

Manual de procedimentos técnicos de gestão de recursos digitais



Coordenação da Biblioteca Nacional Digital / FBN
Laboratório de Digitalização
Otávio Alexandre Jeremias de Oliveira (org.)

2024

Otávio Alexandre Jeremias de Oliveira (org.)

Manual de procedimentos técnicos de gestão de recursos digitais

Fundação Biblioteca Nacional
Rio de Janeiro
2024

REPÚBLICA FEDERATIVA DO BRASIL

PRESIDÊNCIA DA REPÚBLICA

Luiz Inácio Lula da Silva

MINISTÉRIO DA CULTURA

Margareth Menezes da Purificação Costa

FUNDAÇÃO BIBLIOTECA NACIONAL

PRESIDÊNCIA

Marco Americo Lucchesi

DIRETORIA EXECUTIVA

Suely Dias

CENTRO DE PROCESSAMENTO E PRESERVAÇÃO

Gabriela Ayres Ferreira Terrada

BIBLIOTECA NACIONAL DIGITAL

Otávio Alexandre Jeremias de Oliveira

AUTORES

Adriana Pimentel

Flávio Santos

Marcel Carreira

CENTRO DE PESQUISA E EDITORAÇÃO

Naira Christofoletti Silveira

COORDENAÇÃO DE EDITORAÇÃO

Claudio Cesar Ramalho Giolito

ORGANIZAÇÃO

Otávio Alexandre Jeremias de Oliveira

PRODUÇÃO EDITORIAL

Paula Rocha Machado

REVISÃO TÉCNICA

Flávio Santos

REVISÃO E PREPARAÇÃO DE ORIGINAIS

Francisco Madureira

PROJETO GRÁFICO, DIAGRAMAÇÃO E TRATAMENTO DE
IMAGENS

Ana Sanches

REVISÃO DE PROVA

Paula Rocha Machado

Biblioteca Nacional Digital

Av. Rio Branco, 219, 3º andar

Rio de Janeiro – RJ | 20040-008

bndigital@bn.gov.br | www.gov.br/bn

Copyright © 2023 Fundação Biblioteca Nacional (FBN)

DADOS INTERNACIONAIS PARA CATALOGAÇÃO NA PUBLICAÇÃO (CIP)

Pimentel, Adriana

Manual de procedimentos técnicos de gestão de recursos digitais / Otávio Alexandre Jeremias de Oliveira (org.) ; [autores: Adriana Pimentel, Flávio Santos, Marcel Carreira]. - Rio de Janeiro : Fundação Biblioteca Nacional, Coordenação da Biblioteca Nacional Digital, Laboratório de Digitalização, 2023.

53 p. : il. color ; 30 cm.

1. Materiais bibliográficos – Digitalização – Manuais, guias, etc. 2. Documentos eletrônicos – Administração. 3. Preservação Digital – I. Oliveira, Otávio Alexandre Jeremias de. II. Santos, Flávio, 1976-. III. Carreira, Marcel, 1978-. IV. Biblioteca Nacional (Brasil). Coordenação da Biblioteca Nacional Digital. Laboratório de Digitalização. V. Título.

CDD 025.84

Angela Cardoso de Matos - Bibliotecária – CRB -5808

SUMÁRIO

1 - INTRODUÇÃO	5
2 - FLUXO DE ENTRADA DE OBRAS NO LABORATÓRIO DE DIGITALIZAÇÃO	6
2.1 - Fluxo de entrada	6
2.2 - Formulário de encaminhamento de obras (FEO)	7
2.3 - Fluxo de entrada de periódicos	8
2.4 - Direcionamento das obras por equipamento dentro do Laboratório	9
2.5 - Fluxo de saída	10
3 - CODIFICAÇÃO/NOMEAÇÃO DAS PASTAS E ARQUIVOS	11
3.1 - Codificação/nomeação das pastas e arquivos de acervos especiais	11
3.2 - Codificação/nomeação das pastas e arquivos de periódicos	13
4 - CONTROLE DE QUALIDADE DAS IMAGENS DIGITALIZADAS	16
4.1 - Formatos de arquivos: <i>master</i> e derivados	16
4.2 - Processos de controle de qualidade	17
5 - METADADOS NO SISTEMA SOPHIA	20
5.1 - Inserindo metadados da digitalização no sistema SophiA	20
6 - ARQUIVOS DERIVADOS DE DIFUSÃO	23
6.1 - Tipos de obras	23
6.2 - Configurações dos arquivos derivados	24
6.2.1 - Preparo dos arquivos a serem utilizados	24
6.3 - Dragão simples (1 ou 2 páginas)	24
6.4 - Dragão frente-verso	26
6.5 - PDF/PDF OCR	27
6.6 - <i>Book-Reader</i>	27
6.7 - Galeria	29
6.8 - Anexo 1	34
6.9 - Anexo 2	39
6.10 - Anexo 3	41
6.11 - Anexo 4	43
7 - CONFERÊNCIA E GUARDA DOS ARQUIVOS DIGITALIZADOS DE PERIÓDICOS... ..	44
7.1 - Criação e organização de pastas para <i>backup</i>	44
7.2 - Conferência da codificação de pastas e arquivos	46
7.2.1 - Conferência da codificação de pastas	46
7.2.2 - Conferência da codificação de arquivos	46
7.3 - Conferência dos quantitativos de arquivos de imagens nas pastas	47
7.4 - Conferência da qualidade de arquivos de imagens	47
7.5 - Organização final para guarda no <i>Storage</i>	48
8 - ARMAZENAMENTO	51
8.1 - Armazenamento no servidor dentro do Laboratório	51
8.2 - Armazenamento no <i>Storage</i>	51

1 - INTRODUÇÃO

Nesta nova era da informação, muitas mudanças e avanços tecnológicos estão permitindo que o livro impresso seja transformado em mídias digitais. Todas estas transformações vêm alterando a nossa cultura e o nosso cotidiano. Cada vez mais as pessoas estão conectadas a web e querendo acesso fácil e rápido as informações.

Com este pensamento as Bibliotecas Nacionais em todo o mundo iniciaram um processo de digitalização de suas coleções, pensando em novas formas de preservação e dando acesso aos seus valiosos acervos.

Com intuito de facilitar, ampliar e democratizar o acesso ao público e garantir a preservação dos documentos a longo prazo foi criado em 2006 a BNDigital, onde o Laboratório de Digitalização está inserido, digitalizando e disponibilizando on-line a vasta coleção da Biblioteca Nacional, possibilitando ao usuário consultar vários tipos de obras com várias maneiras de visualização.

O Laboratório possui equipamentos de última geração e equipe altamente especializada. Para nortear seu futuro o Laboratório sempre estará buscando e acompanhando a evolução de tecnologias e de metodologias na área de digitalização. Sempre procurando desenvolver ações, com técnica, qualidade e dentro de padrões internacionais.

O presente manual elaborado pela equipe do Laboratório de Digitalização da Fundação Biblioteca Nacional é uma iniciativa de registrar e estabelecer critérios técnicos para a reprodução digital de seu acervo, visando criar diretrizes que possam ser seguidas e compartilhadas. Ele é baseado na larga experiência adquirida até o momento por seus colaboradores no constante desafio de reproduzir digitalmente seu rico e diversificado acervo, buscando sempre as soluções técnicas que sejam adequadas e possíveis para execução desse importante trabalho. Abordaremos aqui os processos, ferramentas, soluções e procedimentos técnicos utilizados pelo Laboratório de Digitalização na gestão de recursos digitais.

Assim esse manual pretende servir de referência para o um melhor desempenho técnico dos processos empregados, norteados não só a equipe atual, mas também futuros colaboradores e demais interessados em se aperfeiçoar para exercer essa nobre função de digitalizador.

2 - FLUXO DE ENTRADA DE DOCUMENTOS NO LABORATÓRIO DE DIGITALIZAÇÃO

2.1 - Fluxo de entrada

O fluxo de documentos do Laboratório de Digitalização é feito para suprir tipos diferentes de demandas que acontecem na Biblioteca Nacional. Os principais são:

Digitalização para preservação e acesso: é o principal fluxo do Laboratório, os documentos são digitalizados e guardados para preservação da memória bibliográfica e documental do País e para serem publicadas no portal da BN Digital, dando acesso a sociedade ao acervo da Biblioteca Nacional.

Pedidos externos: são solicitações de imagens para suprirem uma pesquisa de usuário. Estes pedidos são de “fora da FBN”, oriundos do Setores de Guarda do Acervo ou da DINF, pois são estes que prestam este tipo de atendimento de demanda.

Pedidos internos: são digitalizações de documentos inteiros, ou parte deles, para exposição na FBN (virtual), demandas judiciais, pedidos de outra Instituição e outras solicitações internas da FBN.

Obras encaminhadas pelo Setor de Restauração: a partir do processo de trabalho do Laboratório de Preservação e Restauração, algumas obras são desencadernadas, então são encaminhadas para o Laboratório de Digitalização, antes de serem reencadernadas, por ser muito mais fácil e seguro para o documento a digitalização das folhas soltas. Essas obras são na maioria seculares, pertencentes ao Setor de Obras Raras.

Obras vindas de outras instituições: a Biblioteca Nacional cria parceria com outras instituições, além de dar suporte de trabalho a outras. Com isso, esporadicamente, o Laboratório recebe obras para digitalização que não são do seu acervo.

Obs.: a seleção de documentos a serem digitalizados é feita principalmente a partir da curadoria dos setores responsáveis pela guarda das obras.

- A obra só deve estar inserida no SophiA quando for para “Digitalização para a BN DIGITAL”, caso contrário não necessita de cadastro neste sistema.

- Deve estar descrito no “FEO” todas as características da(s) obra(s) nos respectivos campos, como: localização física no setor de guarda do acervo, autor, título, assunto, número de registro (que é o número do Objeto Digital) e código do título SophiA (se for o caso).

Cada obra é única e deverá corresponder a uma linha do FEO. Em alguns casos existem várias obras com registros diferentes, mas com um único código no SophiA (como por exemplo as coleções - Analíticas).

A falta de qualquer informação ou o não preenchimento do correto do FEO poderá implicar no não recebimento da(s) obra(s) pelo Laboratório de Digitalização.

2.3 - Fluxo de entrada de periódicos

O fluxo de recebimento dos documentos vindos do setor de Publicações Seriadas, para serem digitalizadas no Laboratório de Digitalização, possui algumas rotinas diferenciadas em relação aos fluxos de recebimentos das documentações vindas dos outros setores da BN. Uma das diferenças é em relação ao responsável pelo recebimento. Enquanto toda outra documentação é recebida por um técnico da equipe de gestão documental, essa documentação é recebida por um técnico selecionado da equipe Zeutschel. Esse técnico terá a responsabilidade de estar em contato com o setor de Publicações Seriadas, solicitando, sempre que necessário, documentos para que o fluxo de digitalização dos scanners Zeutschel não parem. Ele também é responsável pela devolução da documentação ao setor após a digitalização.

Com relação a documentação de periódicos selecionada para digitalização, essa seleção é feita para atender principalmente a demanda de publicação dos arquivos digitalizados na Hemeroteca Digital. Então, levando em consideração esse fator, um determinado título de periódico deve ser digitalizado em sua totalidade, ou seja, todos os fascículos desse título devem ser digitalizados. Sendo assim, depois de selecionado um determinado título do acervo, o setor de Publicações Seriadas encaminha os fascículos para o laboratório de digitalização, seguindo a ordem cronológica de data de lançamento, dos fascículos mais antigos até os mais recentes. Alguns títulos de periódicos são tão numerosos em quantidade de fascículos que demoram até meses para terem a sua totalidade digitalizada.

Em relação ao formulário de encaminhamento de obras do setor de Publicações Seriadas, este se difere pois é gerado pelo sistema SophiA, usando a base de dados do setor de guarda deste acervo. É gerado um relatório no sistema com as obras que serão destinadas para o Laboratório de Digitalização. Outros detalhes referentes a esta rotina são:

1 - As obras do setor de Publicações Seriadas não são registradas com o carimbo do setor de guarda do acervo e número de patrimônio (no verso da folha de rosto quando obra encadernada ou verso da própria obra).

2 - Cada formulário encaminha apenas um título de periódico, com um ou vários fascículos.

3 - No formulário de encaminhamento deve constar o número de BIN do título de periódico que está encaminhado.



Fundação BIBLIOTECA NACIONAL

Cartografia

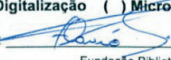
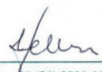
Encaminhados para: Seção do Lab. de Digitalização (DIG P/SOLICITA)

17/11/2023
Página 1 de 1

Localização	Autor	Data	Número de tombo	Código BND	Domínio Público (x)
1 ARC.029,11,020	Brasil. Inspetoria de Obras Contra as Sec	1910	523.621	105.834	

Título: Mappa dos Estados do Ceará, Rio Grande do Norte e Parahyba

Finalidade: ☒ Digitalização ☐ Microfilmagem ☐ Conservação ☐ Restauração ☐ Exposição

Entregue a:  Data: 17/11/2023 Devolvido a:  Data: 24/11/2023

Fundação Biblioteca Nacional - Av. Rio Branco, 219 - Centro - Rio de Janeiro, RJ - CEP: 20040-008 - +55 (21) 2220-3040 - <https://www.bn.gov.br>

17/11/23

2.4 - Direcionamento das obras por equipamento dentro do Laboratório

Uma vez recebidas as obras, elas serão destinadas aos diferentes equipamentos que existem dentro do Laboratório de Digitalização. Essa destinação será feita levando-se em conta as características físicas das obras, quais as necessidades para reproduzi-la da melhor maneira digitalmente e qual o equipamento atenderá melhor essas necessidades.

Câmera de grande formato

Nesse equipamento são digitalizadas as obras que possuem grandes dimensões e que também tenham muitas informações e detalhes, exigindo um grande nível de ampliação do arquivo digitalizado para dar acesso a toda a informação contida no documento. Para isso é necessário uma grande resolução de captura, que é encontrada nesse equipamento.

São digitalizados nesse equipamento obras como cartografias, pergaminhos manuscritos, iconografias de grande dimensão e quaisquer outras obras que ultrapassem dimensões que inviabilizem a digitalização nos outros equipamentos.

Câmera de médio formato

Nesse equipamento são digitalizadas as obras que possuem muitas informações e detalhes, exigindo um grande nível de ampliação do arquivo digitalizado para dar acesso a toda a informação contida no documento. Para isso é necessário uma grande resolução de captura, que é encontrada nesse equipamento.

São digitalizados nesse equipamento obras como desenhos, fotografias, livros ricamente ilustrados com imagens do Acervo Especial, gravuras, algumas obras “efêmeras” que englobam recortes de jornais e revistas, cartazes, cartões postais e calendários. Geralmente todas as obras Iconográficas.

Scanner planetário

Nesse equipamento são digitalizadas as obras com predominância textual, que não precisam de um grande nível de ampliação do arquivo digitalizado para dar acesso a informação contida no documento. Para isso é necessário uma boa resolução de captura, que é encontrada nesse equipamento.

São digitalizados nesse equipamento obras como manuscritos, livros do acervo de Obras Gerais, além das obras de características similares aos tipos mencionados e principalmente as coleções de Periódicos como jornais, revistas, anuários e todas as obras com circulação regular.

2.5 - Fluxo de saída

Após a digitalização(captura) as imagens (objetos digitais) entrarão no fluxo de controle de qualidade das imagens digitalizadas, e finalizado essa etapa, a(s) obra(s) podem ser devolvidas. Deve-se conferir estas nos seus respectivos FEO e devolvê-la(s) de volta ao setor que a(s) encaminhou.



MINISTÉRIO DO TURISMO

Fundação BIBLIOTECA NACIONAL

Periódicos

Devolvidos por: Seção de Lab. de Digitalização (DIG P/HDB/BND)

16/03/2020
Página 1 de 2

Joia **CT: 872237** Total de volumes 19

	Localização	Período	Devolvido em	Número de tomo	Cód. de circui
1	4.350,01.01	1957/11/30 a 1957/12/31	16/03/2020	0001301195	31009
2	4.350,01.02	1958/01/15 a 1958/03/31	16/03/2020	0001301187	31010
3	4.350,01.03	1958/04/15 a 1958/06/30	16/03/2020	0001301179	31011
4	4.350,01.04	1958/07/15 a 1958/09/30	16/03/2020	0001301160	31012
5	4.350,01.05	1958/10/15 a 1958/12/31	16/03/2020	0001301233	31013
6	4.350,01.06	1959/01/15 a 1959/03/31	16/03/2020	0001301225	31014
7	4.350,01.07	1959/04/15 a 1959/05/31	16/03/2020	0001301217	31015
8	4.350,01.08	1959/06/15 a 1959/09/30	16/03/2020	0001301209	31016
9	4.350,01.09	1959/10/15 a 1959/12/31	16/03/2020	0001301276	31017
10	4.350,01.10	1960/01/01 a 1960/03/31	16/03/2020	0001301268	31018
11	4.350,01.11	1960/04/01 a 1960/07/15	16/03/2020	0001301250	31019
12	4.350,01.12	1960/07/01 a 1960/12/31	16/03/2020	0001301241	31020
13	4.350,01.13	1960/10/01 a 1960/12/31	16/03/2020	0002053233	31021
14	4.350,01.14	1961/01 a 1961/04	16/03/2020	0001396935	31022
15	4.350,01.15	1961/05/01 a 1961/08/15	16/03/2020	0001301292	31023

Ocorrências: _____

Recebido por: Alex da Silva

Coordenador

Fundação Biblioteca Nacional - Av. Rio Branco, 219 - Centro - Rio de Janeiro, RJ - CEP: 20040-914

Fundação Biblioteca Nacional

Mat. STAPE 1552305

3 - CODIFICAÇÃO/NOMEAÇÃO DAS PASTAS E ARQUIVOS

3.1 - Codificação/nomeação das pastas e arquivos de acervos especiais

O processo de codificação/nomeação das pastas e arquivos de imagens produzidos no Laboratório de Digitalização é uma das etapas mais importantes de todo o fluxo. Com uma alta produção de arquivos durante o fluxo de trabalho no laboratório, uma pasta e arquivo nomeados erradamente poderá ocasionar a perda de ambos, por dissociá-los com o documento original, criando a necessidade de uma nova digitalização.

Na codificação/nomeação de pastas e arquivos de imagens produzidos no Laboratório de Digitalização, utilizamos códigos alfanuméricos que espelham a codificação dos documentos físicos, que é o mesmo código ao qual eles são lançados nas bases de dados da Instituição.

Essa codificação/nomeação é feita com o uso do código do setor responsável pela guarda do documento na Biblioteca Nacional, mais o número patrimonial da obra dado por esse setor. A exceção dessa regra é em relação as obras provenientes do setor de Periódicos. A codificação/nomeação dos arquivos digitalizados das obras desse setor será previamente abordada no final desse capítulo e com mais propriedade no manual de operação do equipamento Zeutschel, onde é feita a digitalização do fluxo de obras desse setor.

Os códigos dos setores são os seguintes:

Setor de Obras Raras	or
Setor de Iconografia	icon
Setor de Cartografia	cart
Setor de Manuscritos	mss
Setor de Obras Gerais	drg
Setor de Música	mas

O número patrimonial da obra é encontrado no encaminhamento que acompanha a movimentação da obra ao Laboratório de Digitalização e se encontra escrito na mesma (o número de patrimônio tem que estar escrito e carimbado pelo setor responsável pela guarda da obra, caso não esteja, ela deve voltar ao setor para ser inserido antes da digitalização).

Obs.: durante o processo de digitalização e posterior nomeação das pastas e arquivos, é importante saber a proveniência do documento em relação ao setor responsável pela guarda da obra, pois a codificação dos arquivos de imagens digitais leva em consideração esse setor. No documento não existe menção a esse setor, só existe o código patrimonial. A proveniência do setor é verificada através do encaminhamento que acompanha a movimentação da obra.

Exemplificaremos abaixo como são feitas as codificações/nomeações das pastas:

SETOR DE GUARDA DO DOCUMENTO	NÚMERO PATRIMONIAL DO DOCUMENTO	CÓDIGO/NOME DA PASTA
Obras Raras	120213	or120213
Iconografia	1203418	icon1203418
Cartografia	1614	Cart1614
Manuscritos	814612	mss814612
Obras Gerais	5143	drg5143
Música	73213	mas73213

Todas as letras dos códigos dos setores devem ser escritas em minúsculo.

A nomeação dos arquivos de imagens que se encontram dentro das pastas mantém a mesma codificação/nomeação das pastas, mas o acréscimo de _ (*underline*), mais a numeração sequencial do arquivo dentro da pasta.

Exemplificaremos abaixo como são feitas as codificações/nomeações dos arquivos de imagens:

CÓDIGO/NOME DA PASTA	QUANTIDADE DE ARQUIVOS DE IMAGENS DENTRO DA PASTA	CÓDIGO/NOME DOS ARQUIVOS DE IMAGENS
or120213	302	or120213_0001 a or120213_1302
icon1203418	2	icon1203418
cart1614	2	mss814612_01 a mss814612_22
mss814612	22	mss814612
drg5143	325	drg5143_001 a drg5143_325
mas73213	13	mas73213_01 a mas73213_13

Obs.: não pode haver espaços em branco na codificação dos arquivos.

3.2 - Codificação/nomeação das pastas e arquivos de periódicos

Para criação das pastas para cada fascículo de periódico digitalizado, o padrão de nomenclatura das pastas é o seguinte:

per000000_0000_0000
(perBIN_ANODATA_Nº FASCÍCULO)

per (abreviação de periódico)

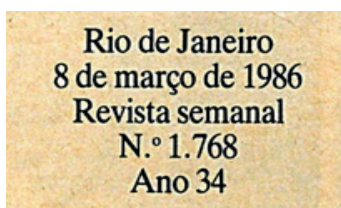
000000 (campo que deve ser usado para o código BIN da publicação seriada no cadastro de periódico, este código é dado pelo Setor de Periódicos. Sempre deve ser usado 6 casas decimais nesse nível de codificação, se o código tiver apenas 5 dígitos ou menos, deve ser complementado com o zero. Ex.: per002633, per015482, per000158)

0000 (campo que deve ser usado para o ano da data de lançamento da publicação, sempre deve ser usado 4 casas decimais)

0000 (campo que deve ser usado para o número do fascículo da publicação, sempre deve ser usado 4 casas decimais)

underline_ (sempre deve ser usado o *underline* para separar os campos dos códigos)

Exemplo:



per004120_1986_1768

Sendo:

4120 o código BIN da <i>Revista Manchete</i>
1986 o ano da data de lançamento
1768 o número do fascículo

Exceções no padrão de codificação das pastas

Quase todas as publicações seriadas já digitalizados no Laboratório de Digitalização tem o seu número de fascículo sequenciado de um ao infinito, sendo o número 1 a publicação inaugural da revista, mas existe algumas publicações que reiniciam a numeração dos fascículos quando mudam o ano da publicação, ou quando criam uma lógica de VOLUME ou TOMO dentro da organização original da publicação. Nesses casos, quando a mudança do ano da revista não é em janeiro, é necessário para que a organização das pastas mantenha a ordem de lançamento das revistas, criar um novo campo para a codificação das pastas.

Para esse caso específico, citado acima, o padrão de nomenclatura das pastas é o seguinte:

per000000_0000_0000_0000
(perBIN_ANODATA_ANO DA REVISTAou VOLUME_Nº FASCÍCULO)

Campo acrescentado no código acima:

0000 (campo que deve ser usado para o ano da revista ou volume, sempre deve ser usado 4 casas decimais)

Exemplo:



per171379_1909_0003_0001

Sendo:

171379 o código BIN da <i>Revista Boletim Policial</i>
1909 o ano da data de lançamento
0003 o ano da revista
0001 o número do fascículo

Um outro caso em que se faz necessário criar mais um campo de codificação ocorre quando o fascículo digitalizado possui mais de uma edição dentro da mesma brochura.

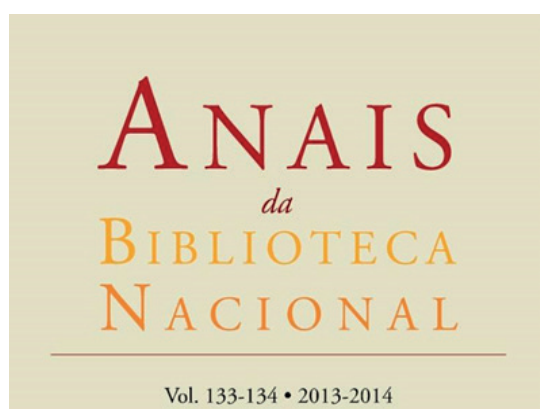


per176648_1930_0015_0004_0005

Sendo:

176648 o código BIN da <i>Revista Informação Goyana</i>
1930 o ano da data de lançamento
0015 o volume
0004_0005 os números dos fascículos

Mais um caso que se faz necessário criar mais um campo de codificação é quando o fascículo digitalizado possui mais de um ano de edição dentro da mesma brochura.



per002344_2013_2014_0133_0134

Sendo:

2344 o código BIN da <i>Revista Anais da Biblioteca Nacional</i>
2013_2014 os anos da data de lançamento
0133_0134 os números dos fascículos

Obs.: sempre que houver um periódico que tenha uma necessidade de codificação diferenciada, em que os exemplos acima não sejam adequados, essa questão deverá ser levada para o grupo, a fim de que a decisão final seja implementada e padronizada por todos.

4 - CONTROLE DE QUALIDADE DAS IMAGENS DIGITALIZADAS

O processo de controle de qualidade feito no Laboratório de Digitalização da Biblioteca Nacional tem como objetivo garantir o máximo de fidelidade entre a captura gerada pelos equipamentos de digitalização e o documento original, levando em consideração todas as suas características físicas.

Todos os processos de conferência descritos aqui correspondem ao fluxo de digitalização dos arquivos digitalizados no processo com as câmeras de alta resolução.

4.1 - Formato de arquivos: *master* e derivados

No processo de trabalho do laboratório de digitalização são usadas três extensões/formatos de arquivos:

- TIFF (*Tagged Image Format*);
- JPEG2000 (*JP2*);
- JPEG (*Photographic Experts Groups*).

Os formatos TIFF e JPEG2000 são utilizados como arquivos *MASTER*, enquanto o JPEG é utilizado para os arquivos derivados – reprodução e difusão.

Os arquivos JPEG2000 são utilizados no fluxo de captura dos scanners planetário, e serão abordados posteriormente. Nesse capítulo, falaremos dos arquivos TIFF e JPEG.

Arquivo *master* (matriz digital)

Formato TIFF, matriz de alta resolução, com compressão sem perda de qualidade aparente, que permita uma melhor visualização e acesso. Deve-se evitar interferências estéticas e o uso de filtros e outras ferramentas que alterem ou distorçam a integridade e autenticidade da captura digital em termos de fidelidade visual ao documento original. O documento nesse Arquivo *Master* deve ser capturado por inteiro, sem nenhum corte. Deve haver sobras em todos os lados de aproximadamente 2 centímetros de toda página ou documento inteiro digitalizado, respeitando e garantindo o aproveitamento máximo das características das câmeras de médio formato que tem o CCD já pré-definido. Nesse arquivo, junto ao documento, deve estar a escala de cor.

Arquivos derivados (reprodução - difusão)

A partir do arquivo *master*, através do uso de um *software* de tratamento de imagem é feita a compressão em formato JPG e resolução 300 dpi para criar um arquivo digital bem mais leve em relação a tamanho, facilitando o seu acesso, disseminação e uso. Os arquivos de difusão serão abordados posteriormente em outro capítulo.

4.2 - Processos de controle de qualidade

Conferência da nomenclatura

Antes do início do processo de controle, é preciso verificar se a nomenclatura dada ao arquivo digital que chamamos de Objeto Digital corresponde ao Registro Patrimonial do documento original. Toda obra tem que ter seu registro, que pode ser consultado no verso da folha de rosto ou contra capa. Havendo alguma divergência de informação a consulta pode ser realizada pelo seu cadastro no SophiA ou informações inseridas no Formulário de Encaminhamento de Obras (FEO), o documento original retorna para a captura ou até mesmo ao setor de guarda para correção.

Conferência da qualidade do arquivo produzido

A imagem digitalizada é conferida com o uso de *software* de tratamento de imagem, em escala 1:1 - 100%, com uso da escala de cor. Todo processo de conferência é realizado com consulta constante ao documento original. São conferidos os seguintes aspectos do arquivo:

Foco

É feita uma varredura em toda extensão do arquivo digital gerado, verificando o foco da imagem, identificando se em algum ponto houve perda de nitidez conferindo e comparado ao original.

Interferências

Ainda no mesmo processo também é observado qualquer interferência na imagem, tais como sujeira, enquadramento, presença de interferência em imagens com linhas - Padrão Moiré, aspecto granulado, reflexos etc.

Cor

Os valores de RGB (RED, GREEN e BLUE - as cores primárias) são conferidos na escala de cor com uso de *software* de tratamento para evitar algum desequilíbrio de cor. A partir da ferramenta “conta gotas” checamos os valores de cinza da escala de cor no arquivo digitalizado. Esses valores tem que ser iguais, pois o cinza é uma cor neutra, caso tenhamos algum desequilíbrio em um dos valores de RGB, essa imagem terá que ser refeita.

Densidade / luminosidade

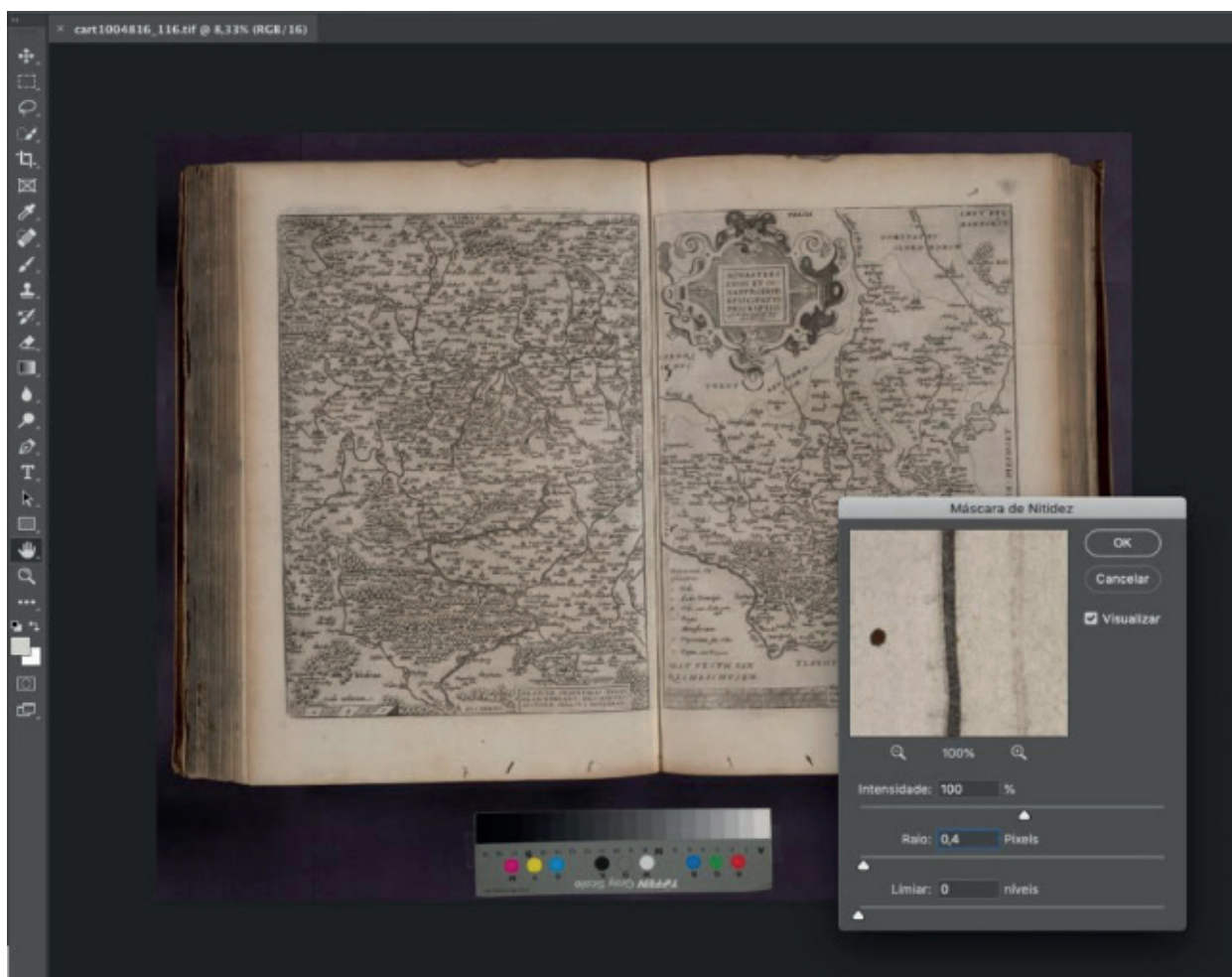
Na escala de cor existe uma gradação de cinzas que vai do preto ao branco, passando por várias tonalidades de cinzas, com uso desse recurso da escala, é conferida a densidade da imagem (se ela está muito clara ou muito escura). Além das referências da escala, é importante, a todo tempo, comparar com o documento original. Em caso de digitalização de imagens de alto contraste, é importante verificar nas zonas mais escuras e mais claras se existe alguma perda de informação no arquivo digitalizado em comparação com o documento.

Montagem de imagens

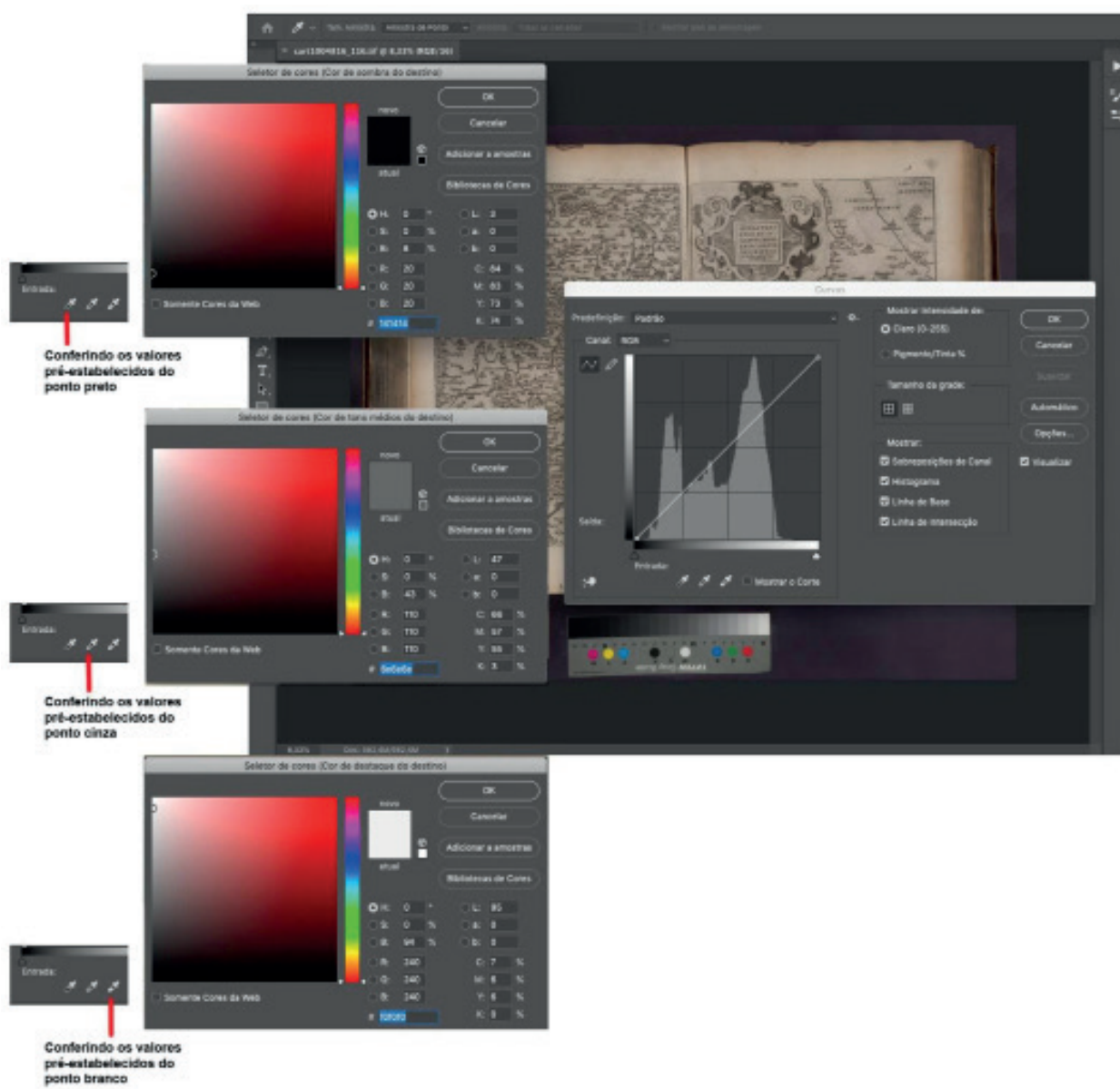
Alguns originais de grandes formatos, em sua maioria Cartografias e Iconografias, são capturados por partes obedecendo os limites de cada equipamento utilizado para sua digitalização. Nesses casos, além dos processos de controle de qualidade, é feita a junção dessas partes. Essa junção pode ser feita de modo automático através das ferramentas disponibilizadas pelo software de tratamento de imagem ou de forma manual através das junções das camadas com cada parte digitalizada. É importante ressaltar que os arquivos sejam visualizados em escala 1:1 – 100% para garantir que a junção obedeça aos critérios da autenticidade e integridade do documento original.

Geração de arquivos derivados - reprodução

A partir das imagens *master*, que já passaram por todo o processo de conferência, geramos os arquivos derivados JPG com compressão em 12 bits. As imagens podem receber alguns tipos de tratamento, como aplicação de filtros de nitidez não ultrapassando o raio de 0,4.



Ou correção de cores, obedecendo os valores pré-estabelecidos para o ponto preto (R=20 G=20 B=20), ponto cinza (R=110 G=110 B=110) e o ponto branco (R=240 G=240 B=240), a fim de permitir melhor visualização ou impressão.



As imagens são disponibilizadas com a ausência da escala de cinza e de margens pretas. Entretanto, critérios éticos devem pautar esse tipo de intervenção para que elas não se tornem dissociadas e não representem corretamente o documento original que as gerou.

Os arquivos derivados são conferidos e comparados aos arquivos *master* para garantir que não exista ausência de documentos digitalizados em relação ao documento original.

5 - METADADOS NO SISTEMA SOPHIA

5.1 - Inserindo metadados da digitalização no sistema SophiA

As obras que são digitalizadas pelo Laboratório de Digitalização, que futuramente estarão disponibilizadas na *web* para consulta, estão registradas no SophiA. Este é um *software* para gestão de bibliotecas e acervos. Nele iremos inserir as informações pertinentes ao processo de digitalização, os metadados dos arquivos *master* de preservação.

Selecionadas as imagens que vamos trabalhar, criamos um arquivo derivado “thumbnail” para cada registro. Este é uma pequena imagem, que aparece quando se pesquisa determinada obra. Esta imagem que será inserida no SophiA é de qualidade reduzida, com 72 dpi e reduzida em 50%. Exemplos: capa do livro, frente da obra, primeira página etc.

Para este processo usamos o “Objeto Digital” como número de registro da obra para acessá-la no sistema. Após introduzir este no campo de “busca” (somente nas abas obras ou analíticas), podemos editar o registro. Para editá-lo, basta clicar no “lápis” e aparecerá uma janela menor, como na imagem abaixo.

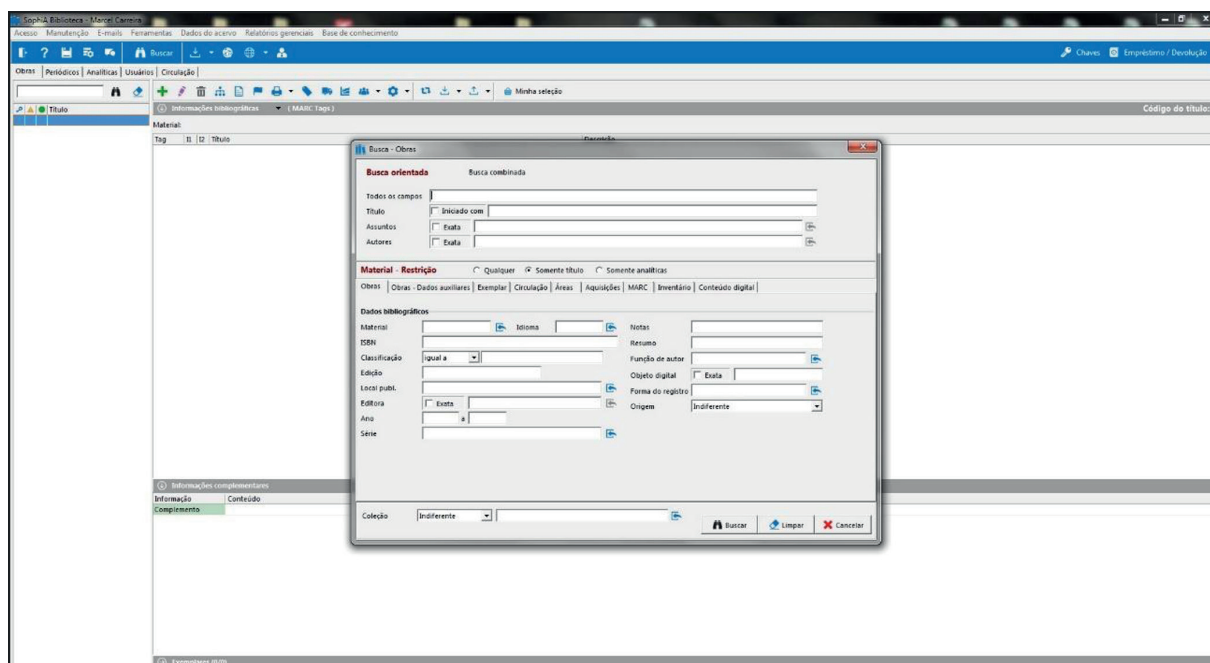


Imagem do SophiA - Resultado da busca

Na janela que apareceu temos, à esquerda, vários campos. Vamos fazer alterações/inserir dados em três deles.

1 - Informações básicas

Selecione esta aba para adicionar o “thumbnail”. Clicamos em “+” abaixo da janela embaixo à direita. Selecionamos o arquivo e damos ok.

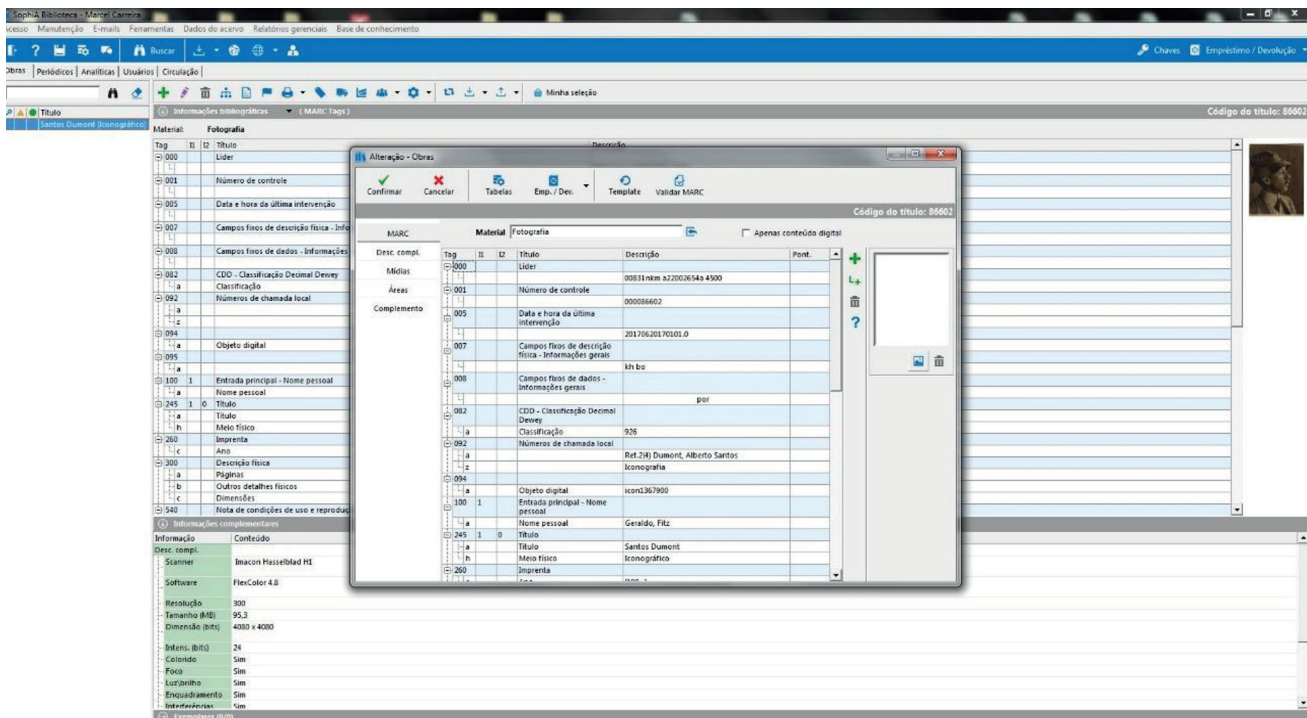


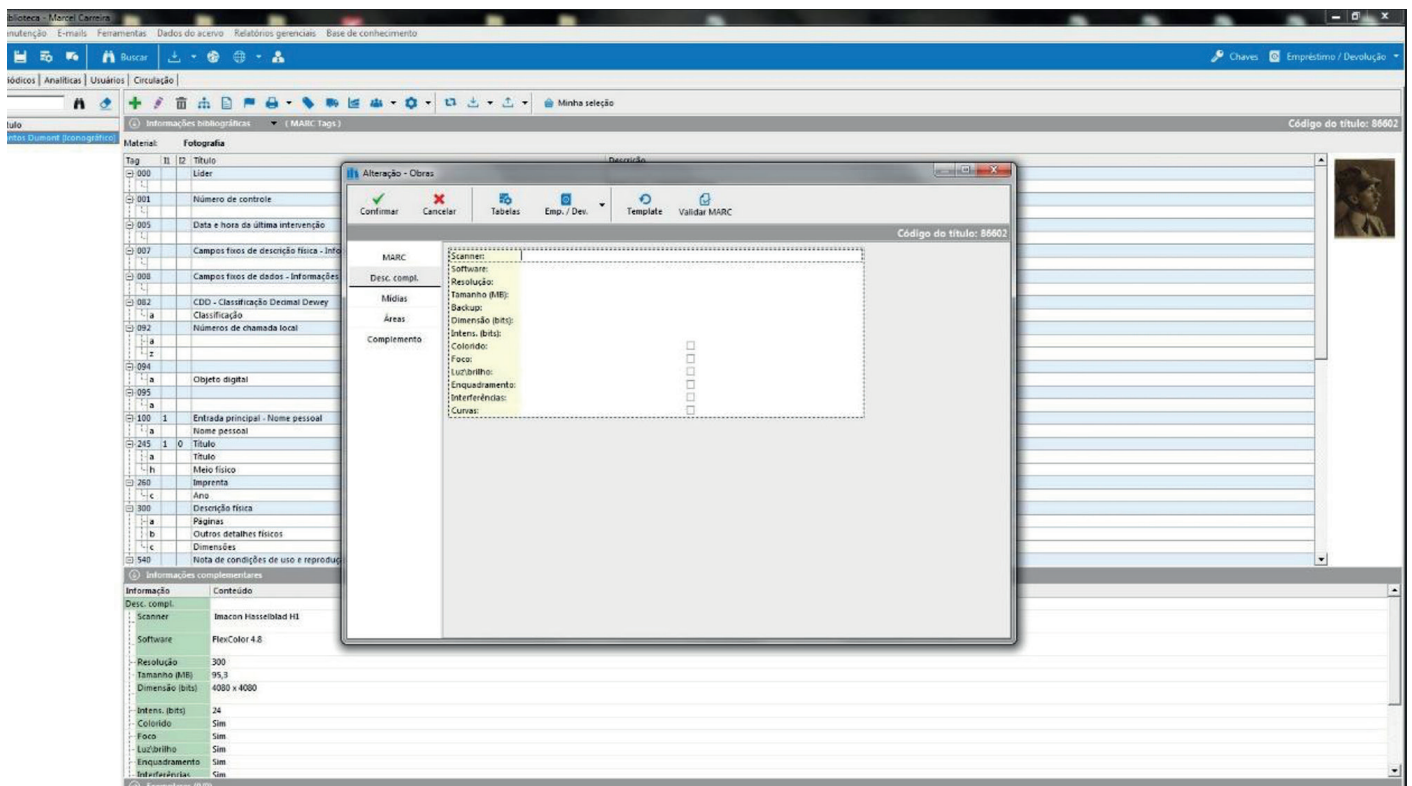
Imagem do SophiA - Imagem após clicar no +

2 - Desc. compl.

Vamos inserir nesta aba informações que são referentes aos arquivos que foram gerados na digitalização e que são específicos para cada registro.

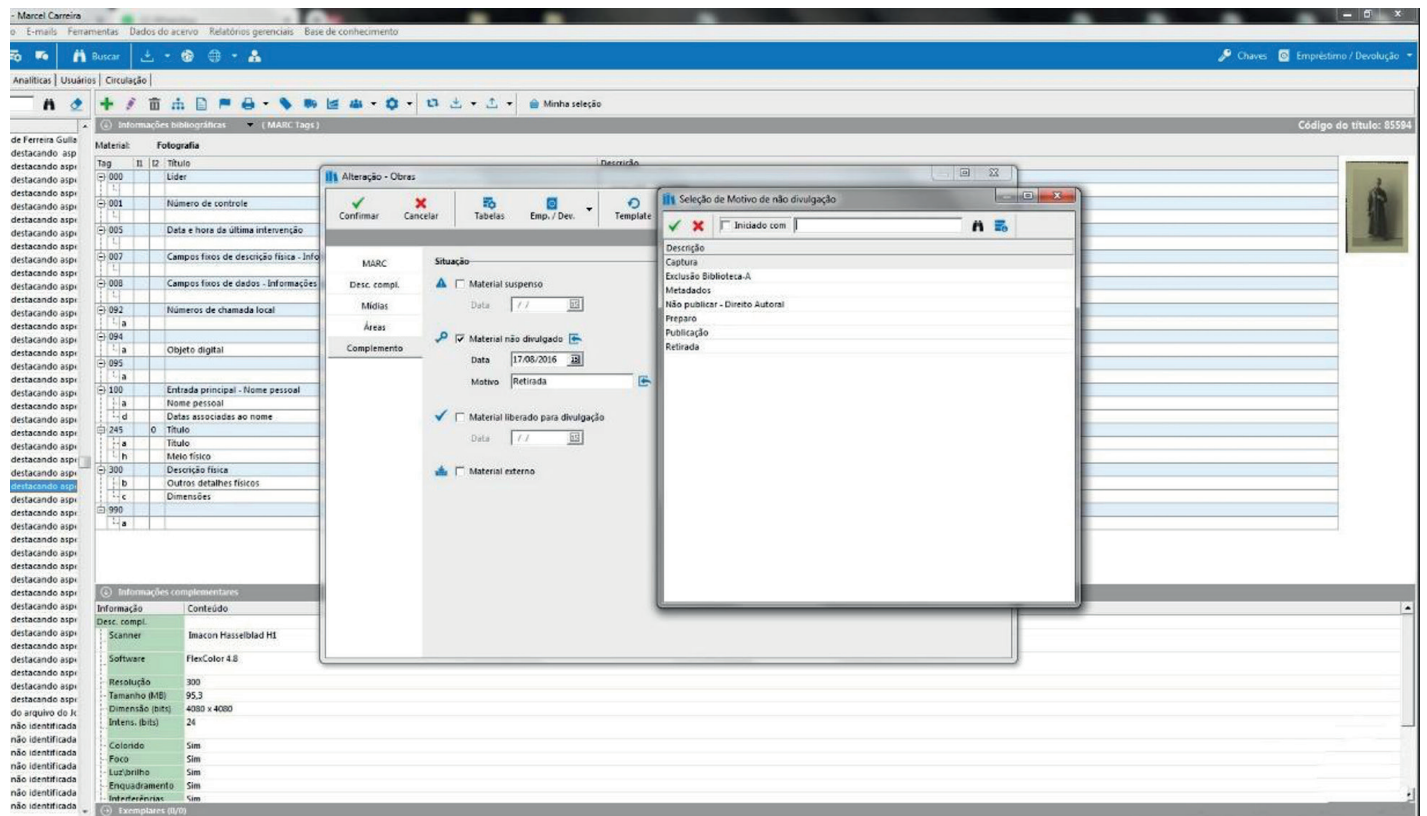
Tipo:

- Scanner: em qual equipamento foi capturado;
- *Software*: qual o programa utilizado na captura;
- Resolução: sempre 300 dpi;
- Tamanho: em “MB” do arquivo;
- *Backup*: se usava para quando o *backup* era feito em DVD, após o implemento do *Storage* este campo fica em branco (vazio);
- Dimensão (bits): tamanho em bits do arquivo gerado;
- Intensidade (bits): por serem sempre três canais de bits na configuração de captura em todos os equipamentos, temos o “Valor x 3” como 24 (3x8) nos scanners planetários ou 48 (3x16) nas câmeras de médio e grande formato;
- Colorido, Foco, Luz/Brilho, Enquadramento, Interferências e Curvas: temos ao lado da descrição “caixas” vazias onde se preenche um “x”, clicando sobre estas. Todas essas opções deverão ser ativadas/marcadas.



3 - Complemento

Nesta aba mudamos o *status* da obra de “retirada” para “metadados”.



Após uma breve conferência dos dados que foram alterados/inseridos, salvamos as alterações clicando em “Confirmar”.

A tela inicial será atualizada com o registro atual e suas novas informações.

6 - ARQUIVOS DERIVADOS DE DIFUSÃO

Procedimentos técnicos para os arquivos derivados de difusão:

6.1 - Tipos de obras

Para cada tipo de obra são utilizados arquivos derivados específicos. Seguem abaixo as especificações.

ICON

Obras com página única, são utilizados os arquivos JPG simples e o Dragão Marinho. Obras com página e verso, são utilizados os arquivos PDF e o Dragão Marinho.

Para álbum de fotografias será feito: uma Galeria em HTML, o Dragão Marinho, JPG simples e um PDF.

Para obras com imagem e texto, encadernadas, são utilizadas a Galeria, o Dragão Marinho, JPG simples e o PDF.

CART

Obra com página única, são utilizados os arquivos JPG simples e o Dragão Marinho. No caso de obras com páginas montadas, deve-se utilizar o PDF e o Dragão Marinho.

Para obras com imagem e texto, encadernadas, são utilizadas a Galeria, o Dragão Marinho, JPG simples e o PDF.

MSS / DRG / OR / MAS

Se a obra tem página única, deve-se utilizar o arquivo JPG simples, para duas páginas, deve-se utilizar o PDF, e acima de 3 páginas, utiliza-se o PDF (ou PDF OCR) e *Book-Reader*. O *Book-Reader* é utilizado somente se as imagens estiverem na mesma proporção de tamanho; caso isso não ocorra, será feito somente o PDF (ou PDF OCR).

Para obras com imagem e texto, encadernadas, utiliza-se a Galeria, o Dragão Marinho, JPG simples e o PDF.

Caso apareça algum texto com letras pequenas, fotos, gravuras, e desenhos, utiliza-se o Dragão Marinho.

Obs.:

- Todas as imagens devem ser posicionadas na horizontal para melhor leitura.
- Quando fizer o Dragão Marinho para mais de uma imagem, não precisa deixar os arquivos JPG soltos na pasta, a menos que seja uma Galeria.

6.2 - Configurações dos arquivos derivados

6.2.1 - Preparo dos arquivos a serem utilizados

Apagar dos arquivos as imagens não usadas, por exemplo, capas não originais, carimbo, páginas em branco, inscrição do registro patrimonial.

Criar JPG reduzido de tudo, aplicando a *bath* JPG 50% (Anexo 2) no *Photoshop CS3*.

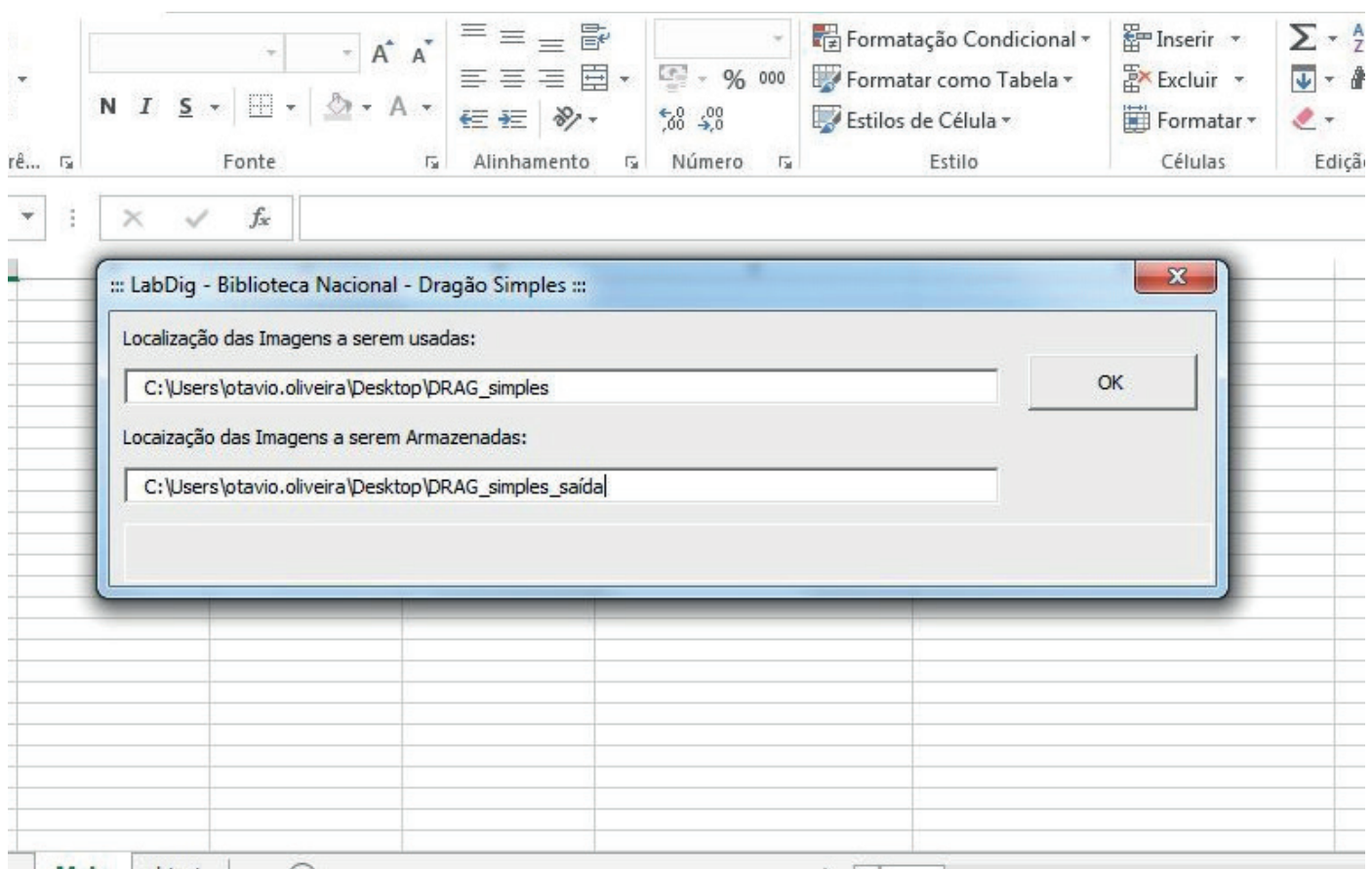
Separar as obras pelo tipo de derivadas a serem feitas, por exemplo, galeria, dragão simples, dragão frente e verso, PDF e *Book-Reader*, e mover para as pastas criadas para cada tipo de obra.

6.3 - Dragão simples (1 ou 2 páginas)

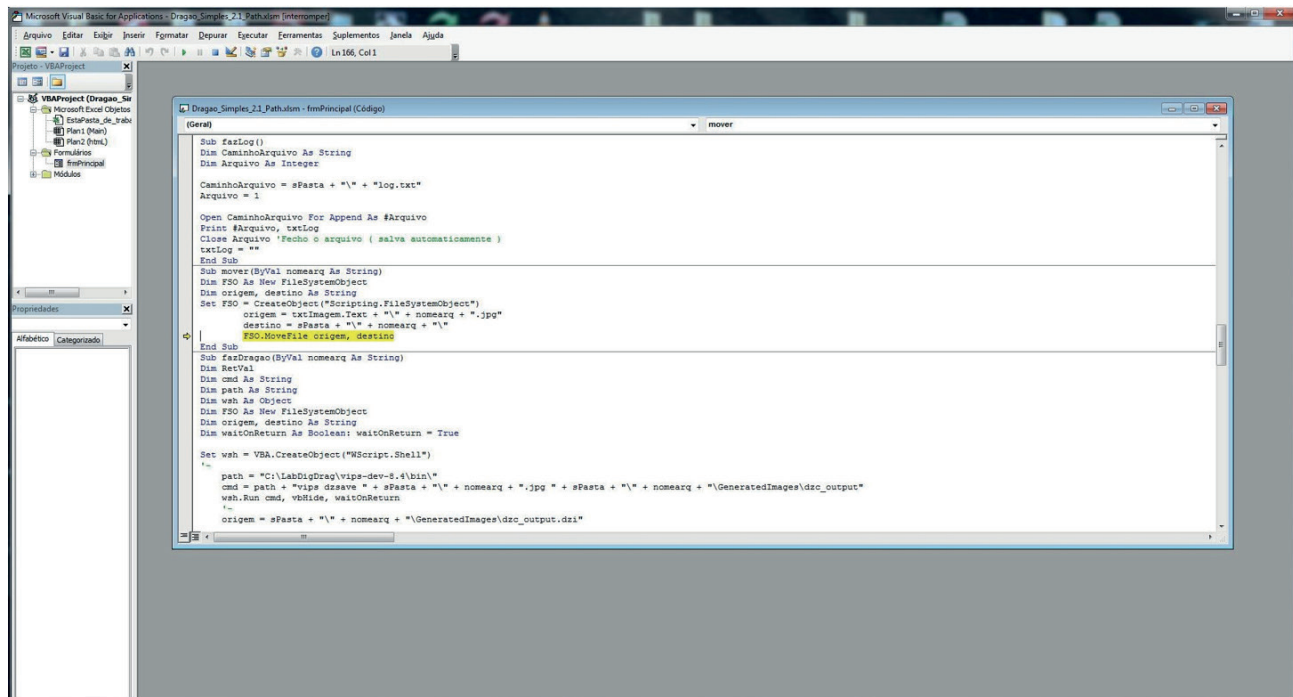
Copiar a imagem *master* da pasta entrada para a pasta DRAG_simples no desktop, e a imagem reduzida para a pasta DRAG_simples_saída.

Aplicar a planilha “**Dragão simples_2.1.xlsm**”.

Incluir os caminhos das pastas onde as imagens *master* e reduzidas se encontram.



Quando a próxima tela surgir, prosseguir clicando na tecla F5 e ok.



Na pasta DRAG_simples será criada a pasta do Dragão (arquivos do dragão e o JPG reduzido). Apagar na pasta raiz os arquivos *master.jpg* e *log.txt*.

Testar o dragão (arquivo html) utilizando o navegador *Firefox*, o único que funciona.

6.4 - Dragão frente-verso

As imagens têm que ter a mesma direção e orientação (vertical/vertical-horizontal/horizontal).

Renomear os arquivos *masters* para o formato frente e verso.

Exemplo:

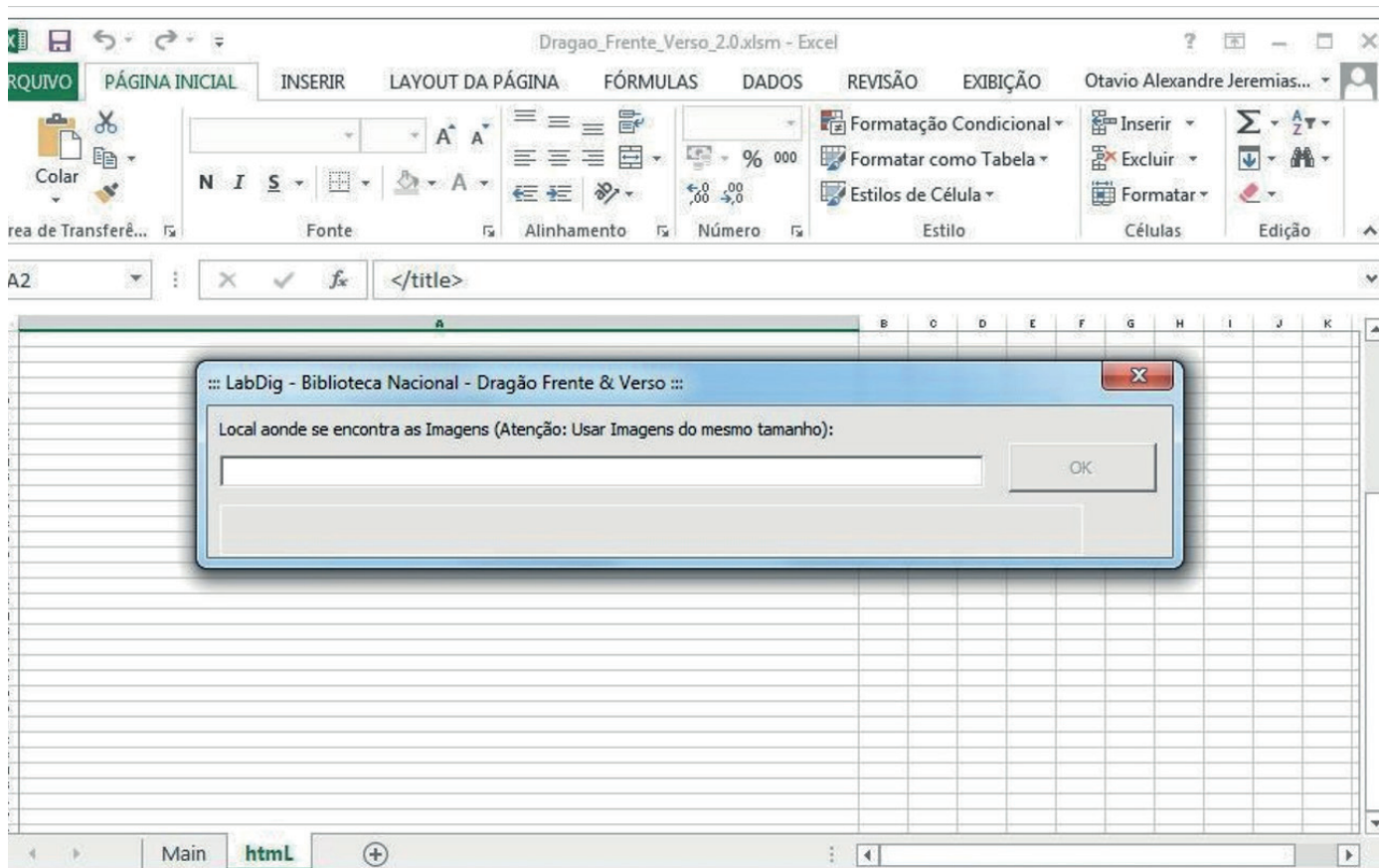
icon12345 e

icon12345 e icon12345v

Copiar a imagem *master* da pasta entrada para a pasta DRAG_frente_verso no desktop.

Aplicar a planilha “**Dragão frente_verso_2.0.xlsm**”.

Incluir o caminho da pasta onde as imagens *masters* se encontram.



Na pasta DRAG_frente_verso será criada a pasta do Dragão (arquivos do dragão e o JPG reduzido - apagar este arquivo). Apagar na pasta raiz os arquivos *master.jpg* e *log.txt*.

Testar o dragão (arquivo html) utilizando o navegador *Firefox*, o único que funciona.

Fazer o arquivo PDF com as imagens *masters* conforme item abaixo.

Quando finalizado, mover para a raiz da pasta do Dragão, em pasta DRAG_frente_verso.

6.5 - PDF/PDF OCR

Usar o JPG simples e *smallest minimum*.

Pode ser necessária uma nova redução das imagens caso o tamanho do PDF seja maior que 70 Mb.

Este é o uso padrão, mas dependendo da qualidade da obra e da quantidade de páginas essas configurações serão ajustadas para uma melhor leitura.

Criando PDF com o aplicativo “**Adobe Acrobat Reader 7.0 Professional**”.

- Com o aplicativo aberto clicar em: *File* (Arquivo) – *Create PDF* (Criar PDF) – *From Multiple Files* (a partir de múltiplos arquivos);
- Em “**Add Files (Adicionar os Arquivos)**” clicar em *Browse...* (Localizar);
- Localizar o local (pasta) com os arquivos, selecionar todos e clicar em *Add* (adicionar), após isto, clicar em ok;
- Abrirá uma janela com o nome “**Save As**” (Salvar Como). Selecionar o local (pasta) onde será criado o arquivo PDF e clicar em Salvar.

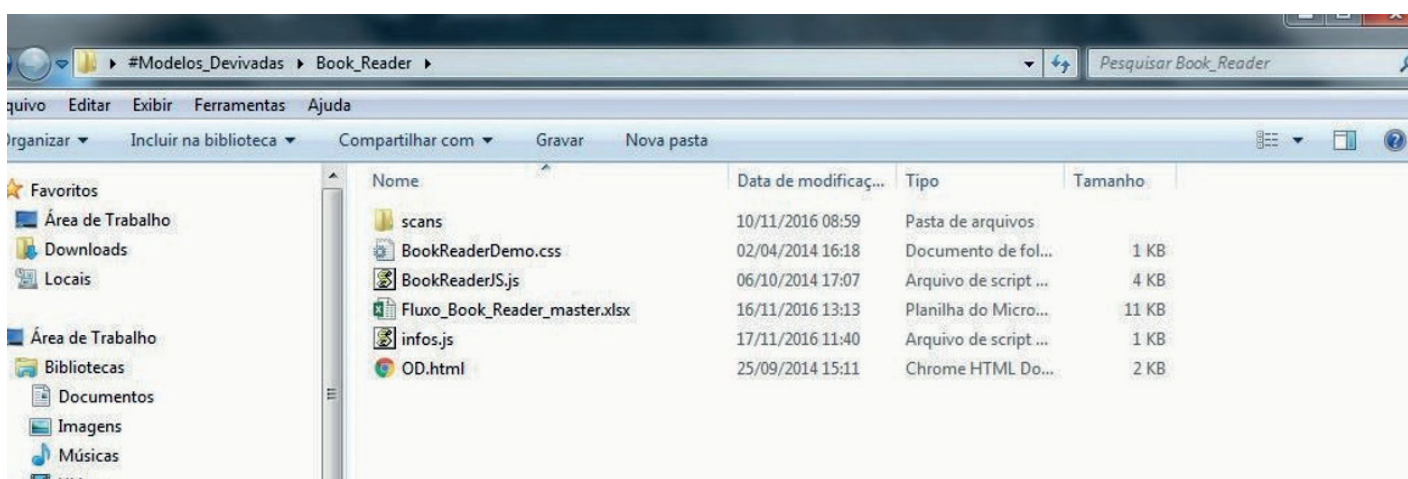
Obs.: o arquivo gerado já vai estar aberto pelo programa para conferência.

6.6 - Book-Reader

É acompanhado de um PDF ou PDF OCR.

Copiar os arquivos *master* para a pasta BR_entrada.

Criar uma pasta com o OD da obra, na pasta BR_saída, onde serão incluir os arquivos modelos da pasta “*Book-Reader*”. Seguem os arquivos abaixo:



Para uma **grande quantidade** de arquivos, utilizar a planilha excel, onde automatiza o fluxo dos processos. Não é obrigatório a sua utilização, apenas uma ajuda, a planilha se chama: **“Fluxo_BookReader”**. Ela funciona de copiar e colar comandos em DOS (Anexo 4).

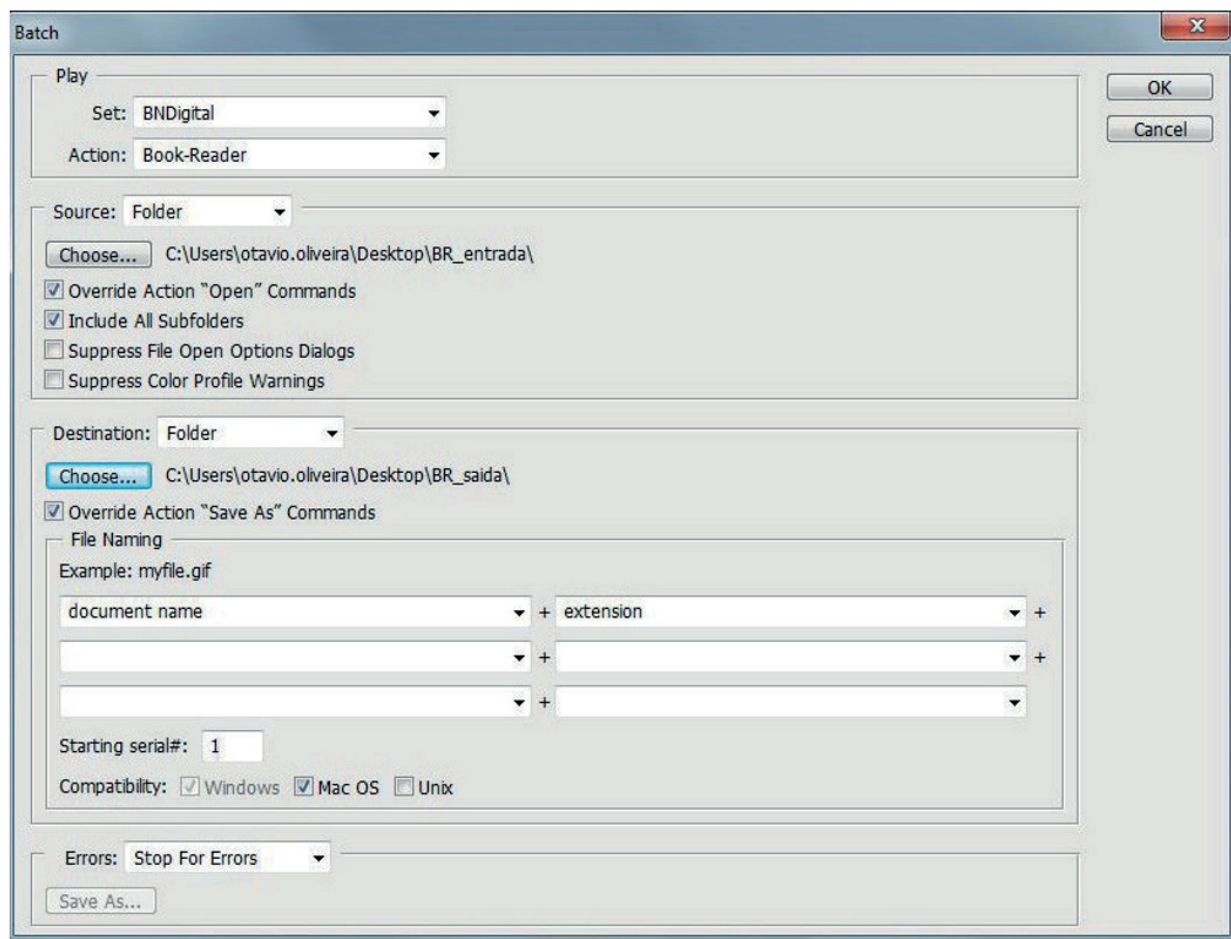
Usar a *batch* **“BookReader”** (Anexo 2) no *Photoshop CS3*.

Para iniciar o processo, selecione **“File”, “Automate”** e **“Batch...”**.

Depois selecionar no campo **“action” “BookReader”**.

Verificar se o caminho de origem e destino dos arquivos estão corretos.

Apertar ok.



Mover os arquivos criados para dentro da pasta *scans*.

Renomear as imagens no formato crescente: “**page0001.jpg; page0002.jpg**” etc.

Sugestão: usar o programa “**Renomear tudo**” ou outro similar (Anexo 3).

Editar o arquivo *infos.js*.

Substituir onde estiver marcado com “**X**” por, “*Número de página*”, “**título do livro**”, no link download “**div_acervo/od/od.pdf**”, tamanho e código do título do SophiA.



Renomear as imagens no formato crescente: “**page0001.jpg; page0002.jpg**” etc.

Sugestão: usar o programa “**Renomear tudo**” ou outro similar (Anexo 3).

6.7 - Galeria

Mover as imagens JPG *master*, um OD de cada vez, para a pasta “**GAL_Dragão**” e excluir as imagens não utilizadas para o Dragão Marinho.

Mover as imagens reduzidas para a pasta “**GAL_reduzida**”.

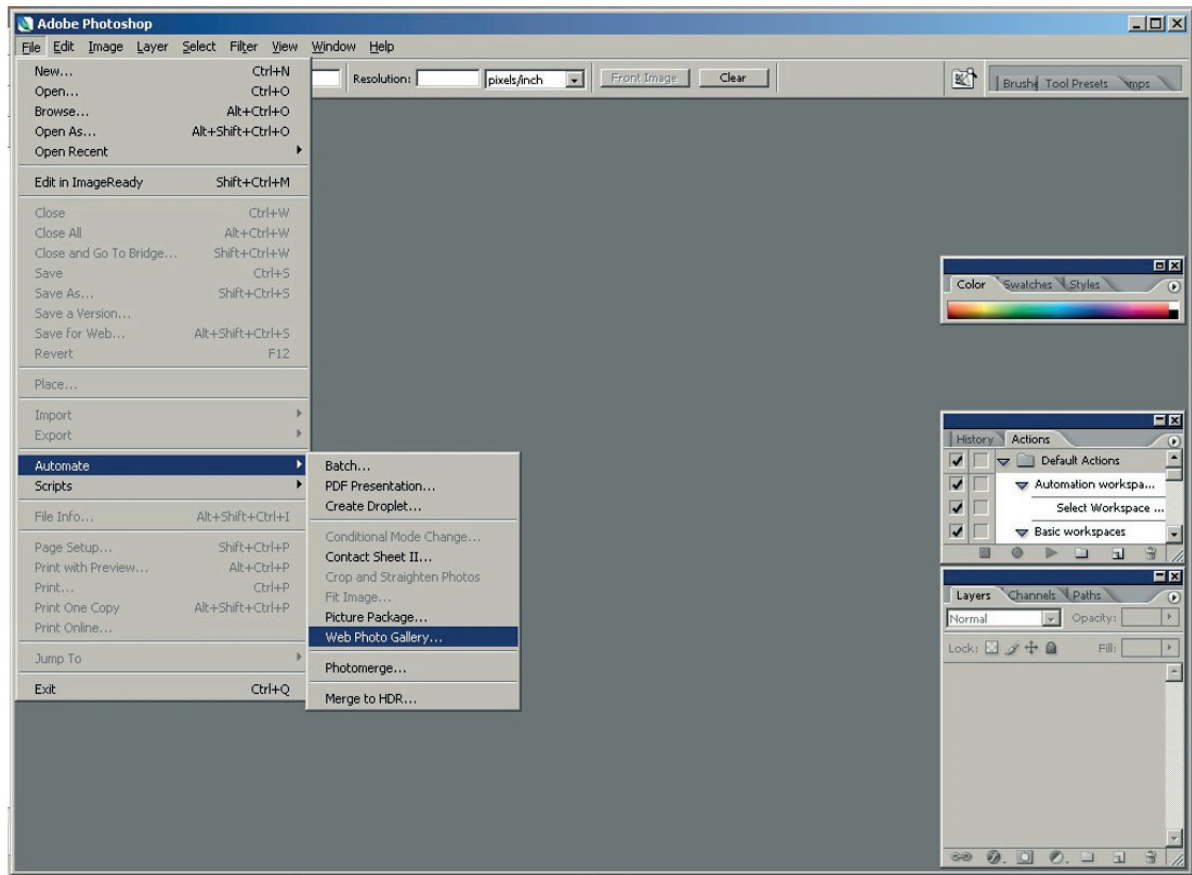
ATENÇÃO:

- Sempre observar se as imagens estão na posição correta (horizontal); se não estiverem, rotacionar e salvar.
- Sempre observar a ordem das imagens geradas pelo *Photoshop*, a posição das imagens deve estar correta, pois quando uma obra tiver **vários ODs diferentes**, nem sempre ele vai manter a ordem desejada. Ex.: álbum.

A galeria é acompanhada de um PDF ou PDF OCR.

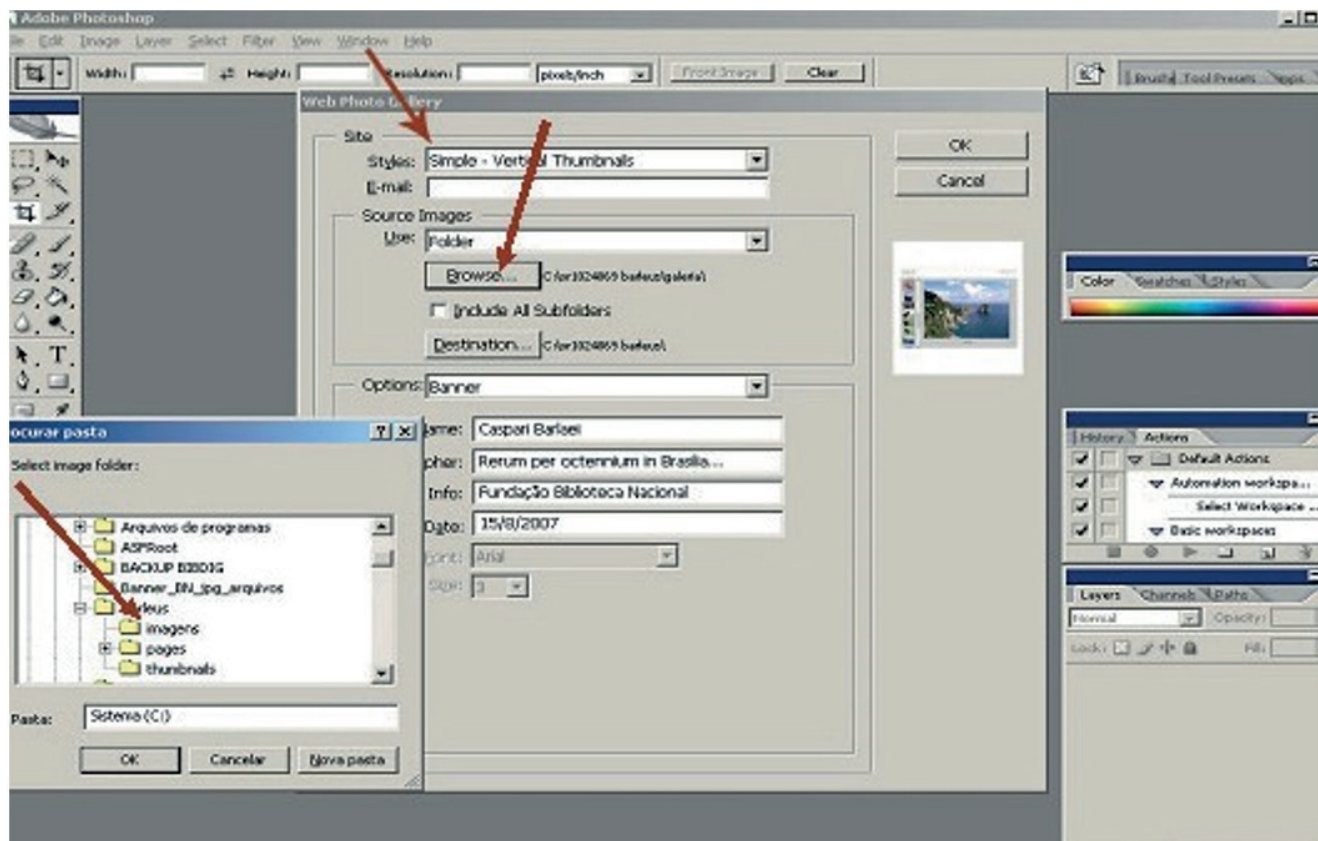
Abrir o *Adobe Photoshop CS3*.

Selecionar “**File**”, “**Automate**” e “**Web photo gallery...**”.

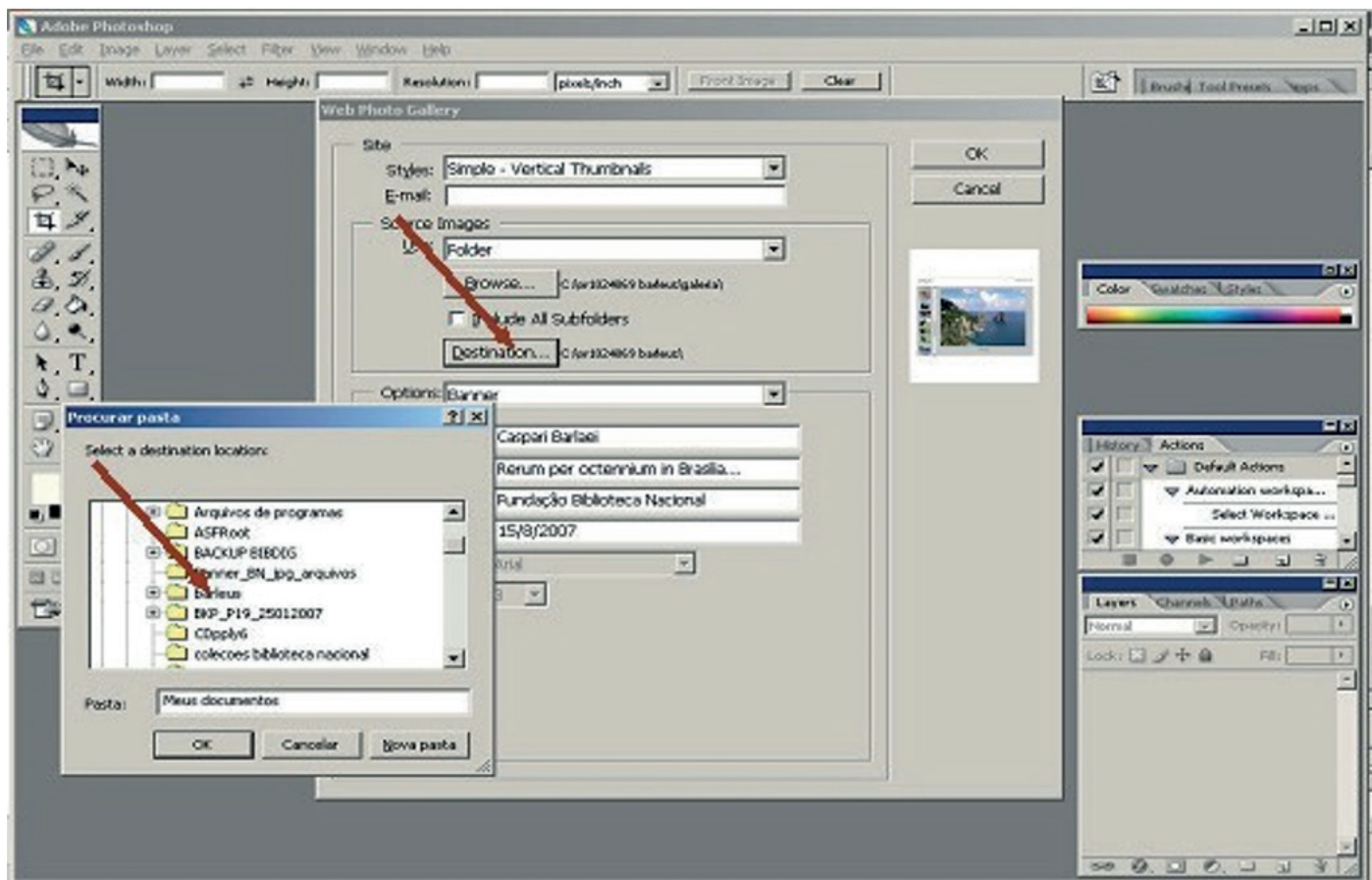


Em “**Style**”, escolher a opção: “**Simple Vertical thumbnails**”.

Depois selecionar em “**Source imagens**” a pasta onde contém as imagens criadas na pasta “**GAL_reduzida**”, clicando no botão “**Browse**”.

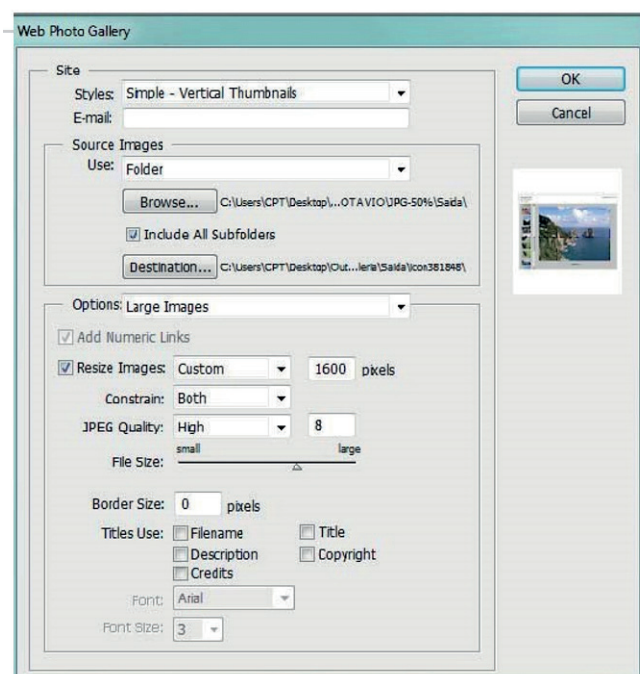


Clicar no botão “**Destination...**” e seleccionar a pasta destino para as imagens da galeria, a pasta “**GAL_saída**”.



Em “**Options**”, escolher a opção “**Large Images**”.

Depois seleccionar em “**Resize Images**”, “**Custom**” e incluir **1600**, nosso padrão (ou 800-1600), marcar em “**Jpg Quality**”, “**High - 8**”; e desmarcar “**Filename**” em “**titles use**” e dar ok.



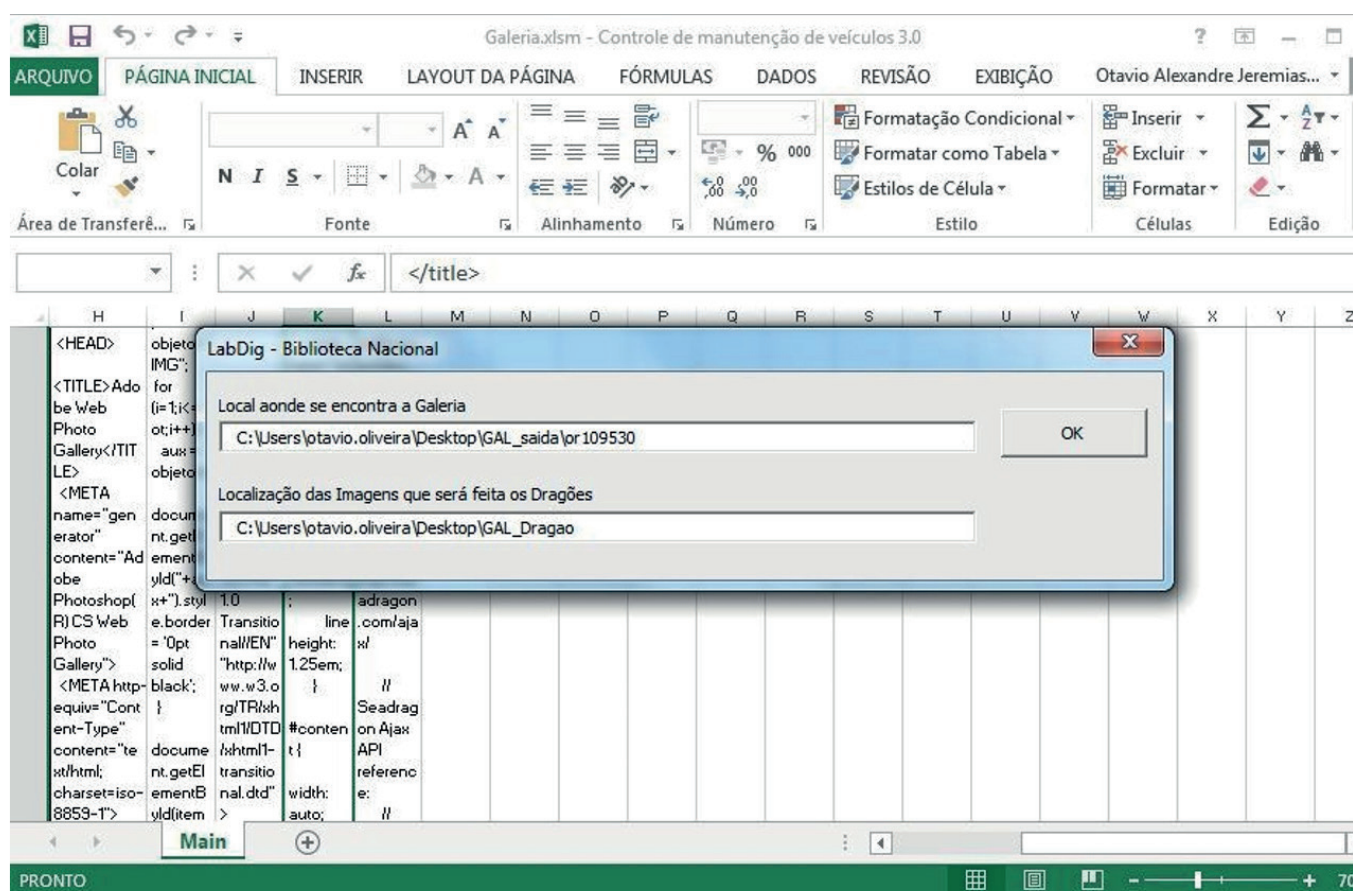
Mover as imagens da pasta “**GAL_reduzida**” para a pasta “**GAL_saída**”, pasta destino do *Photoshop*.

Criar dentro da pasta “**GAL_saída**” a pasta com OD da obra e colocar todos os arquivos dentro.

Aplicar a planilha “**Labdig_2017v0117**” (ou Galeria.xlsm).

Obs.: ao rodar a aplicação não utilizar o computador até o fim da mesma.

Incluir o caminho da pasta onde as imagens da galeria foram criadas, “**GAL_saída/odXXXX/**” e o caminho das imagens a serem feitas o Dragão, pasta “**GAL_Dragão**” e clicar ok.



Editar o arquivo index.htm da galeria (abrir com um editor de texto ou outro *software*).

1 - Inserir título, subtítulo, nome do autor (quando houver) e ano da obra no index na “*div*” abaixo:
Exemplo:

```
<div class="banner">  
<div class="title">Fussball am Megres Strand</div>  
<div class="info">Hesshaimer, Ludwig</div>  
<div class="info">1950</div>  
</div>
```

Para achar o título etc., procurar no catálogo da BNDigital ou BN pelo OD da obra.

2 - Substituir na última linha “*iframe*”, o “XXXXX” pelo OD da primeira imagem, quando a primeira página da galeria NÃO for um Dragão Marinho.

Exemplo:

```
src="pages/icon12345_001.htm"
```

3 - Quando a primeira página da galeria for um Dragão Marinho, substituir na ultima linha "*iframe*", o “*pages/XXXXX*”, pelo OD da primeira imagem, e incluir a letra “*I*” na extensão.

Exemplo:

```
src="icon12345_001.html"
```

Renomear o arquivo index.htm para o OD da obra.

Abrir o arquivo htm renomeado e testar a galeria, usando o navegador *Firefox*.

6.8 - Anexo 1

Dragão Marinho

Procedimento para usar o programa “**Deep Zoom Composer**”.

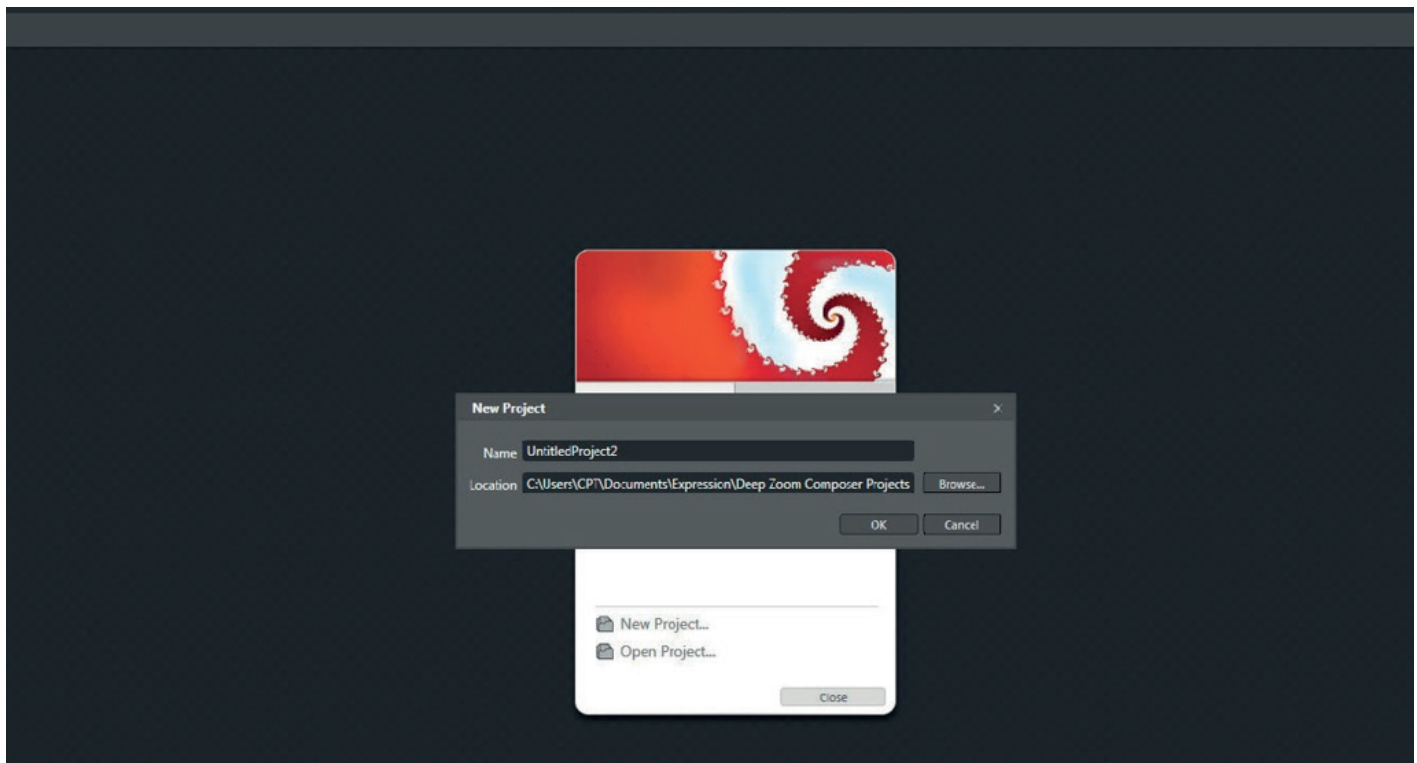
Feito com a imagem original “**master**”, com extensão TIFF ou JPG.

Antes de iniciar o processo fazer uma filtragem, pois não utilizamos os versos das obras que não possuem informação (obs.: anotações, carimbo da BN etc.).

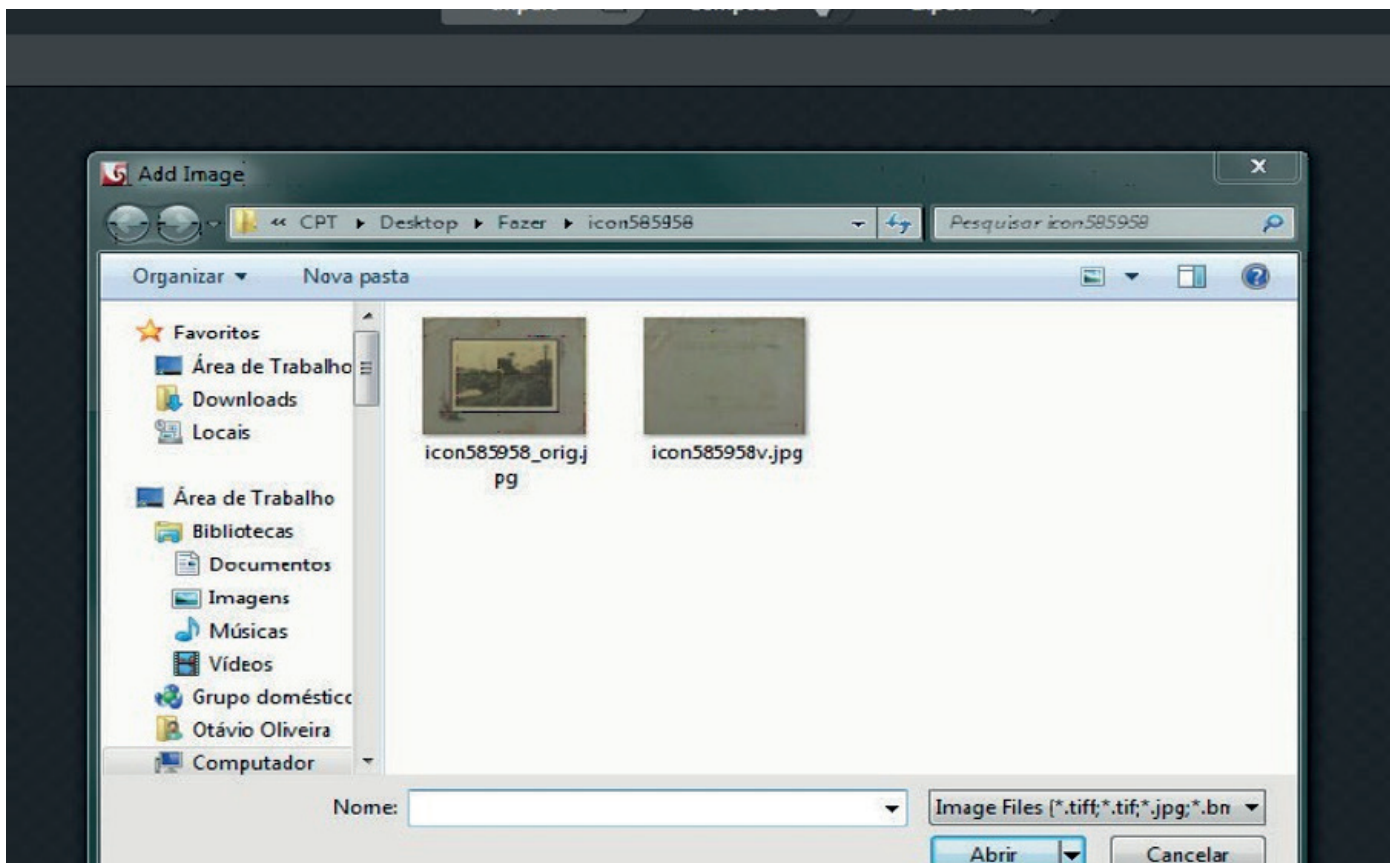
01 - Abrir o programa *Deep Zoom Composer* (Dragão Marinho).

02 - Clicar em “**New Projet...**” (novo projeto).

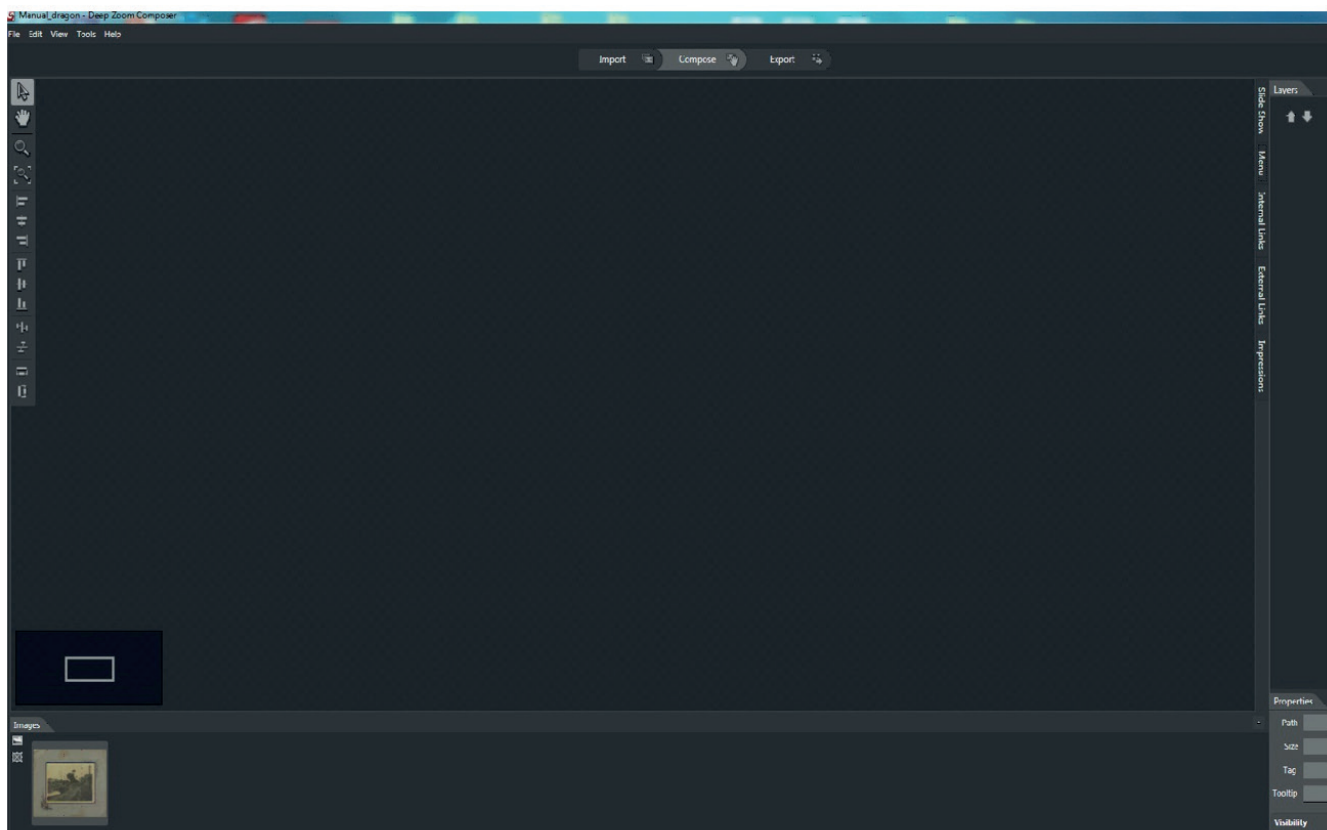
03 - Dar um nome ao projeto.



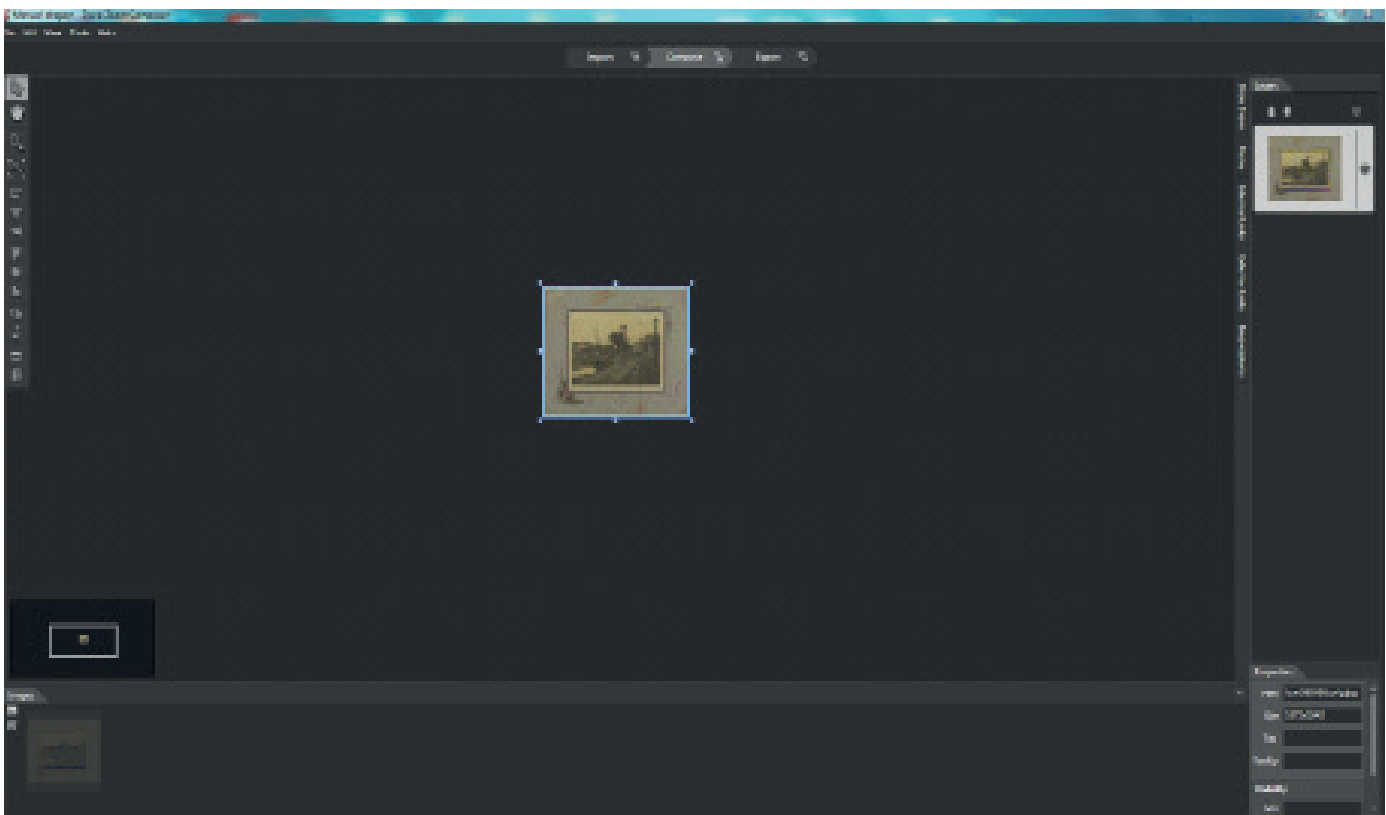
04 - Clicar em “**Add image...**” abrirá uma tela para selecionar o diretório; clicar duas vezes na pasta a ser trabalhada e escolher os arquivos (JPG em alta).



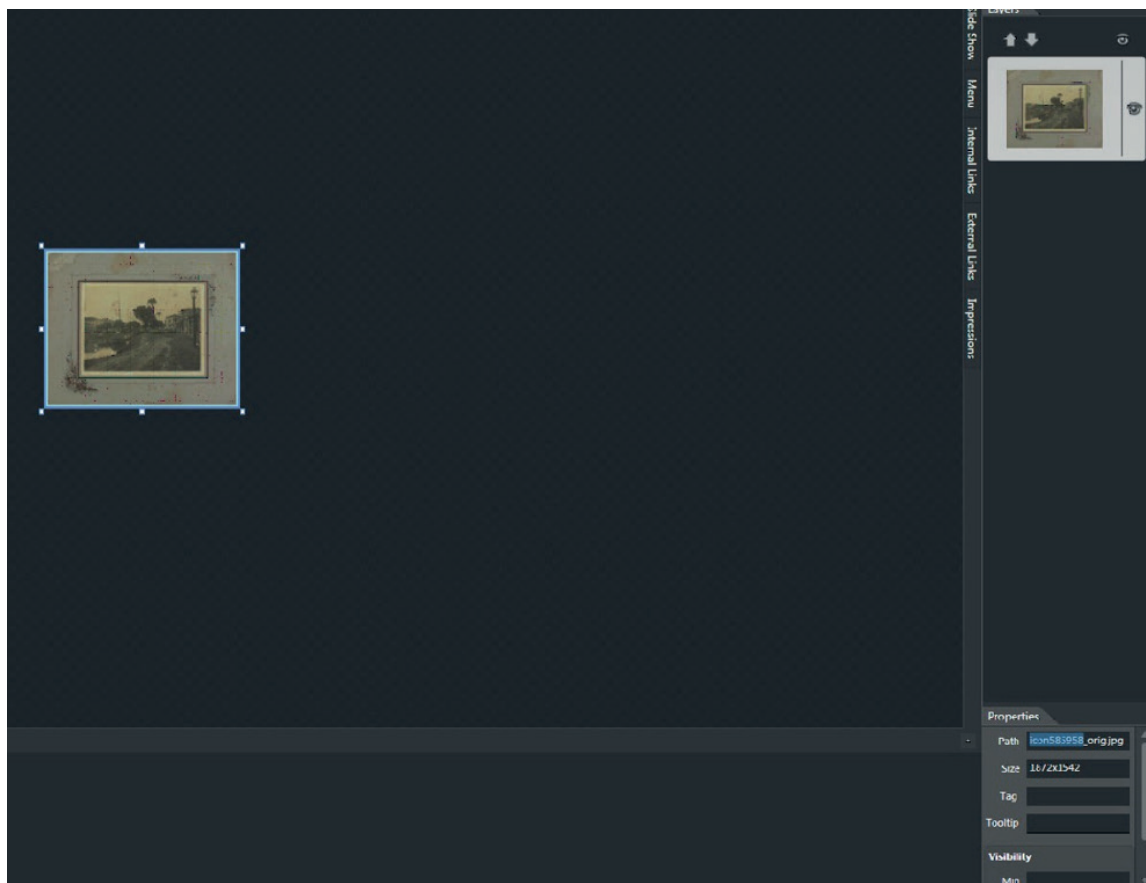
05 - Após a importação, clicar na aba “**Compose**”.



06 - Clique no *thumb* da imagem que está na janela “**Images**” na parte inferior e arraste para o centro da tela (podem ser duas se houver informação no verso).



07 - No canto inferior direito da tela, na aba “**Properties**”, copie do campo “**Path**”, o nome do arquivo sem a extensão, selecione e dê um “**Crtl+c**”.



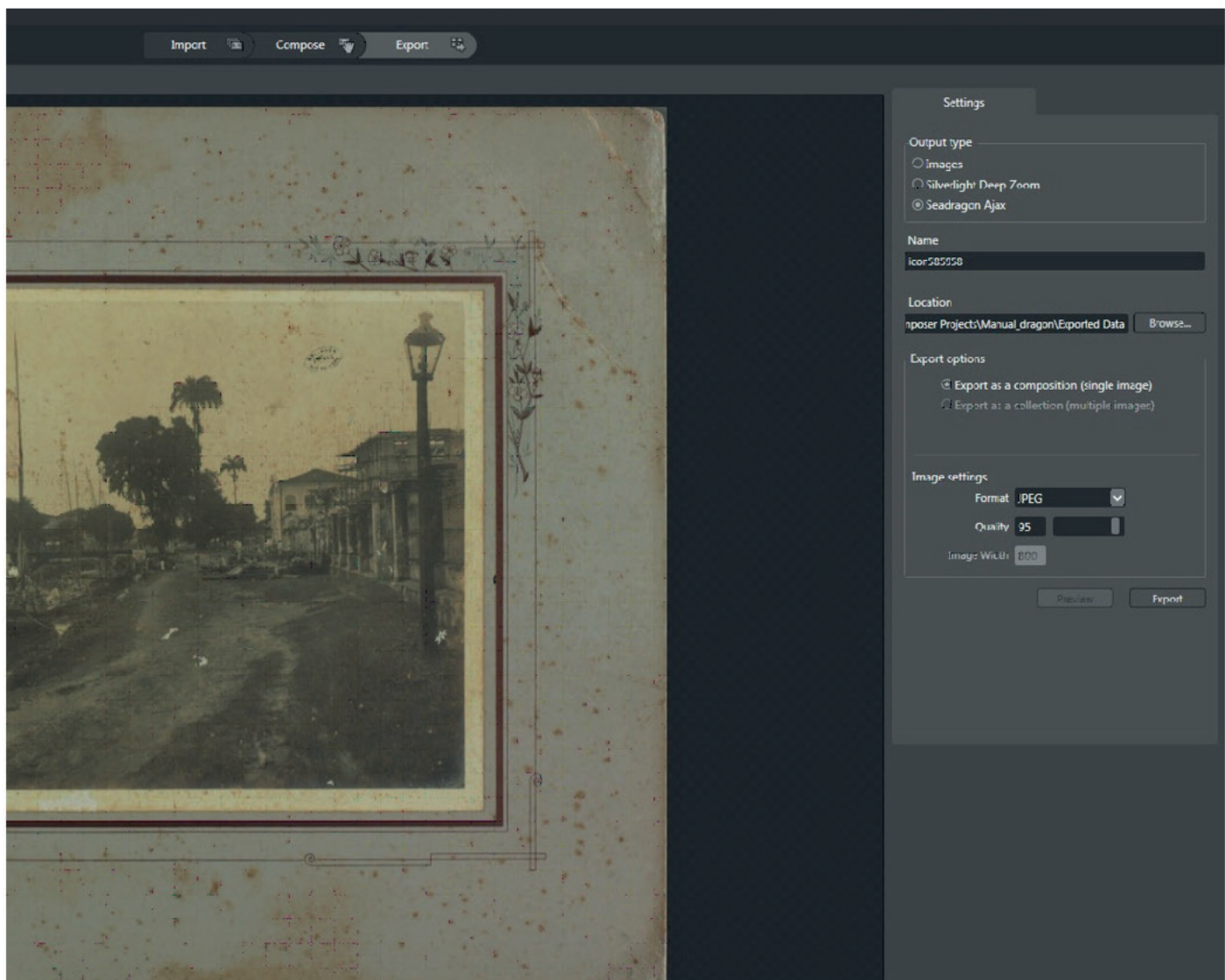
08 - Clique na aba “**Export**”.

09 - Em “**Settings**” selecione a opção “**Seadragon Ajax**”.

10 - Cole o nome do arquivo copiado (“**Crtl+v**”) no campo “**Name**” (tem que ficar apenas o “**OD**” sem a extensão “**jpg**”).

11 - Em “**Location**”, selecionar o diretório de saída.

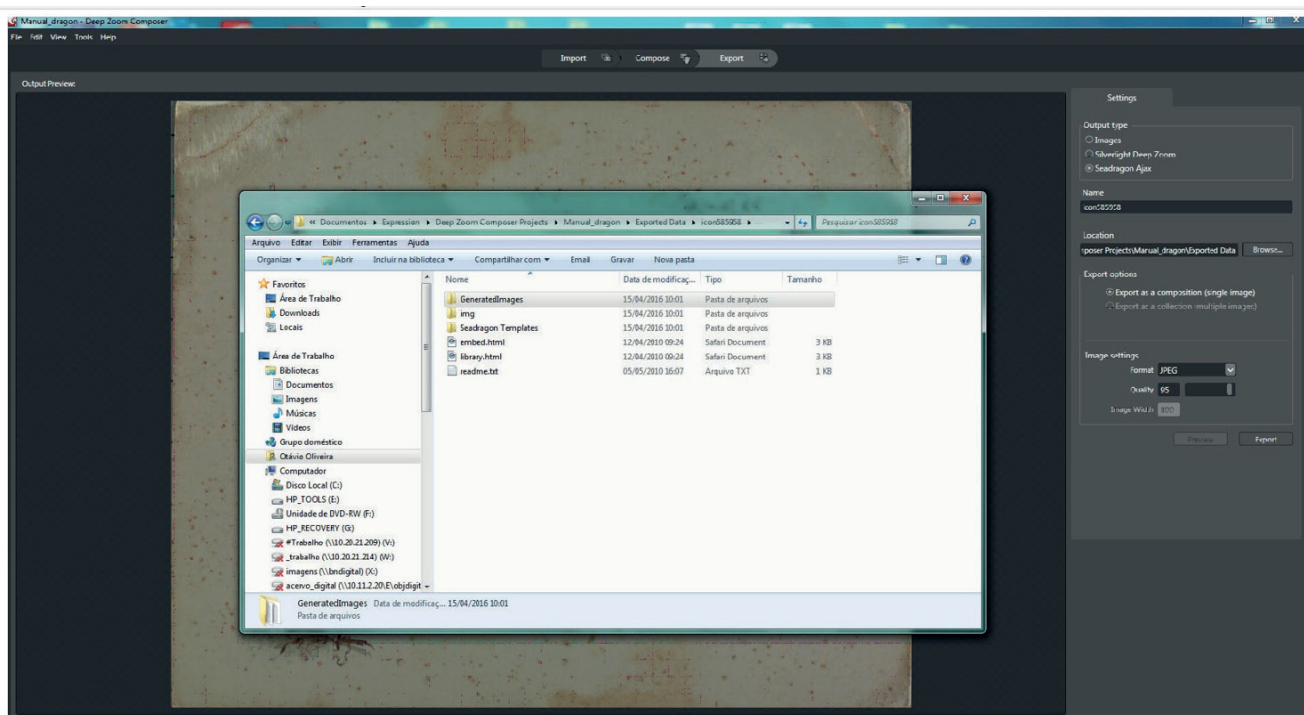
12 - Clicar no botão “**Export**” no lado direito da tela.



13 - Na tela que abrir clicar em **“View Image Folder”**, e volte para pasta anterior à aberta (OD da obra).

14 - Na pasta gerada com o nome da obra, apague a sub pasta **“Seadragon Templates”** e os arquivos **“embed”**, **“library”** e **“readme”**.

Obs.: pode-se apagar os arquivos usando **“apagar.bat”**. Copie este arquivo para dentro da mesma pasta, e clique duas vezes para executar.



15 - Copie o arquivo modelo **“HTML-DM.html”**, para a pasta aberta e renomeie este arquivo html, com o mesmo nome da pasta.

16 - Selecione o arquivo html e clique com o botão direito em abrir com **“bloco de notas”**:

Na linha **“<title>XXX</title>”**, substitua **“XXX”** pelo **“OD”** da obra.

Clique na aba **“arquivo”** e em **“salvar”** e feche o editor.

17 - Abra o arquivo **“html”** com um “navegador” para conferir se está funcionando.

18 - Volte ao programa Dragão Marinho e feche a caixa de diálogo **“Export completed”**.

Para continuar a trabalhar e fazer um novo Dragão Marinho, volte para a aba **“Import”** e comece do item **“04”**.

Para sair, feche o programa e na caixa de diálogo, clique **“Não”**.

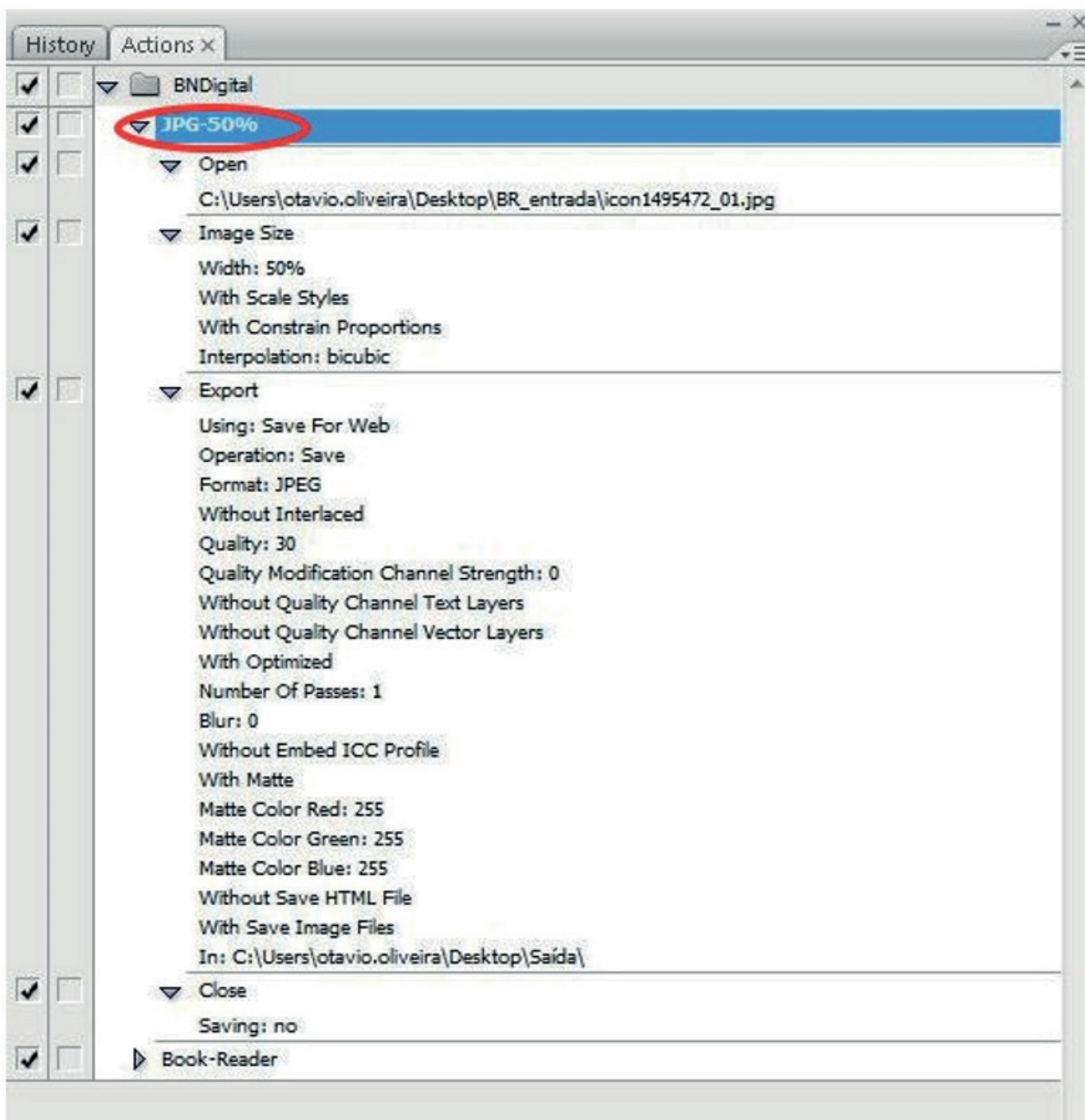
6.9 - Anexo 2

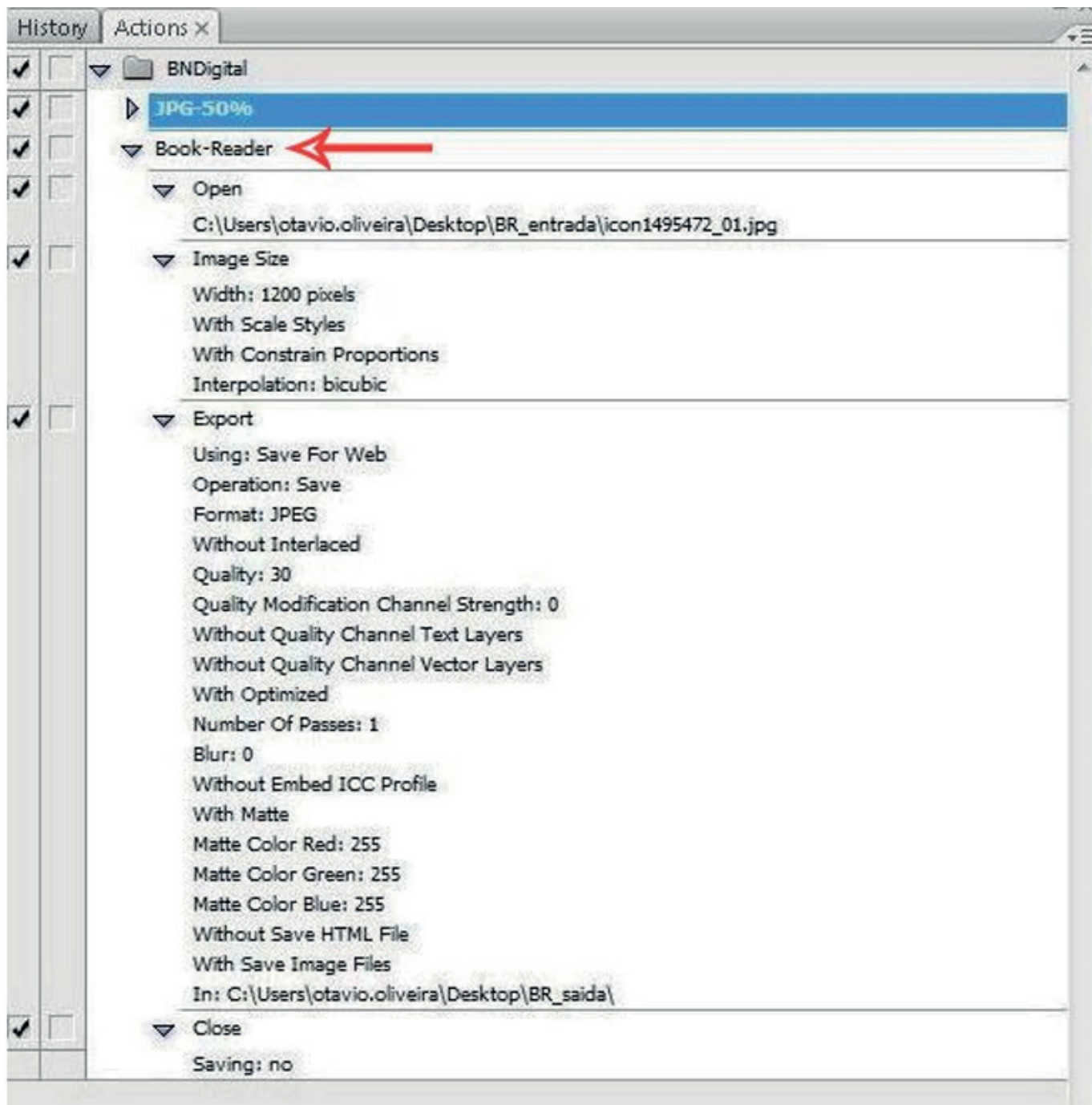
Batch

Estas *batches* são uma maneira de automatizar o processo de trabalho. Elas foram feitas para reduzir imagens em grande quantidade. Todo a programação é feita no *Photoshop CS3*.

As figuras abaixo mostram as ações que serão feitas nas imagens.

JPG 50%

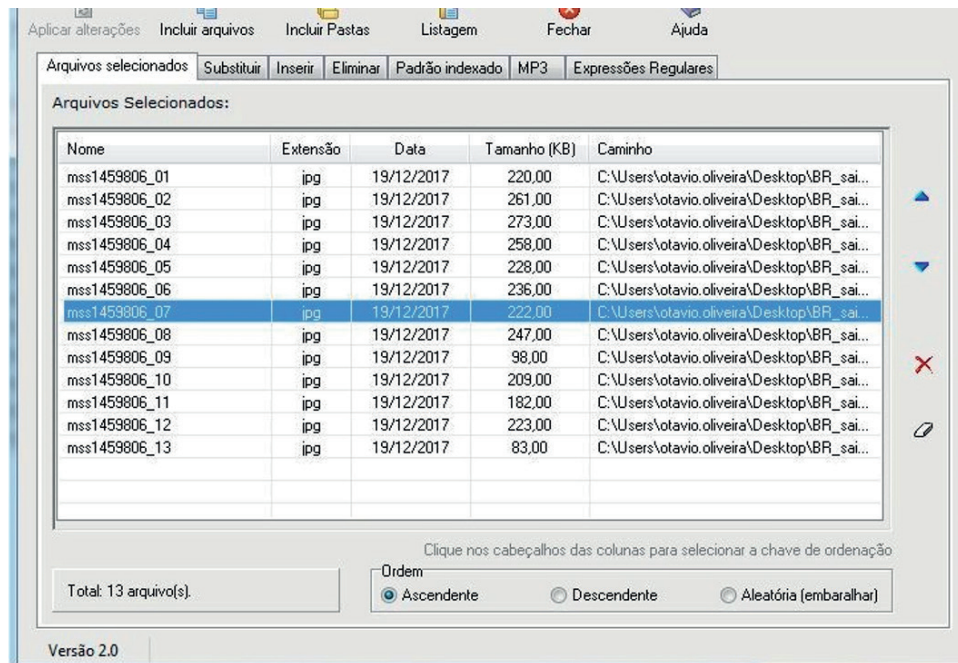


BOOK-READER

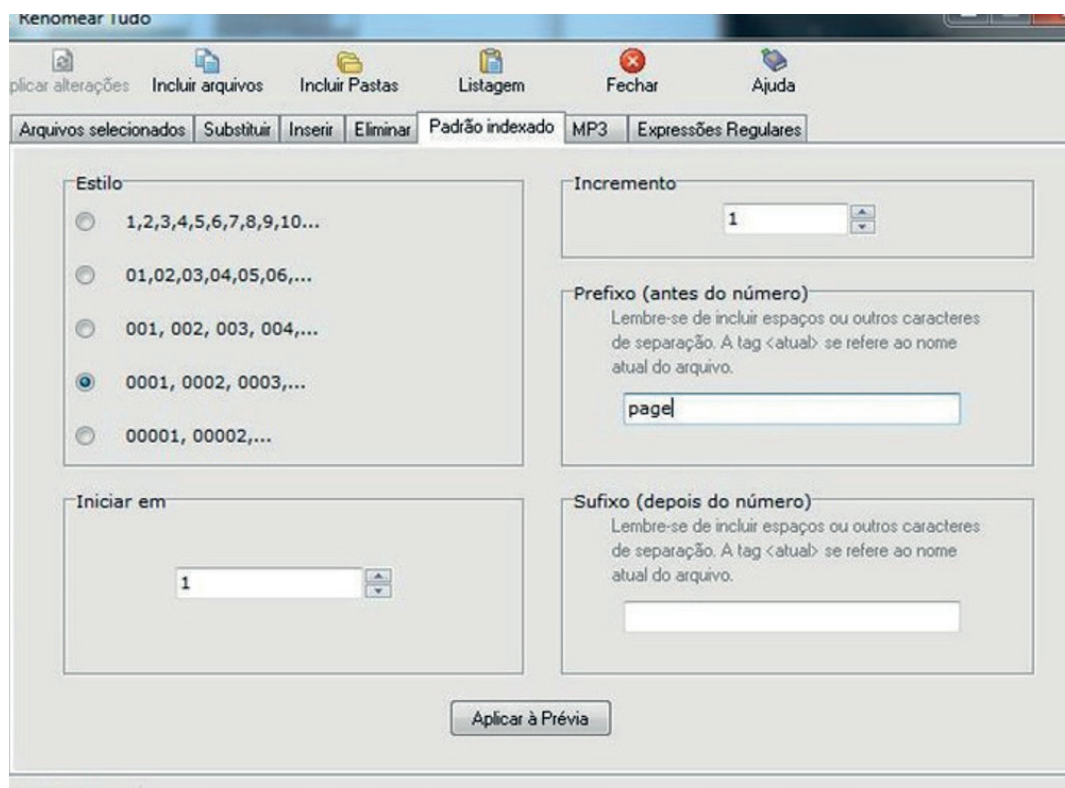
6.10 - Anexo 3

Renomear tudo

Abrir o programa e incluir arquivos que serão renomeados. Pode ser feito de duas maneiras: clicar na tecla “***incluir arquivos***” ou arrastar o arquivo com o *mouse* e soltar na grade do programa “***arquivos selecionados***” do programa.

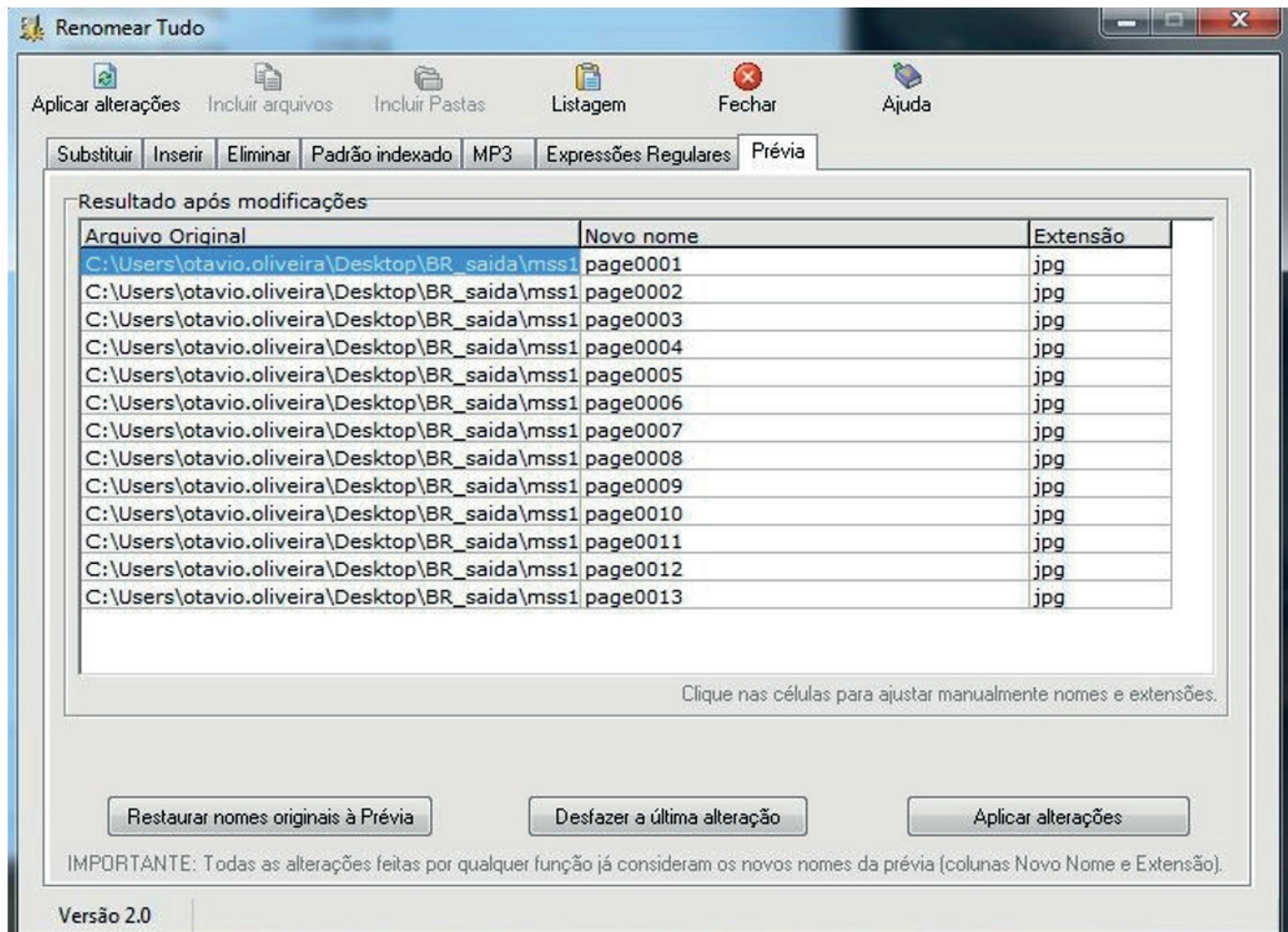


Selecionar a guia “***Padrão indexado***” e marcar os itens conforme tela abaixo:



Em “**Estilo**”, escolher a opção de quatro algoritmos, indicar “**01**” no início e incluir “**page**” no campo prefixo. Clicar em “**aplicar à prévia**”.

Selecionar a guia “**Prévia**” e clicar em “**aplicar alterações**”, finalizando o processo.



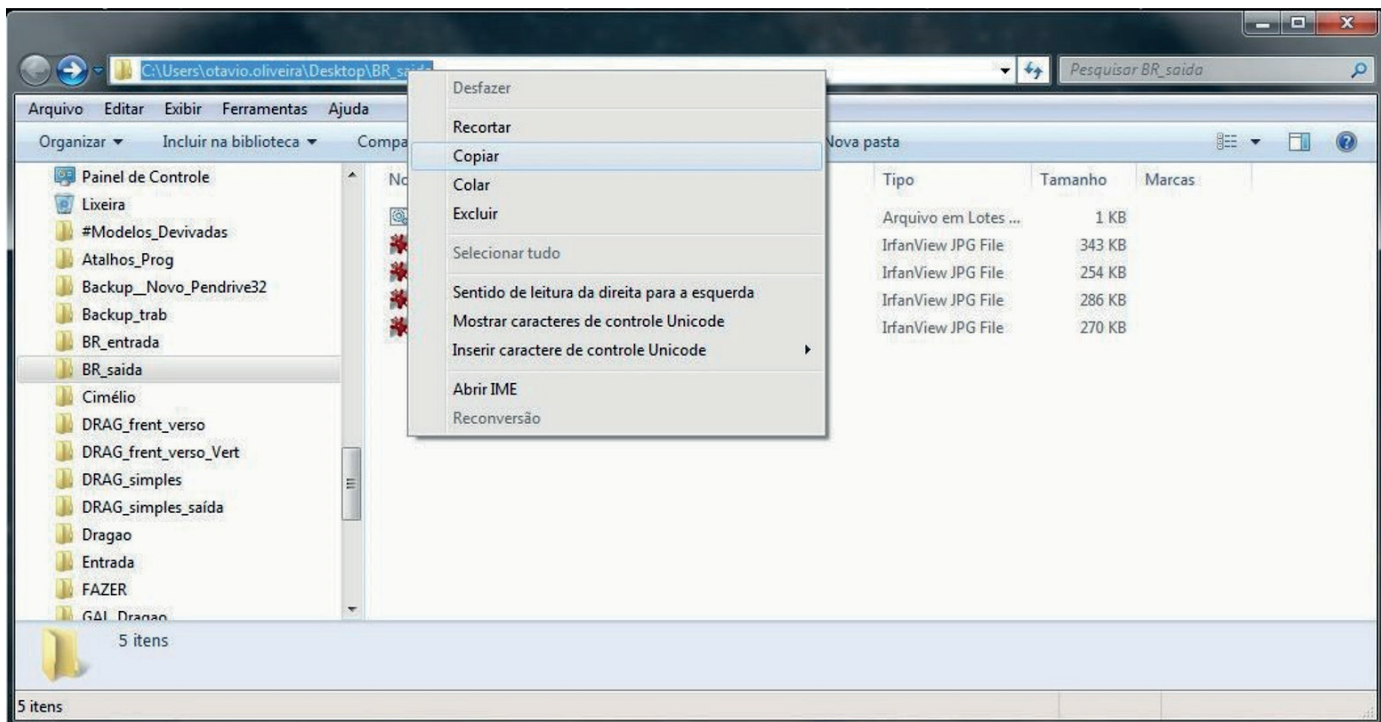
6.11 - Anexo 4

Fluxo_Book_Reader

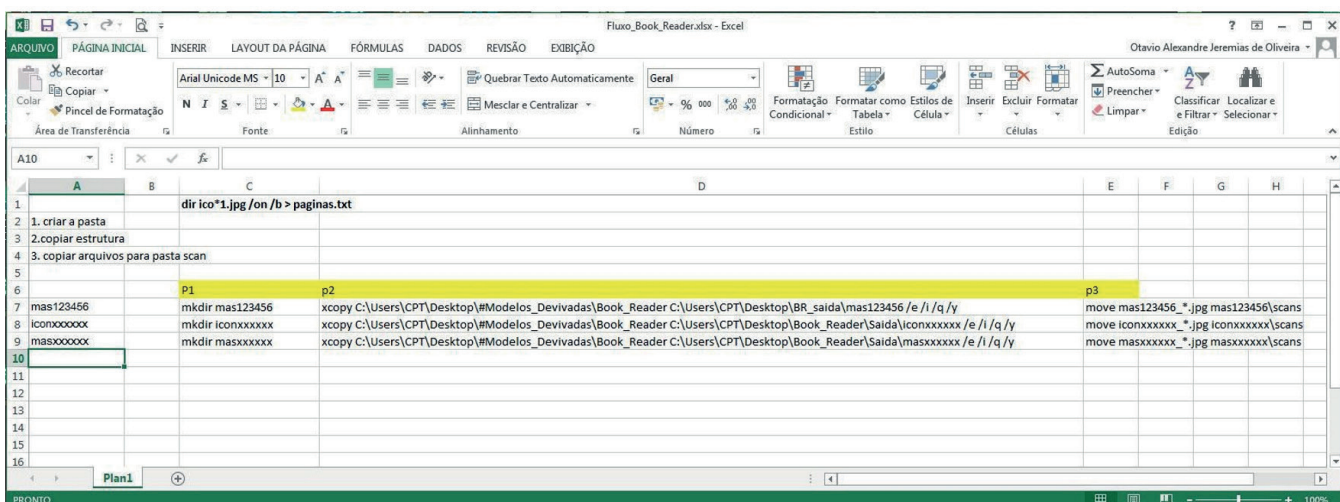
Abrir a planilha e copiar o OD da obra para a coluna A da planilha.

Abrir o “DOS” e voltar para a raiz da pasta com o comando “**cd**”.

Ir para a pasta “**BR_saida**” no *desktop* e copiar (botão direito do *mouse*) o caminho desta pasta.



No “DOS” usar o comando “**cd**”, colar (botão direito do *mouse*) o caminho da pasta “**BR_saida**” e depois dar o “**Enter**”. Pronto, você está dentro da pasta onde irá incluir os comandos P1, P2 e P3 da planilha.



Obs.: o caminho da pasta “**BR_saida**” varia de acordo com sua localização no PC; ficar atento para a mudança na planilha.

7 - CONFERÊNCIA E GUARDA DOS ARQUIVOS DIGITALIZADOS DE PERIÓDICOS

O processo de conferência dos arquivos de imagens produzidos pelos *scanners* Zeutschel difere-se em relação ao dos produzidos pelos outros equipamentos de digitalização de originais do laboratório, e isso se deve principalmente a alta produtividade desses *scanners*. Enquanto no outro processo todos os arquivos são conferidos e comparados com os documentos originais, nos arquivos de imagens produzidos pelos *scanners* Zeutschel a conferência é feita por amostragem.

Deve-se levar em consideração que cada técnico que opera o *scanner* Zeutschel faz uma conferência minuciosa de sua própria produção, então a conferência feita pela equipe de gestão é a segunda feita nesses arquivos de imagens.

A outra diferença na rotina de conferência e guarda dos arquivos produzidos pelos *scanners* Zeutschel são os critérios de criação, codificação e organização das pastas para *backup* dos arquivos de imagens, e isso se deve à diferença que a codificação dos documentos periódicos tem em relação a documentação de todos os outros setores (manuscritos, obras raras, iconografia, cartografia, música e obras gerais). Na codificação dos outros setores, o documento recebe um número patrimonial, que é individual por item documental, enquanto que na documentação de periódicos, a codificação é feita pelo número de BIN do título do periódico, que pode abranger dezenas de itens documentais.

7.1 - Criação e organização de pastas para *backup*

O controle da criação de pastas para organização e guarda dos arquivos digitalizados no fluxo de trabalho do laboratório de digitalização é fundamental, tendo em vista a alta produção de arquivos de imagens digitalizadas pelos *scanners* Zeutschel, e a padronização nessa etapa do trabalho garante não haver perda de material digitalizado. A criação dessas pastas é de responsabilidade dos técnicos da equipe de gestão e dos técnicos de digitalização dos *scanners* Zeutschel, em níveis diferentes.

As pastas criadas pelos técnicos da equipe de gestão são para organização e guarda da produção de todos os arquivos de imagens gerados pela equipe Zeutschel, enquanto a criação de pastas feita por cada técnico de digitalização é para organização e guarda da sua própria produção.

As pastas criadas pelos técnicos da equipe de gestão ficam em HD's nos computadores pré-selecionados na área de gestão. Esses HD's são utilizados como um dos *backups* temporários da produção dos *scanners* Zeutschel, antes dos arquivos serem enviados em definitivo para o *Storage*. O outro *backup* é feito no HD do computador do técnico que produziu as imagens e lá será mantido até que a equipe de gestão informe que ele pode ser deletado. Isso só ocorrerá quando o *backup* definitivo no *Storage* for concluído.

As pastas criadas pelos técnicos da equipe de gestão para *backup* da produção dos *scanners* Zeutschel ficam em dois computadores diferentes na área de gestão. Um computador será usado para *backup* dos Arquivos *Master* e outro computador para *backup* dos Arquivos Derivados. Em ambos computadores serão criadas pastas específicas para esse *backup*.

As pastas serão nomeadas da seguinte maneira:

Computador 01 -  **producao_zeutschel_masters**
Computador 02 -  **producao_zeutschel_derivados**

São criados atalhos dessas duas pastas na área de trabalho de cada técnico que opera os *scanners* Zeutschel, isso para facilitar o acesso a estas pastas.

Dentro de cada uma dessas pastas citadas acima (**producao_zeutschel_masters** e **producao_zeutschel_derivados**), serão criadas pastas para a produção do mês corrente. A criação dessas pastas também fica sob a responsabilidade dos técnicos da equipe de gestão.

Exemplo:  **producao_zeutschel_2020_janeiro**
 **producao_zeutschel_2020_fevereiro**

A partir desse nível as próximas pastas são criadas pelos técnicos de digitalização. Dentro das pastas de produção do mês, cada técnico criará uma pasta com seu nome, onde fará o *backup* de toda sua produção de arquivos de imagens do mês.

Exemplo:  **nome do técnico**

A próxima pasta é referente aos títulos de periódicos que são digitalizados, cada título deve ter uma pasta para concentrar dentro dela todos as brochuras com os fascículos daquele título. Essa pasta deve ser nomeada com o número de BIN daquele título.

Exemplo:  **per004120**

Dentro da pasta com número de BIN são criadas as pastas correspondentes as brochuras que foram digitalizadas, e ela será codificada com o número sequencial daquela brochura naquele mês + o número de patrimônio da brochura.

Exemplo:  **01_per_371_002_2**

Por fim, a última pasta a ser criada é a que corresponde a cada fascículo de periódico que está encadernado na brochura. Vale lembrar que a maioria dos periódicos da BN são encadernados em brochuras e dentro dessas brochuras existe uma quantidade de fascículos variável. Não há um padrão. Deste modo, o número de pastas criadas corresponderá ao número de fascículos existentes em cada brochura.

Essa pasta será codificada com o número do BIN + o ano de lançamento do fascículo + o número do fascículo.

Exemplo:  **per004120_1986_1766**

Obs.: embora os critérios para codificação de pastas para periódicos tenham sido descritos anteriormente, os mesmos serão abordados com profundidade no *Manual de procedimentos técnicos de digitalização do acervo (scanner Zeutschel)*, capítulo 9, que trata especificamente das instruções de codificação e nomeação de pastas para periódicos.

7.2 - Conferência da codificação de pastas e arquivos

7.2.1 - Conferência da codificação de pastas

A primeira etapa da conferência é a da codificação de pastas. Qualquer erro não identificado pode ocasionar a completa dissociação dos arquivos de imagens com o documento original o qual representa.

Inicialmente, vale destacar que para esta etapa de conferência, as pastas correspondentes às brochuras não são muito relevantes. Na verdade, sua criação tem por objetivo associar os arquivos digitais à brochura da qual eles foram digitalizados, bem como dar suporte à confecção do relatório de produção de cada técnico. Finalizados esses processos, essa pasta é descartada, sendo mantidas somente as pastas com os fascículos de periódico, com o número do BIN + o ano de lançamento do fascículo + o número do fascículo. São essas que serão enviadas para o *Storage*.

A conferência dessas pastas segue as seguintes etapas:

- Verificar, junto ao FEO (Formulário de Encaminhamento de Obras) o **número de BIN** do título de periódico e se o mesmo foi empregado corretamente na nomeação da pasta;
- Verificar se o **ano de publicação** foi corretamente transcrito na nomeação da pasta;
- Verificar se o **número do fascículo** foi corretamente transcrito na nomeação da pasta.

As informações relativas ao ano da publicação e número do fascículo, presentes na codificação das pastas, devem obrigatoriamente ser conferidas com a imagem digitalizada da capa do periódico. Ocorrem situações em que essas informações não estão presentes na capa. Nestes casos, recorre-se a páginas no interior do fascículo.

É importante lembrar que essa conferência, e todas as outras, devem ser feitas nas pastas criadas para *backups* dos Arquivos *Master* e dos Arquivos Derivados. As pastas, nesses dois *backups*, devem ter exatamente os mesmos nomes.

7.2.2 - Conferência da codificação de arquivos

A nomeação dos arquivos de imagens dentro das pastas é simples: em cada pasta os arquivos são nomeados sequencialmente de 1 ao infinito. É só comparar a quantidade de arquivos de imagens que tem dentro da pasta com a codificação.

7.3 - Conferência dos quantitativos de imagens nas pastas

Durante essa fase é necessário ter as pastas de Arquivos Derivados e Arquivos *Master* abertas simultaneamente. Inicialmente, verificamos se a quantidade de arquivos de imagens confere com a quantidade de páginas do documento original digitalizado. Essa conferência deve ser feita com os Arquivos Derivados JPEG.

Posteriormente, verificamos se as quantidades de Arquivos Derivados e Arquivos *Master* são semelhantes, pois os Arquivos *Master* são produzidos com uma escala de cor e o documento inteiro aberto, portanto, duas páginas em um único arquivo de imagem. Em contrapartida, os Arquivos Derivados, para cada página é produzido um arquivo de imagem. Neste sentido, confeccionou-se uma espécie de “fórmula” para fornecer suporte a essa conferência. Levando-se em consideração que no Arquivo *Master* o documento é digitalizado por inteiro, aberto, cada arquivo *master* terá duas páginas de documento. A exceção disso é a capa e a folha final, que serão apenas uma página por arquivo. Abaixo, segue exemplo:

- Uma pasta de Arquivos *Master* tem 51 arquivos de imagens. Porém, 2 desses arquivos de imagens (a capa e a folha final) só tem 1 página de documento digitalizada. Sobram os outros 49 arquivos de imagens, que terão 2 páginas cada um. Multiplicamos então $49 \times 2 = 98$, e somamos + 2, um total de 100. A pasta dos Arquivos Derivados terá 100 arquivos de imagem.

Então, de acordo com o exemplo a “fórmula” seria:

quantidade de arquivos derivados $\div$ 2 + 1 = quantidade de arquivos *master*

Cabe ressaltar que na maioria das vezes essa fórmula funciona, mas existem exceções. Uma delas ocorre quando os arquivos de imagens contêm apenas uma página por arquivo, por exemplo no caso de páginas grandes.

7.4 - Conferência da qualidade de arquivos de imagens

A primeira conferência a ser feita na qualidade das imagens é verificar se todas as imagens derivadas passaram pelo processo de corte, para retirada da margem preta. A forma mais fácil de fazê-la é abrir a pasta com os arquivos de imagens e colocar a visualização do computador em ícones grandes (no Windows isso é feito com o comando EXIBIR - ÍCONES GRANDES). Dessa forma é fácil verificar se algum arquivo não teve as margens pretas cortadas. Ainda no que diz respeito ao corte, é necessário verificar também se as imagens sofreram cortes excessivos suprimindo informação. Isto é feito por amostragem.

Em seguida, verificamos se a cor e se o foco dos arquivos de imagens estão de acordo com os padrões estabelecidos. Vale ressaltar que essa verificação é feita somente nos Arquivos Derivados e através do método de amostragem. Se esses padrões estiverem corretos, consequentemente os Arquivos *Master* também estarão.

Por fim, verifica-se a largura da borda preta: se está sendo respeitada a margem de 1 a 3 cm e se a escala de cor está posicionada corretamente na imagem. Neste caso a verificação é feita somente nos Arquivos *Master* e através do método de amostragem. Durante todo fluxo de trabalho, incluindo a fase de conferência, o documento original fica guardado na gaveta do técnico responsável por sua digitalização. Após finalizada essa fase, deve-se informar ao técnico que o documento original está liberado para devolução ao setor de Periódicos.

7.5 - Organização final para guarda no *Storage*

Finalizada a etapa de conferência, inicia-se a organização final das pastas para a guarda no *Storage*. O técnico do laboratório responsável pela conferência também é responsável por reorganizar as pastas em uma nova estrutura para posterior *backup* final dentro do *Storage*.

Deste modo, no interior das pastas **producao_zeutschel_masters** e **producao_zeutschel_derivados** é criada uma nova pasta para o *backup* da produção do mês corrente:

Exemplo:  **producao_zeutschel_2020_janeiro_Storage**

Para essa pasta será transferida toda a produção de imagens do mês. As pastas com o nome do técnico digitalizador e as com o número de patrimônio da brochura serão deletadas nesse momento.

Além disso, é importante acrescentar que essa transferência dos Arquivos *Master* e Arquivos Derivados ocorrem em momentos distintos. Neste sentido, a transferência dos Arquivos *Master* é feita logo após a finalização dessa etapa de conferência. Em contrapartida, nos Arquivos Derivados, a transferência é realizada somente quando a chefia imediata finaliza a verificação dos relatórios de produtividade de cada técnico.

Estrutura das pastas de becape temporário (5 níveis)	Estrutura das pastas de becape final no <i>Storage</i> (3 níveis)
 producao_zeutschel_2020_janeiro  nome do colaborador  per004120  01_per_371_002_2  per004120_1986_1766	 producao_zeutschel_2020_janeiro_Storage  per004120  per004120_1986_1766

Quando os arquivos forem transferidos para o *Storage*, é importante fazer um levantamento de todos os fascículos de cada título, para verificar se existem fascículos de periódicos que estejam faltando. Caso existam, deve-se verificar o motivo.

Possíveis motivos e suas soluções:

1 - Os fascículos vieram para o laboratório para serem digitalizados, mas por algum motivo não foram. Mapear onde estão e proceder com sua digitalização.

2 - Os fascículos não vieram para o laboratório, mas existem no setor de Periódicos. Solicitar ao setor de Periódicos o encaminhamento dos mesmos para proceder com sua digitalização.

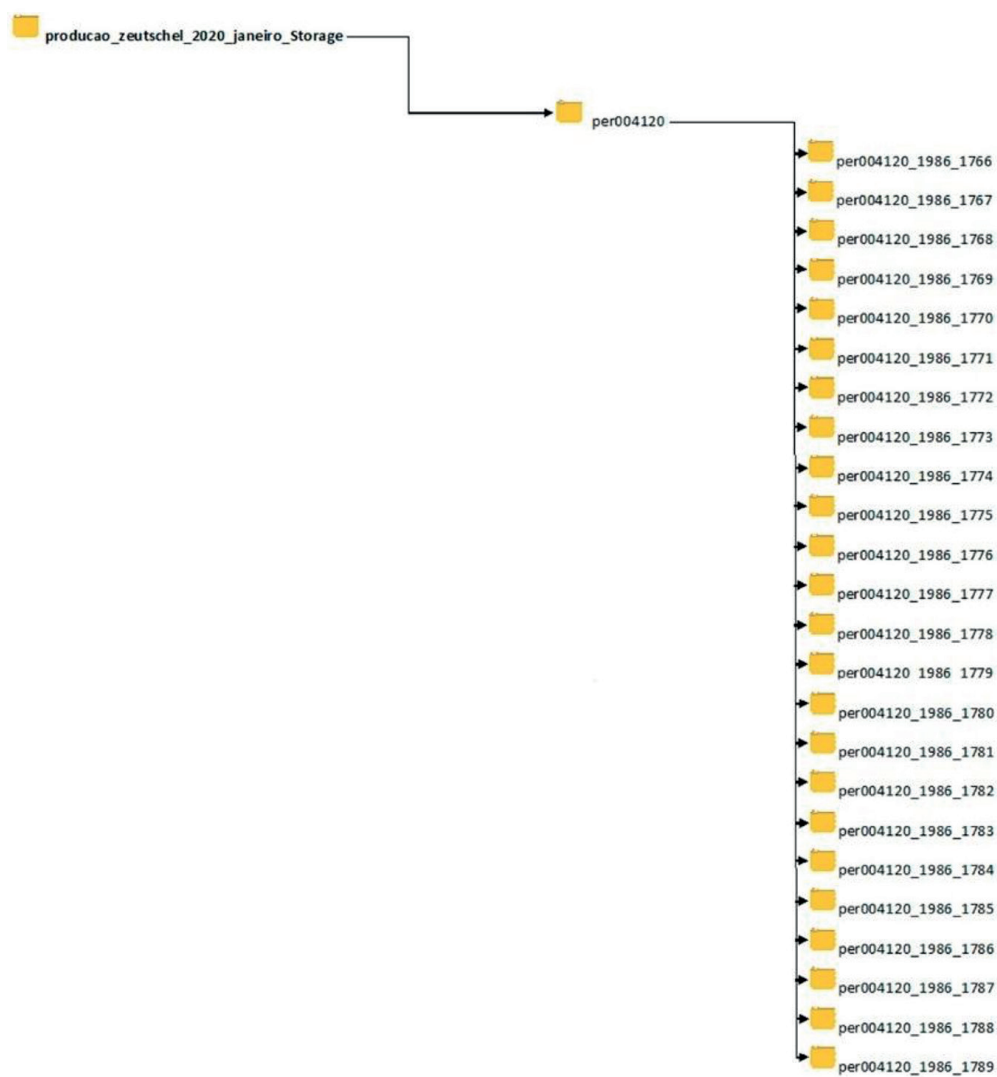
3 - Os fascículos não existem no setor de Periódicos. Não serão digitalizados.

Obs.: quando não existem alguns fascículos de um determinado periódico na coleção da Biblioteca Nacional, a equipe da BN Digital tenta, junta a instituições parceiras, solicitar por empréstimo esses documentos para que eles possam ser digitalizados no Laboratório de Digitalização, e então sejam disponibilizados na Hemeroteca Digital.

Seguem, abaixo, dois esquemas detalhados de como ficam a organização das pastas no *backup* temporário e no *backup* final no *Storage*:



Esquema 2- Organização das pastas de backup final no Storage



8 - ARMAZENAMENTO

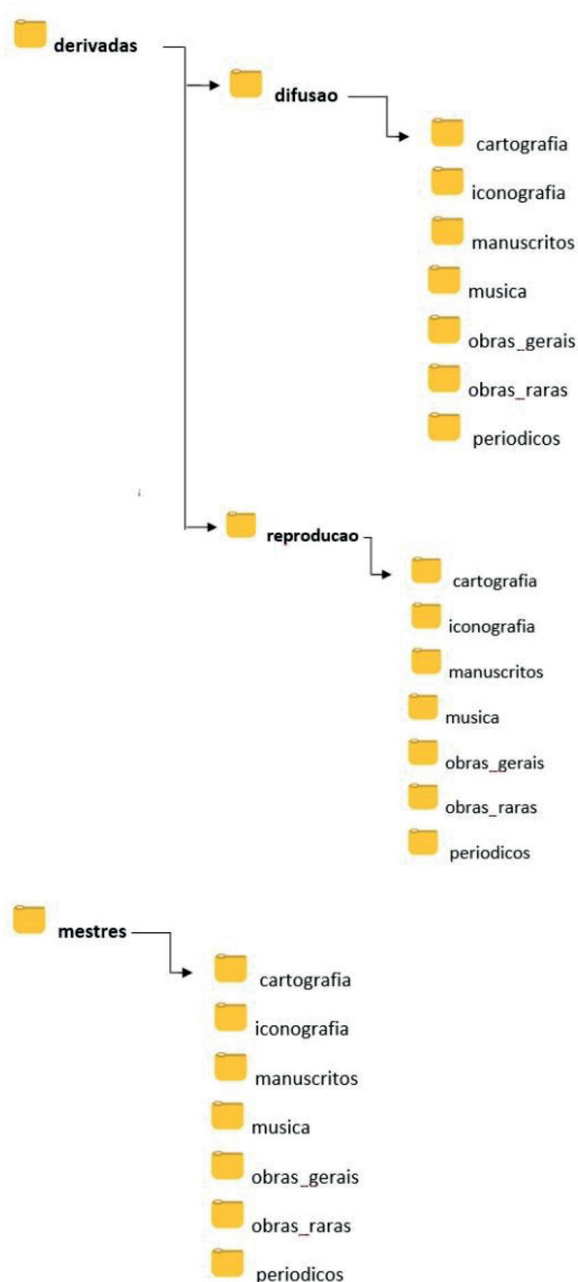
8.1 - Armazenamento no servidor dentro do Laboratório

Os arquivos digitais dentro de suas respectivas pastas são enviados para um servidor, em um espaço reservado para arquivo *master* (TIFF), para serem transferidos em lotes para o *Storage*, assegurando a integridade do material digitalizado.

8.2 - Armazenamento no *Storage*

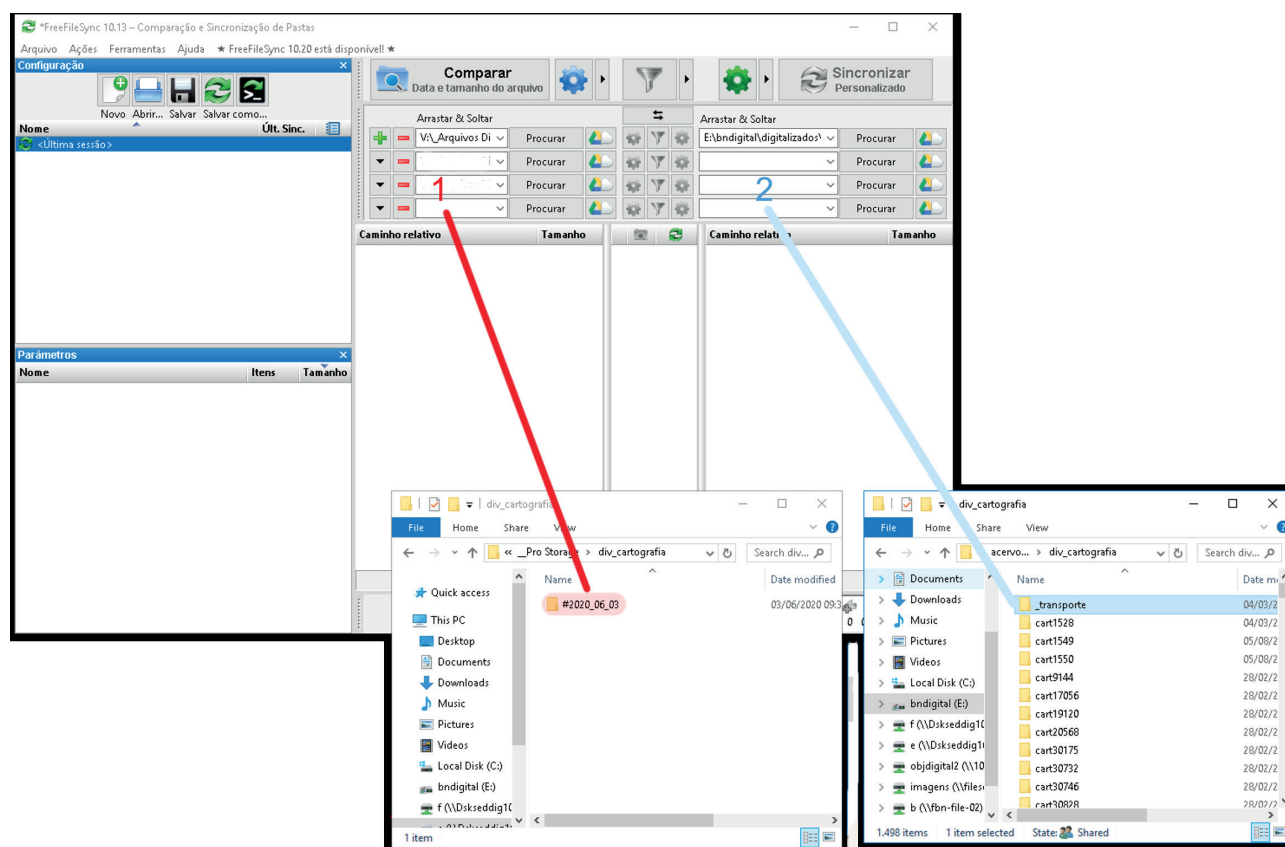
Os arquivos digitais são armazenados em *Storage* altamente protegido, com acessos restritos e exclusivos para os arquivos *master* e perfis de acesso para as derivadas digitais. As pastas são divididas de acordo com os tipos de arquivos digitais gerados e são organizadas conforme as Divisões Estruturais da Biblioteca Nacional, através da divisão dos setores, tais como manuscritos, iconografia, cartografia, música, obras gerais, obras raras e periódicos.

O *Storage* é dividido da seguinte forma:



O espaço designado para os arquivos mestres tem acesso restrito; somente usuários exclusivos podem inserir ou fazer alterações. Já os arquivos derivados – reprodução ou difusão possuem um acesso maior para usuários, porém sem autorização para qualquer tipo de alteração. Assim é assegurado a integridade do material de todo o tipo de material digitalizado.

A transferência dos arquivos digitais é feita através de um *software* de comparação e sincronização de pastas que gerencia cópias de *backup*, além de determinar as diferenças entre uma pasta de origem e uma de destino e transfere apenas a quantidade mínima de dados necessária. Para agilizar o tempo de transmissão e comparação de arquivos foi criada uma pasta (*_transporte*) dentro de cada subpasta, evitando que o programa precise todo conteúdo em cada nova transferência. Ainda assim, ao final de toda transferência, as pastas de origem são comparadas com as respectivas pastas destino garantindo a fidelidade da cópia.



Legenda:

- 1) Local para direcionar a pasta de origem, sinalizando os arquivos que ficam armazenados dentro de um servidor no Laboratório de Digitalização e são transferidos em lotes.
- 2) Local utilizado para direcionar a pasta de destino, que vai receber os arquivos para serem armazenados no *Storage*. Utilizamos a pasta (*_transporte*) para agilizar e garantir o recebimento de novos arquivos, reduzindo o tempo de varredura para sincronização de novos dados e assegurar qualquer alteração nas transferências anteriores.

