



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO
CURSO DE BIBLIOTECONOMIA

Wagner Silva Wessfl

**DIGITALIZAÇÃO DE OBRAS RARAS:
ESTUDO DE CRITÉRIOS DE SELEÇÃO**

Porto Alegre

2015

Wagner Silva Wessfl

**DIGITALIZAÇÃO DE OBRAS RARAS:
ESTUDO DE CRITÉRIOS DE SELEÇÃO**

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção do título de Bacharel em Biblioteconomia, pelo Departamento de Ciências da Informação, da Faculdade de Biblioteconomia e Comunicação, da Universidade Federal do Rio Grande do Sul.

Orientador: Prof. Dr. Rafael Port da Rocha.

Porto Alegre
2015

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Carlos Alexandre Netto

Vice-reitor: Prof. Dr. Rui Vicente Oppermann

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretora: Prof. Dra. Ana Maria Mielniczuk de Moura

Vice-diretor: Prof. Dr. André Iribure Rodrigues

DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Chefe: Prof. Dr. Moisés Rockembach

Chefe substituto: Prof. Dr. Valdir Jose Morigi

COMISSÃO DE GRADUAÇÃO DO CURSO DE BIBLIOTECONOMIA

Coordenador: Prof. Dr. Rodrigo Silva Caxias de Sousa

Coordenador substituto: Prof. Dr. Jackson da Silva Medeiros

CIP - Catalogação na Publicação

Wessfl1, Wagner Silva
Digitalização de obras raras: estudo de critérios
de seleção / Wagner Silva Wessfl1. -- 2015.
77 f.

Orientador: Rafael Port da Rocha.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Faculdade
de Biblioteconomia e Comunicação, Curso de
Biblioteconomia, Porto Alegre, BR-RS, 2015.

1. Digitalização. 2. Obras raras. 3. Critérios de
seleção. 4. Preservação digital. I. Rocha, Rafael Port
da, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Faculdade de Biblioteconomia e Comunicação (FABICO)

Rua Ramiro Barcelos, 2705 – Bairro Santana

CEP 90035-007 – Porto Alegre – RS

Fone: (51) 3308-5067 | Fax: (51) 3308-5435

Site: www.ufrgs.br/fabico | E-mail: fabico@ufrgs.br

Wagner Silva Wessfl

**DIGITALIZAÇÃO DE OBRAS RARAS:
ESTUDO DE CRITÉRIOS DE SELEÇÃO**

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção do título de Bacharel em Biblioteconomia, pelo Departamento de Ciências da Informação, da Faculdade de Biblioteconomia e Comunicação, da Universidade Federal do Rio Grande do Sul.

Aprovado em 09 de dezembro de 2015.

BANCA EXAMINADORA

Prof. Dr. Rafael Port da Rocha
Universidade Federal do Rio Grande do Sul (UFRGS)
Orientador

Prof^a. Dr^a. Sonia Elisa Caregnato
Universidade Federal do Rio Grande do Sul (UFRGS)
Examinadora

Bibl. Dr^a. Caterina Marta Groposo Pavão
Centro de Processamento de Dados (UFRGS)
Examinadora

Dedico à minha família, pelo incessante incentivo ao estudo.

AGRADECIMENTOS

Em primeiro lugar, agradeço à minha família pelo apoio e incentivo em todas as horas.

Aos meus pais, Nádia e Carlos, desde cedo semearam em mim a importância do estudo, sempre com carinho e amor para dar.

À minha irmã, Cyntia Wessfl, que é uma segunda mãe para mim, que sempre me mostrou o caminho certo e se certificou que a minha caminhada fosse a menos pedregosa possível.

Ao meu namorado, Felipe Nogueira (mais conhecido como Felipe Nogs), que sempre foi o meu fã número um, nesta curta e incrível jornada que nós embarcamos.

À minha madrinha, Cássia Selbach, por ser a precursora da profissão na família. E à minha tia Nara Silva, pela torcida de sempre.

Aos meus colegas e amigos desta trajetória de meia-década: Bruno Luce, Edna Dias, Luis Fernando Massoni, Thainá Loureiro e Tuany Schmitt, e em especial à minha colega e amiga de todas as horas Graziela Silveira. E também a todos aqueles que este curto espaço não poderia comportar...

À UFRGS, que ampliou os meus horizontes e os meus conhecimentos sobre o mundo e sobre mim mesmo... E mesmo com seus erros e acertos, acabou sendo uma experiência e tanto!

À FABICO, por ser um espaço tão acolhedor e plural.

Aos Professores da FABICO, que além de seus conhecimentos e ensinamentos técnicos, trouxeram ensinamentos para a vida.

À Biblioteca Pública Municipal Josué Guimarães (BPMJG), minha primeira experiência profissional, que me acolheu muito bem e me fez ter um novo olhar sobre as pessoas. Ao Centro de Processamento de Dados (CPD/UFRGS), especialmente à Biblioteca e a bibliotecária Caterina Pavão, que acreditou no meu potencial e me ofereceu oportunidades em dose dupla: estágio para o meu Curso Técnico em Informática e na Graduação em Biblioteconomia.

Ao meu Orientador, Prof. Dr. Rafael Port da Rocha, pelos ensinamentos e contribuições significativas neste trabalho.

Às componentes da minha banca, Professora Sonia Caregnato e a Bibliotecária Caterina Pavão, pela prontidão e o interesse no que eu tinha a dizer.

A todos aqueles que contribuíram, direta ou indiretamente, para o êxito deste trabalho e deste percurso.

"Textos autobiográficos costumam contar pouco da história das pessoas.

E nem poderiam contar tudo.

Numa vida cabem tantas coisas.

Em poucas linhas é difícil contar sequer uma única experiência, que dirá uma vida inteira."

(O autor).

RESUMO

Apresenta estudo acerca dos critérios de seleção para digitalização de obras raras, especialmente de livros. Discute aspectos relativos à digitalização, como a preservação digital e a produção de metadados, que visam a recuperação e a salvaguarda de informações. Destaca as sugestões sobre seleção para digitalização apresentadas por cinco instituições de diferentes esferas, sendo respectivamente, um conselho, uma biblioteca universitária, uma empresa, uma federação internacional e um pesquisador da área: CLIR – Council on Library and Information Resources, Cornell University Library, EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária, IFLA – International Federation of Library Associations and Institutions e o pesquisador Prof. Dr. Rubens Ribeiro Gonçalves da Silva – da Universidade Federal da Bahia. Realiza um levantamento a respeito das informações disponibilizadas nas fontes deste estudo. Compara os critérios de seleção apontados pelas fontes, com suas características e focos. Propõe um quadro de apoio à seleção de obras raras para digitalização. Avalia que a melhor forma de selecionar obras raras para este fim é observar as características e necessidades da instituição para qual este processo será dedicado e que é possível valer-se da utilização de instrumentos de outras instituições como apoio.

PALAVRAS-CHAVE: Digitalização. Obras raras. Critérios de seleção. Preservação digital.

ABSTRACT

The following paper aims to study the criteria chosen to scan rare pieces of art, especially books. It brings to light the discussion of other aspects of scanning such as digital preservation and the production of metadata, which have the objective of recovering and saving information. This study also highlights suggestions about the selection to scan books given by five institutions, each one from a different field: a council (CLIR – Council on Library and Information Resources), a College library (Cornell University Library), a private company (EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária), an international federation (IFLA – International Federation of Library Associations and Institutions) and a researcher of the field, professor Dr. Rubens Ribeiro Gonçalves da Silva from Universidade Federal da Bahia. Therefore, it lists the information available in the sources, comparing their criteria of selections, their characteristics and focus. This paper comes up with a support framework for the selection of those rare works that are going to be scanned. Finally, it evaluates that the best way to select rare works to this end is to consider the institution's needs and characteristics and that it is possible to make use of another institution's tools as a support.

KEY-WORDS: Scanning. Rare works. Criteria of selection. Digital preservation.

LISTA DE ABREVIATURAS E SIGLAS

BN	Biblioteca Nacional (Brasil)
CEDAP	Centro de Documentação e Acervo Digital da Pesquisa
CLIR	Council on Library and Information Resources
CONARQ	Conselho Nacional de Arquivos
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
FABICO	Faculdade de Biblioteconomia e Comunicação
IFLA	International Federation of Library Associations and Institutions
MARC	Machine Readable Cataloging
METS	Metadata Encoding & Transmission Standard
OAIS	Open Archival Information System
PREMIS	Preservation Metadata
TICs	Tecnologias da Informação e Comunicação
UFBA	Universidade Federal da Bahia
UFRGS	Universidade Federal do Rio Grande do Sul

LISTA DE QUADROS

QUADRO 1	–	Possibilidades de um representante digital.....	25
QUADRO 2	–	Configuração do representante digital.....	27
QUADRO 3	–	Estratégias de preservação digital.....	33
QUADRO 4	–	Metodologia.....	38
QUADRO 5	–	Critérios propostos pelas fontes.....	46
QUADRO 6	–	Resumo dos focos nas fontes.....	49
QUADRO 7	–	Análise comparativa sobre seleção.....	50
QUADRO 8	–	Modelo de apoio ao processo de seleção.....	56

SUMÁRIO

1	INTRODUÇÃO.....	12
1.1	IDENTIFICAÇÃO E JUSTIFICATIVA DO PROBLEMA.....	13
1.2	OBJETIVOS.....	14
1.2.1	Objetivo Geral.....	15
1.2.2	Objetivos Específicos.....	15
2	REVISÃO DE LITERATURA.....	16
2.1	OBRAS RARAS (LIVROS).....	16
2.2	OBJETO DIGITAL.....	18
2.3	DIGITALIZAÇÃO.....	19
2.3.1	Seleção de Materiais.....	22
2.3.2	Conversão.....	24
2.3.3	Controle de Qualidade.....	28
2.3.4	Fluxo de Digitalização.....	29
2.4	PRESERVAÇÃO DIGITAL.....	32
2.5	METADADOS.....	35
3	METODOLOGIA.....	37
3.1	TIPO DE ESTUDO E ABORDAGEM.....	37
3.2	MÉTODOS.....	38
3.3	FONTES UTILIZADAS.....	39
3.3.1	<i>Council on Library and Information Resources</i> (CLIR).....	40
3.3.2	<i>Cornell University Library</i>.....	41
3.3.3	Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA).....	42
3.3.4	<i>International Federation of Library Associations and</i> <i>Institutions (IFLA)</i>.....	42
3.3.5	Manual de Digitalização de Acervos – Silva (UFBA).....	43

4	RESULTADOS.....	45
4.1	CRITÉRIOS DE SELEÇÃO NAS FONTES.....	45
4.2	FOCOS E CARACTERÍSTICAS DAS FONTES.....	47
4.3	CATEGORIAS PARA CRITÉRIOS DE SELEÇÃO.....	50
4.3.1	Categorias Escolhidas.....	52
4.3.2	Sobre as Categorias.....	53
4.4	MODELO DE APOIO AO PROCESSO DE SELEÇÃO.....	55
5	CONSIDERAÇÕES FINAIS.....	59
	REFERÊNCIAS.....	61
	APÊNDICE A – <i>Selecting Research Collections for Digitization</i> (CLIR).....	64
	APÊNDICE B – Moving Theory into Practice, Digital Imaging Tutorial (Cornell University Library).....	67
	APÊNDICE C – Guia para digitalização de documentos versão 2.0 (EMBRAPA).....	69
	APÊNDICE D – Diretrizes para planejamento de digitalização de livros raros e coleções especiais (IFLA).....	71
	APÊNDICE E – Manual de Digitalização de Acervos: textos, mapas e imagens fixas (SILVA).....	72
	ANEXO A – Geração de matrizes digitais.....	75
	ANEXO B – Modelo de verificação da qualidade do representante digital.....	76
	ANEXO C – A seleção para digitalização: matriz de apoio ao processo de decisão.....	77

1 INTRODUÇÃO

O mundo avança em busca de um conforto, muitas vezes, inalcançável. Com o advento das novas tecnologias, principalmente as computacionais, todos os dias, as pessoas buscam facilidade para a realização das suas tarefas diárias, das mais importantes até aquelas que têm uma importância irrisória. Não existem dúvidas de que os computadores mudaram a forma dos seres humanos lidarem consigo e com tudo aquilo que está a sua volta. Computadores pessoais, *smartphones*, *tablets*, TVs digitais, impressoras 3D, dentre outras tecnologias, trouxeram muito mais do que inovação para as casas das pessoas: elas trouxeram um mundo de possibilidades até então desconhecido por estes indivíduos.

Vannevar Bush afirma¹ em seu artigo chamado “*As We May Think*”, datado do ano de 1945, que após a Segunda Guerra Mundial (1939-1945) houve um expressivo aumento do volume e do valor da informação, e que controlar este fluxo era difícil. Poucas décadas após surgiram os primeiros computadores.

Existem algumas ideias sobre quando, onde e como o computador surgiu como conceito. Os computadores pessoais popularizaram-se entre as décadas de 1970 e 1980 com a participação de empresas multinacionais como a IBM, a Apple e a Microsoft.

Durante muitos anos, no decurso de décadas, o ramo da tecnologia procurou avançar sobre as possibilidades até então desconhecidas sobre a informática. Com isto, as seguintes dúvidas começaram a pairar no ar: “Será o fim do livro?”; “Vamos parar de utilizar documentos em papel?”. Do ponto de vista ecológico, é mais vantajoso dizer sim para estes questionamentos. Do ponto de vista sobre a viabilidade destas práticas também, afinal existem diversos suportes necessários para esta transposição. Mas será que estamos prontos de fato, no século XXI, para esta transição?

A preservação de documentos é uma preocupação constante neste mundo digital que vivenciamos. Áreas do conhecimento como a Informática, a Ciência da Informação e suas relacionadas, como a Arquivologia, a Biblioteconomia e a Museologia, preocupam-se cada vez mais com este assunto como forma de guardar a história, a memória, a cultura e todas as informações e conhecimentos sobre todas

¹ Informação disponível em: <<http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>>. Acesso em: 13 dez. 2015.

as áreas. Para o armazenamento de documentos em papel, temos a Digitalização como o caminho que aparenta ser o mais lógico – ao menos, até inventarem alguma nova tecnologia revolucionária. Ou talvez a própria digitalização se reinvente como tecnologia com o passar do tempo.

A digitalização, dentre outras coisas, contribui para a preservação de obras raras, como os livros, visto que ela evita o uso e o manuseio destas obras, permitindo uma maior vida útil das mesmas, e oferece ainda um maior acesso ao conteúdo delas, visando o âmbito digital. Estas justificativas são apenas algumas das diversas possíveis para o que este processo é capaz de realizar.

Este trabalho analisou recomendações de instituições/organizações respeitadas em suas áreas de atuação e realizou um comparativo acerca do que cada uma delas contribuem sobre a seleção para a digitalização de documentos. Por fim, existe a proposição de criar, redigir e apresentar um produto/instrumento relacionado as recomendações existentes por instituições de diversas especializações, que possa ser útil dentro e fora dos muros da nossa Universidade, tanto para o CEDAP, quanto para qualquer instituição ou pessoa que julgá-lo útil.

1.1 IDENTIFICAÇÃO E JUSTIFICATIVA DO PROBLEMA

A Universidade Federal do Rio Grande do Sul (UFRGS), sabendo da necessidade de preservar documentos importantes das suas Áreas de Pesquisa, decidiu criar o Centro de Documentação e Acervo Digital da Pesquisa (CEDAP). Para tanto, investiu grandes recursos financeiros e um grande esforço conjunto para criar este centro de digitalização, que tem por objetivo atender as necessidades da comunidade universitária, bem como preservar a história institucional, contemplando pesquisa e obras raras. A criação deste centro trouxe para a UFRGS a possibilidade de colocar em prática toda a teoria sobre o assunto, além de tornar-se uma referência na área e de abranger um tema tão importante no mundo cada vez mais digital em que estamos situados.

A realização deste trabalho possui quatro pilares principais que podem justificá-lo:

- a) o CEDAP é um importante passo da UFRGS rumo ao seu próprio futuro, não esquecendo bem como reconhecendo as conquistas e a importância do seu passado;
- b) o envolvimento do Professor Orientador, Dr. Rafael Port da Rocha, no assunto e como integrante do CEDAP;
- c) o interesse do aluno pela área das novas tecnologias, explicitado na procura de cursos especializados fora da área da Biblioteconomia (como a formação no Curso Técnico em Informática), mas que possuem uma grande interação e relevância para com a mesma;
- d) a relevância do tema tanto para a Biblioteconomia, quanto para a Ciência da Informação.

Para Rocha (2015 *apud* Santos 2015), é preciso:

Dar suporte à pesquisa científica, tecnológica, artística e cultural através da preservação é o objetivo do Centro de Documentação e Acervo Digital de Pesquisa da UFRGS (CEDAP), como expõe o professor Rafael Port da Rocha, coordenador do Centro. O professor salientou as dificuldades em se tirar do papel a ideia de iniciar um projeto de acervo como este que vem sendo desenvolvido na Universidade. O CEDAP, segundo ele, tem como meta principal garantir que trabalhos de pesquisa possam ser consultados e compartilhados e não se percam com tempo.

Assim sendo, este trabalho pretende responder a seguinte questão: “Como selecionar obras raras para digitalização observando critérios propostos pela literatura da área?”.

1.2 OBJETIVOS

Os objetivos deste trabalho estão nas seções seguintes, em 1.2.1 e 1.2.2.

1.2.1 Objetivo Geral

O objetivo geral é: “Analisar critérios de seleção para a digitalização de obras raras observando critérios propostos pela literatura da área”.

1.2.2 Objetivos Específicos

Os objetivos específicos que este trabalho contempla são os seguintes:

- a) investigar os conhecimentos envolvidos no processo da digitalização;
- b) identificar os critérios de seleção de obras raras no processo da digitalização;
- c) categorizar os principais aspectos abordados na seleção;
- d) comparar os critérios;
- e) propor critérios de seleção, em categorias.

2 REVISÃO DE LITERATURA

Esta seção fará um apanhado teórico acerca dos assuntos pertinentes à esta temática.

2.1 OBRAS RARAS (LIVROS)

O que você lembra quando ouve a expressão “obra rara”? Certamente, vem à sua cabeça um livro antigo, como a Bíblia de Gutenberg (*Mogúncia*), ou quadro de algum pintor renomado, como a Mona Lisa (*La Gioconda*), de Leonardo da Vinci.

Em linhas gerais, em relação a raridade dos livros:

[...] pode-se dizer que livro raro é aquele difícil de encontrar por ser muito antigo, ou por tratar-se de um exemplar manuscrito, ou ainda por ter pertencido a uma personalidade de reconhecida projeção e influência no país e mesmo fora dele (por exemplo: imperadores, reis, presidentes), ou reconhecidamente importantes para determinada área do conhecimento (física, biologia, matemática e outras). (RODRIGUES, 2006, p. 115).

Diversos fatores podem determinar a raridade ou não de algum livro. Entretanto, alguns deles são mais preponderantes do que outros para validar uma obra. Conforme Pinheiro (1989, p. 29-32) *apud* Rodrigues (2006, p. 116), alguns dos principais fatores a serem levados em conta são:

- a) limite histórico: observar, por exemplo, os períodos que caracterizam a produção artesanal de impressos, bem como a fase inicial da imprensa em determinado lugar;
- b) aspectos bibliológicos: observar aspectos como a presença de ilustrações produzidas artesanalmente, os materiais utilizados para a confecção do suporte na impressão, como tipo de papel, emprego de pedras ou materiais preciosos na encadernação;
- c) valor cultural: observar as publicações em pequenas tiragens, personalizadas, censuradas, expurgadas, as primeiras edições etc.;
- d) pesquisa bibliográfica: existem dicionários e enciclopédias bibliográficos especializados nesse tipo de publicação, que apontam certas peculiaridades da obra, como preciosidade e raridade;
- e) características do exemplar: observar as características particulares do exemplar que se tem em mãos, como a presença de autógrafo ou dedicatória de personalidade importante, marcas de propriedade e outros.

É importante ressaltar que o estabelecimento de critérios sobre raridade varia de instituição para instituição. Cada uma realiza uma análise para verificação do seu acervo e então determina como procederá na seleção de obras que irão compor o seu acervo exclusivo de obras raras.

Autoridade no país quando o assunto são as obras raras, a Biblioteca Nacional (BN) brasileira revela sobre o seu acervo que:

O Acervo Especial de Obras Raras é constituído de material diversificado, oriundo de diversas coleções da própria Biblioteca Nacional, de acordo com dois critérios principais de seleção: raridade e preciosidade. Ou seja, não basta que a obra seja antiga, é preciso também que seja única, inédita, faça parte de alguma edição especial ou apresente algum traço de distinção, como uma encadernação de luxo ou o autógrafo de uma celebridade – como D. Pedro II, Coelho Neto, Carlos Drummond de Andrade ou Jorge Amado. Integram também esse acervo periódicos raros publicados até o século XIX. (BIBLIOTECA NACIONAL, ([20--]), *online*).

São exemplos de obras raras do acervo da BN ([20--]):

- a) a edição do poema “A Visita”, autografada por Carlos Drummond de Andrade. A obra é ilustrada com fotografias de Maureen Bisilliat, que homenageia o grande poeta mineiro;
- b) a primeira edição da “Arte da gramática da língua portuguesa”, escrita pelo Padre Anchieta;
- c) a Bíblia de Mogúncia impressa em 1462;
- d) a primeira edição de “Os Lusíadas”, de 1572;
- e) o “Rerum per octennium...Brasília”, de Baerle (1647), com 55 pranchas a cores desenhadas por Frans Post.

Abaixo, na **FIGURA 1**, está a Bíblia de Gutenberg (Mogúncia), uma das obras mais importantes do acervo da BN e que ilustra bem a importância das obras raras.

FIGURA 1 – Bíblia de Gutenberg (Mogúncia)



Fonte: Site da Fundação Biblioteca Nacional (2015)².

2.2 OBJETO DIGITAL

Informação é a palavra-chave no universo biblioteconômico. Ela pode ter vários significados, devido aos vários pontos de vista e vertentes que tratam da sua conceituação. Apesar desta simples palavra de dez letras ser tão significativa em um universo tão amplo, como é o da Biblioteconomia, ela não representa todo o processo que é realizado até que o usuário das bibliotecas saia plenamente satisfeito destes ambientes. Antes existe o dado e após há o conhecimento (embora fale-se ainda em conceitos como sabedoria e competência).

² Imagem disponível em: <<https://www.bn.br/acervo/obras-raras>>. Acesso em: 15 set. 2015.

De acordo com Setzer (1999, *online*):

Definimos dado como uma seqüência de símbolos quantificados ou quantificáveis. [...] Informação é uma abstração informal (isto é, não pode ser formalizada através de uma teoria lógica ou matemática), que representa algo significativo para alguém através de textos, imagens, sons ou animação. [...] Conhecimento é uma abstração interior, pessoal, de alguma coisa que foi experimentada por alguém.

Podemos dizer que o dado é a menor parte de um objeto, o que pode ser decodificado por uma máquina, como o computador, por exemplo. Informação é a união de diversos dados a fim de transmitir uma mensagem. Conhecimento seria a convergência de informações empíricas.

No mundo da Informática, os números binários 0 e 1 definem toda a informação que é possível visualizar na tela. O computador tem a capacidade de traduzir dados codificados em 0 ou 1. A combinação adequada de zeros e um define qual tipo de informação será distribuída, a forma que ela será utilizada pelo usuário irá gerar um conhecimento particular nele. A partir disso, temos o objeto digital como o protagonista do processo.

Segundo Thibodeau (2002) *apud* Ferreira (2006, p. 21), “Um objecto digital pode ser definido como todo e qualquer objecto de informação que possa ser representado através de uma seqüência de dígitos binários”. As possibilidades de objetos digitais são diversas, indo de simples documentos de texto a *softwares* inteiros.

Os objetos digitais podem ser natos ou convertidos. Quando são natos, já são digitais. Quando eles são digitalizados, eles são convertidos de um documento físico, seja ele um livro, uma fotografia, um jornal.

Nos capítulos seguintes veremos qual o papel essencial do objeto digital na preservação digital.

2.3 DIGITALIZAÇÃO

A figura do bebê (**FIGURA 2**) demonstra a primeira imagem digital de que se tem notícia. Ela foi criada por Russell Kirsch em 1957 no seu protótipo de scanner. Ela possuía 176x176 pixels, capacidade da memória do computador à época. (DISCOVERY NEWS, 2010).

FIGURA 2 – Primeira imagem digital



Fonte: Discovery News (2015)³.

Segundo o Conselho Nacional de Arquivos – CONARQ (2010, p. 5-6), Digitalização é:

[...] um processo de conversão dos documentos arquivísticos em formato digital, que consiste em unidades de dados binários, denominadas de bits - que são 0 (zero) e 1 (um), agrupadas em conjuntos de 8 bits (binary digit) formando um byte, e com os quais os computadores criam, recebem, processam, transmitem e armazenam dados.

Um ponto importante a ser destacado sobre digitalizar é que:

De acordo com a natureza do documento arquivístico original, diversos dispositivos tecnológicos (*hardware*) e programas de computadores (*software*) serão utilizados para converter em dados binários o documento original para diferentes formatos digitais. No entanto, o produto dessa conversão não será igual ao original e não substitui o original que deve ser preservado. (CONSELHO NACIONAL DE ARQUIVOS, 2010, p. 6).

A digitalização de um documento, seja qual for a sua natureza e a sua importância, é, de certa forma, conscientemente ou não, uma preservação (em formato) digital de um documento. De acordo com a CONARQ (2010, p. 6): “A digitalização, portanto é dirigida ao acesso, difusão e preservação do acervo documental”.

³ Disponível em: <<http://news.discovery.com/tech/apps/digital-image-pixel.htm>>. Acesso em: 30 set. 2015.

Dentre outros motivos possíveis, de acordo com realidades vivenciadas por diversas instituições, pode-se destacar os seguintes:

- a) contribuir para o amplo acesso e disseminação dos documentos arquivísticos por meio da Tecnologia da Informação e Comunicação;
- b) permitir o intercâmbio de acervos documentais e de seus instrumentos de pesquisa por meio de redes informatizadas;
- c) promover a difusão e reprodução dos acervos arquivísticos não digitais, em formatos e apresentações diferenciados do formato original;
- d) incrementar a preservação e segurança dos documentos arquivísticos originais que estão em outros suportes não digitais, por restringir seu manuseio. (CONARQ, 2010, p. 6).

Com o exponencial crescimento da internet mundialmente desde os anos 90, principalmente na figura do sistema operacional *Microsoft Windows*, a forma de lidar com as leituras realizadas em entidades educacionais difundiu a cultura da digitalização e da fotocópia – conhecido popularmente como “xerox”, nome este advindo da empresa norte-americana Xerox, que popularizou, com sua tecnologia, o procedimento de fotocopiar documentos. Entretanto, de acordo com a IFLA (2015, p. 11): “É sempre preferível a digitalização de uma entidade intelectual completa, em vez de parte dela. Portanto, é melhor a digitalização de um livro inteiro ou documento, em vez de um capítulo ou uma página”. As leis de direito autoral no Brasil, entretanto, vetam a cópia integral de qualquer obra literária, segundo a Lei 9.610 de junho de 1998.

Comercialmente, existem diversas empresas que realizam a digitalização de documentos impressos, de vários formatos e espécies. Entretanto, o processo é bastante simplificado, isto é, um indivíduo ou instituição solicita os serviços de uma determinada empresa e recebe apenas o arquivo digitalizado bruto, não existindo um tratamento técnico para este documento, uma descrição de metadados ou uma disponibilização em diversas mídias, por exemplo. É importante ressaltar que além das empresas que realizam a digitalização de documentos impressos, ainda existem aquelas que oferecem produtos (*hardware* e *software*) e serviços de assistência técnica para outras empresas/instituições, com o objetivo de que elas próprias

realizem seus respectivos projetos de digitalização, adaptados ao seu contexto e a sua realidade. Este é o caso do projeto realizado pela UFRGS.

O processo de digitalização consiste basicamente em cinco passos principais, sendo eles: Seleção, Conversão, Metadados, Fluxo de Digitalização e Preservação Digital. O primeiro deles, a Seleção, é o tema central deste estudo, mas veremos nos próximos capítulos as particularidades de cada um destes processos.

2.3.1 Seleção de Materiais

No que diz respeito à seleção, é necessário analisar a realidade vivenciada na Instituição que optar por realizar a digitalização de seus documentos. Conforme a Empresa Brasileira de Pesquisa Agropecuária – EMBRAPA (2006, p. 15), em seu “Guia de Digitalização”, é preciso verificar inicialmente: “[...] as metas e características da instituição, o perfil e necessidades dos usuários, e as características dos próprios materiais”. Além das metas da instituição e da realidade de seus usuários, é preciso estudar detalhadamente o documento a ser digitalizado: “Coleções, obras, edições e cópias devem ser estudadas e comparadas, em relação ao âmbito da nova coleção digital. Também devem ser considerados aspectos, como cronologia, local de publicação, autor, assunto, formato, possuidor etc.” (IFLA, 2015, p. 11).

Alguns questionamentos devem ser realizados no processo de seleção de materiais, como por exemplo, os propostos pela EMBRAPA:

- a) qual é o propósito de digitalizar material e quais os critérios que serão utilizados para selecioná-los?
- b) quais serão os usuários desses recursos?
- c) quais são suas necessidades?
- d) os materiais selecionados irão requerer cuidados especiais no manuseio?
- e) quais as características dos materiais e como elas podem afetar o processo de digitalização? (EMBRAPA, 2006, p. 15-16).

Já a IFLA (2015, p. 11) propõe os seguintes questionamentos:

- a) queremos digitalizar documentos únicos de distintas tipologias, como por exemplo, “tesouros”?
- b) queremos digitalizar uma coleção existente?
- c) queremos criar uma nova coleção “virtual”? Por exemplo, uma coleção de documentos procedentes de distintas instituições que tenham algo em comum, como sua procedência.

O primeiro aspecto a ser levado em conta na seleção de materiais são as características do material original. Neste sentido, é preciso planejar a digitalização de forma que ela esteja de acordo com as características do documento original e não deverá estar delineada essencialmente pelos instrumentos tecnológicos disponíveis, pois existem diversos equipamentos e técnicas para digitalização, sendo que cada tecnologia contempla melhor um tipo diferente de documento (EMBRAPA, 2006). Questões como originalidade do material, correção da informação contida nele, condições físicas, importância do seu conteúdo são questões a serem analisadas de forma abrangente.

Outro ponto a ser considerado são as propriedades físicas do material. É o que atesta a IFLA (2015, p. 11): “O melhor exemplar físico em mãos deve ser selecionado para digitalização. A seleção deve incluir uma revisão física do material, pois sua condição influirá no processo de digitalização”.

Silva (2005, p. 21) argumenta que existe Seleção para vários fins:

O processo de seleção de documentos ou coleções para conversão digital pode ser bastante similar a outros processos de seleção já bastante conhecidos de profissionais de bibliotecas, arquivos e museus como, por exemplo, seleções prioritárias para conservação, seleção de conteúdos para exibição ou publicação, seleção para tratamento em função de demanda de usuários, seleção de documentos com restrições legais de reprodução ou acesso, etc. Processos de seleção sempre envolvem indicações, avaliações e prioridades, as quais não podem ser decididas por uma única pessoa.

Por último, e não menos importante, estão os direitos autorais da obra. A EMBRAPA (2006, p. 19) recomenda que:

Assim que você identificar o material que deverá ser digitalizado comece a cuidar dos aspectos relacionados a propriedade intelectual. Se o material que você quer digitalizar é de domínio público, ou se a empresa possui os direitos autorais e controla os direitos para digitalização, então você provavelmente poderá prosseguir sem impedimentos.

Se os materiais que você quer digitalizar não são de domínio público, e você não controla os direitos autorais, será preciso identificar o detentor desses direitos e solicitar permissão para digitalizar e publicar o material em questão. Mesmo que você já saiba quem é o detentor dos direitos, e mesmo que você já tenha permissão para uso do material para certos propósitos, você ainda terá que se certificar que a sua permissão cobre a digitalização e publicação em meio digital.

A IFLA (2006, p. 12) alerta para a verificação das leis de propriedade intelectual, que segundo ela, “variam significativamente de país para país”. Como conclusão, fica esclarecido também que, com relação a publicação de uma obra, “[...] um objeto digitalizado pode ser considerado como uma nova edição. Como resultado, a disponibilidade e as condições de uso para cada objeto e coleção digitalizados devem ser claramente indicadas para os usuários”.

2.3.2 Conversão

Após a seleção de materiais, objeto de estudo deste trabalho, está a Conversão dos documentos originais para o meio digital. A Conversão é processo no qual a:

Captura da imagem digital deve levar em consideração os processos técnicos envolvidos na conversão de analógico para representação digital, bem como os atributos dos próprios documentos de origem: tamanho físico e apresentação, nível de detalhe, faixa tonal e presença de cor. Os documentos também podem ser caracterizados pelo processo de produção utilizado para criá-los, incluindo o método manual, automático, de fotografia, e mais recentemente, dos meios eletrônicos. (CORNELL UNIVERSITY LIBRARY, 2003, *online*, tradução nossa).

Neste âmbito, existe o Representante Digital, que é a representação em formato digital de um documento que não é digital genuinamente. A finalidade dos

representantes digitais de um documento que será digitalizado é definida com seu propósito de uso.

O Representante digital possui as Matrizes Digitais (ou Arquivos Mestres) e as Derivadas de Acesso. O Arquivo Mestre é a cópia em formato de imagem mais próxima da originalidade do documento:

Trata-se de arquivos de imagens digitais que contêm o máximo de atributos das imagens originais. O arquivo mestre deve ser de alta qualidade, visto que preserva o conteúdo informacional do original, possibilitando variados usos e formatos alternativos para atendimento às várias demandas, evitando re-trabalho de digitalização. A cópia mestra não deve ser editada nem compactada. Os arquivos de imagens mestras são muito grandes para serem armazenados on-line. (EMBRAPA, 2006, p. 26).

Já as Derivadas de Acesso, como o próprio nome já indica, são as cópias imagéticas que servirão para o acesso dos usuários da instituição/serviço:

Arquivos derivados são criados a partir da imagem digital mestra e são usados geralmente para acesso na Web. Arquivos derivados incluem imagens de acesso, dimensionadas para a tela de um monitor médio, e suas respectivas miniaturas. São geralmente armazenadas on-line. (EMBRAPA, 2006, p. 26).

Um representante digital é capaz de auxiliar: na Preservação, no Acesso, no Índice, na Reprodução e na Recuperação/Busca, como vemos no **QUADRO 1**.

QUADRO 1 – Possibilidades de um representante digital

SERVE PARA	FUNÇÃO
Preservação	<p>Visa o armazenamento a longo prazo. Possui alta qualidade e evita a compressão.</p> <p>São imagens matrizes, pois delas são derivadas imagens de acesso e de navegação.</p>
Acesso	<p>Serve para visualização. Possui média qualidade.</p>

(Continua)

(Conclusão)

Índice	Destina-se a visualização rápida. Apresenta baixa qualidade e tamanho reduzido (em formato de miniatura).
Reprodução	Para reprodução do documento via impressora. Qualidade de reprodução próxima do original.
Recuperação/Busca	Objetiva localizar a ocorrência de palavras no texto do documento digitalizado.

Fonte: O autor (2015).

A imagem abaixo (**FIGURA 3**) apresenta as recomendações propostas pela EMBRAPA, em “Guia para digitalização de documentos. Versão 2.0”, sobre imagens mestras e derivadas.

FIGURA 3 – Características dos diversos arquivos de imagens

Imagem mestra	Imagem de acesso	Imagem em miniatura
<ul style="list-style-type: none"> • Representação mais próxima possível da informação contida na original • Não compactada • Não editada • Serve como fonte para arquivos derivados • Pode servir como substituto do original • Alta qualidade • Arquivo muito grande • Usado para criação de reproduções impressas de alta qualidade • Geralmente armazenado em arquivo de formato TIFF 	<ul style="list-style-type: none"> • Usado no lugar da imagem mestra para acesso via Web • Geralmente cabe dentro da área de visão do monitor médio • Tamanho de arquivo adequado para carregamento rápido; não requer conexão de rede rápida • Qualidade aceitável para pesquisas em geral • Compactada para velocidade de acesso • Geralmente armazenada em arquivo de formato JPEG 	<ul style="list-style-type: none"> • Imagem muito pequena geralmente apresentada como registro bibliográfico • Planejada para amostra rápida on-line; permite ao usuário determinar se ele quer ver a imagem de acesso • Geralmente armazenada em arquivos de formatos GIF ou JPEG • Serve como fonte de arquivos derivados • Não é apropriado para imagens de textos

Fonte: EMBRAPA (2006).

O CONARQ sugere, no documento “Recomendações para Digitalização de Documentos Arquivísticos Permanentes”, sugestões no que tange a geração das matrizes digitais, de acordo com o tipo de documento original, indicando o tipo de

reprodução (modo de cor), formato de arquivo digital e resolução mínima em dpi. O quadro pode ser visualizado no **ANEXO A**.

Existem ainda os aspectos técnicos da Conversão (**QUADRO 2**). Para realizar esta etapa é necessário realizar a configuração da imagem para os tipos de documentos originais identificados. Esta configuração contempla os seguintes itens: a Compactação, o Formato de Arquivo, o Modo e a Profundidade de bit (bitonal, tons de cinza e cor) e a Resolução.

A origem do documento é muito importante para o processo de Conversão, visto que cada tipologia de documento requer cuidados especiais e específicos. A configuração para textos impressos certamente será diferente daquela aplicada para fotografias. Por exemplo, enquanto um texto pode utilizar modo bitonal e resolução e 300 dpi para representar sua parte visual, provavelmente para uma fotografia isso não será suficiente, considerando a grande perda de qualidade do original.

QUADRO 2 – Configuração do representante digital

ITEM	DESCRIÇÃO	EXEMPLO DE CONFIGURAÇÃO POSSÍVEL
Compactação	A imagem tem seu espaço de armazenamento reduzido através de compressão.	Sem compactação; Compactação sem perda; Compactação com perda.
Formato de Arquivo	Formato de representação do arquivo.	Imagem Matriz TIFF; PNG; JPEG2000. Derivada JPEG; GIF; PDF.
Modo e Profundidade de bit	Indica o número de tons diferentes (cores ou tonalidades de cinza) que podem ser utilizadas na imagem.	Modos Bitonal: 1 bit; Tons de cinza: 8, 16 bits; Cor: 8, ..., 24 bits.

(Continua)

(Conclusão)

Resolução	Indica quantos pontos (pixels) são usados para representar a área de imagem. São medidos em pontos por polegada (do inglês <i>dots per inch – dpi</i>).	300 dpi; 600 dpi.
------------------	--	----------------------

Fonte: O autor (2015).

2.3.3 Controle de Qualidade

Na guerra de produtos e serviços oferecidos por diversas empresas no século XXI, não basta realizar a promoção de um produto ou serviço, ele precisa estar dotado de qualidade no seu processo de fabricação para se tornar durável. No processo de digitalização, o Controle de Qualidade é essencial para que as imagens derivadas estejam com qualidade suficiente para a sua posterior disponibilização aos usuários. O Controle de Qualidade (QC – de *Quality Control*) é:

[...] o componente integrante de uma iniciativa de produção de imagem digital que tem como objetivo garantir que as expectativas de qualidade foram cumpridas. Engloba procedimentos e técnicas para verificar a qualidade, precisão e consistência dos objetos digitais produzidos. (CORNELL UNIVERSITY LIBRARY, 2003, *online*, tradução nossa).

A EMBRAPA (2006) afirma que esta verificação deve classificar a qualidade de uma imagem gerada pela digitalização como aceitável ou inaceitável:

A inspeção de arquivos de imagens digitais finais deve ser incorporada ao fluxo de trabalho nas proporções desejadas de acompanhamento. Recomenda-se que procedimentos de controle de qualidade sejam implementados e documentados e que sejam definidos claramente os defeitos inaceitáveis numa imagem. (EMBRAPA, 2006, p. 33).

Além disso, a qualidade de uma imagem é avaliada nos aspectos subjetivos e visuais: “A qualidade é avaliada tanto subjetivamente pela equipe responsável pelas atividades, através da inspeção visual, como objetivamente no software de imagens”. (EMBRAPA, 2006, p. 33).

Itens como resolução incorreta, nome do arquivo incorreto, formato de arquivo incorreto, falta de nitidez, imagem distorcida, equilíbrio de cores incorreto, dentre outros diversos fatores analisados como insuficientes pelo controle de qualidade, são passíveis de reavaliação e correção.

O **ANEXO B** traz um modelo de controle de qualidade. Trata-se do “Modelo de Verificação da Qualidade do Representante Digital”, proposto pelo CONARQ.

2.3.4 Fluxo de Digitalização

O Fluxo de Digitalização, ou Fluxo de Trabalho, é o processo pelo qual um documento digitalizado passará desde sua preparação (Seleção) até a sua disponibilização em repositório digital (para sua efetiva preservação digital).

Consoante a isso, a Cornell University Library (2003, *online*, tradução nossa) afirma que no fluxo de digitalização (Cadeia de Digitalização):

Todo o processo é, às vezes, chamado de cadeia de digitalização, o que sugere uma série de passos logicamente ordenadas. Na prática, a cadeia de digitalização pode ter ramos laterais, *loops* e passos recorrentes [...].

É necessário ter um olhar atento e vasto para todo o processo do fluxo/cadeia:

A tecnologia necessária para navegar a partir de uma extremidade da cadeia de digitalização para a outra consiste principalmente em *hardware*, *software* e rede. [...]. Uma visão verdadeiramente abrangente de infraestrutura técnica também inclui protocolos e padrões, políticas e procedimentos (para fluxo de trabalho, manutenção, segurança, atualizações, etc.), e os níveis de habilidade e responsabilidades do trabalho do pessoal de uma organização. (CORNELL UNIVERSITY LIBRARY, 2003, *online*, tradução nossa).

A tecnologia selecionada para digitalizar é crucial para o sucesso do empreendimento, visto que ela costuma ser terceirizada:

As decisões técnicas da infraestrutura necessitam de um planejamento cuidadoso, porque a tecnologia de imagem digital muda rapidamente. A melhor maneira de minimizar o impacto da depreciação e obsolescência é através de uma avaliação cuidadosa, e a necessidade de evitar soluções únicas e proprietárias. Se as escolhas de equipamentos estão bem combinadas para as utilizações previstas e os resultados esperados e sincronizado com calendários

realistas, o retorno sobre o investimento será maximizado. (CORNELL UNIVERSITY LIBRARY, 2003, *online*, tradução nossa).

Em dez passos, um fluxo de digitalização ideal contemplaria:

- a) preparação do material;
- b) digitalização;
- c) geração de derivadas;
- d) processamento da imagem;
- e) avaliação da qualidade do material digitalizado;
- f) redigitalização, recalibração do scanner, quando necessário;
- g) OCR, se planejado;
- h) produção dos metadados;
- i) empacotamento;
- j) armazenamento em repositório (ingestão em um repositório).

A **FIGURA 4** apresenta os passos da digitalização em um esquema:



Fonte: O autor (2015).

A **Preparação do material** consiste na identificação das condições do documento recebido, como, por exemplo, em que situação estão suas condições físicas (como páginas faltantes, letras ilegíveis ou registro de recebimento).

Neste modelo, a **Digitalização** é o ato de colocar o material no scanner apropriado e utilizar suas ferramentas para a captura de imagens digitais do mesmo.

A **Geração de derivadas** são as imagens que estarão, posteriormente, disponíveis para o público em geral, geralmente via *Web*. A tendência é que elas estejam em um formato de arquivo popular, de fácil acesso e que ocupe pouco espaço em disco. A imagem original é guardada em segurança em outro local, sem ter o acesso dos usuários.

O **Processamento da imagem** é a normalização das imagens de acordo com as normas e políticas estabelecidas previamente pela instituição.

A **Avaliação do material digitalizado** nada mais é que o teste do controle de qualidade das imagens geradas pelo scanner e processadas em *software* específico depois.

A **Redigitalização/Recalibração do scanner** ocorre quando as imagens geradas por ele não estão de acordo com o padrão de qualidade desejado.

Segundo a Wikipédia (2015, *online*):

OCR é um acrônimo para o inglês *Optical Character Recognition*, [que] é uma tecnologia para reconhecer caracteres a partir de um arquivo de imagem ou mapa de bits, sejam eles escaneados, escritos à mão, datilografados ou impressos. Dessa forma, através do OCR é possível obter um arquivo de texto editável por um computador.

Este procedimento depende do grau de especialização almejado pela instituição que está digitalizando um documento textual, visto que ele depende de *software* especializado.

A **Produção de metadados** está fortemente relacionada com a disponibilização do documento em meio digital, visto que os metadados são dados sobre dados, ou seja, informações importantes referentes às características do material e suas especificidades.

O **Empacotamento, ou Encapsulamento**, é a preservação do objeto digital conjuntamente com informações que permitam a recuperação, exibição e processamento do conteúdo do documento no futuro.

O **Armazenamento em repositório** é o ponto de chegada de todo o processo, no qual o objeto torna-se disponível para os usuários após a ingestão de um determinado documento no sistema. O modelo Open Archival Information System (OAIS) é o recomendado, visto sua interoperabilidade e a sua capacidade de preservação digital. O modelo OAIS será aprofundado na seção 2.5 METADADOS.

É preciso salientar que o fluxo de digitalização é algo específico, variando drasticamente de uma instituição para a outra, pois cada uma vive uma realidade diferente, seja financeira, seja de material humano, seja de objetivos, dentre outra série de questões peculiares à cada uma delas.

O fluxo de digitalização faz parte do processo de Produção de Imagens, onde há escolhas técnicas acerca de itens como o scanner a ser utilizado, a determinação

dimensionamento de imagens (tamanho em disco), formato e extensões de arquivos, dentre outras decisões pertinentes.

Apesar de todo este processo ser demorado e minucioso, ele é a garantia de uma preservação digital completa e que atenderia os usuários por ela beneficiados satisfatoriamente.

2.4 PRESERVAÇÃO DIGITAL

A criação do computador, o advento da internet e o avanço contínuo da tecnologia são fatores que inegavelmente mudaram a forma das pessoas se relacionarem entre si e com o mundo em que vivem. Estes fatores contribuíram para a explosão de informação criada e disseminada ao redor do planeta. O aumento crescente de informação, difundida principalmente via *Web*, é um dos temas mais latentes em artigos científicos de áreas como a Ciência da Computação e da Ciência da Informação. Tecnologias da Informação e Comunicação (TICs), Web 2.0 e 3.0, Redes Sociais, são assuntos que frequentemente estão em artigos e rodas informais de discussão. Apesar da importância da Preservação Digital, pouco se discute este tópico, visto a sua complexidade e volubilidade.

A definição para preservação digital aqui adotada é a proposta por Ferreira (2006, p. 20):

Designa-se, assim, por preservação digital o conjunto de atividades ou processos responsáveis por garantir o acesso continuado a longo-prazo à informação e restante patrimônio cultural existente em formatos digitais. A preservação digital consiste na capacidade de garantir que a informação digital permanece acessível e com qualidades de autenticidade suficientes para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação.

A seleção de ferramentas que possibilitem a preservação digital de múltiplos documentos impressos e digitais talvez sejam os fatores mais decisivos em todo o processo. É o que aponta Márdero Arellano (2004, p. 17), visto que:

Na preservação de documentos digitais, assim como na dos documentos em papel, é necessária a adoção de ferramentas que protejam e garantam a sua manutenção. Essas ferramentas deverão servir para reparar e restaurar registros protegidos, prevendo os danos e reduzindo os riscos dos efeitos naturais (preservação prospectiva),

ou para restaurar os documentos já danificados (preservação retrospectiva).

O cumprimento de todos os requisitos (acesso contínuo, acessível, autenticidade e suporte mais avançado) não é uma tarefa considerada fácil. Dúvidas surgem a todo momento, seja da parte técnica, como o formato de arquivo e do novo suporte, seja até a experiência de interação que o usuário final desta tecnologia vai ter no fim do extenso e delicado processo. Conforme Bullock (1999) *apud* Márdero Arellano (2004, p. 18), os requisitos para a preservação de documentos digitais devem ser os seguintes:

- a) fixar os limites do objeto a ser preservado;
- b) preservar a presença física;
- c) preservar o conteúdo;
- d) preservar a apresentação;
- e) preservar a funcionalidade;
- f) preservar a autenticidade;
- g) localizar e rastrear o objeto digital;
- h) preservar a proveniência;
- i) preservar o contexto.

É preciso ressaltar que tais requisitos seriam realizados em situações propícias e ideais. Entretanto, a obtenção de alguns deles já é algo para ser exaltado, visto a complexidade de todo o processo de alcance de todos estes requisitos.

Algumas estratégias para a efetivação da preservação digital são (**QUADRO 3**):

QUADRO 3 – Estratégias de preservação digital

ESTRATÉGIA	DESCRIÇÃO
Conservação/Preservação de Tecnologia	É a conservação da tecnologia (<i>software</i> e <i>hardware</i>) em que o documento se insere.
Emulação	Imitar em <i>software</i> um pedaço de <i>hardware</i> ou <i>software</i> .

(Continua)

(Conclusão)

Encapsulamento	Preservar o objeto digital juntamente com informações que permitem o desenvolvimento no futuro de mecanismos (<i>software</i>) para visualizar o documento.
Identificador de Objeto Persistente (Web)	Identificador dado a um objeto digital na <i>Web</i> que persiste ao longo do tempo, suportando, por exemplo a mudança do local (site) em que o objeto está armazenado. Um exemplo disso é o <i>Handle System</i> (https://www.handle.net/).
Migração	Converter um documento que está representado em um formato ultrapassado para outro, representado em um formato mais atual.
Normalização de formatos	Estabelecer um grupo pequeno de formatos que serão aceitos na base de dados, evitando o uso de mais de um formato para representar documentos de mesmas características.
Refreshamento de suporte	Copiar a informação digital de um suporte que está obsoleto ou fisicamente deteriorado para um suporte mais novo.
Transferência	Migrar para suporte analógico. Transferência para suportes não digitais mais estáveis, como impressão em papel ou microfilme.

Fonte: O autor (2015).

Quando se trata de preservação digital é sempre necessário estar atento a políticas e diretrizes. Embora não exista uma obrigatoriedade de seguir certos padrões e regras quando o assunto é preservação digital, o acompanhamento do que instituições e organizações especializadas recomendam é importante para que o projeto em questão tenha mais sucesso e esteja mais próximo daquilo que seria considerado o ideal. Neste sentido, existem esforços para que o maior número de instituições siga o que é aconselhado, principalmente em iniciativas como eventos, congressos e tratados. Uma destas iniciativas é o Open Archival Information System (OAIS):

[...] o modelo de referência OAIS (Open Archival Information System), um modelo conceptual que visa identificar os componentes funcionais que deverão fazer parte de um sistema de informação dedicado à preservação digital. O modelo descreve ainda as interfaces internas e externas do sistema e os objectos de informação que são manipulados no seu interior. (FERREIRA, 2006, p. 27).

A importância de usar um modelo de referência como o OAIS se dá na conformidade de ações e processos. Utilizar a mesma linguagem que outras comunidades é importante para a solução de dúvidas, a discussão de ideias e o compartilhamento de informações.

2.5 METADADOS

“Dados sobre dados”. Esta expressão, simples e, talvez, até curiosa, define, em termos gerais, o conceito de Metadados – que podem ser a catalogação de documentos digitais, fazendo uma comparação com o universo da Biblioteconomia. De acordo com a EMBRAPA (2006, p. 37):

Metadado significa “informações sobre os dados”, informações criadas sobre o material e a versão digital, o registro de sua identidade, criação, uso e estrutura. O propósito do metadado é facilitar a pesquisa, o uso, a administração e reutilização de material digital.

Os metadados podem ser de três categorias distintas: descritivo, administrativo e estrutural:

O metadado descritivo descreve e identifica as informações sobre os recursos digitais, facilitando a busca, acesso e administração do repositório. As informações são do tipo bibliográficas tais como: o criador/autor, título, data da criação, palavras-chave, entre outras.

O metadado administrativo é usado para facilitar o rastreamento, migração e reuso dos elementos digitais. As informações típicas dessa categoria são: informações sobre a criação, controle de qualidade, direitos, entre outras. O termo “metadado técnico” também é utilizado para indicar a data de captura da imagem e as características técnicas da imagem.

O metadado estrutural descreve a estrutura interna do recurso digital e sua relação com suas partes. É usado para possibilitar a navegação e a apresentação. (EMBRAPA, 2006, p. 37).

Existem diversos padrões de metadados em cada uma destas categorias. Um deles é o Machine Readable Cataloging (MARC), ou em português, catalogação

legível por máquina, padrão de descrição bastante utilizado em bibliotecas. Outros importantes padrões de metadados são o Preservation Metadata (PREMIS) e o Metadata Encoding & Transmission Standard (METS), classificados como administrativos (com enfoque na preservação digital) e estruturais, respectivamente.

O Preservation Metadata (PREMIS), que faz parte da categoria administrativa de metadados, visa dar suporte à preservação digital, buscando a continuidade do seu uso em longo prazo. (LIBRARY OF CONGRESS, 2015, *online*, tradução nossa).

Segundo a Library of Congress (2015, *online*, tradução nossa), o Metadata Encoding & Transmission Standard (METS) tem o objetivo de codificar metadados administrativos, descritivos e estruturais de um recurso em uma biblioteca digital.

Os metadados são ferramentas úteis no objetivo de administrar um repositório digital e todo o seu conteúdo. Pelo fato de metadados serem dados sobre outros dados, eles contribuem para um maior entendimento sobre os dados gerados por um *software* em um determinado contexto.

3 METODOLOGIA

Esta seção fará a elucidação acerca da metodologia (métodos, técnicas e procedimentos) que será empregada no trabalho.

3.1 TIPO DE ESTUDO E ABORDAGEM

Este estudo realizará uma pesquisa de natureza básica, pesquisa que: “Objetiva gerar conhecimentos novos úteis para o avanço da ciência sem aplicação prática prevista. Envolve verdades e interesses universais”. (SILVA, 2001, p. 20). Será básica, porque trará a literatura da área como fonte de estudo, não objetivando alguma prática posterior.

Do ponto de vista de sua abordagem, ela será qualitativa, visto que:

Considera que há uma relação dinâmica entre o mundo real e o sujeito, isto é, um vínculo indissociável entre o mundo objetivo e a subjetividade do sujeito que não pode ser traduzido em números. A interpretação dos fenômenos e a atribuição de significados são básicas no processo de pesquisa qualitativa. Não requer o uso de métodos e técnicas estatísticas. O ambiente natural é a fonte direta para coleta de dados e o pesquisador é o instrumento-chave. É descritiva. Os pesquisadores tendem a analisar seus dados indutivamente. O processo e seu significado são os focos principais de abordagem. (SILVA, 2001, p. 20).

O trabalho consistirá no estudo de bibliografias sobre os assuntos pertinentes à pesquisa, baseando-se em suas ideias, portanto não seria possível quantificar e sistematizar estes dados, descartando a abordagem quantitativa.

Foram realizados dois tipos de pesquisa neste estudo: pesquisa bibliográfica e pesquisa exploratória. A pesquisa bibliográfica consiste na literatura já conhecida da área, sendo baseada prioritariamente em livros e artigos científicos (GIL, 2002). Já a pesquisa exploratória tem o intuito de “[...] proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses. Pode-se dizer que estas pesquisas têm como objetivo principal o aprimoramento de ideias ou a descoberta de intuições” (GIL, 2002, p. 42).

Em geral, pesquisa exploratórias possuem um caráter mais flexível em seu planejamento, possibilitando a análise de vários os ângulos acerca do problema estudado.

3.2 MÉTODOS

Para atender os objetivos deste trabalho, consta o **QUADRO 4**. Nele estão os cinco objetivos específicos desta pesquisa, além das estratégias para solucionar cada um deles e as fontes que serão utilizadas para dar sustentação às mesmas.

QUADRO 4 – Metodologia

	OBJETIVO	MÉTODO	FONTES
A	investigar os conhecimentos envolvidos no processo da digitalização.	Revisão bibliográfica (desenvolvido no Cap. 2) .	Livros, artigos e periódicos da área.
B	identificar os critérios de seleção de obras raras no processo da digitalização;	Realizar o levantamento dos critérios de seleção. (Seção 3.3 e Seção 4.1) . Identificar como os critérios de seleção são categorizados nas fontes selecionadas. (Seção 4.1) .	Publicações com recomendações de critérios de seleção: CLIR⁴, Cornell University Library⁵, EMBRAPA⁶, IFLA⁷, SILVA⁸ (Seção 3.3) .

(Continua)

⁴ *Selecting Research Collections for Digitization (CLIR)*

<http://www.clir.org/pubs/reports/reports/hazen/pub74.html>

⁵ *Moving Theory into Practice, Digital Imaging Tutorial (Cornell University Library)*

<https://www.library.cornell.edu/preservation/tutorial/contents.html>

⁶ Guia para digitalização de documentos versão 2.0 (EMBRAPA)

<https://www.embrapa.br/documents/1355746/1441198/GuiaDigitalizacao.pdf/6394c2e2-be46-4c72-8e81-07247b5f2a38>

⁷ Diretrizes para planejamento de digitalização de livros raros e coleções especiais (IFLA)

http://www.ifla.org/files/assets/rare-books-and-manuscripts/rbms-guidelines/ifla_guidelines_for_planning_the_digitization_portuguese_translation.pdf

⁸ Manual de Digitalização de Acervos: textos, mapas e imagens fixas (SILVA)

<http://www.restaurabr.org/siterestaurabr/CICRAD2011/M8%20Aulas/Manual%20de%20digitalizacao%20de%20acervos.pdf>

(Conclusão)

OBJETIVO		MÉTODO	FONTES
C	categorizar os principais aspectos abordados na seleção.	Identificar o foco e a característica nas fontes. (Seção 4.2).	Baseado nos dados da pesquisa.
D	comparar os critérios.	Analisar as categorias encontradas nas fontes, e seus critérios, identificando semelhanças, diferenças, focos. (Seção 4.3).	
E	propor critérios de seleção, em categorias.	Propor categorias e critérios para elas. (Seção 4.4).	

Fonte: O autor (2015).

3.3 FONTES UTILIZADAS

Sabe-se que o ato de realizar uma pesquisa requer alguns cuidados para que não existam enganos, sejam em números, em dados ou em informações. Neste sentido, a validação de uma pesquisa está muito vinculada às referências utilizadas pelo pesquisador para embasar suas ideias e concepções.

A internet constituiu-se em uma facilitadora do ato da pesquisa, dando maior variabilidade para qualquer conteúdo, além da evidente facilidade para encontrar e utilizar diversos tipos de documentos. Entretanto, independentemente do tipo de suporte do documento que seja utilizado, o ideal é que a organização/instituição por trás deles seja reconhecida em sua área de atuação, seja nacional ou internacionalmente.

Para a validação desta pesquisa, procurou-se trazer fontes produzidas por instituições respeitadas em seus ramos de atuação e que elas estivessem em esferas distintas. Além disso, um pesquisador da área também foi utilizado como fonte.

Villaseñor Rodríguez (1998, p. 33, tradução nossa) afirma que fontes institucionais de informação: “São aquelas fontes que proporcionam informação sobre

uma instituição. Esta instituição se converte em objeto de interesse e é ela mesma quem proporciona a informação que se deseja”.

Desta forma, foram analisadas as recomendações de instituições/organizações respeitadas em suas áreas de atuação e foi realizado um comparativo acerca do que cada uma delas contribui para a digitalização de documentos. São elas: o Conselho de Biblioteca e Recursos de Informação (*Council on Library and Information Resources*), a Biblioteca da Universidade Cornell (*Cornell University Library*), a Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), a Federação Internacional de Associações e Instituições Bibliotecárias (*International Federation of Library Associations and Institutions/IFLA*) e o pesquisador Prof. Dr. Rubens Ribeiro Gonçalves da Silva (Universidade Federal da Bahia – UFBA). Todos os links de acesso destes documentos encontram-se na Metodologia, seção 3.

3.3.1 Council on Library and Information Resources (CLIR)

Formado no fim dos anos 1990, após a união de outras instituições, como o Conselho em Recursos de Biblioteca e a Comissão sobre Preservação e Acesso, o Conselho em Biblioteca e Recursos de Informação (*Council on Library and Information Resources – CLIR*) é:

[...] uma organização independente, sem fins lucrativos que promove estratégias para reforçar a investigação, o ensino e ambientes de aprendizagem em colaboração com as bibliotecas, instituições culturais, e as comunidades de ensino superior. (CLIR, [201-], *online*, tradução nossa).

A colaboração e as parcerias são os principais objetivos do CLIR:

A organização avança em sua missão através de iniciativas de projetos e parcerias, publicações, o programa DLF, e concessão de bolsas de estudo e oportunidades. Através do CLIR Connect, o CLIR proporciona um fórum para a discussão, intercâmbio e colaboração. (CLIR, [201-], *online*, tradução nossa).

O seu documento intitulado “Selecting Research Collections for Digitization” (Coleção de Pesquisas Seleccionadas para Digitalização), datado de 1998, trata do aspecto da seleção de documentos para digitalização, focando no processo de tomada de decisão, principalmente aquelas realizadas pelo gestor da unidade, como

é o caso do bibliotecário nas bibliotecas, e, desta forma, evocando a importância das políticas de seleção, que são tão necessárias em qualquer tipo de biblioteca.

3.3.2 *Cornell University Library*

A Biblioteca da Universidade Cornell (*Cornell University Library*) traz para este trabalho as considerações da área acadêmica, visto que ela faz parte de uma das universidades mais respeitadas do mundo. Segundo Campos (2015, *online*):

Fundada em 1865, a Cornell University possui 21.593 alunos e 1.628 professores, divididos em suas 14 escolas. Dentre ex-alunos e professores, 43 já foram agraciados com o prêmio Nobel. A universidade é também membro da Ivy League, o grupo das oito instituições de ensino superior mais prestigiadas dos Estados Unidos (além de Cornell, Brown, Columbia, Dartmouth, Harvard, Princeton, Yale e Universidade da Pensilvânia).

Sua visão é: “Cornell capacita a pesquisa e a aprendizagem da comunidade com profundo conhecimento, serviços inovadores e excelentes coleções reforçadas por parcerias estratégicas” (CORNELL UNIVERSITY LIBRARY, [201-], *online*, tradução nossa).

Já sua missão visa:

Cornell University Library promove uma cultura de inquérito amplo e apoia a missão da Universidade de descobrir, preservar e disseminar o conhecimento e expressão criativa. [...] E, ela serve como um partido neutro e confiável de suporte ao acesso à informação e comunicação científica. (CORNELL UNIVERSITY LIBRARY, [201-], *online*, tradução nossa).

Seu tutorial de digitalização, intitulado “*Moving Theory into Practice, Digital Imaging Tutorial*”, em Português, “Movendo Teoria em Prática, Tutorial de Imagem Digital”, é um dos instrumentos utilizados como uma das recomendações de digitalização de documentos. Posteriormente, veremos o que é indicado pela Biblioteca para a Seleção.

3.3.3 Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA)

Em seu site, a empresa afirma que: “Somos uma empresa de inovação tecnológica focada na geração de conhecimento e tecnologia para agropecuária brasileira”. (EMBRAPA [201-], *online*). Criada em 1973, a EMBRAPA contribui para esta pesquisa todo o *know-how* de uma empresa especializada em agropecuária e vinculada ao governo brasileiro, através do Ministério da Agricultura, Pecuária e Abastecimento (MAPA).

A sua missão é:

[...] desenvolver, em conjunto com nossos parceiros do Sistema Nacional de Pesquisa Agropecuária (SNPA), um modelo de agricultura e pecuária tropical genuinamente brasileiro, superando as barreiras que limitavam a produção de alimentos, fibras e energia no nosso País. (EMBRAPA, [201-], *online*).

Sua colaboração na investigação se dá através do seu guia de digitalização de documentos, cujo título é “Guia para digitalização de documentos versão 2.0”, do ano de 2006. Segundo a empresa, o guia surgiu pela necessidade de criar uma orientação acerca da digitalização de seus próprios documentos institucionais. Além disso, procurou “estimular e orientar as Unidades da Embrapa no processo de organização da informação, assim como integrá-las e torná-las parte de uma grande rede de fornecedores de informações digitais”. (EMBRAPA, 2006, *online*). Por fim, é enfatizado que o documento é de natureza básica, mas que pode ser considerado um instrumento de partida apropriado quando o assunto é digitalização.

3.3.4 *International Federation of Library Associations and Institutions (IFLA)*

A Federação Internacional de Associações e Instituições Bibliotecárias (*International Federation of Library Associations and Institutions – IFLA*) é:

[...] o principal organismo internacional que representa os interesses das bibliotecas e serviços de informação e seus usuários. É a voz global da biblioteca e dos profissionais da informação. (IFLA, 2015, *online*, tradução nossa).

Sua participação nesta pesquisa é por meio das suas “Diretrizes para planejamento de digitalização de livros raros e coleções especiais”, datado em janeiro de 2015. Dentre os três documentos apresentados até então, o oferecido pela IFLA é que abrange o maior recorte deste trabalho, ou seja, suas recomendações são direcionadas para os livros raros e as coleções especiais, tema central deste trabalho.

Sobre suas diretrizes, de acordo com a IFLA (2015, *online*):

Estas diretrizes são destinadas a quem está envolvido no planejamento de projeto de digitalização de materiais raros e especiais, o que inclui gestores de bibliotecas que lideram projetos, bibliotecários e pesquisadores que planejam e executam projetos e financiadores de organizações que consideram o apoio à digitalização.

Além disso, é ressaltado que:

Ao invés de concentrar-se em questões de tecnologia, métodos específicos de captura ou preservação digital, estas orientações se concentram no planejamento conceitual e na colaboração com potenciais utilizadores a fim de alcançar os resultados desejados e sustentáveis. (IFLA, 2015, *online*).

Em seu escopo, a organização afirma que as informações contidas no documento são abrangentes em alguns aspectos e básicas em outros.

3.3.5 Manual de Digitalização de Acervos – Silva (UFBA)

Além das quatro fontes institucionais anteriores, o documento de um pesquisador da área foi utilizado como fonte para as análises comparativas. Trata-se do “Manual de Digitalização de Acervos: textos, mapas e imagens fixas” (2005), do Prof. Dr. Rubens Ribeiro Gonçalves da Silva, da Universidade Federal da Bahia (UFBA). Silva é Professor Titular do Instituto de Ciência da Informação (ICI) e do Programa de Pós-Graduação em Ciência da Informação (PPGCI) da Universidade Federal da Bahia – UFBA. Sua tese de doutorado chama-se “Digitalização de acervos fotográficos públicos e seus reflexos institucionais e sociais: tecnologia e consciência no universo digital”, assunto pertinente a este trabalho, com obtenção em 2002.

No **ANEXO C** encontra-se o fluxograma proposto por Silva para apoiar o processo de decisão em seleção. “A Seleção para Digitalização: Matriz de Apoio ao Processo de Decisão”, de 1997, foi adaptado dos autores Chapman, Hazel, Horrel e

Merrill-Oldham, da Harvard University Library. Estes autores são os mesmos do documento do CLIR (*Selecting Research Collections for Digitization*), uma das fontes deste trabalho.

4 RESULTADOS

Esta seção analisará todas as variáveis apresentadas até então e fará a correlação entre todas elas.

4.1 CRITÉRIOS DE SELEÇÃO NAS FONTES

Após o Objetivo A ter sido apresentado nos capítulos anteriores, na forma de revisão de literatura, o Objetivo B procurou identificar os critérios de seleção no processo da digitalização. Os critérios de seleção foram coletados e organizados em quadros (APÊNDICES entre A e E). Estes quadros contêm critérios organizados em categorias, além de questões e ponderações apontadas por cada instituição/pesquisador em seus respectivos estudos:

- a) O quadro com as recomendações coletadas do *Council on Library and Information Resources* (CLIR), dos autores Dan Hazen, Jeffrey Horrell, Jan Merrill-Oldham, denominado “*Selecting Research Collections for Digitization*” (1997), está demonstrado no **APÊNDICE A**.
- b) O quadro com as recomendações coletadas da Cornell University Library, em “*Moving Theory into Practice, Digital Imaging Tutorial*” (2003), está representado no **APÊNDICE B**.
- c) O quadro com as recomendações coletadas da EMBRAPA, no “Guia para digitalização de documentos versão 2.0” (2006), está no **APÊNDICE C**.
- d) O quadro com as recomendações coletadas da IFLA, nas “Diretrizes para planejamento de digitalização de livros raros e coleções especiais” (2015), estão no **APÊNDICE D**.
- e) O quadro com as recomendações coletadas do pesquisador Prof. Dr. Rubens Ribeiro Gonçalves da Silva (UFBA), no “Manual de Digitalização de Acervos: textos, mapas e imagens fixas” (2005), aparecem no **APÊNDICE E**.

Com base nas informações coletadas (APÊNDICES A à E), os critérios de seleção propostos pelas fontes foram levantados, conforme a ordem de aparecimento do critério nos seus respectivos textos. O **QUADRO 5** apresenta os critérios de cada fonte.

QUADRO 5 – Critérios propostos pelas fontes

CLIR	CORNELL UNIVERSITY LIBRARY	EMBRAPA	IFLA	SILVA
Direitos do autor (Copyright)	Restrições legais (Direitos autorais)	Características do material original	Estados a serem considerados (Cronologia, local de publicação, autor, assunto, formato, possuidor)	Constituir uma comissão de seleção
Natureza intelectual dos materiais originais	Atributo do documento	Propriedades físicas dos materiais	Questões essenciais (Obras raras, coleção existente, coleção digital)	Cautela e experiência
Usuários atuais e potenciais	Considerações de preservação	Objetivos da digitalização	Revisão física do material	Motivações para a seleção (Motivação dirigida à coleção original; Motivação dirigida à versão digital)
Natureza do uso real e antecipado	Organização e documentação disponível	Direitos autorais	Direitos de autor	Valores: informacional, evidencial, intrínseco, administrativo, associacional e monetário.
O formato e natureza do produto digital	Utilizações previstas		Coleção digital	Risco
Descrevendo, entregando e mantendo produtos digitais	Aumento da coleção digitalizada			Uso
Relacionamentos com outros esforços digitais	Duplicação de esforços			
Custos e benefícios	Capacidade institucional			
	Recursos financeiros			

Fonte: O autor (2015).

4.2 FOCO E CARACTERÍSTICAS DAS FONTES

No texto a seguir, cada fonte é analisada considerando seu foco e suas características, atendendo ao Objetivo C:

Os critérios apontados pelo **CLIR**, na primeira coluna, no **QUADRO 5**, evidenciam uma caracterização bem ampla acerca da digitalização de documentos, pensando em todos os processos. A ideia deste trabalho é orientar a tomada de decisão pelos gestores de bibliotecas, principalmente o posicionamento do bibliotecário em projetos de digitalização.

Este documento apresenta os critérios na forma de questões e procura responder estes questionamentos. Estas perguntas e algumas reflexões do texto podem ser visualizadas no **APÊNDICE A**.

O conteúdo deste documento tem como preocupação relevante a questão das coleções digitais, inquietação esta que estava em voga perto da virada do ano 2000, tendo em vista a popularização do computador e da internet e o contexto temporal desta publicação.

Os autores Hazen, Horrell e Merrill-Oldham destacam ainda que há uma diferença entre a seleção de documentos para compra em uma biblioteca, por exemplo, e a seleção visando a digitalização. Embora muitos dos processos em ambas sejam parecidos, ou até mesmo idênticos em alguns casos, é preciso pensar em cada caso isoladamente, verificando esta série de critérios posteriormente.

A **Cornell University Library** traz um conteúdo abrangente sobre a digitalização de documentos, assim como o CLIR. Entretanto, este conteúdo é mais enxuto e generalista. As restrições legais (direitos autorais ou *copyright*) possuem um tópico específico dentre os critérios apontados pelo autor, colocando este aspecto como o primeiro (e talvez mais importante) passo a ser considerado quando uma obra é selecionada para digitalização.

Obviamente, os direitos autorais devem ser respeitados e verificados por qualquer instituição que queira utilizar obras de autores e instituições que estejam fora do seu círculo autoral, mas acredita-se que este tipo de cuidado deva estar muito latente em uma biblioteca universitária, como é o caso desta instituição. O meio

acadêmico valoriza muito a citação de fontes e o reconhecimento de autoria. Portanto, é razoável acreditar que o foco deste documento esteja levando isso em conta.

Além destes critérios, todos os demonstrados no **QUADRO 5** são abordados em “*Moving Theory into Practice, Digital Imaging Tutorial*” de uma forma sucinta.

A proposta da **EMBRAPA** volta-se muito para a própria empresa, fazendo uso deste material como um manual de procedimentos interno. Porém, segundo a empresa, ele pode servir como uma guia básico para outras organizações que requeiram orientações básicas para iniciar um projeto de digitalização próprio. Como o próprio título do texto sugere, ele serve como um guia para orientar as práticas de digitalização, sendo genérico por um lado, mas específico por outro, principalmente em questões técnicas como a escolha dos *hardwares* a serem empregados.

A **IFLA** apresenta um instrumento abrangente e básico ao mesmo tempo. Baseado em três questões essenciais no item Seleção, esta publicação demonstra diretrizes bem definidas: elas estão voltadas para a raridade da obra, para os documentos impressos existentes e para os repositórios digitais.

O conteúdo deste documento, dentre as cinco fontes utilizadas neste trabalho, é o que menos desmembra seus tópicos para discussão, provavelmente não sendo tão adequado a leitores leigos no assunto, pois pode acabar sendo difícil de se ter um maior entendimento sem alguns conhecimentos prévios do assunto.

Por fim, o **pesquisador Rubens Ribeiro Gonçalves da Silva** traz um estudo pensando em questões técnicas e éticas acerca da digitalização de documentos como textos, mapas e imagens fixas.

O trabalho de Silva diferencia-se dos demais quando manifesta que a criação de uma comissão de seleção é uma das fases mais importantes no planejamento da seleção de um documento. O trabalho desta comissão é realizado por profissionais de diferentes áreas, desde profissionais especializados em digitalização, passando pelos bibliotecários e indo até mesmo aos advogados. Cada profissional, dentro de sua área de atuação, pode contribuir para a digitalização de um documento. Profissionais especializados podem trazer todo o seu conhecimento técnico para um determinado projeto, já os bibliotecários podem contribuir na organização desta

informação e, finalmente, os advogados podem lidar com questões jurídicas, como os direitos autorais de uma obra.

O autor também traz a valorização que uma obra pode ter se for digitalizada, baseada em aspectos como valor institucional (conteúdo da obra para a instituição), valor evidencial (valor legal ou histórico da obra), intrínseco (estado físico da obra), administrativo (uso da obra), associacional (conteúdo relevante da obra) e monetário (valor financeiro da obra).

Em suma, o **QUADRO 6**, após explanação acima, ratifica e simplifica a visualização dos focos nas fontes:

QUADRO 6 – Resumo dos focos nas fontes

FONTE	FOCO
CLIR	Foco na tomada de decisão e na biblioteca.
Cornell University Library	Abrangente, valoriza direito autoral.
EMBRAPA	Abrangente, possui foco na empresa e é específico em questões de <i>hardware</i> .
IFLA	Foco nas obras raras, sendo voltado para conhecedores da área, tendo em vista a organização das informações no texto.
SILVA	Foco no planejamento e na gestão, principalmente na comissão de seleção, com destaque para os valores possíveis de uma obra rara.

Fonte: O autor (2015).

4.3 CATEGORIAS PARA CRITÉRIOS DE SELEÇÃO

O **QUADRO 7**, com base em uma categorização proposta pelo autor, objetiva observar semelhanças, diferenças e focos nas fontes utilizadas, solucionando o Objetivo D.

QUADRO 7 – Análise comparativa sobre seleção

CATEGORIA	CLIR	CORNELL UNIVERSITY LIBRARY	EMBRAPA	IFLA	SILVA
Direito autoral	✓ Direitos do autor (Copyright)	✓ Restrições legais (Direitos autorais)	✓ Direitos autorais	✓ Direitos de autor	✗
Valor intelectual	✓ Natureza intelectual dos materiais originais	✗	✓ Características do material original	✓ Estados a serem considerados (Cronologia, local de publicação, autor, assunto, formato, possuidor) ✓ Questões essenciais (obras raras)	✓ Risco ✓ (Motivação dirigida à coleção original) ✓ Valor informacional, intrínseco e associacional
Aspecto físico	✗	✓ Atributo do documento	✓ Propriedades físicas dos materiais	✓ Revisão física do material	✓ Cautela e experiência ✓ Valor evidencial

(Continua)

(Conclusão)

CATEGORIA	CLIR	CORNELL UNIVERSITY LIBRARY	EMBRAPA	IFLA	SILVA
Uso	✓ Natureza do uso real e antecipado	✓ Utilizações previstas	✓ Objetivos da digitalização	✓ Questões essenciais (coleção existente)	✓ Motivações para a seleção ✓ Uso
Representação digital	✓ O formato e natureza do produto digital ✓ Descrevendo, entregando e mantendo produtos digitais ✓ Relacionamentos com outros esforços digitais	✓ Considerações de preservação ✓ Organização e documentação disponível ✓ Aumento da coleção digitalizada	✗	✓ Coleção digital ✓ Questões essenciais (coleção virtual)	✓ Uso ✓ (Motivação dirigida à versão digital)
Recursos (Pessoal, Financeiro e Colaboração)	✓ Custos e benefícios	✓ Recursos financeiros	✗	✗	✓ Valores administrativo e monetário
Capacidades para o empreendimento (Gestão e Capacitação)	✗	✓ Duplicação de esforços ✓ Capacidade institucional	✗	✗	✓ Constituir uma comissão de seleção

Fonte: O autor (2015).

A verificação dos critérios propostos foi realizada através da correspondência entre os critérios de seleção nas fontes e as categorias sugeridas pelo autor. Este quadro foi realizado conforme categorização aplicada pelo autor observando semelhanças nas suas respectivas características nas fontes e nos contextos apresentados. Portanto, as categorias apresentadas são consideradas observando o entendimento do autor sobre a possível reunião e a importância delas. O que está com “✓” corresponde a um item correspondente na fonte. O que possui o “✗” não está contemplado no texto.

4.3.1 Categorias Escolhidas

A escolha das categorias consideradas importantes foi realizada pela proposição de uma categorização que abrangesse os critérios das fontes. Então, na seção 4.3.2, justificativas para o uso de cada categoria são desenvolvidas, por meio da definição da categoria e da sua importância.

As categorias definidas por esta análise foram:

- a) direito autoral;
- b) originais (valor intelectual);
- c) originais (aspecto físico);
- d) representação digital;
- e) uso digital;
- f) recursos (pessoal, financeiro e colaboração);
- g) capacidades para o empreendimento (gestão e capacitação).

A escolha destas sete categorias foi possível após a leitura nas principais fontes deste trabalho, seguida da criação dos apêndices e dos **QUADROS 5 e 7**, visto que ambos ofereceram recursos para legitimar a importância da preferência destes itens como aspectos relevantes para a seleção de obras raras para digitalização.

4.3.2 Sobre as Categorias

Abaixo veremos uma breve explicação sobre o significado das categorias, no âmbito deste estudo, no processo de seleção e o motivo para escolha como critério relevante:

a) **Direito autoral**

O que é? O direito autoral (conhecido também por *copyright* ou direitos do autor), em linhas gerais, é a propriedade intelectual e/ou jurídica de uma determinada obra rara, no qual seus direitos de uso e reprodução podem variar em diferentes contextos, podendo ser digital, físico ou em ambos.

Por que é importante? Existem diversas leis, nacionais e internacionais, que asseguram os direitos dos autores. É preciso estar atento para os tipos de reprodução que os autores autorizam efetuar, pois elas variam de país para país e conforme a forma de disponibilização (parcial ou total).

b) **Valor intelectual**

O que é? Os originais, neste caso, são as obras raras em sua forma natural, no qual é verificado o seu valor intelectual, sua importância para a literatura da área do conhecimento e do autor no mercado editorial e histórico.

Por que é importante? Livros, para serem considerados obras raras, precisam passar por uma série de avaliações antes de serem classificados como tal. Analisar o renome do autor, sua tiragem de exemplares, seu valor histórico-cultural, por exemplo, são alguns itens importantes a serem averiguados antes de atestar que se trata ou não de uma obra rara. Além disso, a importância de uma obra institucional para alguma instituição pode, em alguns casos, só ser validada pela própria instituição.

c) **Aspecto físico**

O que é? Os originais (aspecto físico), neste contexto, referem-se ao estado físico de uma obra rara. Alguns itens a serem analisados são: agentes biológicos e

químicos atuantes no livro, rabiscos que ele tenha, se há folhas caindo ou não, dentre outros.

Por que é importante? O estado físico de uma obra rara é fator determinante para a digitalização ou não da mesma. Se há a possibilidade de realização do processo, é preciso ver qual aparelho e forma de digitalização agridem menos, a fim de garantir a integridade física do material.

d) **Uso**

O que é? No cenário desta pesquisa, uso está voltado às possibilidades, motivações e benefícios que os documentos digitalizados podem oferecer na pré (livro físico) ou pós (livro digital) digitalização.

Por que é importante? Saber a importância e a forma como uma obra rara é usada antes da sua digitalização pode facilitar este processo e a sua posterior disponibilização digital.

e) **Representação digital**

O que é? Neste âmbito, representação digital é o formato digital de um documento originalmente não digital. A representação digital seria focada principalmente nos repositórios digitais, no qual as obras raras seriam difundidas na Web através das suas derivadas de acesso, como uma das formas de preservação digital.

Por que é importante? Porque nos garante a fidedignidade com o original e permite gerar cópias com as características necessárias tanto para o acesso do usuário, como para a preservação em longo prazo.

f) **Recursos (Pessoal, Financeiro e Colaboração)**

O que é? Os Recursos aqui foram classificados em três esferas: Pessoal (funcionários), Financeiro (monetário) e Colaboração (cooperação dentro da instituição, seja de pessoas ou de patrocínios financeiros).

Por que é importante? Para que um projeto de digitalização saia do papel é vital que existam pessoas engajadas nisso e capital financeiro para a sua execução.

g) Capacidades para o empreendimento (Gestão e Capacitação)

O que é? Capacidades para o empreendimento compreende a capacitação dos funcionários para lidar com a seleção dos documentos e a gestão depende da habilitação profissional dos mesmos.

Por que é importante? Cada vez mais é exigido pelas empresas e instituições profissionais especializados e que saibam empreender e gerir projetos.

Evidentemente, cada autor, em seus respectivos trabalhos, dá uma ênfase maior para um determinado ponto e outros consideram outros aspectos mais fundamentais. A título de exemplo, o CLIR não toca no que tange aos aspectos físicos do material, enquanto isso é citado em todas as outras fontes. Outra exemplificação possível seria o caso da comissão de seleção, enfatizado por Silva, e lembrado apenas por mais uma fonte dentre as outras quatro.

Independentemente dos focos utilizados, é sempre preciso estar ciente de tudo que compõe uma seleção para obras raras, para que a instituição interessada neste procedimento tenha suas necessidades atendidas plenamente.

Para que um planejamento seja bem executado, profissionais de diferentes áreas devem contribuir com suas experiências e uma política de seleção deve ser criada visando estas obras especiais. Estes itens são indispensáveis para que um projeto de digitalização de obras raras obtenha o sucesso desejado.

4.4 MODELO DE APOIO AO PROCESSO DE SELEÇÃO

Esta seção, com base na comparação e categorização de critérios apresentadas anteriormente, propõe critérios e questionamentos para a seleção de obras raras para digitalização, na forma de um quadro, solucionando assim, o Objetivo E deste trabalho.

O modelo a seguir (**QUADRO 8**) apresenta os sete critérios evidenciados na seção anterior como os mais relevantes para a seleção de obras raras visando a digitalização: Direito autoral, Originais (Valor Intelectual), Originais (Aspecto Físico), Uso, Representante digital, Recursos (Pessoal, Financeiro e Colaboração) e Capacidades para o empreendimento (Gestão e Capacitação).

O modelo proposto pelo autor tem como objetivo servir como base para as ações do Centro de Documentação e Acervo Digital da Pesquisa (CEDAP) da UFRGS, bem como servir a qualquer instituição ou indivíduo que o considere útil.

A escolha dos critérios foi sustentada em leituras na literatura da área, conhecimentos empíricos e adquiridos ao longo da pesquisa e está fundamentalmente embasada nas principais fontes deste trabalho: CLIR, Cornell University Library, EMBRAPA, IFLA e Silva.

QUADRO 8 – Modelo de apoio ao processo de seleção

CRITÉRIO	DESCRIÇÃO	QUESTÕES A SEREM OBSERVADAS
Direito autoral	O direito autoral (conhecido também por copyright ou direitos do autor), em linhas gerais, é a propriedade intelectual e/ou jurídica de uma determinada obra rara, no qual seus direitos de uso e reprodução podem variar em diferentes contextos, podendo ser digital, físico ou em ambos.	<p>A obra possui direitos autorais reservados? S () N () P* ()</p> <p>O autor/detentor dos direitos da obra autoriza algum tipo de reprodução? S () N () P* ()</p> <p>* Os direitos de reprodução de uma obra podem ser totais, parciais ou restritos.</p>
Valor intelectual	O valor intelectual, neste caso, são as obras raras em sua forma natural, no qual é verificado o valor intelectual dela, sua importância para a literatura da sua área do conhecimento e do autor no mercado editorial e histórico.	<p>O conteúdo da obra é relevante para os objetivos da instituição? S () N ()</p> <p>O conteúdo da obra está íntegro (legível, sem danos nas informações contidas na obra)? S () N ()</p>
Aspecto físico	O aspecto físico, neste contexto, refere-se ao estado físico de uma obra rara. Alguns itens a serem analisados são: agentes biológicos e químicos atuantes no livro, rabiscos que ele possa vir a ter, se há folhas caindo ou não, dentre outros.	<p>A obra está em bom estado físico? S () N ()</p> <p>A obra possui poucos exemplares disponíveis? S () N ()</p> <p>As folhas que compõem a obra são sensíveis à luz do scanner? S () N ()</p>

(Continua)

(Conclusão)

Uso	No cenário desta pesquisa, uso está voltado às possibilidades, motivações e benefícios que os documentos digitalizados podem oferecer na pré (livro físico) ou pós (livro digital) digitalização.	Existe público para utilizar da obra após a sua digitalização? S () N () Pensando no uso que os usuários farão da obra, o investimento financeiro e de pessoal no projeto realizado pela instituição possui um bom custo-benefício? S () N ()
Representação digital	Neste âmbito, representante digital é a representação em formato digital de um documento originalmente não digital. O representante digital seria focado principalmente nos repositórios digitais, no qual as obras raras seriam difundidas na Web.	O representante digital perdeu poucas características da obra física original? S () N () Visando os repositórios digitais e a preservação digital, a obra permite visualização em alta qualidade? S () N ()
Recursos (Pessoal, Financeiro e Colaboração)	Os Recursos aqui foram classificados em três esferas: Pessoal (funcionários), Financeiro (monetário) e Colaboração (cooperação dentro da instituição, seja de pessoas ou de patrocínios financeiros).	Existem recursos financeiros para a execução do projeto de digitalização? S () N () Existe engajamento da instituição no projeto de digitalização? S () N ()
Capacidades para o empreendimento (Gestão e Capacitação)	Capacidades para o empreendimento compreende a capacitação dos funcionários para lidar com a seleção dos documentos e a gestão depende da habilitação profissional dos mesmos.	Existem funcionários habilitados para executar estas tarefas? S () N () Existe uma política de seleção de obras raras para digitalização na instituição? S () N ()

Marcar a opção adequada com um "X" nos parênteses acima. Mais respostas **S** indicam que a obra possui melhores condições de ser digitalizada.

Legenda:

S: Sim.

N: Não.

P: Parcial.

Fonte: O autor (2015).

Na primeira coluna, estão os critérios selecionados. Na segunda, estão as descrições acerca do que cada critério trata, baseadas no “O que é?” da seção 4.3. Na terceira coluna, encontram-se algumas perguntas a serem respondidas para julgar se uma obra deve ser selecionada ou não pela instituição para digitalização.

Quando o assunto é seleção, diversas outras perguntas seriam pertinentes para avaliar a necessidade e a possibilidade da digitalização de uma obra rara. Entretanto, as perguntas sugeridas acima foram consideradas pelo autor como as mais significativas quando se trata do assunto.

Além do quadro proposto pelo autor (**QUADRO 8**), outros modelos pertinentes de serem analisados acerca da seleção estão nos **ANEXOS C e D**, respectivamente: No “Quadro de Apoio ao Processo de Seleção para Digitalização”, que demonstra um modelo de apoio ao processo de seleção. E em “A Seleção para Digitalização: Matriz de Apoio ao Processo de Decisão”, que apresenta um fluxograma que esquematiza as perguntas-chave no processo da tomada de decisões.

5 CONSIDERAÇÕES FINAIS

O avanço da tecnologia impacta a vida das pessoas mundo a fora de diferentes formas. Se para alguns isso é algo positivo, para outros, mais céticos, ela é prejudicial. Discussões à parte, o certo é que a tecnologia, com suas invenções úteis ou até mesmo mirabolantes, está, cada vez mais, adentrando as casas, os trabalhos, as áreas de lazer da população mundial. Seguindo este caminho, está a necessidade latente de preservar documentos digitalmente. Se há duas décadas, com a popularização da figura do computador e da Internet, isso era uma pauta possível, mas ainda um tanto desconhecida, hoje, isso é algo frequente e até mesmo óbvio. Existem diversas ferramentas disponíveis para a digitalização de documentos, indo desde *scanners* altamente especializados até os domésticos, que qualquer usuário com o mínimo de conhecimento pode operá-lo.

Este trabalho procurou demonstrar as possibilidades do uso da digitalização na Preservação Digital de documentos, sobretudo no aspecto dos critérios referentes à Seleção. Como vimos no seu decurso, para a efetiva preservação digital de um documento, especialmente uma obra considerada rara, existem uma série de processos necessários na digitalização destes materiais, desde a escolha do que é importante de ser digitalizado (Seleção), na configuração técnica das máquinas necessárias (Conversão), na construção de algoritmos de recuperação da informação (Metadados), na criação de uma rotina de trabalho (Fluxo de Digitalização), até a disponibilização destas obras em repositórios digitais, com a preocupação de ter os arquivos originais sempre em segurança, além da constante avaliação da qualidade pela qual o processo todo está sendo submetido.

A seleção, especialmente de obras raras, é a primeira e, talvez, a etapa mais decisiva de todo o processo de Digitalização, pois em muitos casos estas obras são antigas, geralmente são únicas e não podem ser danificadas de nenhuma forma. Então avaliar bem qual obra deve passar por um processo “estressante”, como é o da digitalização, e que pode trazer riscos para a obra original, é muito delicado e minucioso.

A escolha das fontes para comparação de critérios de seleção foi realizada de forma equilibrada, com diferentes tipos de instituição, procurando observar as semelhanças e as diferenças entre os documentos sobre esta temática. O CLIR, a Cornell University Library, a EMBRAPA, a IFLA e o pesquisador da UFBA, Rubens R.

G. da Silva, foram essenciais para a construção deste trabalho, oferecendo auxílio para o entendimento do autor sobre o tema e demonstrando que é sempre necessário olhar qualquer conteúdo de diferentes perspectivas. Cada fonte, dentro da sua proposta e do seu foco de público, trouxe uma importante contribuição para o presente trabalho.

A melhor forma de selecionar obras raras visando a digitalização e a sua posterior preservação digital é observar as características e necessidades da instituição para qual este processo será dedicado e que é possível valer-se da utilização de instrumentos de outras instituições como apoio.

Os critérios apresentados poderiam valer para qualquer obra bibliográfica. Desde as obras consideradas “comuns”, até as obras raras, pois os detalhes e cuidados a serem tomados são os mesmos, possuindo uma diferenciação apenas no grau de exigência, sendo ele maior para as obras raras.

É importante salientar que a seleção para digitalização, de obras raras ou não, possui semelhanças com a seleção de obras para o acervo de uma biblioteca, mas que são processos distintos e com suas peculiaridades e devem ser realizados separadamente.

Os critérios de seleção apresentados nas fontes são muito semelhantes entre si. Portanto, há um certo consenso sobre o que é importante levar em conta quando trata-se desse tipo de seleção. Questões como direitos autorais parecem ser universais, enquanto a criação de uma comissão de seleção é algo facultativo. Os demais critérios se mesclam e são abordados de diferentes formas.

Evidentemente, cabem ainda outras pesquisas acerca da seleção de obras raras, além desta e das já existentes sobre o assunto, afinal obras raras possuem um valor inestimável. Para preservar este valor, elas requerem um cuidado totalmente especializado, visando não só a manutenção do seu estado físico atual, mas também a preservação do seu conteúdo intelectual.

REFERÊNCIAS

BIBLIOTECA NACIONAL (Brasil). **Obras raras**. Disponível em: <<https://www.bn.br/acervo/obras-raras>>. Acesso em: 15 set. 2015.

CAMPOS, Carolina. **Universidade Cornell**: pouco famosa no Brasil, muito respeitada no mundo. Disponível em: <<http://www.estudarfora.org.br/conheca-a-universidade-cornell-pouco-famosa-no-brasil-muito-respeitada-no-mundo/>>. Acesso em: 18 set. 2015.

CONSELHO NACIONAL DE ARQUIVOS – CONARQ. **Recomendações para digitalização de documentos arquivísticos permanentes**. Disponível em: <http://www.conarq.arquivonacional.gov.br/media/publicacoes/recomenda/recomendaes_para_digitalizao.pdf>. Acesso em: 11 jun. 2015.

CORNELL UNIVERSITY LIBRARY. **Vision and mission**. Disponível em: <<https://www.library.cornell.edu/about/inside/mission>>. Acesso em: 18 set. 2015.

_____. **Moving theory into practice, Digital imaging tutorial**. Disponível em: <<https://www.library.cornell.edu/preservation/tutorial/intro/intro-01.html>>. Acesso em: 03 set. 2015.

COUNCIL ON LIBRARY AND INFORMATION RESOURCES. **About Us**. Disponível em: <<http://www.clir.org/about>>. Acesso em: 05 out. 2015.

_____. **Selecting research collections for digitization-Full report**: selecting research collections for digitization. Disponível em: <<http://www.clir.org/pubs/reports/reports/hazen/pub74.html>>. Acesso em: 05 out. 2015.

DISCOVERY NEWS. **Digital image founder smooths out pixels**. Disponível em: <<http://news.discovery.com/tech/apps/digital-image-pixel.htm>>. Acesso em: 30 set. 2015.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA – EMBRAPA. **Guia para digitalização de documentos versão 2.0**. Disponível em: <<https://www.embrapa.br/documents/1355746/1441198/GuiaDigitalizacao.pdf/6394c2e2-be46-4c72-8e81-07247b5f2a38>>. Acesso em: 18 jun. 2015.

_____. **Quem somos.** Disponível em: <<https://www.embrapa.br/quem-somos>>. Acesso em: 18 set. 2015.

LIBRARY OF CONGRESS (Estados Unidos). **Mets.** Disponível em: <<http://www.loc.gov/standards/mets/>>. Acesso em 23 jun. 2015.

_____. **Premis.** Disponível em: <<http://www.loc.gov/standards/premis/>>. Acesso em 23 jun. 2015.

FERREIRA, Miguel. **Introdução à preservação digital: conceitos, estratégias e actuais consensos.** Guimarães: Escola de Engenharia da Universidade do Minho, 2006. Disponível em: <<http://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf>>. Acesso em: 06 maio 2015.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa.** 4. ed. São Paulo: Atlas, 2002.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS – IFLA. **About.** Disponível em: <<http://www.ifla.org/about>>. Acesso em: 18 set. 2015.

_____. **Diretrizes para planejamento de digitalização de livros raros e coleções especiais.** Disponível em: <http://www.ifla.org/files/assets/rare-books-and-manuscripts/rbms-guidelines/ifla_guidelines_for_planning_the_digitization_portuguese_translation.pdf>. Acesso em: 18 set. 2015.

MÁRDERO ARELLANO, Miguel Angel. Preservação de documentos digitais, **Ciência da Informação**, Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004. Disponível em: <<http://www.scielo.br/pdf/ci/v33n2/a02v33n2.pdf>>. Acesso em: 11 maio 2015.

RECONHECIMENTO ÓTICO DE CARACTERES. *In*: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2015. Disponível em: <https://pt.wikipedia.org/w/index.php?title=Reconhecimento_%C3%B3tico_de_caracteres&oldid=42389773>. Acesso em: 23 set. 2015.

RODRIGUES, Márcia Carvalho. Como definir e identificar obras raras? Critérios adotados pela Biblioteca Central da Universidade de Caxias do Sul, **Ciência da Informação**, Brasília, v. 35, n. 1, p. 115-121, jan./abr. 2006. Disponível em: <<http://www.scielo.br/pdf/ci/v35n1/v35n1a12.pdf>>. Acesso em: 15 set. 2015.

SANTOS, Júlia Costa dos. **Mesa Memória, documentação e acervo mostra espaços que visam à preservação da memória.** *In*: ENCONTRO NACIONAL DE HISTÓRIA DA MÍDIA (ALCAR), 10., 2015, Porto Alegre. Disponível em: <<http://www.ufrgs.br/alcar2015/news/mesa-memoria-documentacao-e-acervo-mostra-espacos-que-visam-a-preservacao-da-memoria>>. Acesso em: 11 jun. 2015.

SETZER, Valdemar Waingort. Dado, informação, conhecimento e competência, **DataGramaZero - Revista de Ciência da Informação**, Rio de Janeiro, n. zero, dez. 1999. Disponível em: <http://www.dgz.org.br/dez99/Art_01.htm>. Acesso em: 18 jun. 2015.

SILVA, Edna Lúcia da; MENEZES, Estera Muszkat. **Metodologia da pesquisa e elaboração de dissertação.** 3. ed. rev. atual. Florianópolis: Laboratório de Ensino a Distância da UFSC, 2001. 121 p. Disponível em: <<http://cursos.unipampa.edu.br/cursos/ppgcb/files/2011/03/Metodologia-da-Pesquisa-3a-edicao.pdf>>. Acesso em: 30 jun. 2015.

SILVA, Rubens Ribeiro Gonçalves da. **Manual de digitalização de acervos:** textos, mapas e imagens fixas. Salvador: EDUFBA, 2005. Disponível em: <<http://www.restaurabr.org/siterestaurabr/CICRAD2011/M8%20Aulas/Manual%20de%20digitalizacao%20de%20acervos.pdf>>. Acesso em: 20 out. 2015.

TASK FORCE ON ARCHIVING OF DIGITAL INFORMATION, COMMISSION ON PRESERVATION AND ACCESS AND RESEARCH LIBRARIES GROUP. **Preserving digital information:** report of the Task Force on Archiving of Digital Information. Washington, D.C.: Commission on Preservation and Access, 1996. Disponível em: <<http://www.clir.org/pubs/reports/pub63watersgarrett.pdf>>. Acesso em: 19 set. 2015.

VILLASEÑOR RODRÍGUEZ, Isabel. Los instrumentos para la recuperación de la información: las fuentes. *In*: TORRES RAMIREZ, Isabel de. **Las fuentes de información:** estudios teórico-prácticos. Madrid: Síntesis, 1998. Cap. 2, p. 42.

APÊNDICE A – *Selecting Research Collections for Digitization (CLIR)*

<p>Direitos do autor (Copyright)</p>	<p>“[...] se um projeto de digitalização proposto envolve materiais no domínio público, a obra pode prosseguir. Se as matérias-primas estão protegidas por copyright, mas os direitos são detidos pela instituição ou permissões adequadas pode ser assegurada, a obra pode avançar. Se as permissões não são próximas de fontes de direitos autorais, no entanto, os materiais não podem ser reproduzidos e o foco do projeto deve mudar” (CLIR, 1998, <i>online</i>).</p> <p>“Os direitos de desempenho associados com partituras musicais, por exemplo, ou direitos de exibição de filmes, diferem de direitos para os materiais de inadimplência, tais como revistas eletrônicas ou fotografias documentais”. (CLIR, 1998, <i>online</i>).</p>
<p>Natureza intelectual dos materiais originais</p>	<ul style="list-style-type: none"> a) Será que a qualidade intelectual do material de origem garante o nível de acesso tornada possível através da digitalização? b) Será que a digitalização vai aumentar o valor intelectual do material? c) Será que o acesso eletrônico a um conjunto de informações vai aumentar significativamente o seu potencial para esclarecer, ou são os originais de livros, manuscritos, fotografias ou pinturas suficientes para essa tarefa? d) Até que ponto vai a combinação ou agregação de fontes originais aumentar o seu valor?
<p>Usuários atuais e potenciais</p>	<ul style="list-style-type: none"> a) São os estudiosos que consultarão os materiais das bases propostas? Os materiais estão sendo utilizados tanto quanto eles poderiam estar? b) O acesso atual para os materiais propostos é tão difícil que a digitalização irá criar um novo público? c) Será que a condição física dos materiais originais irá limitar a sua utilização? d) São materiais relacionados tão dispersos que eles não podem ser estudados em contexto? e) Será que os arquivos digitais propostos são de tamanho e formato gerenciável? f) Será que a digitalização atende as necessidades de estudantes e estudiosos locais?

(Continua)

(Continuação)

<p>Natureza do uso real e antecipada</p>	<p>a) Como os estudiosos irão usar os materiais originais existentes? Que abordagem à digitalização vai facilitar o seu trabalho?</p> <p>b) Será a digitalização vai aumentar a utilidade dos materiais originais? Será que vai permitir novos tipos de ensino ou investigação? Não estudiosos concordam que o produto proposto será útil?</p> <p>c) Existem outros acadêmicos, bibliotecários, arquivistas que possam colaborar para criar um produto útil?</p>
<p>O formato e natureza do produto Digital</p>	<p>a) O que é crítico nas características do material de origem que devem ser capturados no produto digital? São cópias de alta resolução, uma apresentação exata das cores, uma combinação perfeita de imagens e texto, ou outras qualidades consideradas essenciais?</p> <p>b) As fontes originais devem ser retidas, elas podem suportar o processo de digitalização?</p> <p>c) O tipo de equipamento deveria ser utilizado para a conversão?</p> <p>d) Será que uma amostra digitalizada atende as necessidades dos usuários? Se sim, como deve a amostra ser construída?</p> <p>e) Será que os recursos de informação sobre o qual o projeto se baseia continuam a crescer?</p> <p>f) Como os usuários irão navegar dentro e entre as coleções digitais?</p>
<p>Descrevendo, entregando e mantendo produtos digitais</p>	<p>a) Como os usuários sabem que o arquivo digital existe?</p> <p>b) Como o produto digital pode ser melhor entregue aos usuários?</p> <p>c) Quem será autorizado a usar o recurso digital e em que circunstâncias?</p> <p>d) Como será assegurada a integridade dos dados digitalizados?</p> <p>e) Para os produtos digitais criados para atender à demanda local, a infraestrutura de tecnologia existente é adequada?</p> <p>f) Quais são as intenções de longo prazo para o arquivo digital?</p> <p>g) É a preservação a longo prazo de materiais deteriorados um objetivo do projeto?</p>

(Conclusão)

<p>Relacionamentos com outros esforços digitais</p>	<p>a) Os materiais propostos para a digitalização já foram convertidos para o formato eletrônico?</p> <p>b) Pode esforços de digitalização de cooperação reunir um corpo coeso de material que caso contrário ficaria desassociado?</p>
<p>Custos e Benefícios</p>	<p>a) Quem se beneficiará com o produto digital proposto?</p> <p>b) É o valor intelectual do produto proposto compatível com a despesa?</p> <p>c) Poderia um produto aceitável ser criado a um custo menor?</p> <p>d) Como o projeto proposto abordará os custos de longo prazo associados aos arquivos digitais?</p> <p>e) Financiamento externo pode ser assegurado para apoiar o projeto proposto?</p>

Fonte: Adaptado de CLIR (1998, tradução nossa).

APÊNDICE B – *Moving Theory into Practice, Digital Imaging Tutorial* (Cornell University Library)

<p align="center">Restrições Legais (Direitos Autorais)</p>	<p>a) Comece o seu processo de seleção, considerando as restrições legais. O material é restrito por causa de privacidade, conteúdo ou de preocupações dos doadores?</p> <p>b) É copyright protegido?</p> <p>c) Se assim for, você tem o direito de criar e divulgar reproduções digitais?</p>
<p align="center">Atributo do documento</p>	<p>a) Será que o material serve para a digitalização?</p> <p>b) Pode o conteúdo informacional ser devidamente colocado em formato digital?</p> <p>c) Será que os formatos físicos e condição do material representam grandes impedimentos?</p> <p>d) São intermediários, como microfimes ou slides, disponíveis e em boas condições?</p> <p>e) Como grande e complexa, em termos de variedade documento, é a coleção?</p>
<p align="center">Considerações de Preservação</p>	<p>a) Será que o material está colocado em risco no processo de digitalização?</p> <p>b) Seriam substitutos digitais reduzir o uso dos originais, oferecendo-lhes assim proteção contra manipulação?</p> <p>c) É a reprodução digital visto como um meio para substituir os originais?</p>
<p align="center">Organização e documentação disponível</p>	<p>a) O material está em uma ordem coerente, logicamente estruturado?</p> <p>b) É paginado ou está à disposição sugerida por outros meios?</p> <p>c) É completo?</p> <p>d) Existe descrição, navegação ou informações estruturais adequadas sobre o material, tais como registros bibliográficos ou um instrumento de pesquisa detalhada?</p>

(Continua)

(Conclusão)

Utilizações previstas	<ul style="list-style-type: none"> a) Que tipos, níveis e frequência de utilização estão previstos? b) Existe uma compreensão clara das necessidades dos utilizadores? c) Pode apoiar a digitalização desses usos? d) Ter acesso ao material pode ser significativamente reforçada pela digitalização? e) Pode a sua instituição apoiar uma gama de utilizações, por exemplo, impressão, navegação, detalhada revisão? f) Existem questões em torno da segurança ou o acesso que devem ser levados em conta (por exemplo, acesso restrito a determinadas pessoas ou utilizar, em determinadas condições?).
Aumento da coleção digitalizada	<ul style="list-style-type: none"> a) Não há incentivo para digitalizar material com base na disponibilidade de recursos digitais complementares (incluindo dados e metadados?) b) É uma oportunidade para a cooperação multi-institucional? c) Para a construção de coerência temática ou "massa crítica"?
Duplicação de esforços	<ul style="list-style-type: none"> a) O material foi digitalizado por outra fonte confiável? b) Se assim for, os arquivos digitais possuem qualidade suficiente, documentação e funcionalidade para servir aos seus propósitos? Que condições regulam o acesso e a utilização desses arquivos?
Capacidade Institucional	<ul style="list-style-type: none"> a) A sua instituição tem a infraestrutura técnica necessária para gerir, fornecer e manter materiais digitalizados? b) Faça que seus principais usuários tenham computação e conectividade adequada para fazer uso efetivo desses materiais? c) Veja infraestrutura técnica para obter informações específicas sobre os componentes técnicos a considerar em tal avaliação.
Recursos financeiros	<ul style="list-style-type: none"> a) Você pode determinar o custo total de aquisição de imagem (seleção, preparação, captura, indexação e controle de qualidade)? b) Este custo justificado com base em benefícios reais ou percebidos provenientes da digitalização? c) Existem fundos para apoiar este esforço? d) Existe um compromisso institucional para a gestão em curso e preservação desses arquivos?

Fonte: Adaptado de Cornell University Library/Research Department (2003, tradução nossa).

APÊNDICE C – Guia para digitalização de documentos versão 2.0 (EMBRAPA)

<p>Características do material original</p>	<ul style="list-style-type: none"> a) Qual a relevância intelectual do material no formato analógico? b) É original? c) Há outro exemplar na instituição em melhores condições? d) O conteúdo está completo ou está faltando alguma parte? e) A informação está atualizada? f) A informação está correta? g) O material é um interesse efêmero dos dias de hoje, ou ele ainda será de interesse com o passar do tempo?
<p>Propriedades físicas dos materiais</p>	<ul style="list-style-type: none"> a) Tipo e categoria do recurso (é um livro impresso, texto escrito a mão, programa de TV em VHS, é um original ou uma reprodução?) b) Data (quantos anos tem? Sabe quando foi feito?) c) Tamanho e dimensões físicas (é um livro de tamanho regular? É uniforme? Qual é o tamanho em cm? Qual a duração da fita de vídeo em horas/minutos/segundos?) d) Tipo de mídia (papel, Madeira, couro, vídeo cassete magnético; combinação de materiais?) e) Formato (disco de 78 rpm, fita cassete analógica, DAT (Digital Audio Tape), Betacam SP tape, NTSC ou PAL formato de vídeo) f) Estrutura do material (o material é encadernado?) g) Condições de preservação (qual o estado de conservação? Dever ser preservado? Requer manuseio especial?)

(Continua)

(Conclusão)

<p>Objetivos da digitalização</p>	<ul style="list-style-type: none"> a) Melhorar o acesso; b) Ajudar na preservação dos materiais originais reduzindo seu manuseio; c) Facilitar atividades de pesquisa; d) Atender as necessidades do público; e) Complementar outras iniciativas de digitalização, de dentro e fora da instituição.
<p>Direitos autorais</p>	<p style="text-align: center;">Sim</p> <ul style="list-style-type: none"> a) O material que você quer digitalizar é de domínio público? b) A empresa possui os direitos autorais e controla os direitos para digitalização? <p style="text-align: center;">Não</p> <ul style="list-style-type: none"> a) Os materiais que você quer digitalizar não são de domínio público? b) A sua permissão (sobre o uso do material) cobre a digitalização e publicação em meio digital?

Fonte: Adaptado de EMBRAPA (2006).

APÊNDICE D – Diretrizes para planejamento de digitalização de livros raros e coleções especiais (IFLA)

Estados a serem considerados	<p>Coleções, obras, edições e cópias devem ser estudadas e comparadas, em relação ao âmbito da nova coleção digital.</p> <p>[...]</p> <p>Também devem ser considerados aspectos, como cronologia, local de publicação, autor, assunto, formato, possuidor etc.</p>
Questões essenciais	<ul style="list-style-type: none"> a) Queremos digitalizar documentos únicos de distintas tipologias, como por exemplo, “tesouros”? b) Queremos digitalizar uma coleção existente? c) Queremos criar uma nova coleção “virtual”?
Revisão física do material	<p>A seleção deve incluir uma revisão física do material, pois sua condição influirá no processo de digitalização. É essencial a comunicação entre a equipe que seleciona e a equipe que realiza as demais tarefas.</p>
Direitos de autor	<p>A identificação dos itens que são de domínio público e aqueles que não são é uma atividade que deve ser realizada quando novas coleções físicas são adquiridas para a instituição.</p>
Coleção digital	<ul style="list-style-type: none"> a) Exame e preparação dos materiais para sua digitalização: condição física e existência de metadados; b) O processo de digitalização (Escolha do equipamento, Qualidade da imagem, Fidedignidade ao original, Aspectos relativos à conservação); c) Processamento de imagens e alimentação do sistema.

Fonte: Adaptado de IFLA (2015).

**APÊNDICE E – Manual de Digitalização de Acervos: textos, mapas e imagens
fixas (SILVA)**

<p>Constituir uma Comissão de Seleção</p>	<p>Responsabilidades:</p> <ul style="list-style-type: none">a) As indicações iniciais de coleções ou itens para a conversão digital;b) Posteriormente a comissão avaliará quais documentos deverão de fato permanecer no conjunto inicialmente destacado para conversão;c) Em seguida deverão ser estabelecidas prioridades baseadas no valor, uso e risco dos materiais. <p>Membros da comissão:</p> <p>Membros da instituição, doadores, pesquisadores especialistas, educadores, especialistas em digitalização, especialistas em preservação e conservação, bibliotecários, arquivistas e museólogos/curadores, pesquisadores em recursos online, advogados.</p>
--	--

(Continua)

(Continuação)

Cautela e Experiência	<p style="text-align: center;">Evitar documentos:</p> <ul style="list-style-type: none">a) De grandes dimensões ou de formatos variados/não usuais;b) Com amplas variações tonais;c) A cores;d) Documentos manuscritos complexos ou pouco legíveis. <p style="text-align: center;">Dar prioridade para documentos:</p> <ul style="list-style-type: none">a) Com tamanhos e tipos relativamente padronizados;b) Com contrastes normais;c) Com dimensões máximas próximas ao formato A4, como textos impressos, ou mesmo ampliações fotográficas originais com alta definição e contrastes;d) Sobre os quais a instituição detenha os direitos de propriedade ou as licenças necessárias à sua utilização/reprodução/disponibilização em suportes/formatos diversos.e) Sobre os quais não venham a recair restrições devidas a aspectos relativos a privacidade, publicidade, difamação, obscenidade, ou outros fatores declarados pelos doadores; f) Que estejam submetidos a altos riscos de perda, mas que sejam suficientemente estáveis, ou que já tenham sido reproduzidos fotograficamente de forma adequada, para serem escaneados sem danos;g) Mais frequentemente solicitados;h) Únicos na instituição.i) Formulários de avaliação podem ser utilizados, com aceitação ou não de responsabilidade da comissão.
------------------------------	---

(Conclusão)

<p>Motivações para a Seleção</p>	<p>Motivação dirigida à coleção original:</p> <ul style="list-style-type: none">a) Desenvolvimento da coleção (necessidades dos usuários);b) Preservação/arquivamento (risco de perda ou dano). <p>Motivação dirigida à versão digital:</p> <ul style="list-style-type: none">a) Desenvolvimento da coleção (valor informacional e maior acesso);b) Preservação/arquivamento (preservação digital a longo prazo, acompanhando as inovações tecnológicas e a preservação do conteúdo informacional do documento);c) Acesso (conteúdo do documento integrado aos serviços da biblioteca, maior distribuição, maior possibilidade de acesso visto as fragilidades do documento, metadados, autenticação digital);
---	--

Fonte: Adaptado de SILVA (2005).

ANEXO A – Geração de matrizes digitais

Tipo de documento	Tipo de Reprodução	Formato de arquivo digital	Resolução mínima, modo de cor e observações
Textos impressos, sem ilustração, preto e branco. (*) sem manchas	Bitonal (**)	TIFF ³⁸ sem compressão	Resolução mínima de 300 dpi ³⁹ , escala 1:1, com margem preta de 0,2 cm ao redor do documento, 4 bits, modo bitonal(**)
Textos impressos, com ilustração e preto e branco.(*) com manchas	Tons de cinza (***)	TIFF sem compressão	Resolução mínima de 300 dpi, escala 1:1, com margem preta de 0,2 cm ao redor do documento, 8 bits, modo tons de cinza (***)
Textos impressos, com ilustração e cor	Cor	TIFF sem compressão	Resolução mínima de 300 dpi, escala 1:1, com margem preta de 0,2 cm ao redor do documento, 24 bits (8 bits por canal de cor), modo RGB (****)
Manuscritos sem a presença de cor	Tons de cinza	TIFF sem compressão	Resolução mínima de 300 dpi, escala 1:1, com margem preta de 0,2cm ao redor do documento, 8 bits, modo tons de cinza (***)
Manuscritos com a presença de cor	Cor	TIFF sem compressão	Resolução mínima de 300 dpi, escala 1:1, com margem preta de 0,2 cm ao redor do documento, 24 bits (8 bits por canal de cor), modo RGB (****)
Fotografias (Preto e Branco e Cor) (**)	Cor	TIFF sem compressão	Resolução mínima de 300 dpi, escala 1:1, com margem preta de 0,2 cm ao redor do documento, 24 bits (8 bits por canal de cor), modo RGB, com carta de cinza para ajuste de níveis (preferencialmente)
Negativos fotográficos e diapositivos (a)	Cor	TIFF sem compressão	Resolução mínima de 3000 dpi, 24 bits (8 bits por canal de cor), modo RGB (****)
Documentos cartográficos	Cor	TIFF sem compressão	Resolução mínima de 300 dpi, escala 1:1, com margem preta de 0,2cm ao redor do documento, 24 bits (8 bits por canal de cor), modo RGB, com carta de cinza para ajuste de níveis (preferencialmente) (****)
Plantas	Preto e branco	TIFF	Resolução mínima de 600 dpi, 8 bits, com possibilidade de modo tons de cinza (****)
Microfilmes e microfichas	Tons de cinza (***)	TIFF sem compressão	Resolução mínima de 300 dpi, 8 bits, modo tons de cinza (***)
Gravuras, cartazes e desenhos (Preto e Branco e Cor)	Cor	TIFF sem compressão	Resolução mínima de 300 dpi, escala 1:1, com margem preta de 0,2cm ao redor do documento, 24 bits (8 bits por canal de cor), modo RGB, com carta de cinza ou cores para ajuste de níveis (preferencialmente) (****)

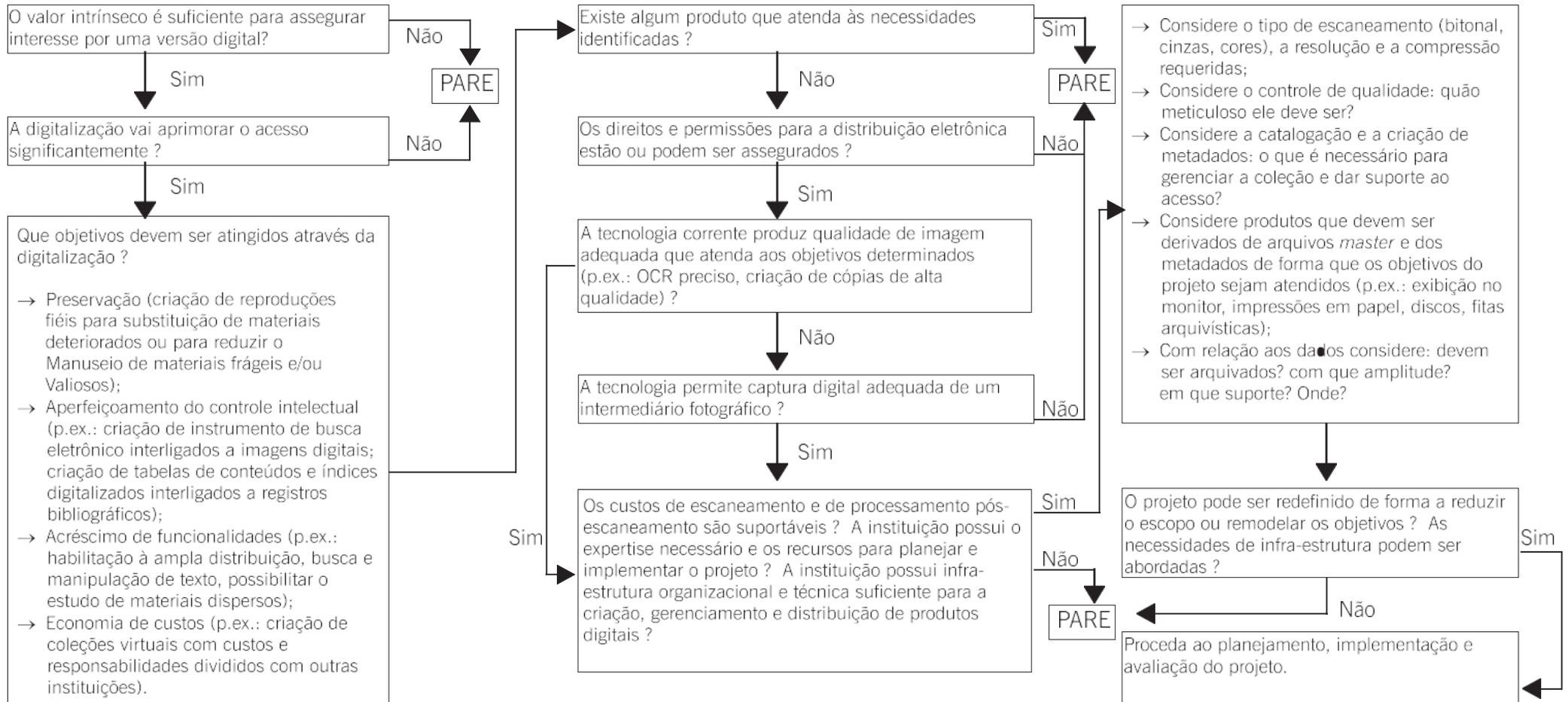
Fonte: CONARQ (2010).

ANEXO B – Modelo de verificação da qualidade do representante digital

1. A imagem tem o tamanho/resolução pretendido?
a. Matriz Digital
b. Matriz Digital com Processamento da Imagem (opcional)
c. Formatos de arquivo digitais de Acesso (Derivadas de Acesso)
d. Thumbnail (Derivadas de Navegação)
e. Outros formatos de arquivo digitais derivadas de acesso
2. O nome da imagem digital esta correto?
a. Matriz Digital
b. Matriz Digital com Processamento da Imagem (opcional).
c. Formatos de arquivo digitais de Acesso
d. <i>Thumbnail</i> (Derivadas de Navegação)
e. Outros formatos derivados de acesso
3. O Formato Digital da imagem está correto?
a. Matriz Digital
b. Matriz Digital com Processamento de Imagem (opcional)
c. Formatos de arquivos digitais de Acesso
d. <i>Thumbnail</i> (Derivadas de Navegação)
e. Outros formatos de arquivo digitais derivadas de acesso
4. Verificação de Qualidade Digital com comparação com o documento original
a. A imagem esta correta no modo de cor
b. Recorte correto
c. Sem rotação
d. Sem inversão
e. Sem inclinação
f. Perda de nitidez/ excesso de nitidez
g. Presença de interferência em imagens com linhas - Padrão Moiré ⁵²
h. Não pixelado
i. Predominância de uma das cores da imagem
j. Histograma ⁵³
k. Sem alto contraste / sem alta luminosidade
l. Valores tonais desiguais / reflexos
n. Aspecto granulado
m. Observações gerais

Fonte: CONARQ (2010).

ANEXO C – A seleção para digitalização: matriz de apoio ao processo de decisão



Fonte: Silva adaptado a partir de Chapman, Hazel, Horrell e Merrill-Oldham (1997).