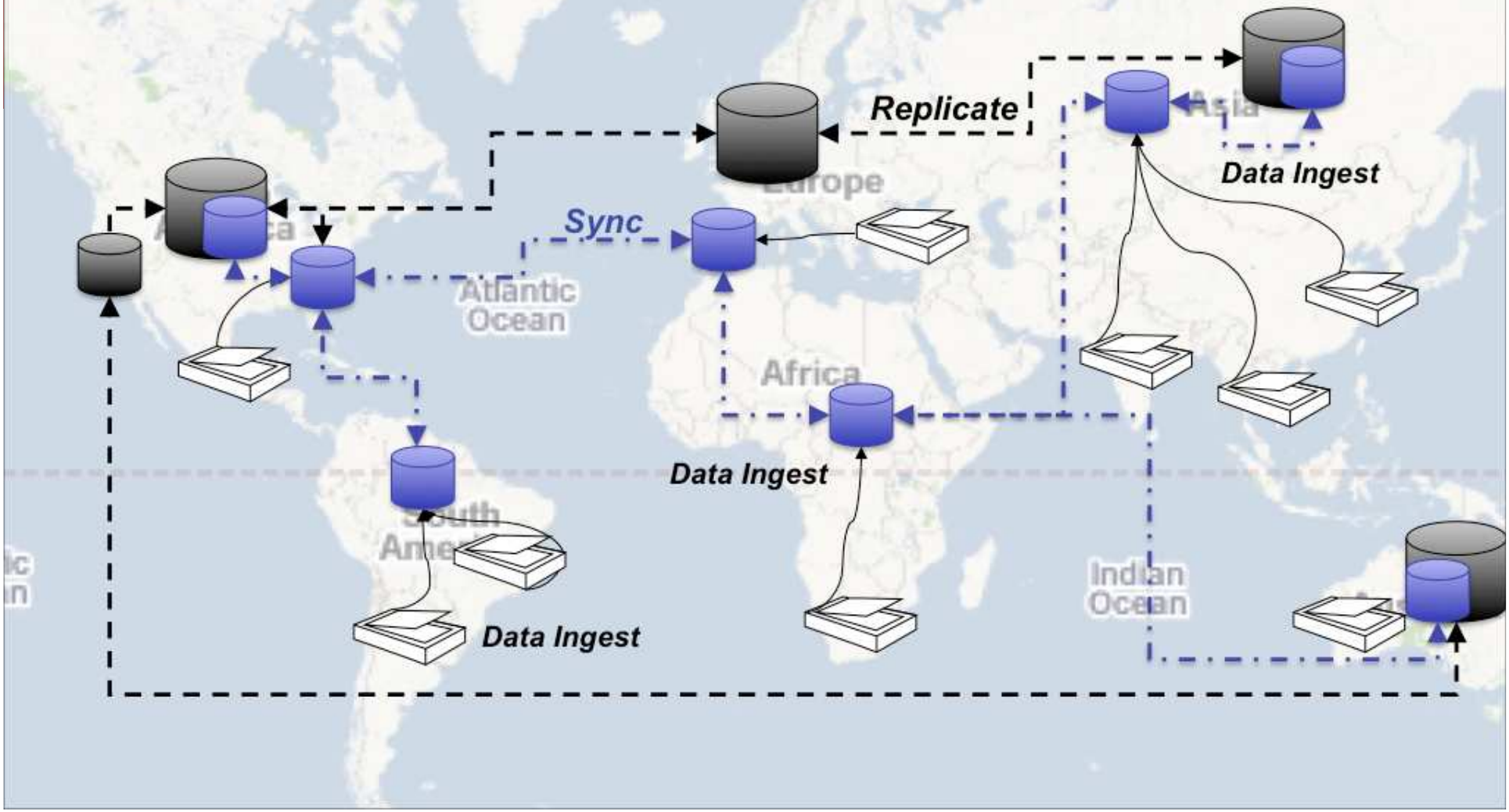
Distributed (Somewhat)

Global Acquisition and Storage

Preservation System – multiple redundant copies of all digitized content.

Access System – files, metadata & services needed to deliver content.

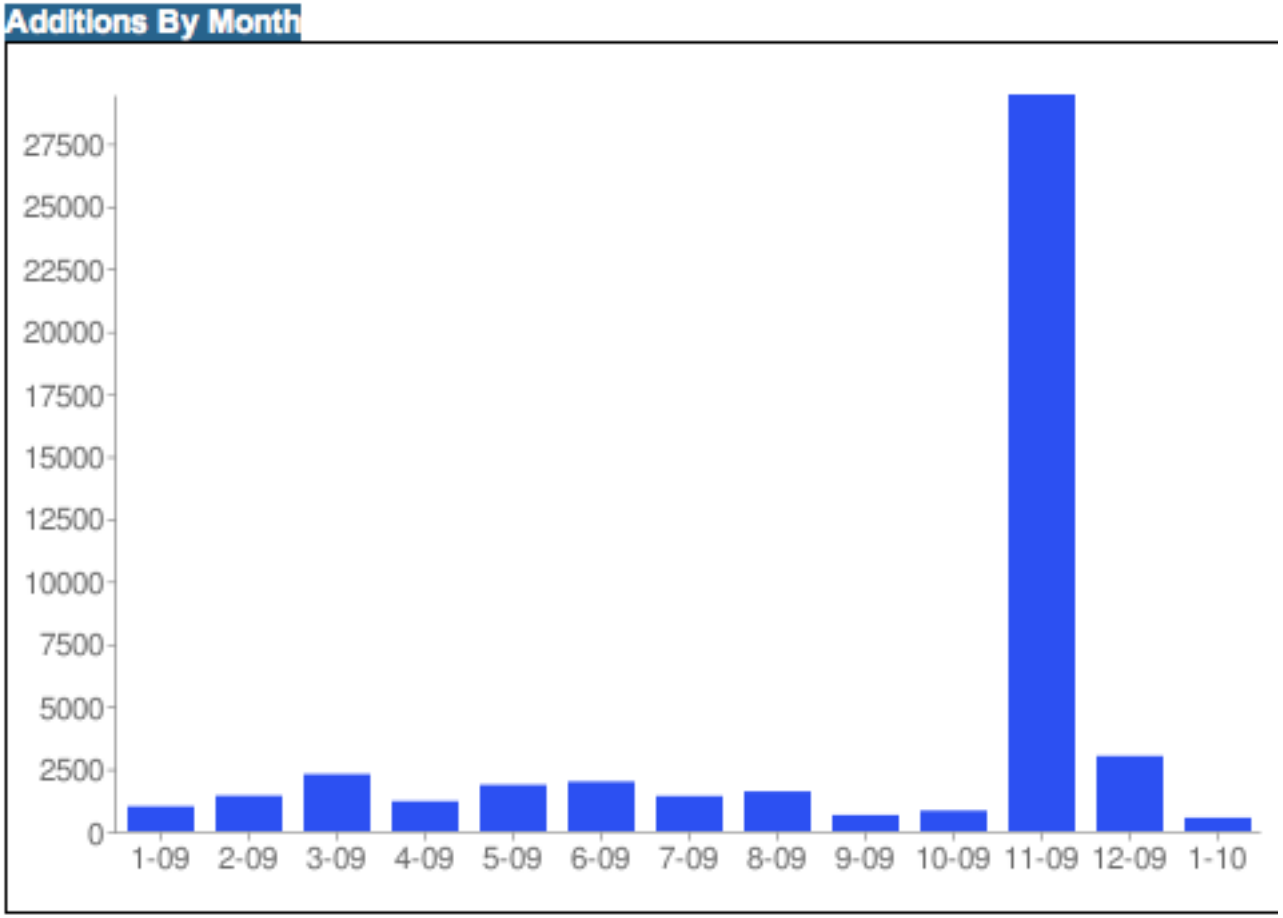
Data Providers – organizations & projects that contribute content.



Size of BHL

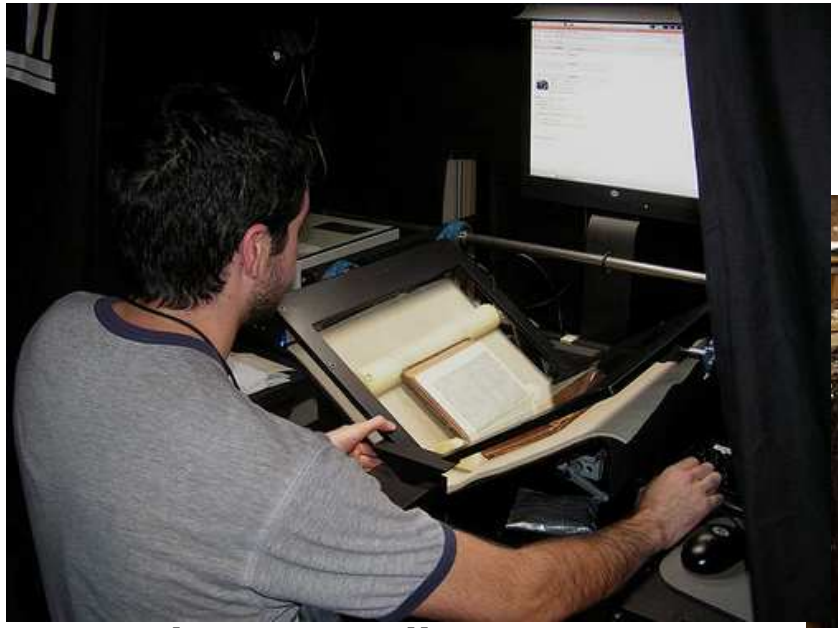
- 24TB
  & growing!

Now Online Details
38,019 titles
73,531 volumes
27,723,795 pages
Recent additions



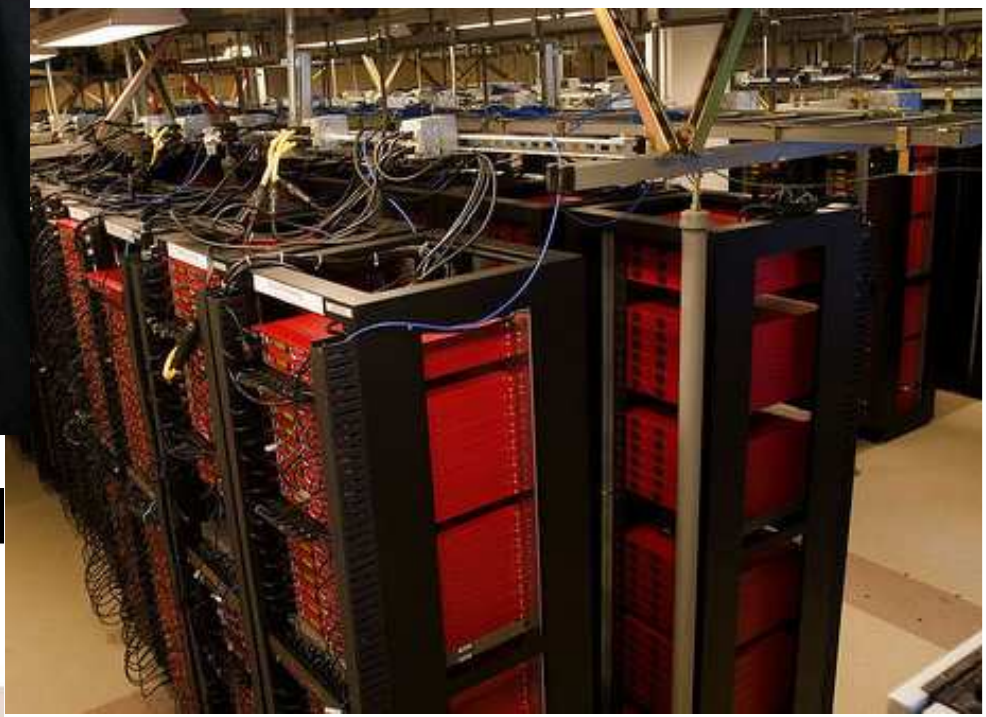
Biodiversity Heritage Library: http://biodiversitylibrary.org

Scan & Store: Internet Archive



Scanning on Scribes

Storage in Petaboxes



Biodiversity Heritage Library: http://biodiversitylibrary.org

Scanning & Derivatives

Master

- XML
- JP2

Derivatives

- PDF
- JPG
- TXT
- DJVu

PDF

OCR

JP2

XML

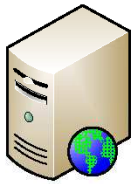
Name	Last Modified	Size	Type
Parent Directory/		-	Directory
mushroomsofameri00palm.djvu	2007-Oct-10 00:13:27	704.9K	image/vnd.djvu
mushroomsofameri00palm.gif	2007-Oct-09 23:56:53	253.9K	image/gif
mushroomsofameri00palm.pdf	2007-Oct-10 00:15:56	2.0M	application/pdf
mushroomsofameri00palm_abbyy.gz	2007-Oct-10 00:11:19	367.7K	application/octet-stream
mushroomsofameri00palm_bw.pdf	2007-Oct-10 00:36:19	1.6M	application/pdf
mushroomsofameri00palm_dc.xml	2007-Oct-04 13:19:41	0.4K	application/xml
mushroomsofameri00palm_djvu.txt	2007-Oct-10 00:36:28	24.9K	text/plain
mushroomsofameri00palm_djvu.xml	2007-Oct-10 00:11:35	225.1K	application/xml
mushroomsofameri00palm_files.xml	2008-Apr-30 16:10:51	4.4K	application/xml
mushroomsofameri00palm_flippy.zip	2007-Oct-09 23:57:09	848.8K	application/zip
mushroomsofameri00palm_jp2.zip	2007-Oct-09 23:56:36	10.5M	application/zip
mushroomsofameri00palm_marc.xml	2007-Oct-04 13:19:41	2.1K	application/xml
mushroomsofameri00palm_meta.mrc	2007-Oct-04 13:19:41	0.6K	application/octet-stream
mushroomsofameri00palm_meta.xml	2008-Feb-05 18:47:48	1.4K	application/xml
mushroomsofameri00palm_metasource.xml	2007-Oct-04 13:19:41	0.4K	application/xml
mushroomsofameri00palm_orig_jp2.tar	2007-Oct-09 19:27:16	24.2M	application/x-tar
mushroomsofameri00palm_scandata.xml	2007-Oct-09 19:27:15	24.5K	application/xml

Biodiversity Heritage Library: http://biodiversitylibrary.org

BHL Hardware Infrastructure

Web/App

MBGServ17
www.biodiversitylibrary.org
beta.biodiversitylibrary.org



MBGServ16
Images.mobot.org

DB

MBGDevServ01
BHLDevDB

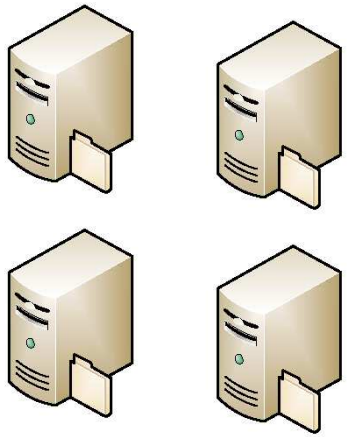
MBGServ11 (Prod)
BHLImport DB
BHL DB
Botanicus DB

FileSys

MBGServ14
30+TB

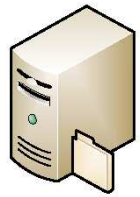
MOBOT

Petabox cluster

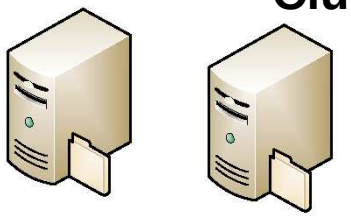


Internet Archive

Image Server



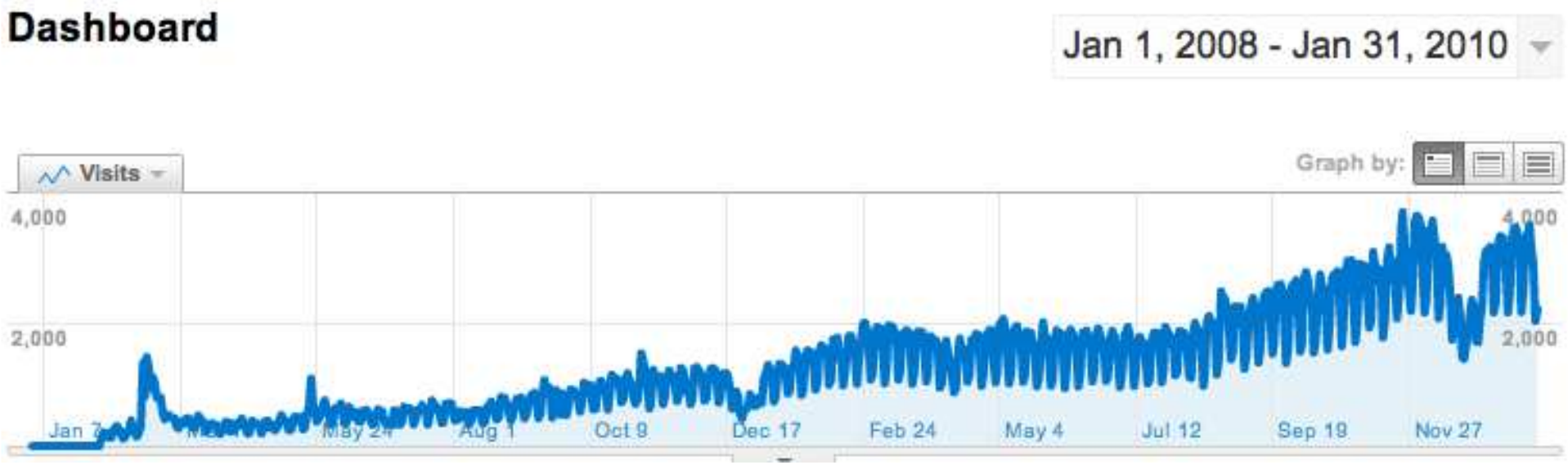
Cluster



MBL

Stats: Usage

- Daily average
  - 1,026 visitors
  - 1,680 visits / day
  - 8,200 pageviews / day



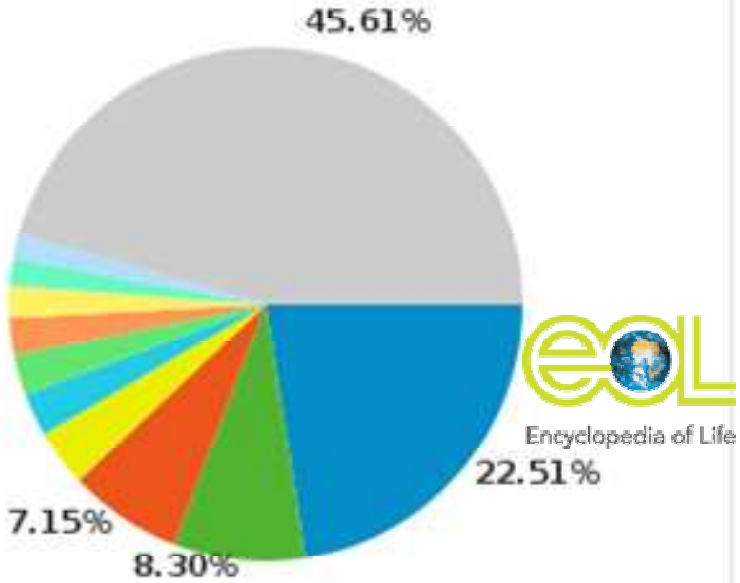
Biodiversity Heritage Library: http://biodiversitylibrary.org

Referrers

Visits	Pages/Visit	Avg. Time on Site	% New Visits	Bounce Rate
230,385	**4.90**	**00:05:32**	**56.99%**	**43.44%**
% of Site Total: 23.73%	Site Avg: 4.71 (3.98%)	Site Avg: 00:05:04 (9.26%)	Site Avg: 62.13% (-8.27%)	Site Avg: 49.74% (-12.67%)

	Source	None	Visits	Visits
1.	eol.org		51,856	22.51%
2.	rzbbx1.uni-regensburg.de		19,120	8.30%
3.	en.wikipedia.org		16,477	7.15%
4.	sil.si.edu		8,289	3.60%
5.	tropicos.org		6,192	2.69%
6.	bioone.org		5,921	2.57%
7.	zipcodezoo.com		5,446	2.36%
8.	animalbase.uni-goettingen.de		4,594	1.99%
9.	biodiversitylibrary.blogspot.com		3,892	1.69%
10.	bibdigital.rjb.csic.es		3,530	1.53%

Source contribution to total: Visits



Jan 1, 2008 – Jan 31, 2010

Biodiversity Heritage Library: http://biodiversitylibrary.org

BHL Development Methodology

Software

Systems

Technical Staff

- Chris Freeland, Technical Director
  - chris.freeland@mobot.org
- Mike Lichtenberg, Application Developer
  - mike.lichtenberg@mobot.org
- Phil Cryer, Systems Developer
  - phil.cryer@mobot.org
  - http://twitter.com/fak3r

BHL Architecture Meetings

Global, coordinated development

- New functionality from BHL-Europe
  - Improved deduplication tools
  - Semantic interface
  - OAIS-compliant preservation infrastructure
- Building a community of developers
  - Funded & volunteer
  - RubyBHL: http://github.com/mjy/rubyBHL
  - PyBHL: http://linux.softpedia.com/get/Programming/Libraries/pybhl-51612.shtml
- New partners, new content

Biodiversity Heritage Library: http://biodiversitylibrary.org

Open Software & Development

- BHL Bits:
  - Portal code, utilities, services
  - http://code.google.com/p/bhl-bits/
- Taxonomic Literature Group
  - Google Group for discussion of "taxonomic literature & the services required to make literature interoperable within biodiversity research and biodiversity informatics."
  - http://groups.google.com/group/taxonlit

Open Data

- Downloads
  - Simple tab-delimited exports of core data
  - http://www.biodiversitylibrary.org/data/BHLExportSchema.pdf

- Data model
  - DB schema as ERD
  - http://bhl-bits.googlecode.com/files/20090930_BHLDataModel.pdf

BHL 中国节点

http://bhl-china.org

BHL-China

首页 数据检索 联系我们 BHL主站 相关项目 Flickr组 数字化图书馆 用户登录

学名或中文名: 搜索

典型图书推荐

1 2 3 停止

欢迎访问BHL中国节点

由 xuzp 于 星期五, 09/18/2009 - 16:03 发表

数据状况（截止2010-1-21）：

（1）书本：267册；

（2）页码：149,444页；

（3）名称-页码词条记录：中文名-页码记录：178,668个； 拉丁名-页码词条：187,188个。

关于BHL-China

中国生物多样性历史文献图书馆（BHL-China）是由中国科学院生物多样性委员会资助的预研项目，旨在通过与BHL（**Biodiversity Heritage Library**）合作，联合其他生物学研究（院）所，共同建立BHL-China网络平台，对重要生物多样性（前期重点是植物学）文献进行全面收集、扫描、重要生物学信息提取和系统整理，建立可供查

最新blog文章

- About BHL-China
- 镜像合作、加深了解－BHL总部与BHL－China座谈成功举行
- ABBYY与图书馆合作经典案例之一
- 版权说明
- 文献数字化版权问题
- 关于BHL-China
- 联系我们
- 相关项目

更多

语言

- 简体中文
- English

搜索

Google™ 自定义搜索 搜索

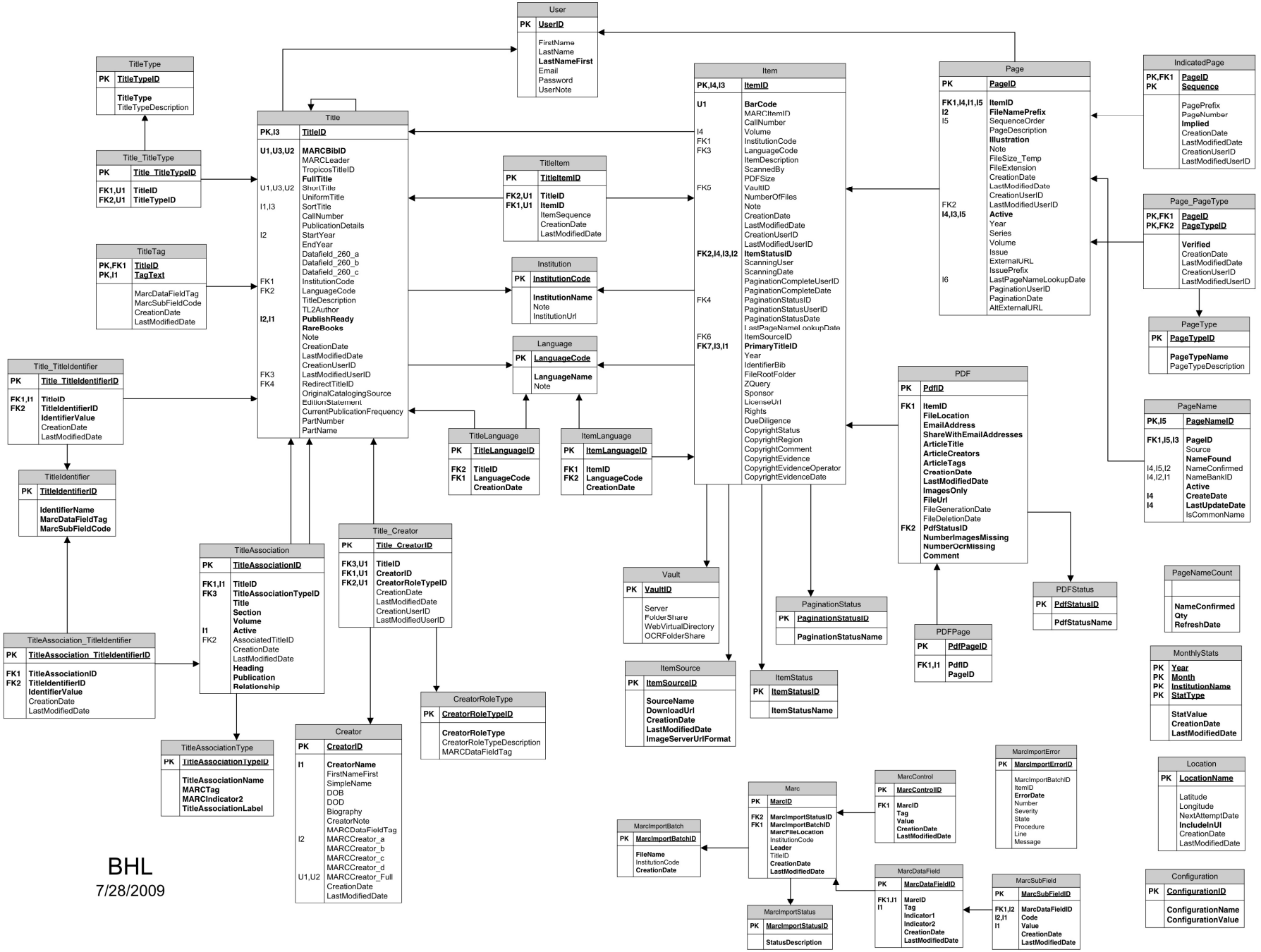
新增文献

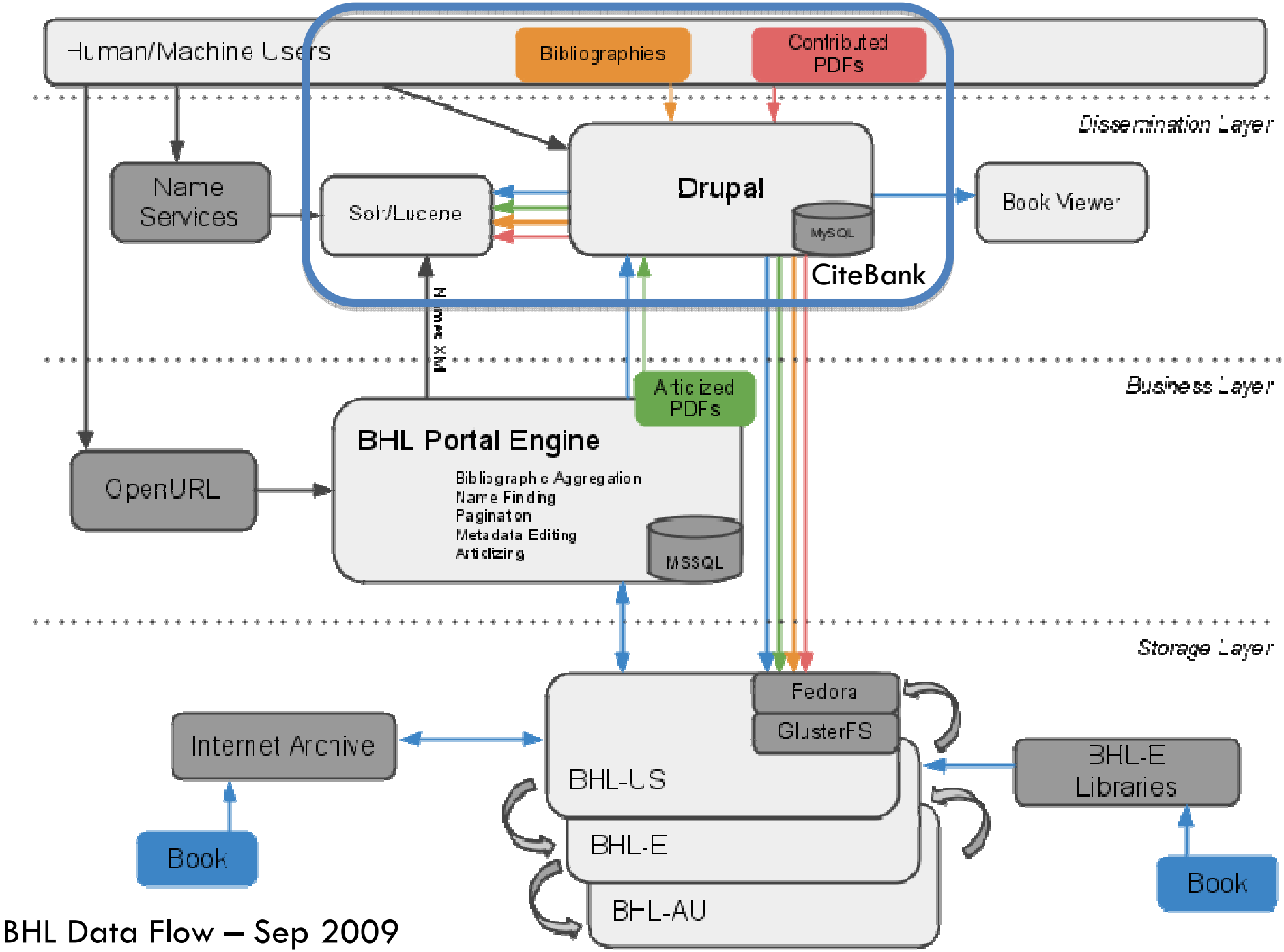
- 上海植物志
- 湖北植物志
- Musa chunii Hakkinen, a new

Data Models

Purpose-driven

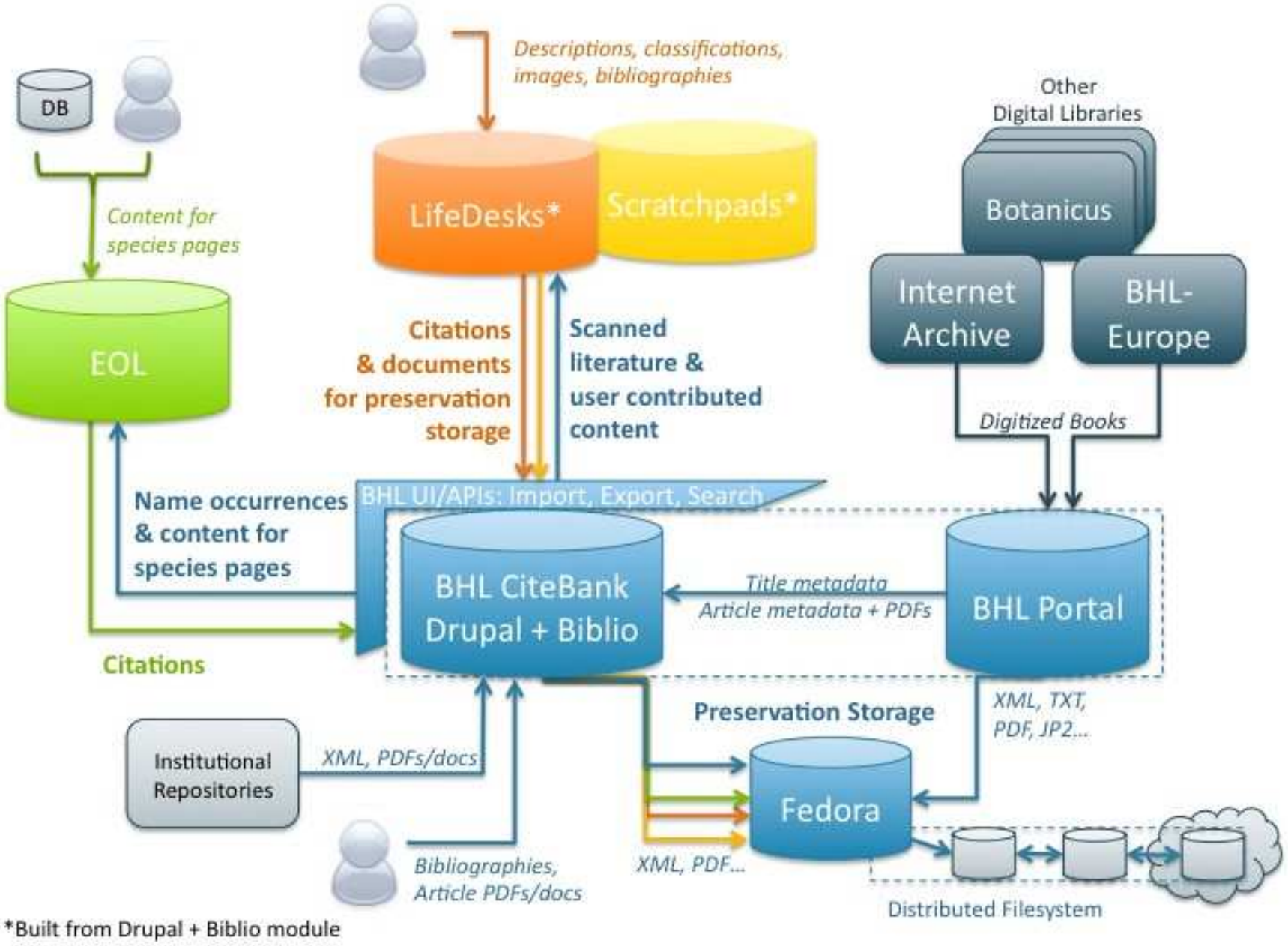
Standards


User
PK UserID
FirstName
LastName
LastNameFirst
Email
Password
UserNote
TitleType
PK TitleTypeID
TitleType
TitleTypeDescription
Title_TitleType
PK Title_TitleTypeID
FK1,U1 TitleID
FK2,U1 TitleTypeID
Title
PK,I3 TitleID
U1,U3,U2 MARCBibID
MARCLeader
TropicosTitleID
FullTitle
U1,U3,U2 ShortTitle
UniformTitle
I1,I3 SortTitle
CallNumber
PublicationDetails
I2 StartYear
EndYear
Datafield_260_a
Datafield_260_b
Datafield_260_c
FK1 InstitutionCode
FK2 LanguageCode
TitleDescription
TL2Author
I2,I1 PublishReady
RareBooks
Note
CreationDate
LastModifiedDate
CreationUserID
FK3 LastModifiedUserID
FK4 RedirectTitleID
OriginalCatalogingSource
EditionStatement
CurrentPublicationFrequency
PartNumber
PartName
TitleItem
PK TitleItemID
FK2,U1 TitleID
FK1,U1 ItemID
ItemSequence
CreationDate
LastModifiedDate
Institution
PK InstitutionCode
InstitutionName
Note
InstitutionUrl
Language
PK LanguageCode
LanguageName
Note
Item
PK,I4,I3 ItemID
U1 BarCode
MARCItemID
CallNumber
I4 Volume
FK1 InstitutionCode
FK3 LanguageCode
ItemDescription
ScannedBy
PDFSize
FK5 VaultID
NumberOfFiles
Note
CreationDate
LastModifiedDate
CreationUserID
LastModifiedUserID
FK2,I4,I3,I2 ItemStatusID
ScanningUser
ScanningDate
PaginationCompleteUserID
PaginationCompleteDate
FK4 PaginationStatusID
PaginationStatusUserID
PaginationStatusDate
LastPageNameLookupDate
FK6 ItemSourceID
FK7,I3,I1 PrimaryTitleID
Year
IdentifierBib
FileRootFolder
ZQuery
Sponsor
LicenseUrl
Rights
DueDiligence
CopyrightStatus
CopyrightRegion
CopyrightComment
CopyrightEvidence
CopyrightEvidenceOperator
CopyrightEvidenceDate
Page
PK PageID
FK1,I4,I1,I5 ItemID
I2 FileNamePrefix
I5 SequenceOrder
PageDescription
Illustration
Note
FileSize_Temp
FileExtension
CreationDate
LastModifiedDate
CreationUserID
FK2 LastModifiedUserID
I4,I3,I5 Active
Year
Series
Volume
Issue
ExternalURL
IssuePrefix
I6 LastPageNameLookupDate
PaginationUserID
PaginationDate
AltExternalURL
IndicatedPage
PK,FK1 PageID
PK Sequence
PagePrefix
PageNumber
Implied
CreationDate
LastModifiedDate
CreationUserID
LastModifiedUserID
Page_PageType
PK,FK1 PageID
PK,FK2 PageTypeID
Verified
CreationDate
LastModifiedDate
CreationUserID
LastModifiedUserID
PageType
PK PageTypeID
PageTypeName
PageTypeDescription
PageName
PK,I5 PageNameID
FK1,I5,I3 PageID
Source
NameFound
I4,I5,I2 NameConfirmed
I4,I2,I1 NameBankID
Active
I4 CreateDate
I4 LastUpdateDate
IsCommonName
PageNameCount
NameConfirmed
Qty
RefreshDate
TitleTag
PK,FK1 TitleID
PK,I1 TagText
MarcDataFieldTag
MarcSubFieldCode
CreationDate
LastModifiedDate
Title_TitleIdentifier
PK Title_TitleIdentifierID
FK1,I1 TitleID
FK2 TitleIdentifierID
IdentifierValue
CreationDate
LastModifiedDate
TitleIdentifier
PK TitleIdentifierID
IdentifierName
MarcDataFieldTag
MarcSubFieldCode
TitleLanguage
PK TitleLanguageID
FK2 TitleID
FK1 LanguageCode
CreationDate
ItemLanguage
PK ItemLanguageID
FK1 ItemID
FK2 LanguageCode
CreationDate
PDF
PK PdfID
FK1 ItemID
FileLocation
EmailAddress
ShareWithEmailAddresses
ArticleTitle
ArticleCreators
ArticleTags
CreationDate
LastModifiedDate
ImagesOnly
FileUrl
FileGenerationDate
FileDeletionDate
FK2 PdfStatusID
NumberImagesMissing
NumberOcrMissing
Comment
PDFStatus
PK PdfStatusID
PdfStatusName
PDFPage
PK PdfPageID
FK1,I1 PdfID
PageID
PaginationStatus
PK PaginationStatusID
PaginationStatusName
Vault
PK VaultID
Server
FolderShare
WebVirtualDirectory
OCRFolderShare
ItemSource
PK ItemSourceID
SourceName
DownloadUrl
CreationDate
LastModifiedDate
ImageServerUrlFormat
ItemStatus
PK ItemStatusID
ItemStatusName
TitleAssociation
PK TitleAssociationID
FK1,I1 TitleID
FK3 TitleAssociationTypeID
Title
Section
Volume
I1 Active
FK2 AssociatedTitleID
CreationDate
LastModifiedDate
Heading
Publication
Relationship
Title_Creator
PK Title_CreatorID
FK3,U1 TitleID
FK1,U1 CreatorID
FK2,U1 CreatorRoleTypeID
CreationDate
LastModifiedDate
CreationUserID
LastModifiedUserID
CreatorRoleType
PK CreatorRoleTypeID
CreatorRoleTypeDescription
MARCDataFieldTag
TitleAssociation_TitleIdentifier
PK TitleAssociation_TitleIdentifierID
FK1 TitleAssociationID
FK2 TitleIdentifierID
IdentifierValue
CreationDate
LastModifiedDate
TitleAssociationType
PK TitleAssociationTypeID
TitleAssociationName
MARCTag
MARCIndicator2
TitleAssociationLabel
Creator
PK CreatorID
I1 CreatorName
FirstNameFirst
SimpleName
DOB
DOD
Biography
CreatorNote
MARCDataFieldTag
I2 MARCCreator_a
MARCCreator_b
MARCCreator_c
MARCCreator_d
U1,U2 MARCCreator_Full
CreationDate
LastModifiedDate
Marc
PK MarcID
FK2 MarcImportStatusID
FK1 MarcImportBatchID
MarcFileLocation
InstitutionCode
Leader
TitleID
CreationDate
LastModifiedDate
MarcImportBatch
PK MarcImportBatchID
FileName
InstitutionCode
CreationDate
MarcImportStatus
PK MarcImportStatusID
StatusDescription
MarcControl
PK MarcControlID
FK1 MarcID
Tag
Value
CreationDate
LastModifiedDate
MarcDataField
PK MarcDataFieldID
FK1,I1 MarcID
I1 Tag
Indicator1
Indicator2
CreationDate
LastModifiedDate
MarcSubField
PK MarcSubFieldID
FK1,I2 MarcDataFieldID
I2,I1 Code
I1 Value
CreationDate
LastModifiedDate
MarcImportError
PK MarcImportErrorID
MarcImportBatchID
ItemID
ErrorDate
Number
Severity
State
Procedure
Line
Message
MonthlyStats
PK Year
PK Month
PK InstitutionName
PK StatType
StatValue
CreationDate
LastModifiedDate
Location
PK LocationName
Latitude
Longitude
NextAttemptDate
IncludeInUI
CreationDate
LastModifiedDate
Configuration
PK ConfigurationID
ConfigurationName
ConfigurationValue
BHL
7/28/2009



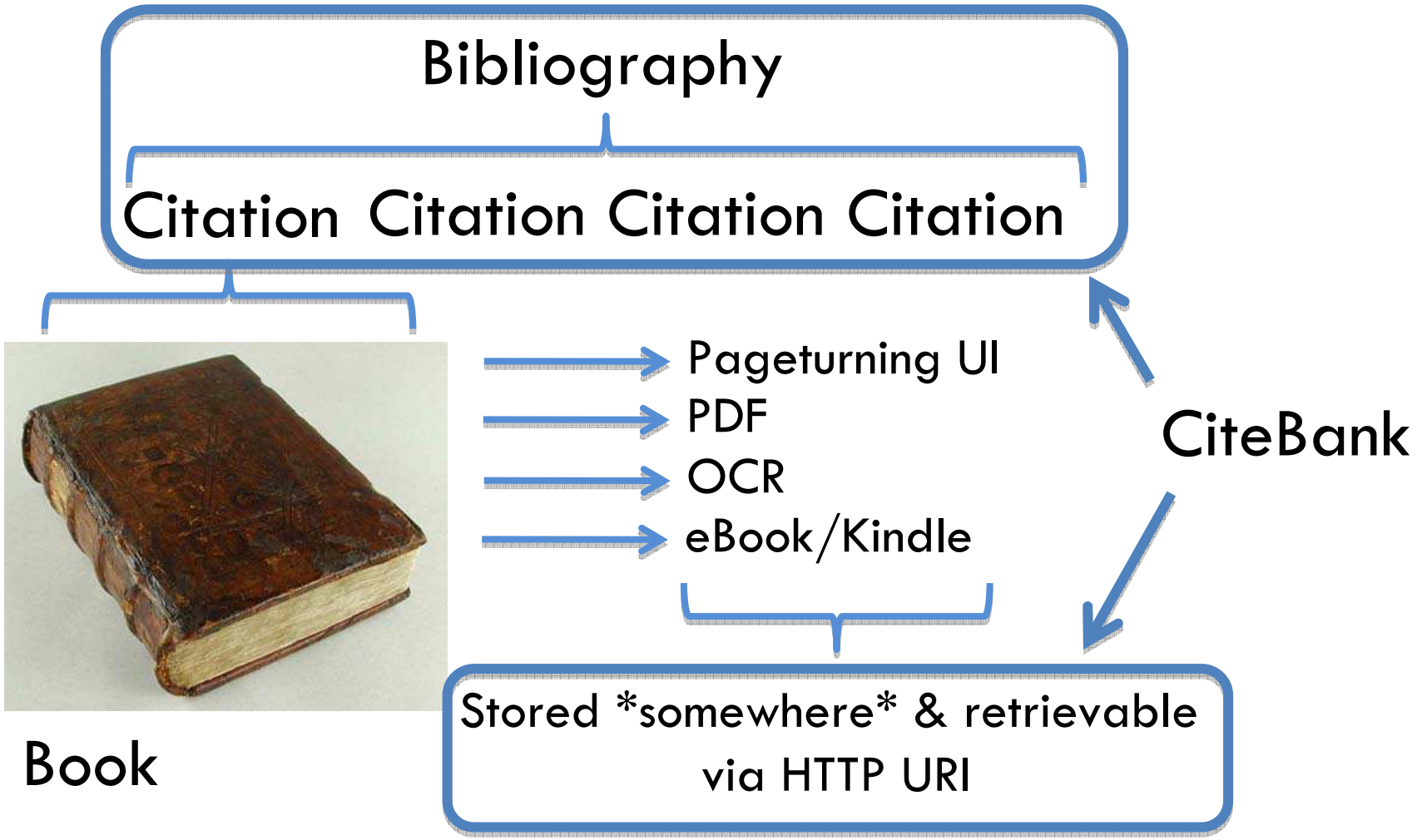
BHL Data Flow – Sep 2009

Bibliographic metadata sharing between BHL and EOL/LifeDesks/Scratchpads



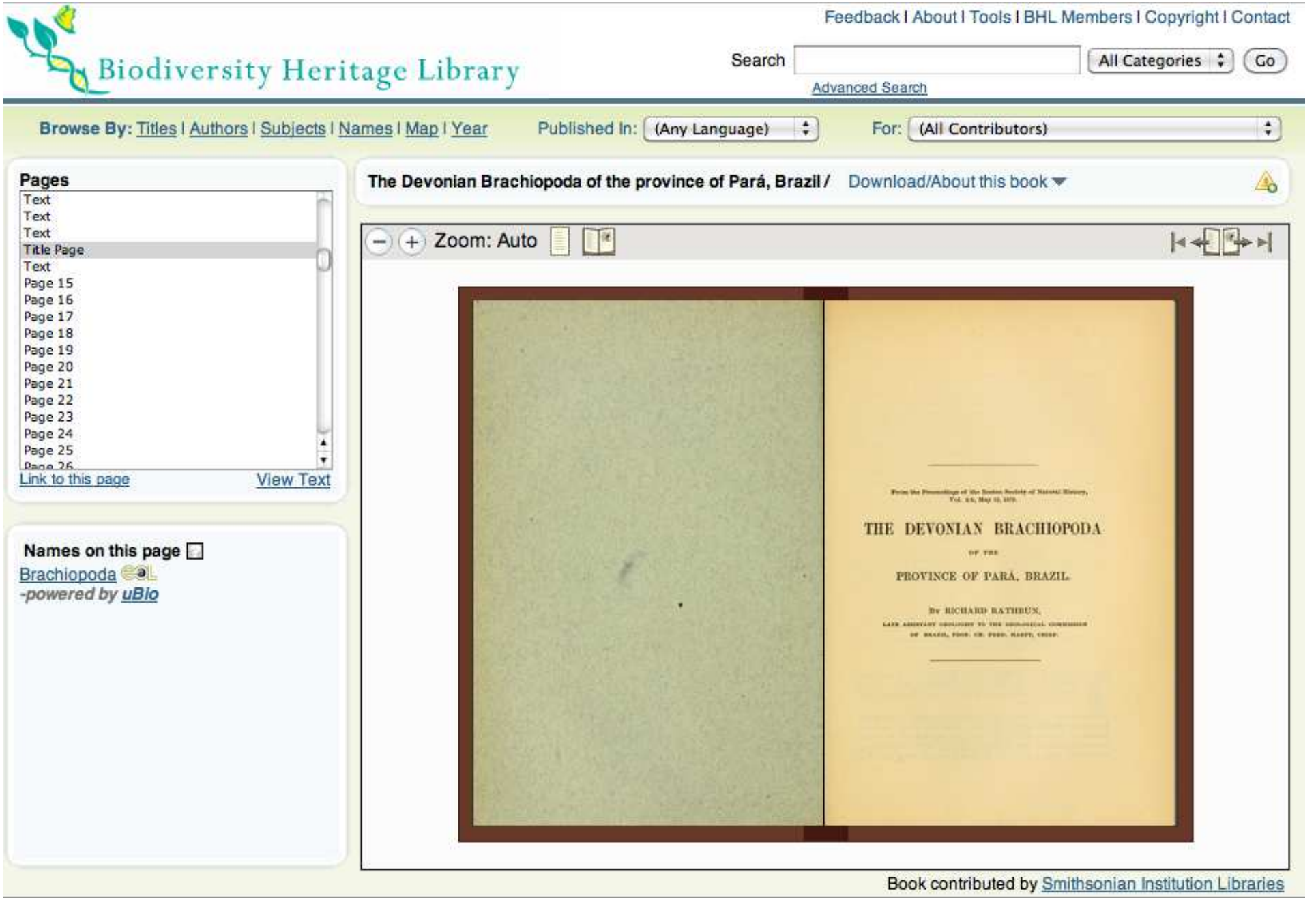
*Built from Drupal + Biblio module

CiteBank boundaries



Biodiversity Heritage Library: http://biodiversitylibrary.org

BHL Portal: User Interface(s)



http://www.biodiversitylibrary.org/item/54491

UI Features

- Blog integration
- Open Source Page Turning application
- OCR text behind the image
- Subject heading tag cloud
- Geocoding subject headings for map-based inquiry
- Taxonomic Name Finding
  - Links to EOL
- Discovered Bibliographies

Tools

Deduplication & Bidding

Admin Interfaces

“Paginator”

“Articlizing”

Complexities of distributed, mass scanning

Flora of Colorado,
by Per Axel Rydberg *from NYBG*

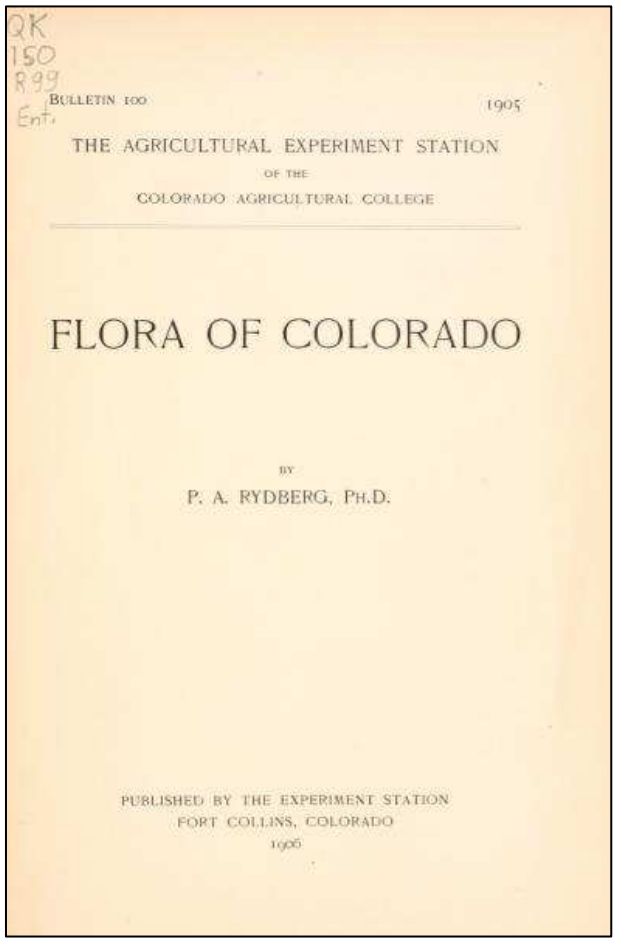
☆☆☆☆☆
(not yet rated)

Type:	Book; English
Publisher:	Fort Collins, Experiment Station, 1906.
Editions:	3 Editions
OCLC:	13927851
Related Subjects:	Plants -- Colorado.

Flora of Colorado
by Per Axel Rydberg *from Smithsonian*

☆☆☆☆☆
(not yet rated)

Type:	Book : State or province government publication; English
Publisher:	Fort Collins, Colo. : Agricultural Experiment Station of the Colorado Agricultural College, 1906.
Editions:	3 Editions
OCLC:	1577518
Related Subjects:	Botany -- Colorado.



Biodiversity Heritage Library: http://biodiversitylibrary.org

Articles

http://www.biodiversitylibrary.org/item/17298

THE

BIOLOGICAL BULLETIN

PUBLISHED BY

THE MARINE BIOLOGICAL LABORATORY

Associate Editors

James A. Black, ENSR Marine & Coastal Center, Woods Hole

Lucy E. Burnett, Grice Marine Biological Laboratory, College of Charleston

William D. Cohen, Hunter College, City University of New York

Charles D. Derby, Georgia State University

Shinya Inoué, Marine Biological Laboratory

Rudolf A. Raff, Indiana University

Editorial Board

Peter B. Armstrong, University of California, Davis

Andrew R. Cameron, California Institute of Technology

Thomas H. Dietz, Louisiana State University

Richard B. Emlet, Oregon Institute of Marine Biology, University of Oregon

David Epel, Hopkins Marine Station, Stanford University

Daphne Gail Fautin, University of Kansas

William F. Gilly, Hopkins Marine Station, Stanford University

Roger T. Hanlon, Marine Biological Laboratory

Gregory Hinkle, University of Massachusetts, Dartmouth

Makoto Kobayashi, Hiroshima University of Economics

Michael LaBarbera, University of Chicago

Donal T. Manahan, University of Southern California

Margaret McFall-Ngai, Kewalo Marine Laboratory, University of Hawaii

Mark W. Miller, Institute of Neurobiology, University of Puerto Rico

Tatsuo Motokawa, Tokyo Institute of Technology

Yoshitaka Nagahama, National Institute for Basic Biology, Japan

Sherry D. Painter, Marine Biomedical Institute, University of Texas Medical Branch

K. Ranga Rao, University of West Florida

Baruch Rinkevich, Israel Oceanographic & Limnological Research Ltd.

Richard Strathmann, Friday Harbor Laboratories, University of Washington

Steven Vogel, Duke University

J. Herbert Waite, University of Delaware

Sarah Ann Woodin, University of South Carolina

Richard K. Zimmer-Faust, University of California, Los Angeles

Generate a PDF (Step 1 of 2) BETA

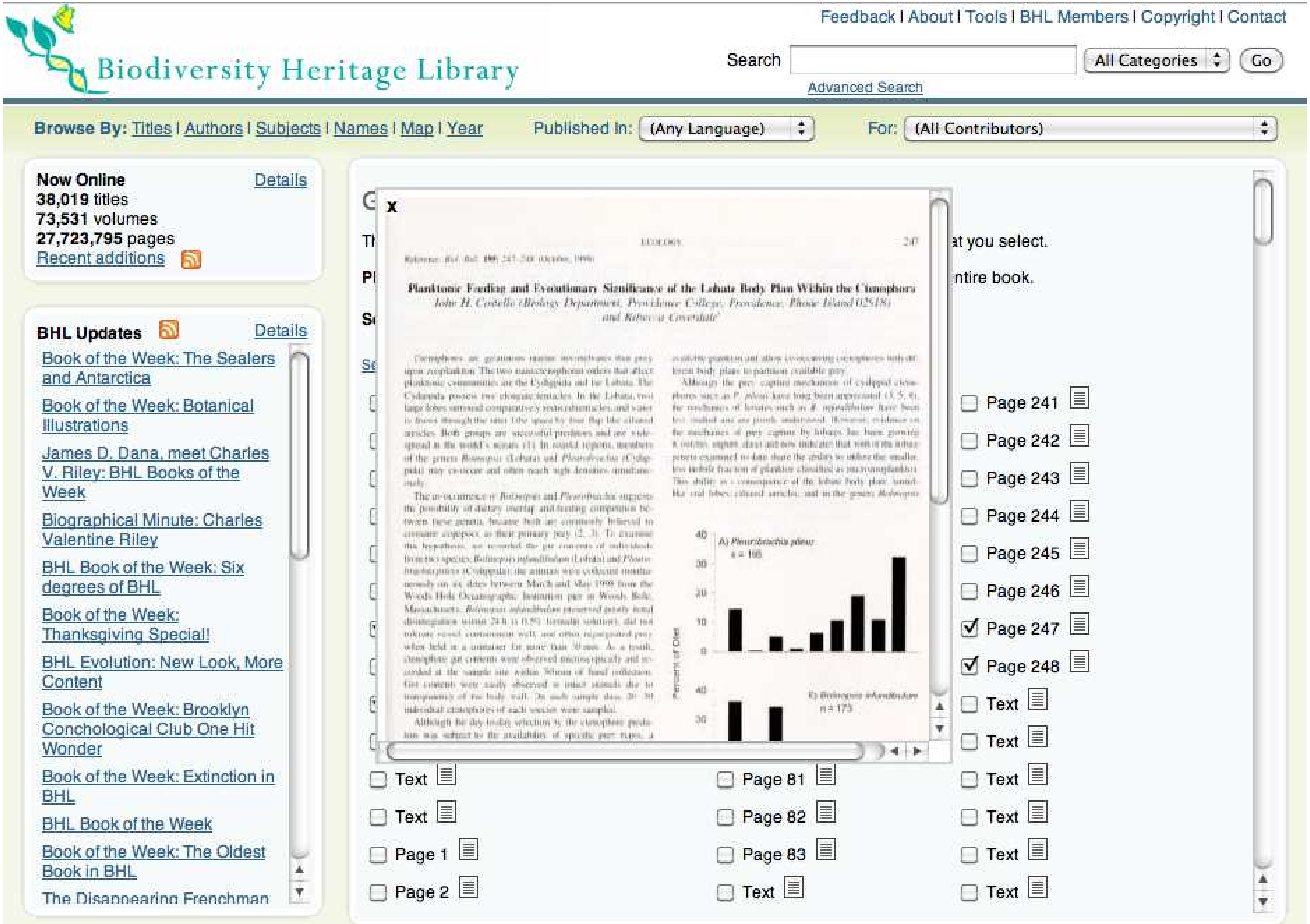
This page allows you to generate a downloadable PDF from only the pages that you select.

PDF generation is limited to 50 pages. Click here to download a PDF of the entire book.

Select page images Next

Select All Clear All

- [ ] Cover
- [ ] Text
- [ ] Text
- [ ] Text
- [ ] Text
- [ ] Text
- [x] Title Page
- [ ] Text
- [x] Table of Contents
- [ ] Text
- [ ] Text
- [ ] Text
- [ ] Page 1
- [ ] Page 2
- [ ] Page 71
- [ ] Page 72
- [ ] Page 73
- [ ] Page 74
- [ ] Page 75
- [ ] Page 76
- [ ] Page 77
- [ ] Page 78
- [ ] Page 79
- [ ] Page 80
- [ ] Page 81
- [ ] Page 82
- [ ] Page 83
- [ ] Text
- [ ] Page 241
- [ ] Page 242
- [ ] Page 243
- [ ] Page 244
- [ ] Page 245
- [ ] Page 246
- [x] Page 247
- [x] Page 248
- [ ] Text
- [ ] Text
- [ ] Text
- [ ] Text
- [ ] Text
- [ ] Text


Feedback | About | Tools | BHL Members | Copyright | Contact
Biodiversity Heritage Library
Search
All Categories
Go
Advanced Search
Browse By: Titles | Authors | Subjects | Names | Map | Year
Published In: (Any Language)
For: (All Contributors)
Now Online
Details
38,019 titles
73,531 volumes
27,723,795 pages
Recent additions
BHL Updates
Details
Book of the Week: The Sealers and Antarctica
Book of the Week: Botanical Illustrations
James D. Dana, meet Charles V. Riley: BHL Books of the Week
Biographical Minute: Charles Valentine Riley
BHL Book of the Week: Six degrees of BHL
Book of the Week: Thanksgiving Special!
BHL Evolution: New Look, More Content
Book of the Week: Brooklyn Conchological Club One Hit Wonder
Book of the Week: Extinction in BHL
BHL Book of the Week
Book of the Week: The Oldest Book in BHL
at you select.
ntire book.
Page 241
Page 242
Page 243
Page 244
Page 245
Page 246
Page 247
Page 248
Text
Text
Text
Text
Text
Text
Text
Text
Page 1
Page 2
Page 81
Page 82
Page 83
Text
ECOLOGY
247
A) Pleurobrachia pileus
Percent of Diet

Generate a PDF (Step 2 of 2) BETA

Previous Done

Please enter the email address to which you would like a link to the PDF to be sent. If you want to share the link with others, enter their email addresses as well.

Email Address: chris.freeland@mobot.org

Share With:

Include: ○ Page images only ◉ Page images and OCR text

Are you generating a PDF containing the text of a single journal article or book chapter? If so, please help us out by providing the following information!

Article/Chapter Title: Planktonic feeding and evolutionary significance of the lobate body plan withi

Author(s): J. Costello, Coverdale R.

Subject(s): Enter subjects, separated by commas

Done

PDF Generation Stats

Summary Statistics

Status	# Of PDFs	With OCR	With Article Info	With Missing Images	With Missing OCR
Processing	1	0	0	0	0
Generated	20514	2688	8303	388	74
Declined	13	7	4	0	0
Total	**20528**	**2695**	**8307**	**388**	**74**

Expanded Statistics

Year	Week	Status	# Of PDFs	PDFs With OCR	PDFs With Article Info	PDFs Missing Images	PDFs Missing OCR	Total Missing Images	Total Missing OCR	Avg Minutes To Generate
2010	6	Generated	87	4	15	0	0	0	0	9.57
2010	5	Processing	1	0	0	0	0	0	0	0
2010	5	Generated	792	74	276	23	1	187	37	17.2
2010	4	Generated	565	66	207	23	0	38	0	9.6
2010	3	Generated	613	84	245	13	3	67	93	9.08
2010	3	Declined	1	0	0	0	0	0	0	0
2010	2	Generated	487	58	200	19	1	210	15	8.25
2010	1	Generated	128	0	54	2	0	8	0	7.33
2009	53	Generated	306	21	111	2	1	5	1	8.98
2009	52	Generated	528	44	208	3	0	6	0	11.71
2009	51	Generated	896	47	549	3	1	3	2	14.96
2009	50	Generated	648	79	179	3	2	25	74	10.1
2009	49	Generated	636	69	197	12	9	58	410	8.72
2009	48	Generated	656	76	291	9	12	17	593	16.34
2009	47	Generated	652	53	276	5	2	41	56	14.52
2009	47	Declined	1	0	1	0	0	0	0	0

Storage

Internet Archive

BHL clusters

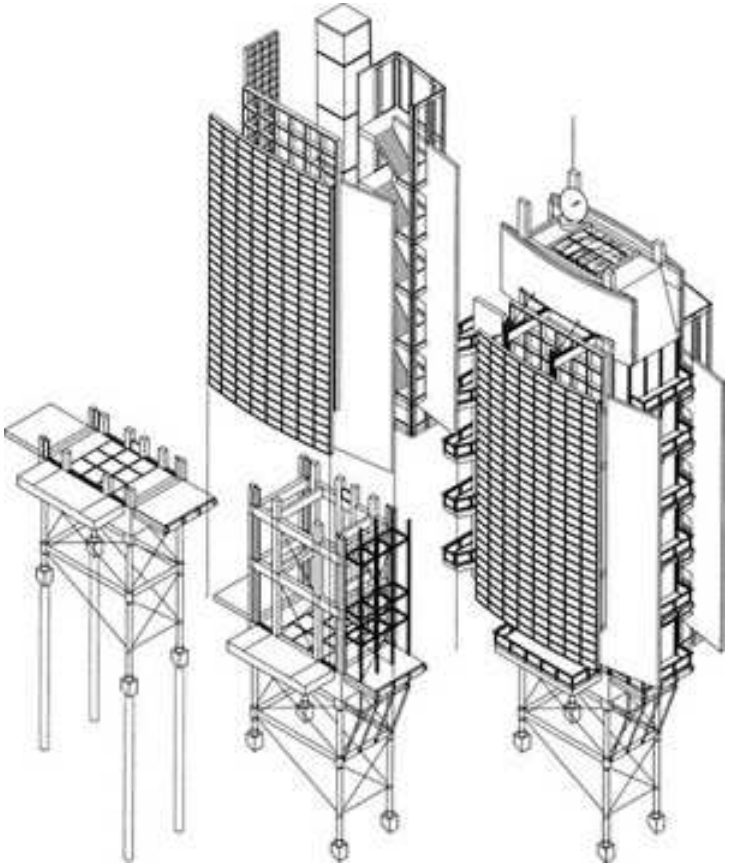
Cloud

Slides courtesy Phil Cryer, BHL Developer

http://twitter.com/fak3r

> BHL hardware architecture

- storage issues and solutions
- hardware risks and configuration
- cloud computing possibilities
- code and communication



Biodiversity Heritage Library: http://biodiversitylibrary.org

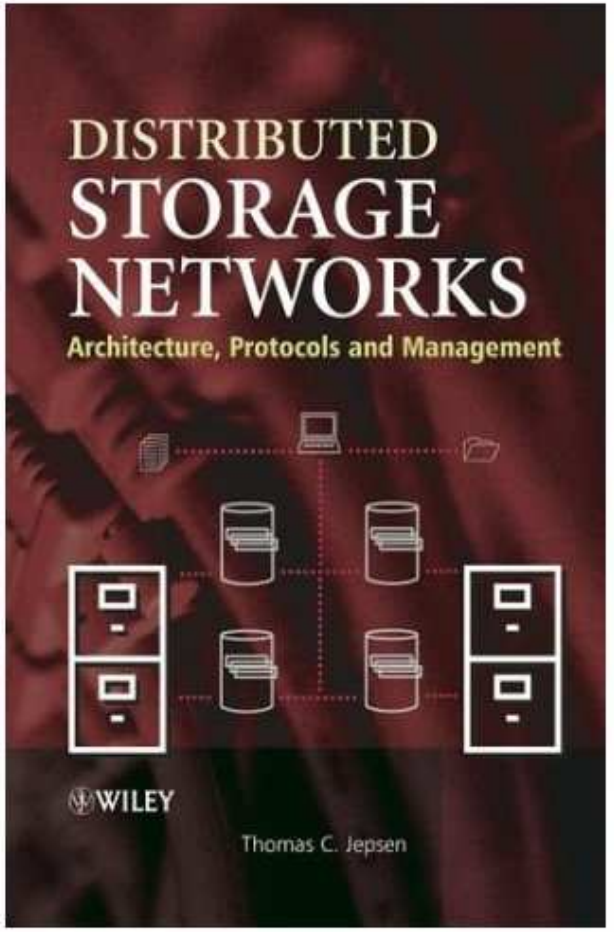
> Storage issues

- more data being created and saved
- expanding SANs can be expensive
- storage has not kept up with Moore's Law
- backups are usually for disaster recovery, not redundancy or failover

Biodiversity Heritage Library: http://biodiversitylibrary.org

> Storage solutions: distributed storage

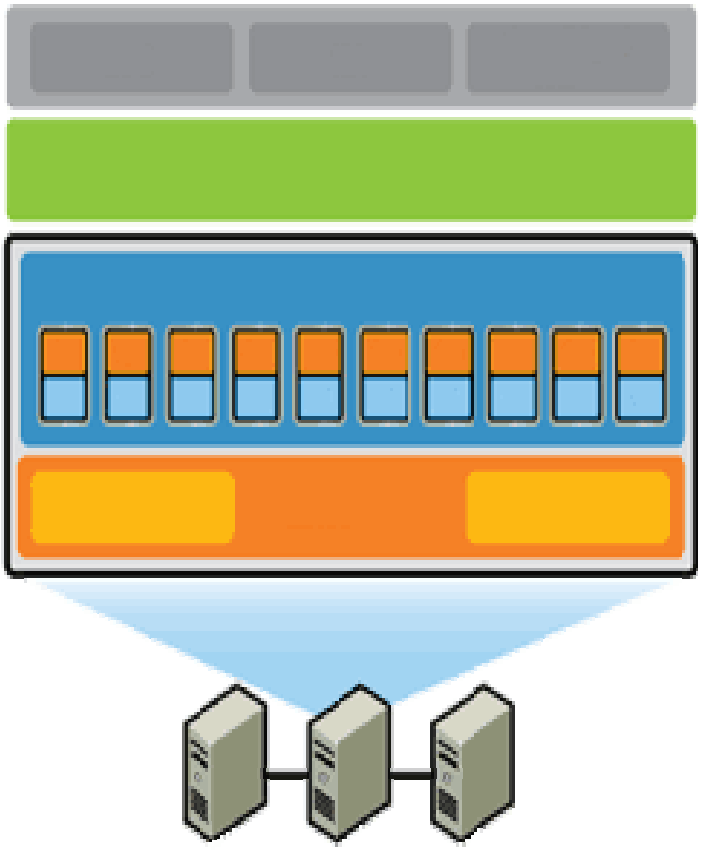
- write once, read anywhere
- **replication and fault tolerance**
- error correction
- automatic redundancy
- scalable horizontally



Biodiversity Heritage Library: http://biodiversitylibrary.org

> Distributed storage - Options

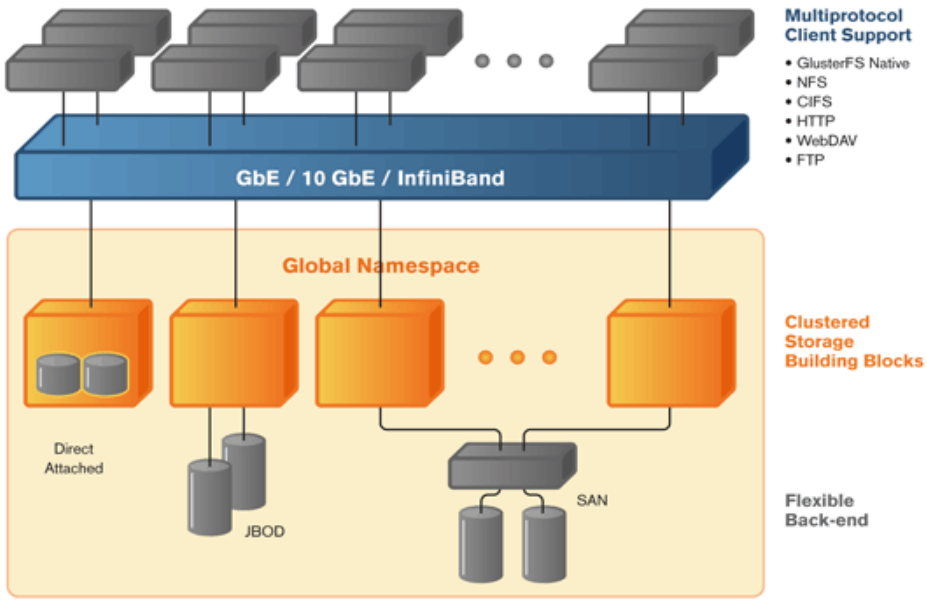
- hosted storage (cloud, Amazon S3)
- self hosted with proprietary hardware and software (Sun Thumper)
- self hosted with commodity hardware and open source distributed filesystem (GlusterFS, Lustre)



Biodiversity Heritage Library: http://biodiversitylibrary.org

> Distributed storage - GlusterFS

- GlusterFS: a cluster file-system capable of scaling to several petabytes*
- open source software on commodity hardware
- simple to install and manage
- very customizable
- offers seamless expansion



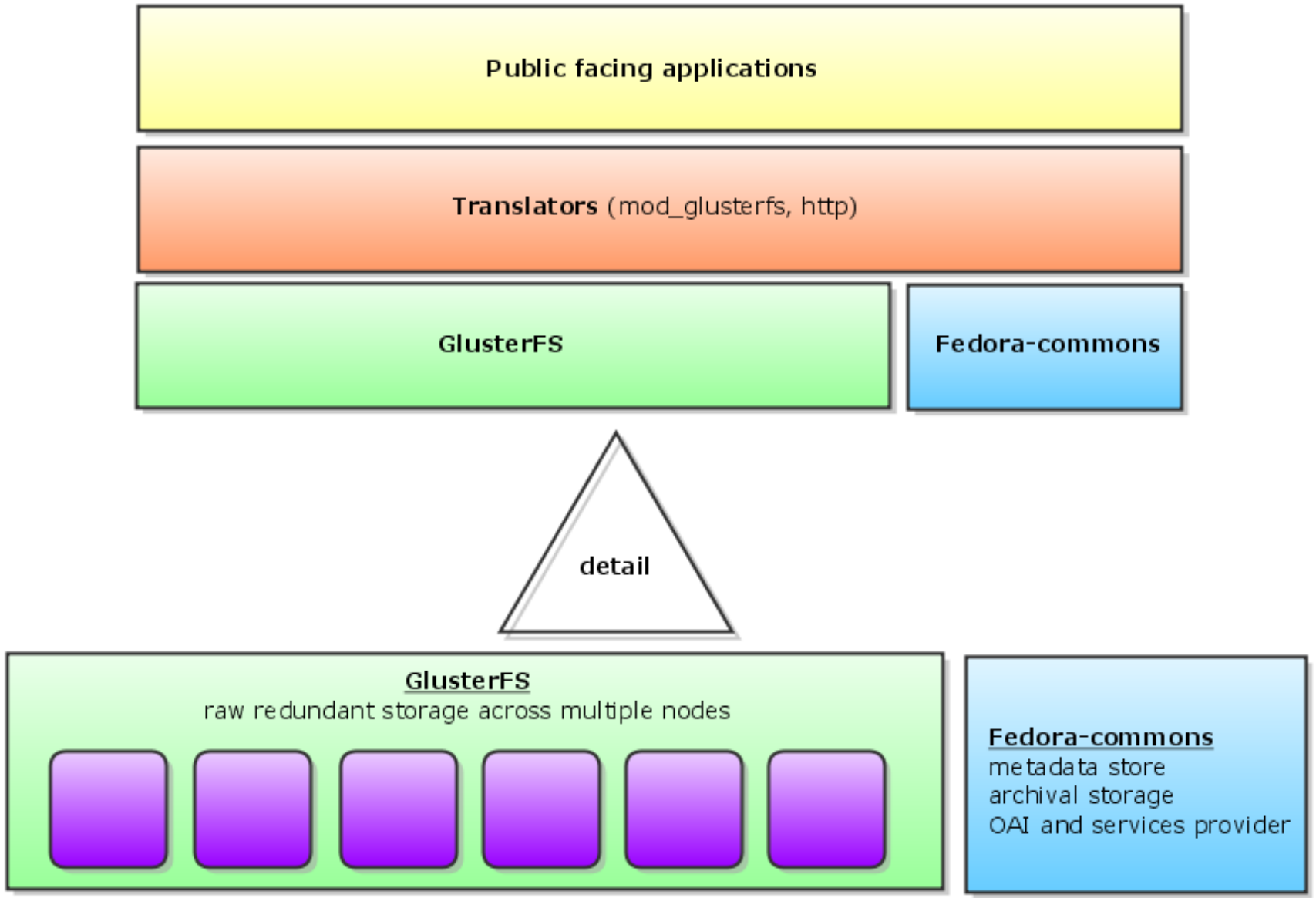
*(1,024 terabytes = 1 petabyte)

Biodiversity Heritage Library: http://biodiversitylibrary.org

> Distributed storage - Archival

- Fedora-commons is an open source repository (not a Linux distribution)
- accounts for all changes, so it provides built-in version control
- provides disaster recover
- open standards to mesh with future file formats
- provides open sharing services such as OAI-PMH

Biodiversity Heritage Library: http://biodiversitylibrary.org


Public facing applications
Translators (mod_glusterfs, http)
GlusterFS
Fedora-commons
detail
GlusterFS
raw redundant storage across multiple nodes
Fedora-commons
metadata store
archival storage
OAI and services provider

> Distributed storage - Mirrored data

- now we have redundancy
- in fact, **multiple redundant copies**
- provides fault tolerance
- offers load balancing
- **gives us future geographical distribution**

Biodiversity Heritage Library: http://biodiversitylibrary.org

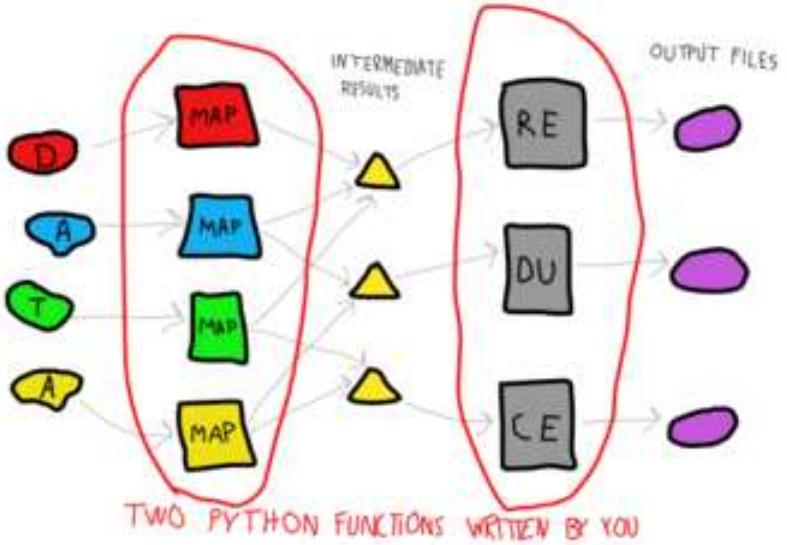
> Distributed processing

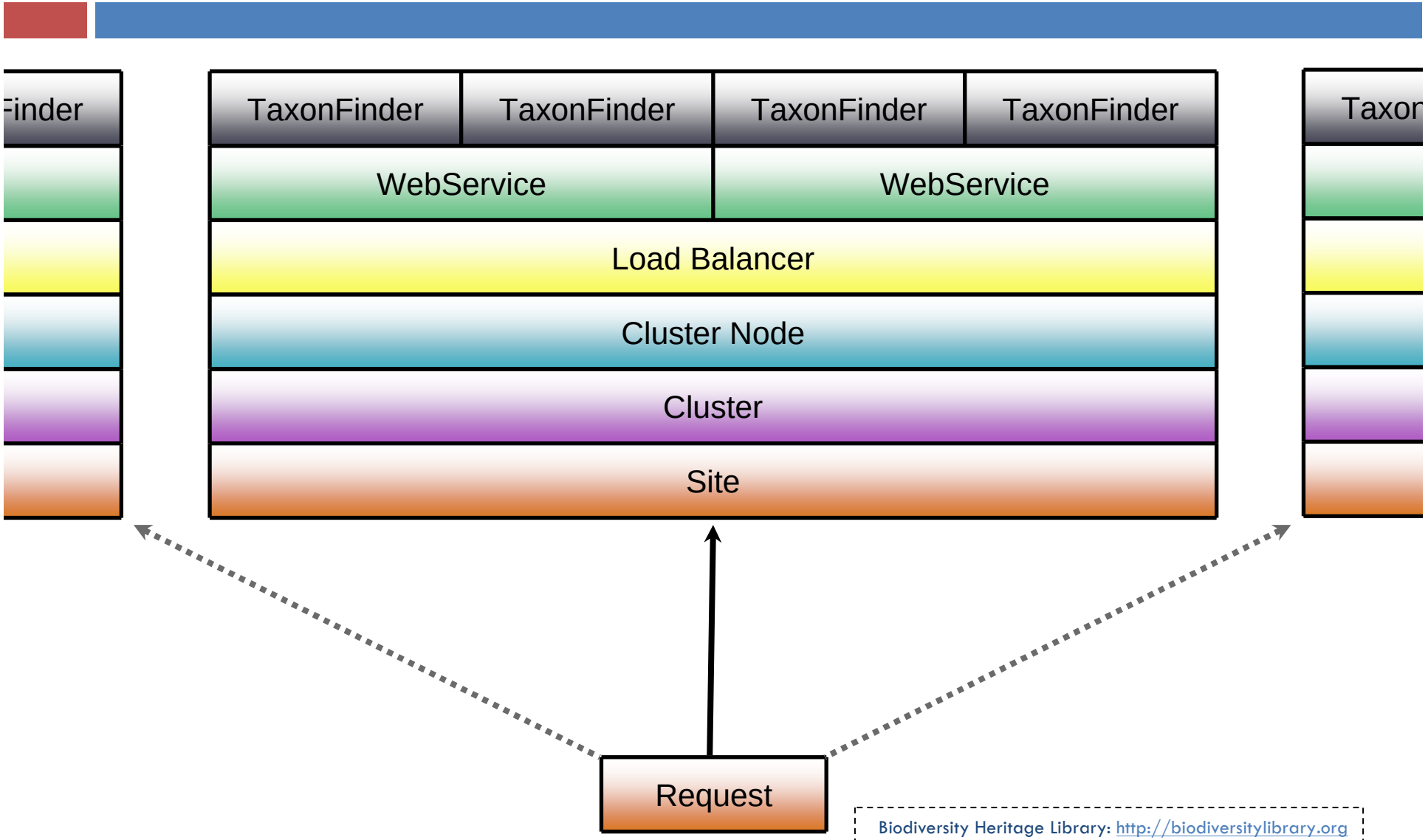
- more abilities available than just storing data
- with distributed storage comes distributed processing
- distributed processing means faster answers
- faster answers mean new questions
- lather, rinse, repeat

Biodiversity Heritage Library: http://biodiversitylibrary.org

> Distributed processing

- make existing data more useful
- image and OCR re-processing
- distributed web services
- identifier resolution pools
- map/reduce frameworks
- generate new visualizations, text mining, NLP and ???





Biodiversity Heritage Library: http://biodiversitylibrary.org

> Hardware configuration: risks

- computers crash
- hard drives die
- networks fail
- natural disasters occur

but...

Biodiversity Heritage Library: http://biodiversitylibrary.org

...so plan for it.

Biodiversity Heritage Library: http://biodiversitylibrary.org

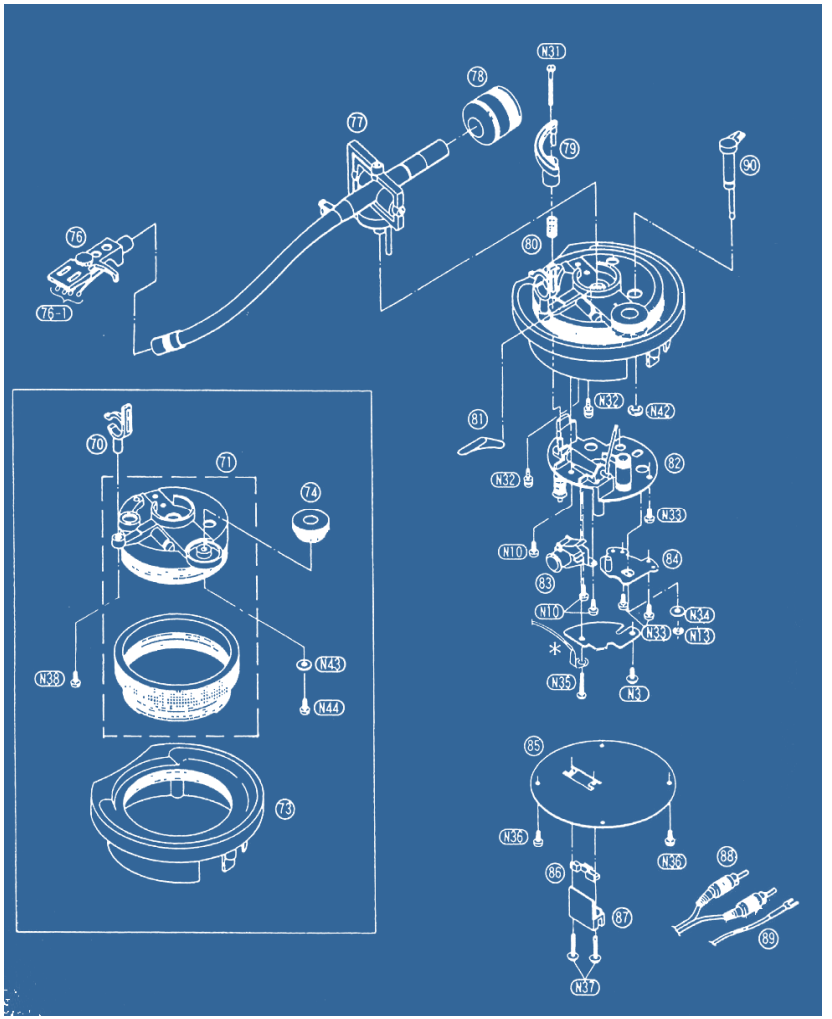
> Hardware configurations

- six 5U sized servers hosted at MBL in Woods Hole
- 8 Gigs RAM in each server
- 24 / 1.5 TB drives in each server
- 100 TB storage overall
- fast connection, far greater bandwidth than ever before
- mirror sets (3 and 3) will be stored in different locations

but...

> Some assembly required (optional)

- while our example uses new, faster commodity hardware...
- it could run on any hardware that can run Linux
- you could **chain together old "out dated" computers**
- build your own cluster for next to nothing (host it in your basement)
- solves some infrastructure funding issues and provides a working proof of concept
- hardware vendor neutrality



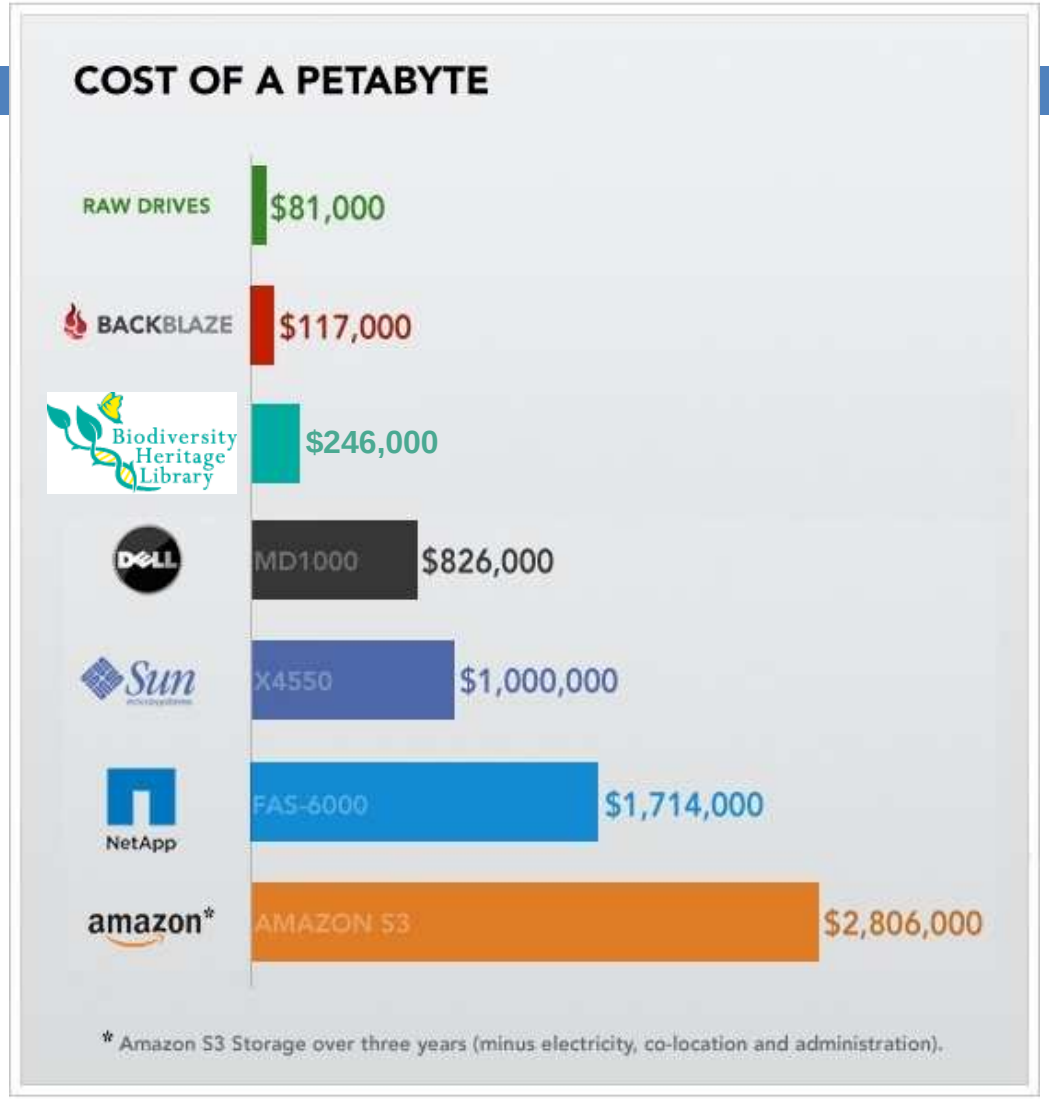
Biodiversity Heritage Library: http://biodiversitylibrary.org

> Our proof of concept

- we ran a six box cluster to demonstrate GlusterFS clustered filesystem
- ran Debian/GNU Linux
- simulated hardware failures
- synced data with a remote cluster (STL to MBL/WH)
- ran map/reduce jobs (distributed computing)
- defined procedures, configurations and build scripts

Biodiversity Heritage Library: http://biodiversitylibrary.org

> Distributed storage - Projected costs



Graph from Backblaze (http://www.backblaze.com)

Biodiversity Heritage Library: http://biodiversitylibrary.org

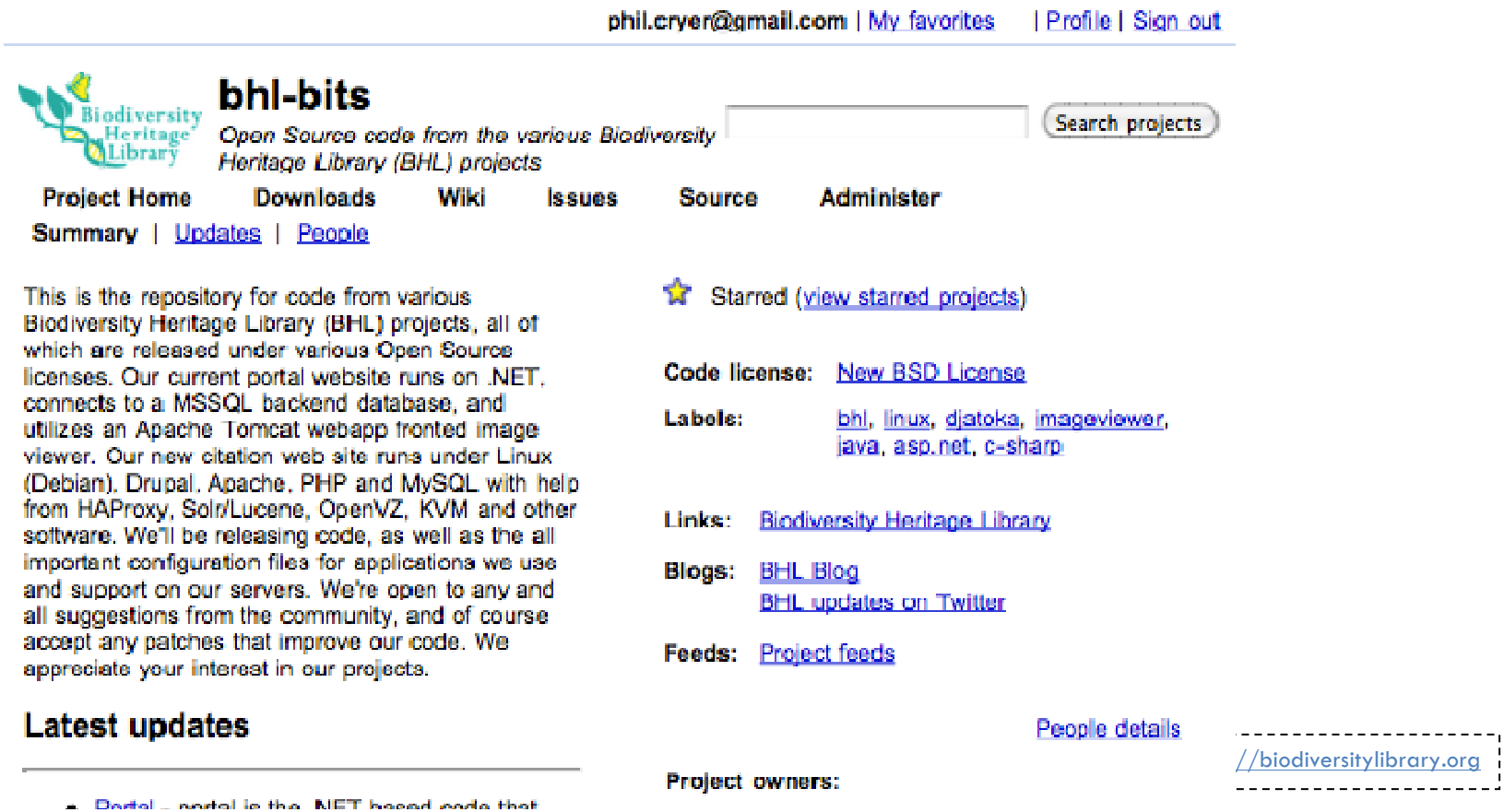
> Cloud computing

- BHL is participating in a pilot for Duraspace with The New York Public Library
- Duraspace would provide a link to cloud providers
- pilot to show feasibility of hosting
- testing use of image server, other services in the cloud
- cloud could seed new clusters

> Code (63 6f 64 65)

- our **code** and **configurations** are open source, hosted on Google Code

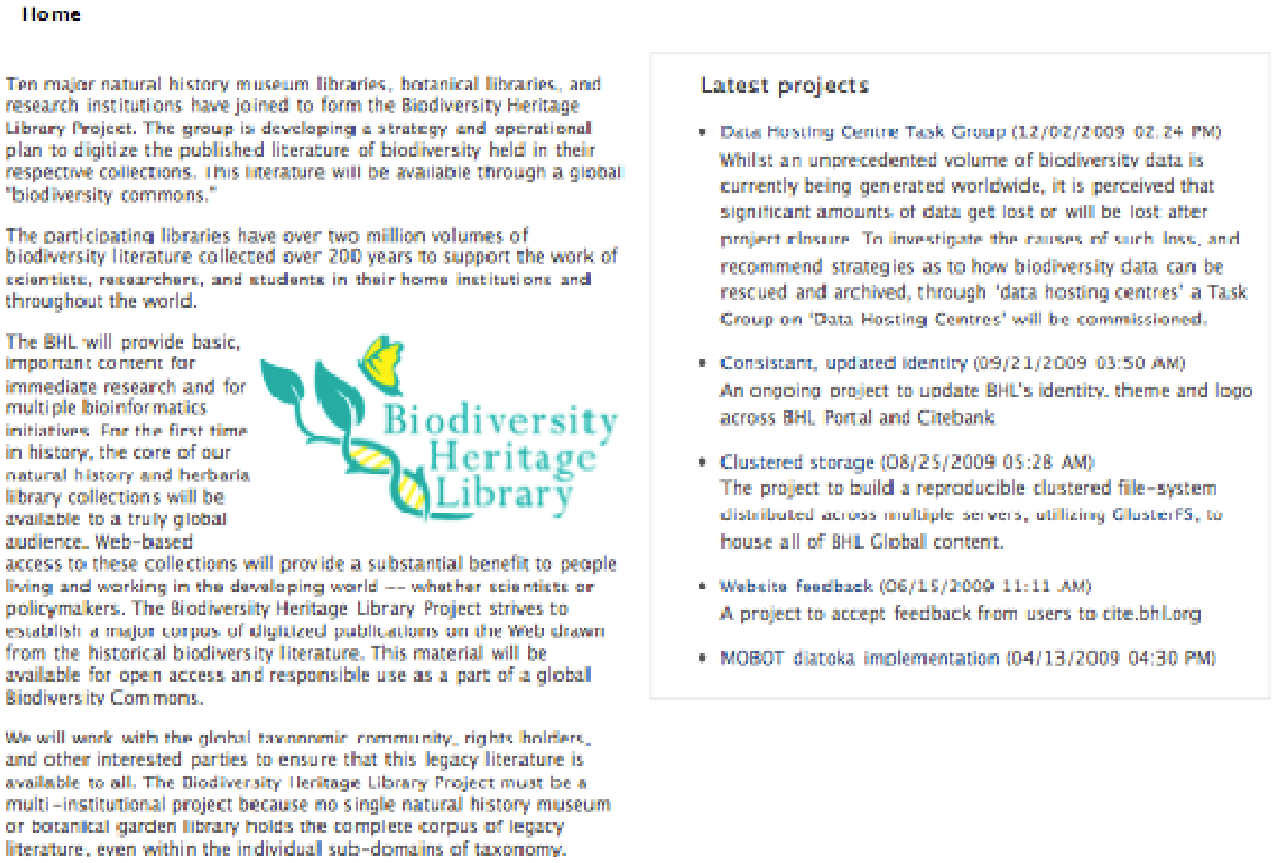
http://code.google.com/p/bhl-bits/



> Communication

- our **Projects** server is a portal for various projects

http://projects.biodiversitylibrary.org/



Home

Ten major natural history museum libraries, botanical libraries, and research institutions have joined to form the Biodiversity Heritage Library Project. The group is developing a strategy and operational plan to digitize the published literature of biodiversity held in their respective collections. This literature will be available through a global "biodiversity commons."

The participating libraries have over two million volumes of biodiversity literature collected over 200 years to support the work of scientists, researchers, and students in their home institutions and throughout the world.

The BHL will provide basic, important content for immediate research and for multiple bioinformatics initiatives. For the first time in history, the core of our natural history and herbaria library collections will be available to a truly global audience. Web-based access to these collections will provide a substantial benefit to people living and working in the developing world -- whether scientists or policymakers. The Biodiversity Heritage Library Project strives to establish a major corpus of digitized publications on the Web drawn from the historical biodiversity literature. This material will be available for open access and responsible use as a part of a global Biodiversity Commons.

We will work with the global taxonomic community, rights holders, and other interested parties to ensure that this legacy literature is available to all. The Biodiversity Heritage Library Project must be a multi-institutional project because no single natural history museum or botanical garden library holds the complete corpus of legacy literature, even within the individual sub-domains of taxonomy.

Latest projects

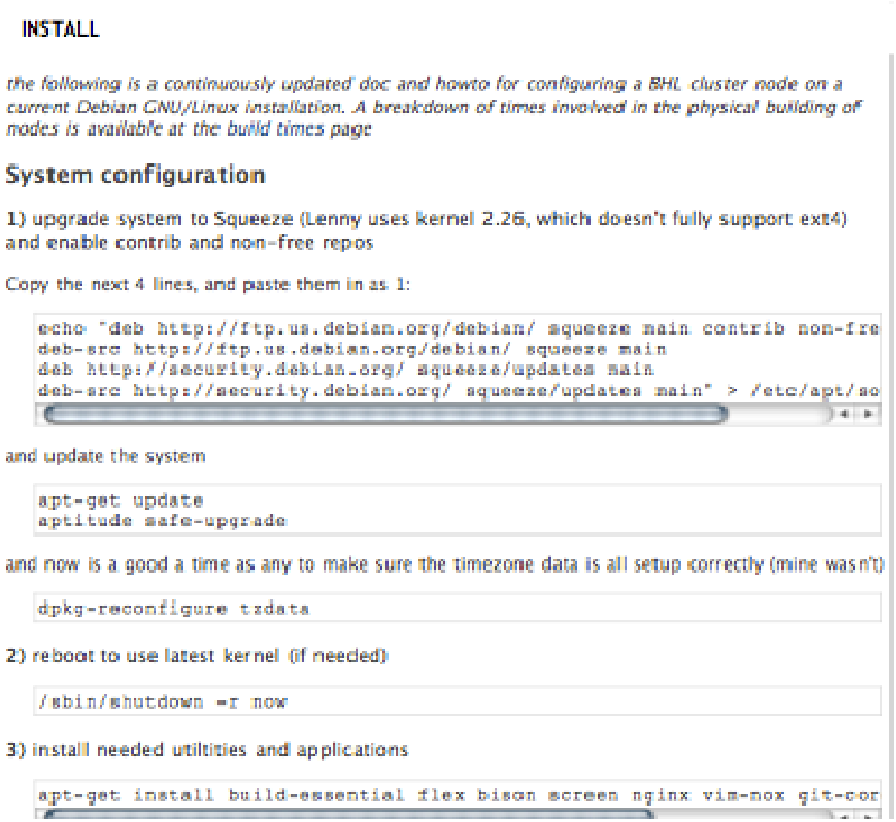
- Data Hosting Centre Task Group (12/02/2009 02:24 PM)
  Whilst an unprecedented volume of biodiversity data is currently being generated worldwide, it is perceived that significant amounts of data get lost or will be lost after project closure. To investigate the causes of such loss, and recommend strategies as to how biodiversity data can be rescued and archived, through 'data hosting centres' a Task Group on 'Data Hosting Centres' will be commissioned.
- Consistant, updated identity (09/21/2009 03:50 AM)
  An ongoing project to update BHL's identity, theme and logo across BHL Portal and Citebank
- Clustered storage (08/25/2009 05:28 AM)
  The project to build a reproducible clustered file-system distributed across multiple servers, utilizing GlusterFS, to house all of BHL Global content.
- Website feedback (06/15/2009 11:11 AM)
  A project to accept feedback from users to cite.bhl.org
- MOBOT diatoka implementation (04/13/2009 04:30 PM)

biodiversitylibrary.org

> Communication

- each project has its own wiki, with detailed, how-to instructions

  http://projects.biodiversitylibrary.org/wiki/clusteredfilesystem/



INSTALL

the following is a continuously updated doc and howto for configuring a BHL cluster node on a current Debian GNU/Linux installation. A breakdown of times involved in the physical building of nodes is available at the build times page

System configuration

1) upgrade system to Squeeze (Lenny uses kernel 2.26, which doesn't fully support ext4) and enable contrib and non-free repos

Copy the next 4 lines, and paste them in as 1:

```
echo "deb http://ftp.us.debian.org/debian/ squeeze main contrib non-fre
deb-src http://ftp.us.debian.org/debian/ squeeze main
deb http://security.debian.org/ squeeze/updates main
deb-src http://security.debian.org/ squeeze/updates main" > /etc/apt/so
```

and update the system

```
apt-get update
aptitude safe-upgrade
```

and now is a good a time as any to make sure the timezone data is all setup correctly (mine wasn't)

```
dpkg-reconfigure tzdata
```

2) reboot to use latest kernel (if needed)

```
/sbin/shutdown -r now
```

3) install needed utilities and applications

```
apt-get install build-essential flex bison screen nginx vim-nox git-cor
```

biodiversitylibrary.org

> Code (63 6f 64 65)

- our code and configurations are open source and hosted on Google Code
- our projects server shares detailed instructions to get you get started
- get involved
- join our mailing-lists (bhl-tech, bhl-bits)
- ask us questions
- you can do it, and we will help

Phil Cryer, Open Source Development Lead

code http://code.google.com/p/bhl-bits
projects http://projects.biodiversitylibrary.org

questions phil.cryer@mobot.org

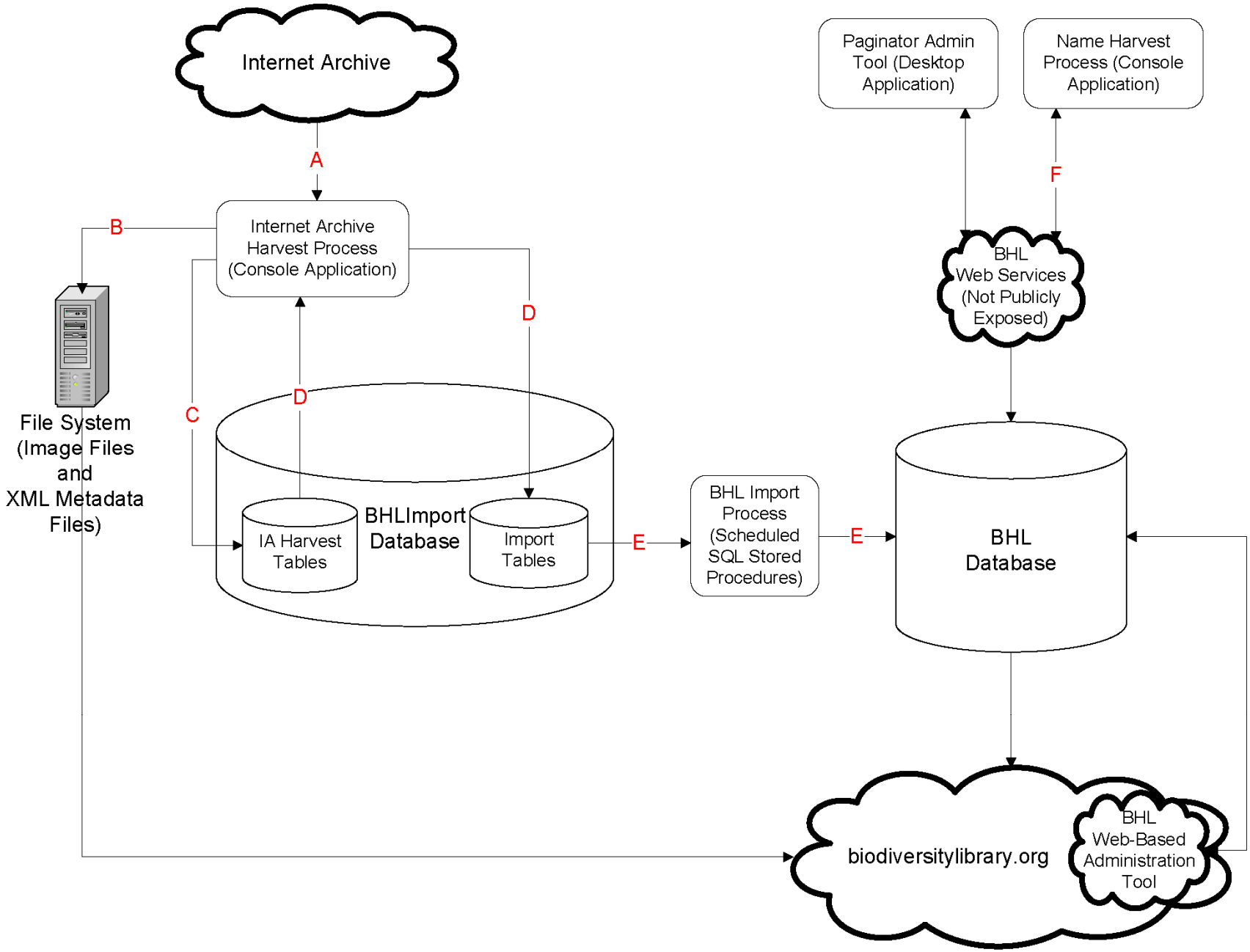
Workflow / Ingest

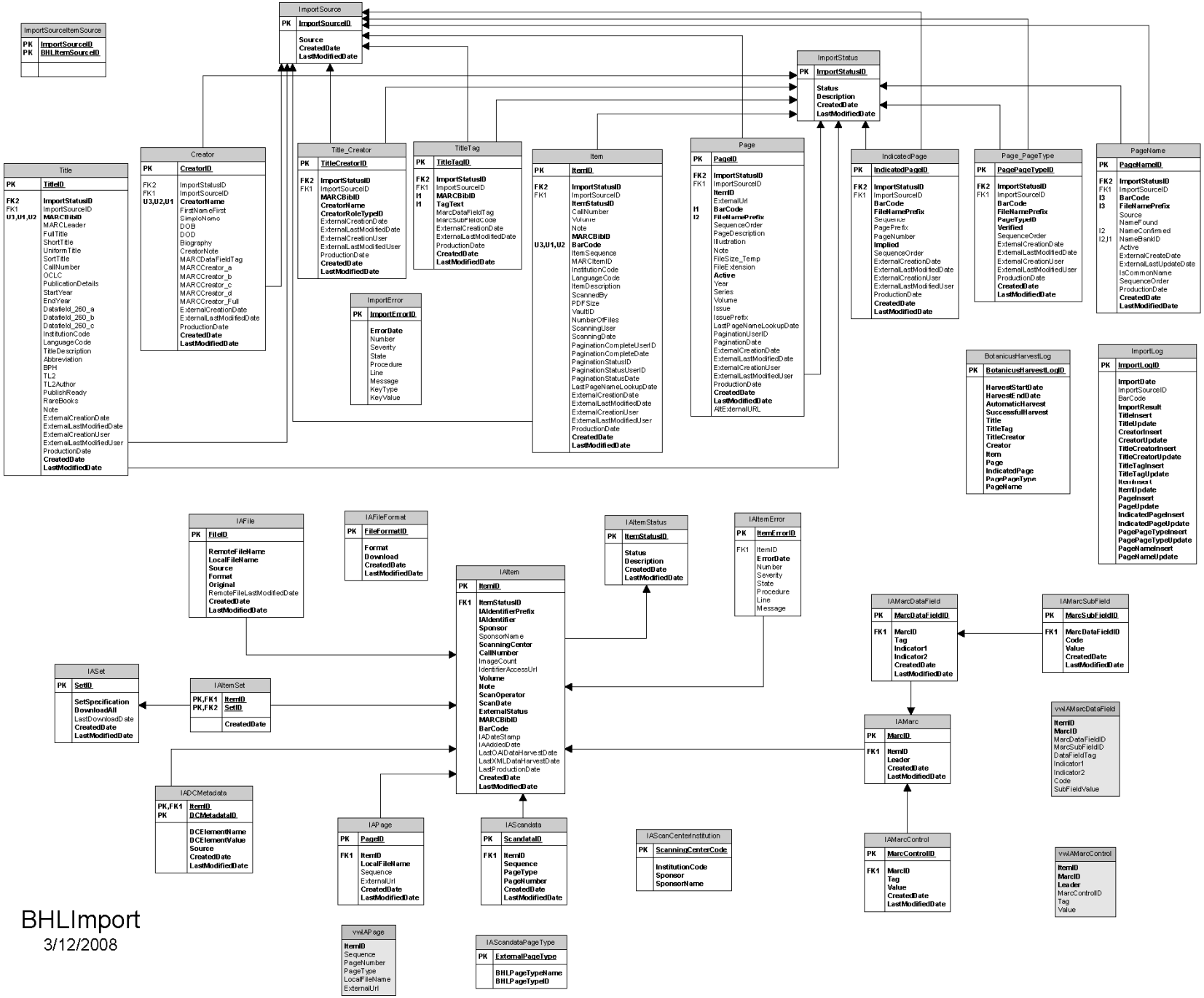
Harvest from IA

Extract, Transform, Load (ETL)

- Custom scripts to extract content via IA's APIs
- Database scripts to transform to relational data structure
- Load into database

IA Harvesting End-to-End Architecture



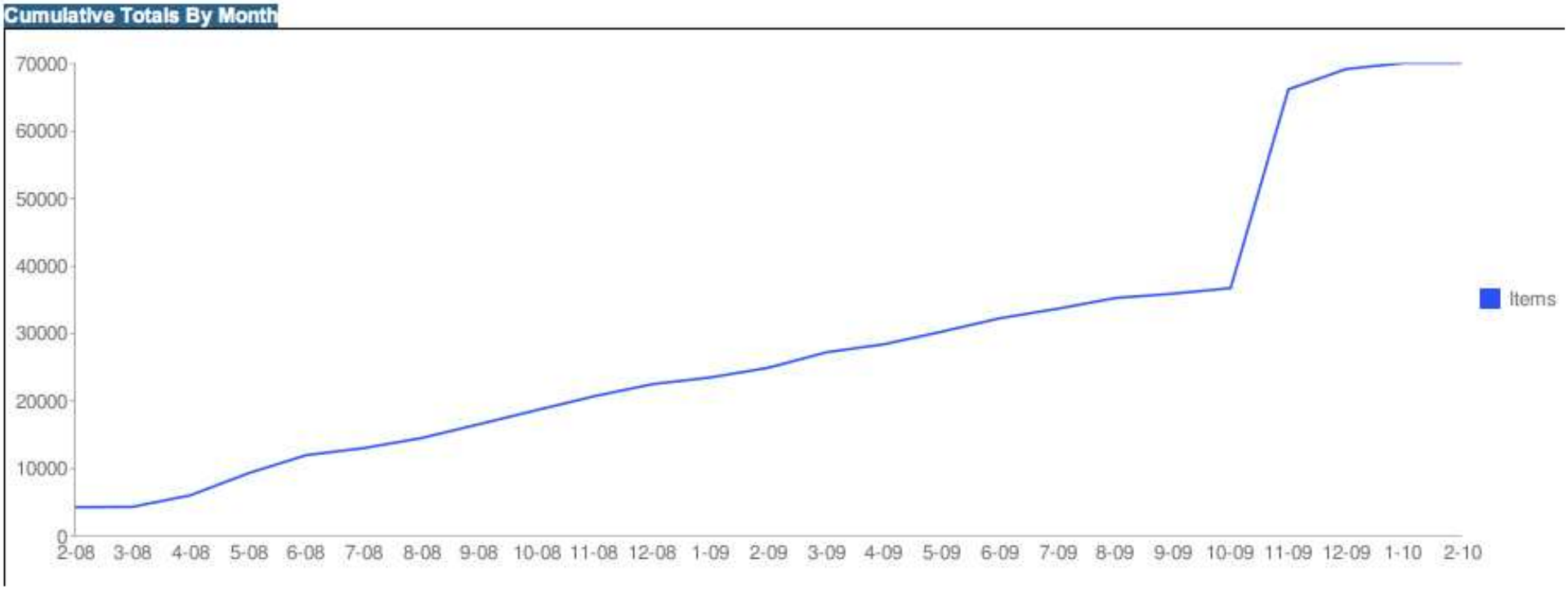


BHLImport

3/12/2008

Ingest from other IA Partners

- Used mixture of subject analysis & other bibliographic metadata to identify content for inclusion in BHL



Data Mining: Taxon Name Finding

Name Finding via TaxonFinder

Discussion.—Pigmy antechinuses are the smallest of the dasyurids and are among the smallest-sized mammals. They are not well-represented in collections and are poorly known as to their dental morphology, their ecology, and their taxonomy.

A comparison of the Madura Cave specimens with those of Recent species shows many similarities, such as size and the tendency to crowd the premolars. Comparison also reveals a number of differences that indicate that the Madura Cave form is different from the described Recent species. In general, the cheek teeth of the Madura Cave form are slightly larger than their counterparts in Recent *Planigale ingrami*, especially $P_{\overline{4}}$ and $M_{\overline{2}-\overline{4}}$ (figs. 12, 13; tables 1, 2). Comparison with $M_{\overline{1}-\overline{3}}$ of *Antechinus maculatus* is very close (fig. 12). However, $P_{\overline{4}}$ is much longer in the Madura Cave form than it is in either *Planigale ingrami* or *Antechinus maculatus*. The $M_{\overline{4}}$

Biodiversity Heritage Library: http://biodiversitylibrary.org

Name Finding in action with Taxonomic Intelligence…

SOAP response

Name finding via TaxonFinder

Submit to NameBank

Extract names

Image from Scan Converted to text via OCR


```xml
<?xml version="1.0" encoding="UTF-8"?>
<allNames>
    <entity>
        <nameString>SOBRALIA amplexicaulis</nameString>
        <parsedName canonical="SOBRALIA amplexicaulis">
            <component type="name" rank="genus">SOBRALIA</component>
            <component type="name" rank="species">amplexicaulis</component>
        </parsedName>
        <score>1</score>
        <namebankID>8652923</namebankID>
    </entity>
    <entity>
        <nameString>SoBRALIA liliastrum</nameString>
        <parsedName canonical="SoBRALIA liliastrum">
            <component type="name" rank="genus">SoBRALIA</component>
            <component type="name" rank="species">liliastrum</component>
        </parsedName>
        <score>1</score>
        <namebankID>3493662</namebankID>
    </entity>
    <entity>
        <nameString>Epidendrum liliastrum</nameString>
        <parsedName canonical="Epidendrum liliastrum">
            <component type="name" rank="genus">Epidendrum</component>
            <component type="name" rank="species">liliastrum</component>
        </parsedName>
        <score>1</score>
        <namebankID>8764188</namebankID>
    </entity>
```


SoBRALIA liliastrum
Epidendrum liliastrum
SOBRALIA
Serapias
Cymbidium hirsutum
SOBRALIA amplexicaulis

(hab. s. sp.)
examini ulteriori sub-

icatis terminalibus.

233.
libus, racemo terminali."
us et locis meridionalibus,

ispersum. Germen ca-
que evaginata. R. et P.
Presl. Reliq. Hoenk p.

plicata. Spica radicalis, triflora, caule multo br...
inferioribus foliaceis, vaginatus. Flores 2 poll...

Hab. in Peruvia.

Name Finding Evaluation

- Structured and performed by Qin Wei
  - Ph.D. student at UIUC, working with Bryan Heidorn
- Methodology
  - Scholarly volunteers manually identified scientific names on random sample of **392 pages** in BHL corpus
  - Compared those against **OCR**,then two name finding algorithms (**TaxonFinder & FAT**)
- Goals
  - Spark discussion, set baseline for future work

Characteristics of sample

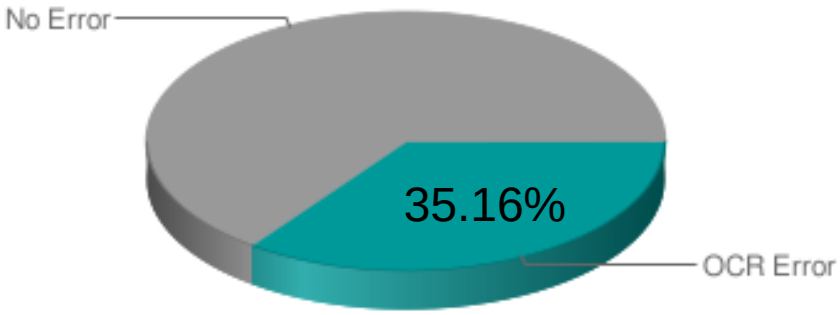
Number of Pages	392
Average Number of Words per Page	446.8
Average Number of Names per Page	7.7
Total Number of Names	**3003**
Total Number of Unique Names	**2610**

= 86.91%

Biodiversity Heritage Library: http://biodiversitylibrary.org

OCR error rate *for names only*

Of the 3,003 names, 1,056 were incorrectly transcribed by OCR.



Top OCR errors

1	Insert Space	8	n->v
2	Omit Space	9	l->i
3	e->c	10	r->i
4	u->l	11	u->ii
5	u->n	12	h->l
6	i->l	13	h->ii
7	c->e	14	e->o

Biodiversity Heritage Library: http://biodiversitylibrary.org

Performances of algorithms

	TaxonFinder	FAT
Precision	40.32%	28.20%
Recall	36.62%	23.34%
F-score	**38.47%**	**25.77%**

Excluding names with OCR errors

Precision	43.77%	32.25%
Recall	25.82%	17.21%
F-score	**34.80%**	**24.73%**

Including names with OCR errors

Biodiversity Heritage Library: http://biodiversitylibrary.org

Considerations

- Improving OCR software is out of scope
  - Google's Tesseract is only viable open source option
  - Flurry of activity in 2006-2007, quiet since
- Rekeying is expensive given size of corpus
  - Will not scale

Recommendations

- Enhance "fuzzy" retrieval in algorithms
  - Exception rules to overcome OCR errors

- More work needed in this space
  - More evaluations & experiments
  - Robust training sets
    - *reCAPTCHA* for names?

Biodiversity Heritage Library: http://biodiversitylibrary.org

Name finding statistics

- 27.7 million pages scanned
- 70.4 million name strings found
- 56.2 million names verified with a NameBankID
- 1.4 million unique names with a NameBankID
- 3.3 million unique names *without* a NameBankID
  - This is where the interesting data live!!!

Biodiversity Heritage Library

Feedback | About | Tools | BHL Members | Copyright | Contact

Search [ ] All Categories Go

Advanced Search

Browse By: Titles | Authors | Subjects | Names | Map | Year Published In: (Any Language) For: (All Contributors)

Bibliography for "Physeter catodon" by Title

As of 3 Feb 2010 1:31PM (New Search)

269 pages found in **102** titles

View NameBank record

- Allgemeines Polyglotten-Lexicon der Naturgeschichte mit erklaerenden Anmerkungen. (2)
- Animal biology / (1)
  - **Page 4**
- The Annals and magazine of natural history; (2)
- Annual report of the Board of Regents of the Smithsonian Institution. (1)
- Annual report of the Marine Mammal Commission : (2)
- Archiv für Naturgeschichte. Abteilung B. (1)
- Arsberattelse om Zoologiens Framsteg...till Kongl. Vetenskaps Akademien afgiven af Zoologiae Intendenterna. Andra Delen (Insecta. Linn.) (1)
- Beached marine birds and mammals of the North American West Coast : (1)
- The Biological bulletin. (1)
- The biology of marine animals. (2)
- Boletín de la Real Sociedad Española de Historia Natural. (1)
- Boletín de la Sociedad Española de Historia Natural. (1)
- Boletin de pescas / (1)
- A book of whales. (2)
- The brain in hominid evolution / (1)
- Brehms Thierleben : (2)
- Brehms thierleben, allgemeine kunde des thierreichs. (2)
- Brehms Tierleben. (2)
- Bulletin - United States National Museum. (24)
- Bulletin of the British Museum (Natural History). Zoology. (5)

Zoom: 25% View Page in Book

4. Size of Animals.—The smallest animals are invisible without the aid of the microscope and the largest are the 100-foot whales weighing about 150 tons. These huge water animals are probably the largest that have ever existed and far exceed in size the largest terrestrial 6½-ton elephants.

5. Distribution of Animals.—Animals are found everywhere on the earth's surface, except perhaps on the glaciated tops of the highest mountains and a very few at the poles. On these mountains creeping and flying forms pass the margins of the snow fields, and the areas of ice and snow at the poles are constantly invaded by such forms as are able to venture into them. Animal life is found throughout the waters of the sea and even penetrates to the deepest parts of the oceans. Animals burrow below the surface of the ground to considerable depths and also follow fissures still deeper to reach the farthest recesses of the most extensive caverns. Finally, myriads of living creatures live within the bodies of other living things, both plant and animal.

FIG. 1.—Diagram showing the relative size of a small sperm whale, *Physeter catodon*, of 30 tons and a herd of 24 horses, *Equus caballus*. The largest whales may be four to five times this size.

6. Relations of Animals.—Animals are related in various ways to other animals, to plants, and to their physical environment. Between parents and offspring the relation is that of descent. Between other animals the relation may be nutritive, one living upon the other; reproductive, where two join in the production of young; locomotor, where one attaches itself to another for the purpose of being transported from one place to another; or any one of many other relations which might also be named. Plants serve as food for animals, afford them concealment, and are useful to them in other ways. A solitary existence, in which one animal lives without any relationship to any other, is possible but rarely occurs in nature. Animals possessing sex associate together for a longer or shorter time as mates. Many of the same kind live together, forming a colony or, as in the case of ants, bees, and wasps, are organized into a society. The relations of animals to their physical environments are manifold. Some are confined to the land, others to the water, and still others may be at home in both. Aquatic forms may be restricted to fresh waters, others may be only marine, while there

http://www.biodiversitylibrary.org/name/Physeter_catodon

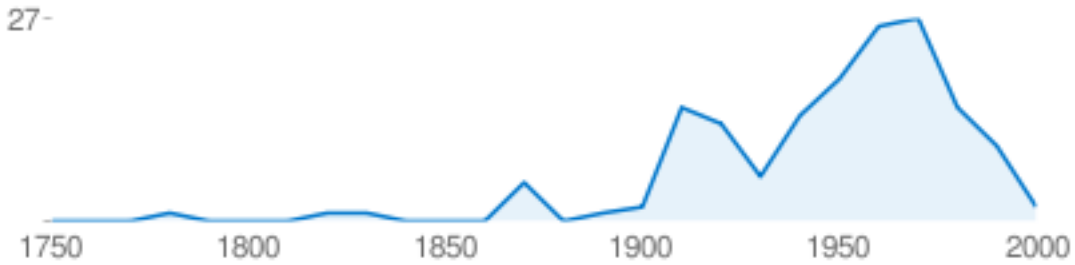
Biodiversity Heritage Library: http://biodiversitylibrary.org

http://bioguid.info/bhl/compare.php?name1=Physeter+catodon&name2=Physeter+macrocephalus

Physeter catodon | Physeter macrocephalus | Search

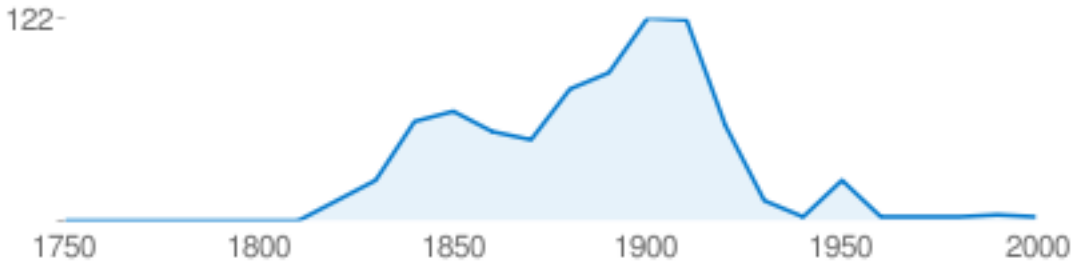
Physeter catodon

View timeline



Physeter macrocephalus

View timeline

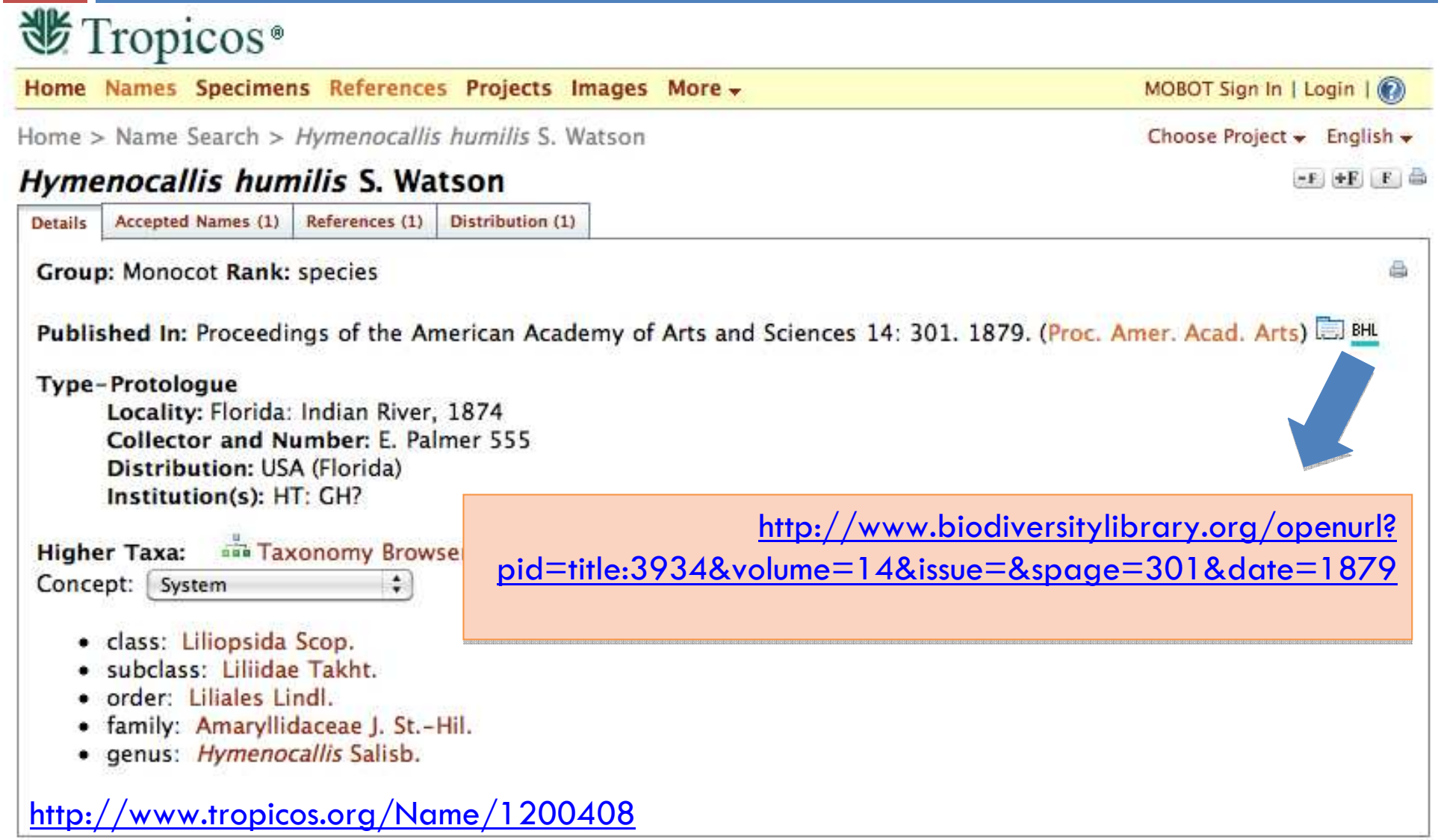


OpenURL & Other Services

Services

- OpenURL
  - Facilitate links to citations: protologues, articles, references
    - Documentation: http://www.biodiversitylibrary.org/openurlhelp.aspx
  - Useful to Nomenclators, Reference Systems
    - IPNI
    - Tropicos
- Names Service
  - Return all occurrences of a name throughout BHL digitized corpus
    - Documentation: http://bit.ly/2e6sg9
  - Access to 51million name strings using TaxonFinder
    - 1.4million unique names
  - Working out a strategy for obscure species
  - Algorithm improvements to detect nomenclatural & taxonomic acts

Services: OpenURL



http://www.biodiversitylibrary.org/openurl?pid=title:3934&volume=14&issue=&spage=301&date=1879

Services: OpenURL Disambiguation

- Looking for:

(Proc. Amer. Acad. Arts) 14: 301. 1879.

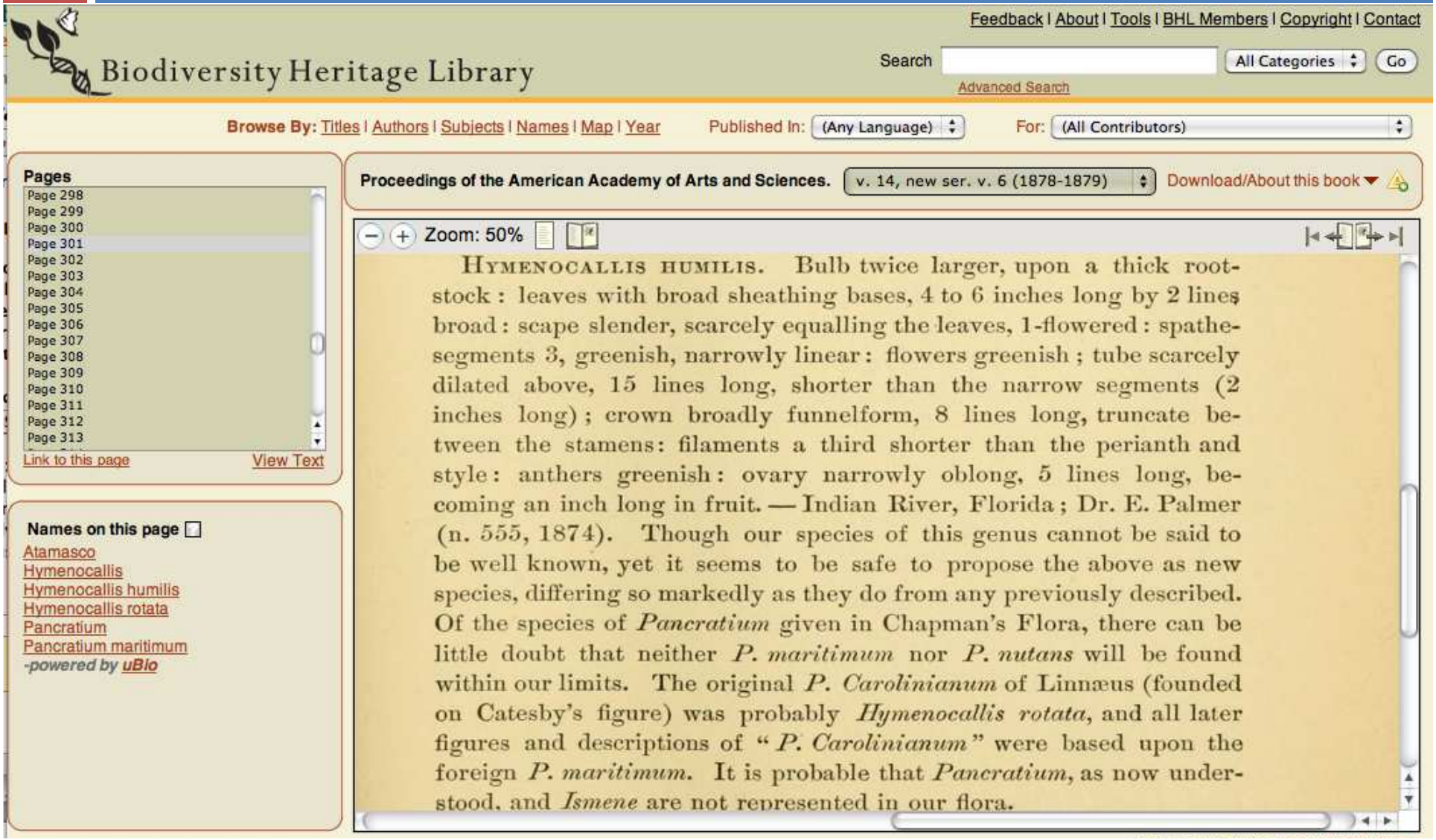
- BHL returns:

OpenUrl Results

Select one of the items below to find the desired citation.

Title	Volume	Issue	Year	Page
Proceedings of the American Academy of Arts and Sciences.	v. 14, new ser. v. 6 (1878-1879)			Page 301
Proceedings of the American Academy of Arts and Sciences.	v. 22, new ser. v. 14 (1886)			Page 301

Services: OpenURL Results



How?

- Tropicos maintains internal authority list of publications: http://www.tropicos.org/Publication/775
- Each protologue/reference tied to authority:

(Proc. Amer. Acad. Arts) 14: 301. 1879.

- Matched Tropicos TitleIDs to BHL TitleIDs:

http://www.tropicos.org/Publication/775 = http://www.biodiversitylibrary.org/title/3934

- Throw citations at resolver at regular intervals & cache data in Tropicos

http://www.biodiversitylibrary.org/openurl?pid=title:3934&volume=14&issue=&spage=301&date=1879

BHL Name Services

http://www.biodiversitylibrary.org/services/name/NameService.asmx


```
- <Name>
    <NameBankID>8764188</NameBankID>
    <NameConfirmed>Epidendrum liliastrum</NameConfirmed>
  - <Titles>
    - <Title>
        <TitleID>499</TitleID>
        <MarcBibID>b12026323</MarcBibID>
        <PublicationTitle>The genera and species of orchidaceous plants</PublicationTitle>
        <PublicationDetails>London : Ridgways, 1830-1840.</PublicationDetails>
        <Author>Lindley, J.</Author>
        <TL2>4643</TL2>
        <Abbreviation>Gen. Sp. Orchid. Pl.</Abbreviation>
        <TitleUrl>http://www.biodiversitylibrary.org/title/b12026323</TitleUrl>
      - <Items>
        - <Item>
            <ItemID>9889</ItemID>
            <BarCode>31753002698485</BarCode>
            <MarcItemID>i12594374</MarcItemID>
            <CallNumber>QK495.F50 L5543 1830 -40</CallNumber>
          - <ItemUrl>
              http://www.biodiversitylibrary.org/item/31753002698485
            </ItemUrl>
          - <Pages>
            - <Page>
                <PageID>393438</PageID>
                <Year>1830-1840</Year>
                <Prefix>Page</Prefix>
                <Number>177</Number>
                <PageUrl>http://www.biodiversitylibrary.org/page/393438</PageUrl>
              - <ThumbnailUrl>
                  http://images.mobot.org/viewer/viewerthumbnail.asp?cat=botanicus3&client=b12026323/31753002698485/jp2&image=31753002698485_01
                </ThumbnailUrl>
```

Codebase:

Overview

Mandate for new development

- display / manage articles
- meet community demands for bibliography / citation management
- build from more open source tools

Biodiversity Heritage Library: http://biodiversitylibrary.org

Development goals re: citations

- Create a repository for community-vetted taxonomic bibliographies.
- Ability to ingest, display, download, and index articles so that the BHL can operate as an article repository.
- Build from existing community of work around Drupal / Biblio.
  - In use by collaborators

http://citebank.biodiversitylibrary.org/

Biodiversity Heritage

Biodiversity Heritage Library an open access repository for biodiversity publications

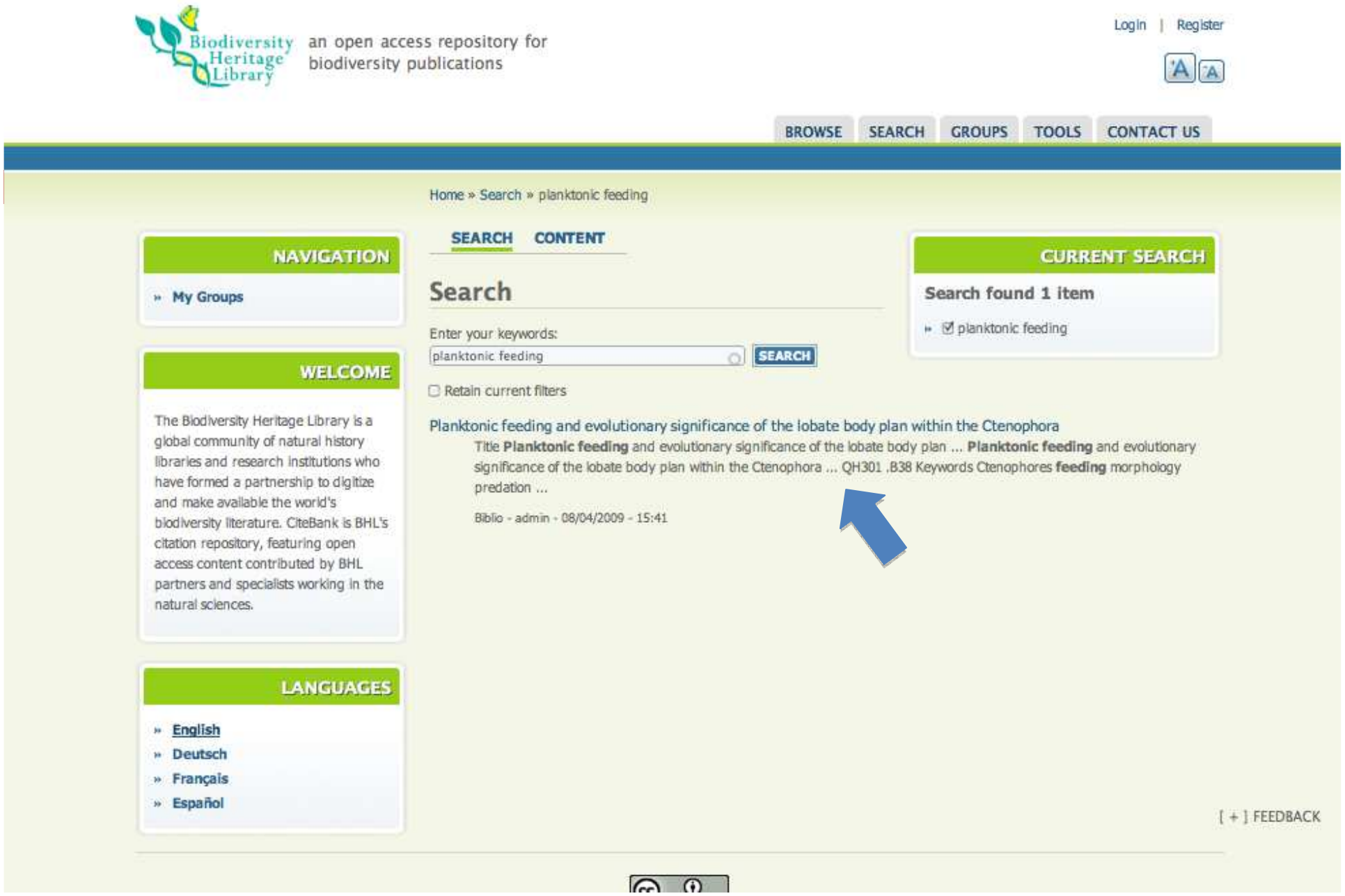
Login | Register

BROWSE SEARCH GROUPS TOOLS CONTACT US


Search...

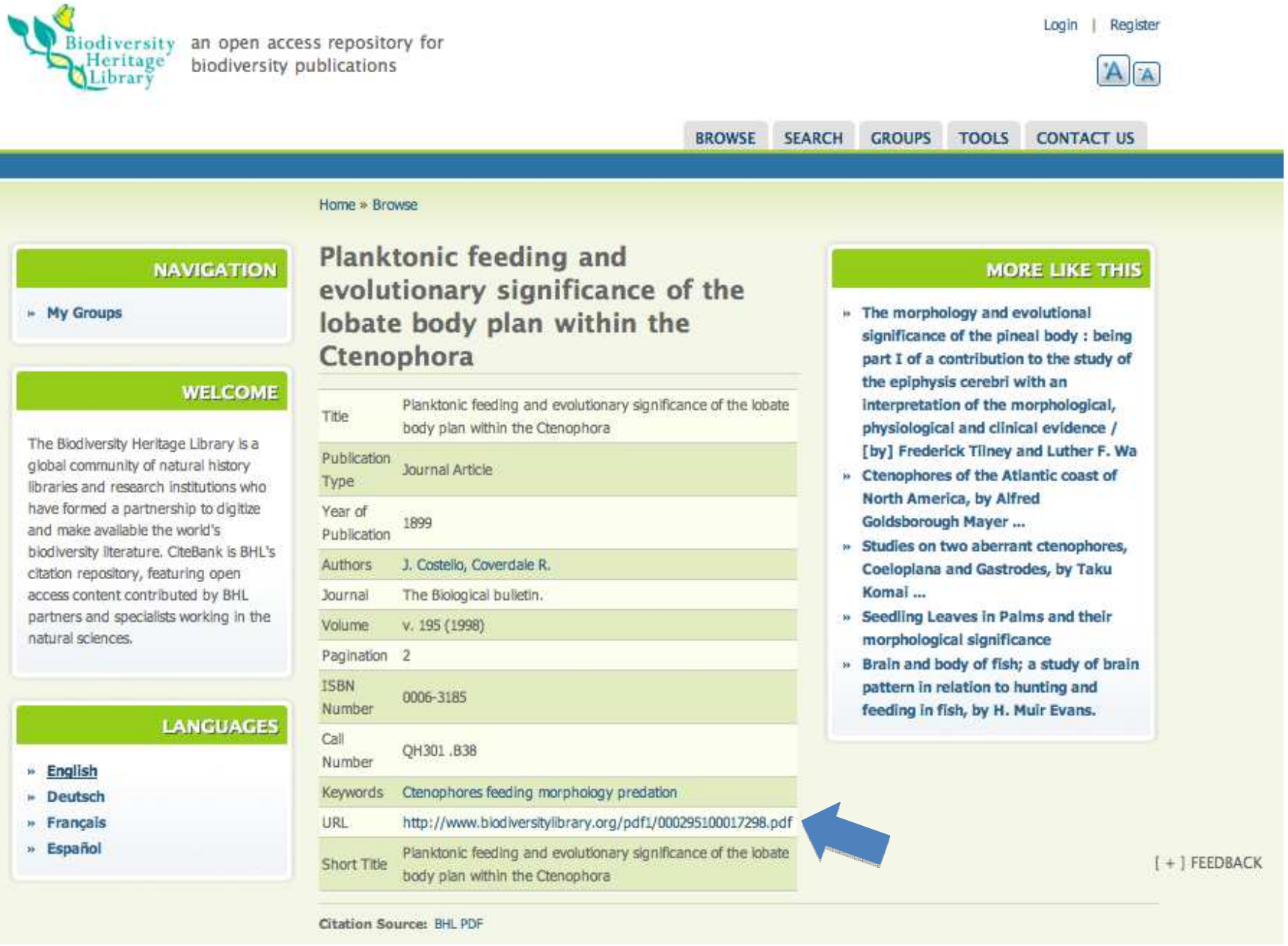

Explore...


Collaborate...



http://citebank.biodiversitylibrary.org/search

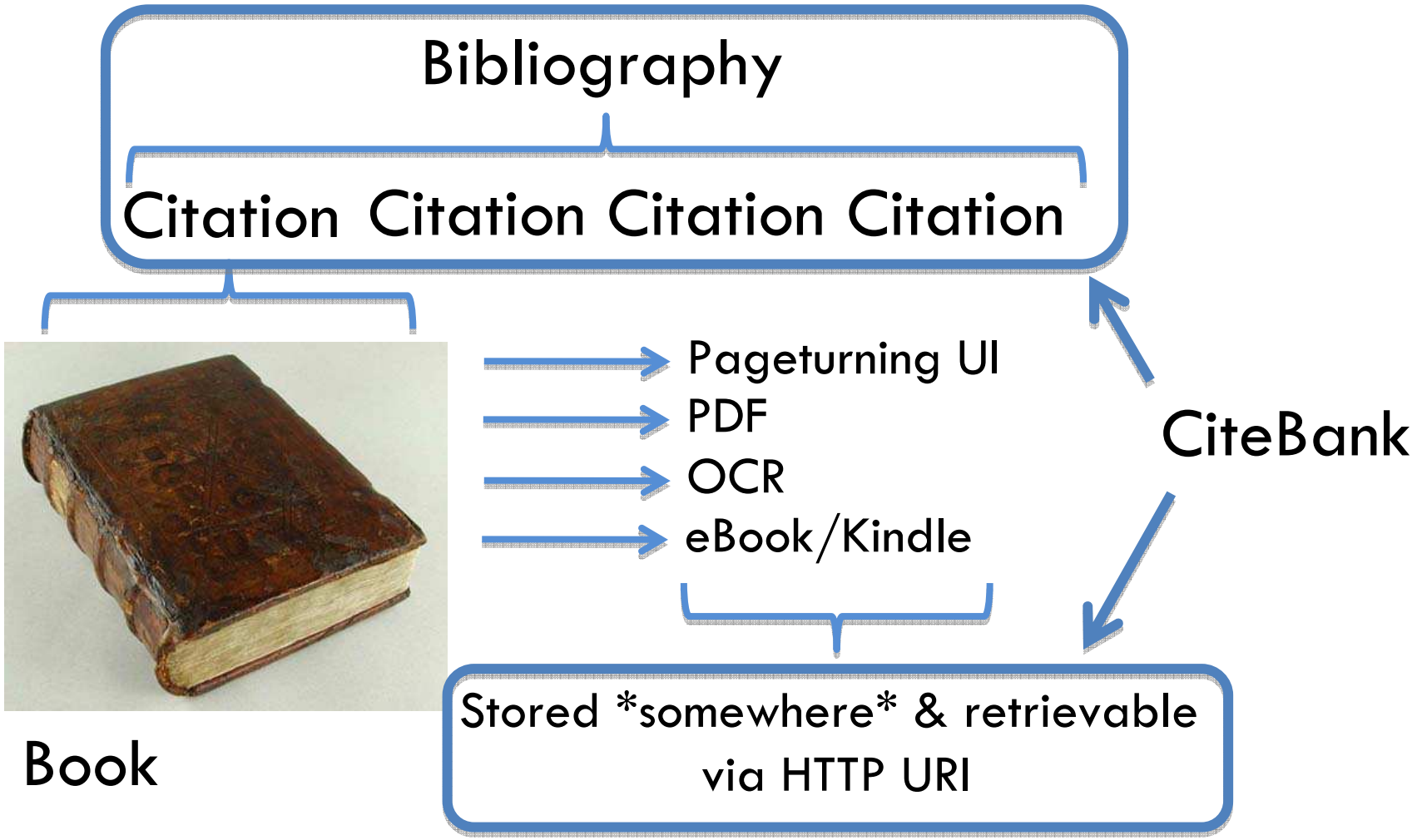
Biodiversity Heritage Library: http://biodiversitylibrary.org



http://citebank.biodiversitylibrary.org/node/47423

Biodiversity Heritage Library: http://biodiversitylibrary.org

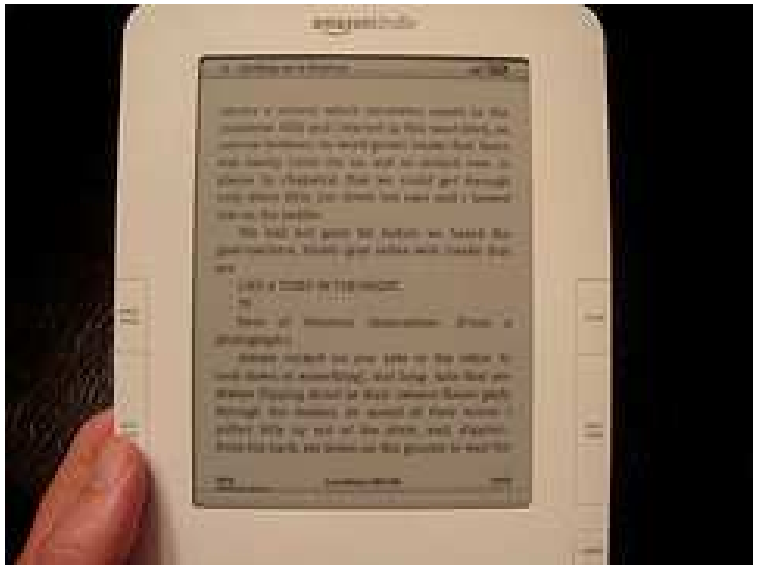
CiteBank boundaries



Biodiversity Heritage Library: http://biodiversitylibrary.org

New devices & hardware

- Content now available in EPUB & Kindle format
  - Used by Stanza, Kindle
- Blog post by John Mignault (NYBG):
  - http://john.mignault.net/blog/2009/10/28/first-bhl-e-book-experiments/



Biodiversity Heritage Library: http://biodiversitylibrary.org

Links & such

Biodiversity Heritage Library

http://biodiversitylibrary.org

CiteBank beta

http://citebank.biodiversitylibrary.org

CiteBank sandbox

http://sandcite.biodiversitylibrary.org

Go play!

Follow BHL on

twitter

http://twitter.com/BioDivLibrary

Obrigado!

Chris Freeland

Technical Director, BHL

Director, Center for Biodiversity Informatics, Missouri Botanical Garden

chris.freeland@mobot.org

http://twitter.com/chrisfreeland