

Laboratório de Digitalização

*Projeto Digitalização e Publicação
Online de uma Coleção de Obras Raras Essenciais em
Biodiversidade das Bibliotecas Brasileiras*

Marcia Y Barretto
ITI/GA, Coordenadora
BIREME/OPAS/OMS
Junho 14, 2010

Visitas em dezembro/2009

• Harvard College Library, Imaging Services

- equipamentos não robotizados, customizados localmente, visando as necessidades específicas.
- arquivos eletrônicos (formato JPEG2000), juntamente com arquivos XML de metadados, enviados para a “central de processamento de dados” da Harvard e responsável pela preservação e disponibilização. O PDF é gerado dinamicamente quando solicitado.

•Library of Congress, Adams Building

- 11 estações “Internet Archive scribe” (<http://www.archive.org/scanning>) e 1 equipamento para documento plano
- 13 tipos de arquivos, entre eles: JPEG2000, OCR em XML, XML com metadados de registros MARC, *searchable* PDF
- custo: US\$0.10 por página padrão, US\$2.00 por página com mapa/imagem com tratamento diferenciado
- 72 horas para disponibilização dos arquivos após envio ao Internet Archive
- processados cerca de 1000 volumes por

semana, com média de 300 páginas por volume

•George Washington University, The Melvin Gelman Library

- Equipamentos Kirtas utilizados: Kabis III e SkyView
- camera fotográfica (Canon, 21MP) é a mesma
- arquivos JPEG2000 são processados pelo software BSE (Book Scan Editor Software) cujas licenças podem ser adquiridas à parte

•Biodiversity Heritage Library / Smithsonian Institute

–Obras raras: fornece cópias digitalizadas manualmente e *on-demand*

- Minolta PS7000C MKII: *full color digital book scanner* (

<http://kmb.s.konicaminolta.us/content/products/models/PS>

)

– Digitalização para preservação:

- 1 Estação “Internet Archive scribe” no Natural History Museum para documentos que não podem ser retirados

- envio periódico de documentos para o FEDLINK Federal Scanning Center – FedScan, localizado na Library of Congress

Termos de referência para digitalização



BIREME | OPAS | OMS

Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde
Organização Pan-Americana da Saúde | Organização Mundial da Saúde
Rua Botucatu, 862 - Vila Clementino - 04023-901 São Paulo, Brasil. Tel.: 55-11-5576-9800 -
Fax: 55-11-5575-8868 - info@bireme.br | <http://www.bireme.org>

Termos de referência para a contratação de serviços de digitalização de documentos

05.03.2010

Esse documento descreve os termos de referência para a digitalização de documentos no âmbito da BVS e redes associadas com base em standards internacionais de resolução e formato de arquivos.

Os documentos digitalizados deverão ser entregues no formato de imagem de acordo com as seguintes especificações:

1. Resolução, profundidade de cor (bit depth)

Tipo de documento	Resolução (mínima)	Resolução (preferida)	Formato de arquivo e profundidade de cor
Bitonal	300dpi	600dpi	1-bit TIFF
Grayscale	300dpi	600dpi	8-bit uncompressed TIFF or lossless compressed image (e.g. LZW, JPEG2000 [* .jp2])
Color	300dpi	400dpi	24-bit uncompressed TIFF, or lossless compressed images (e.g. LZW, JPEG2000 [* .jp2])

2. Arquivos

2.1 Arquivos originais

a) Os arquivos imagem originais devem ser entregues sem pós-processamento. Os arquivos não devem ser editados e devem incluir, sempre que possível, o perfil de cores usado durante o processo de digitalização. Esses arquivos servirão para o propósito de preservação.

b) Os arquivos imagem devem ser armazenados e organizados em pastas identificadas, sendo uma pasta para cada objeto físico.

2.2 Arquivos pós-processados

O pós-processamento dos arquivos deve incluir a produção de:

- Arquivo PDF com OCR (*Optical Character Recognition*) embutido
- Arquivo tipo MS-Word ou texto com o OCR
- Os arquivos devem ser armazenados e organizados em pastas identificadas, sendo uma pasta para cada objeto físico.



Workshop Essential Rare Works Collection in
Biodiversity: governance, operation and digitization
3 - 5 February 2010

ATA da Reunião realizada em 04 de fevereiro de 2010

Participantes:

BIREME: Abel Packer, Marcia Barretto
Museu de Zoologia / USP: Hussam Zaher
Ministério do Meio Ambiente: Bráulio Dias
Biodiversity Heritage Library: Tom Garnett, Martin Kalfatovic,
Chris Freeland

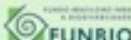
Objetivo: Análise de proposta para implementação do laboratório de digitalização no âmbito BHL Brasil

Martin apresentou um documento onde apresentava o desenvolvimento do projeto referente à digitalização em duas fases: a primeira com aquisição de 2 equipamentos de menor porte e que poderiam ser, ao final da 1ª fase, direcionados a outras instituições participantes. Na 2ª fase seriam adquiridos os equipamentos de maior porte e que comporiam o laboratório de digitalização propriamente dito.

O Anexo – Proposta para implementação do laboratório de digitalização mostra a proposta apresentada e aceita por todos os participantes.



CAIXA



Ministério
do Meio Ambiente



Anexo – Proposta para implementação do laboratório de digitalização

BHL Brasil

Phase I: +3 months

- 1 centralized scanning center
 - Centralized scanning center in Sao Paulo
 - 2 staff (librarian and technician)
 - 1 (or 2) Indus-style planetary book scanner
 - Creates JP2/TIFF image files; title, item, page metadata creation
 - Internet Archive serves as host for files and derivative file creation
 - BHL portal engine creates title aggregation, taxonomic intelligence, etc.
 - BHL Brasil (option): local storage, Portuguese language portal, etc.

BHL Brasil

Phase II: +6-9 Months

- 1 centralized mass scanning center
 - Centralized scanning center in Sao Paulo
 - 1-2 librarians
 - 2-4 technicians
 - 1 (or 2) Kirtas, Scribe, Phase I type book scanner
 - Creates JP2/TIFF image files; title, item, page metadata creation
 - All sub-modules as in Phase I
- 1 (or more) distributed/local scanning centers
 - As described in Phase I

Equipamentos

- Laboratório
 - 2 tipo SkyView para documentos de tamanho grande tipo folio, mapas
 - 2 book scanners + 2 workstations
 - 1 Internet Archive Scribe
- Kits móveis
 - 3 book scanners
 - 3 workstations



Obrigada.

marcia.barretto@bireme.org