

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

TAMIRES LISIANE PEREIRA DUARTE

ANÁLISE DA DIGITALIZAÇÃO DE DOCUMENTOS TEXTUAIS DATILOGRAFADOS

Porto Alegre

2016

TAMIRES LISIANE PEREIRA DUARTE

ANÁLISE DA DIGITALIZAÇÃO DE DOCUMENTOS TEXTUAIS DATILOGRAFADOS

Trabalho de Conclusão de Curso apresentado ao Curso de Biblioteconomia da Faculdade de Biblioteconomia e Comunicação, da Universidade Federal do Rio Grande do Sul, como pré-requisito para a obtenção do título de Bacharel em Biblioteconomia.

Prof. Dr. Rafael Port da Rocha

Porto Alegre

2016

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Rui Vicente Oppermann

Vice-Reitora: Prof. Dra. Jane Fraga Tutikian

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Direção: Prof. Dra. Ana Maria Mielniczuk de Moura

Vice-Direção: Prof. Dr. André Iribure Rodrigues

DEPARTAMENTO DE CIENCIAS DA INFORMAÇÃO

Chefe: Prof. Dr. Moisés Rockembach

Chefe Substituto: Prof. Dr. Valdir José Morigi

COMISSÃO DE GRADUAÇÃO DO CURSO DE BIBLIOTECONOMIA

Coordenador: Prof. Dr. Rodrigo Silva Caxias de Souza

Coordenador Substituto: Prof. Dr. Jackson da Silva Medeiros

CIP - Catalogação na Publicação

Duarte, Tamires Lisiane Pereira

Análise da digitalização de documentos textuais datilografados / Tamires Lisiane Pereira Duarte. -- 2016.

76 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, 2016.

1. Documentos textuais. 2. Digitalização. 3. CONARQ. 4. Configuração do documento digital. I. Rocha, Rafael Port da, orient. II. Título.

TAMIRES LISIANE PEREIRA DUARTE

ANÁLISE DA DIGITALIZAÇÃO DE DOCUMENTOS TEXTUAIS DATILOGRAFADOS

Trabalho de Conclusão de Curso apresentado ao Curso de Biblioteconomia da Faculdade de Biblioteconomia e Comunicação, da Universidade Federal do Rio Grande do Sul, como pré-requisito para a obtenção do título de Bacharel em Biblioteconomia.

Prof. Dr. Rafael Port da Rocha

Examinado em 09 de dezembro de 2016

Banca Examinadora:

Prof. Dr. Rafael Port da Rocha

Prof. Dra. Sonia Elisa Caregnato

Me. Rozi Mara Mendes

Aprovado em 09 de dezembro de 2016.

AGRADECIMENTOS

Primeiramente, a Deus, simplesmente por tudo.

A minha mãe, uma mulher incrível, paciente, corajosa e forte. Sem o apoio dela, creio que nem teria começado essa caminhada, ou muitas vezes teria desistido. Sua fé inabalável em Deus e sua crença em mim fizeram com que eu me levantasse todas as vezes que cai.

A todos os demais membros da minha família, irmãos, pai, sobrinhos, tios e primos, por me apoiarem.

Ao Prof. Rafael Port da Rocha, meu orientador. Obrigada pelo apoio, pelas orientações. Mesmo quando eu chegava sem avisar, você sempre me recebia, esclarecia e me entendia.

Ao pessoal do IST Couro e Meio Ambiente. Vocês me acolheram de braços abertos, e me sinto imensamente feliz por fazer parte desta equipe.

Ao pessoal da Biblioteca Pública Municipal Josué Guimarães, que me acolheu como estagiária. Aprendi muito com vocês.

A Catarina e a Adriana, da Escola Técnica Estadual Parobé, que me ajudaram muito no meu estágio obrigatório. Vocês são pessoas que levarei sempre comigo.

Aos amigos de longa data (Nina, Carol, Felipe, Vanessa, Rogério, Camila, entre tantos outros), que entenderam quando eu não podia encontra-los, pois estava "enrolada" com a faculdade.

As Moemas, amigas incríveis que a Josué Guimarães me deu. As Fionas, as amigas queridas que a Biblioteconomia me deu. Mesmo eu estando meio ausente, vocês são pessoas que moram no meu coração para sempre. Sou muito grata pela amizade de vocês.

E ao grupo mais querido da Biblioteconomia: Aline, Ingra, Dedy e Leticia. Gurias, muito obrigada por me acolherem, me ajudarem, me apoiarem em todos os momentos. Vocês estiveram do meu lado em momentos muito importantes, em que com certeza sozinha eu não conseguiria suportar. Obrigada pelo apoio, pelas risadas, os chocolates, os vídeos de gatos e de corujas fofas. Obviamente, esse agradecimento se estende aos pais, mães e tios de vocês, que me acolheram como uma filha. Vocês todos moram no meu coração, e espero mantê-los para sempre na minha vida.

**Two roads diverged in a wood, and I-
I took the one less traveled by,
And that has made all the difference.**

(The Road not Taken – Robert Frost)

RESUMO

A pesquisa apresenta uma análise da digitalização de documentos textuais com características especiais com relação as recomendações do Conselho Nacional de Arquivos (CONARQ). O estudo mostra um comparativo através de seis amostras de documentos datilografados com as configurações de modo e resolução indicadas pelo CONARQ. O objetivo da pesquisa é investigar as configurações para digitalização de documentos textuais datilografados em suas diferentes apresentações com diferentes características. Utiliza como metodologia a pesquisa exploratória com aspectos qualitativos e comparação entre as seis amostras digitalizadas e as recomendações para modo e resolução. Conclui que as recomendações não englobam todas as configurações do documento a ser digitalizado, e sugere uma ampliação.

Palavras-chave: Documentos textuais. Digitalização. CONARQ. Configurações do documento digital.

ABSTRACT

The research presents an analysis of the digitization of textual documents with special characteristics in relation to the recommendations of the Conselho Nacional de Arquivologia (CONARQ). The study shows a comparative across six sample typed documents with the mode and resolution settings indicated by CONARQ. The purpose of the research is to investigate the settings for scanning typed textual documents in their different presentations with different characteristics. It uses as methodology the exploratory research with qualitative aspects and comparison between the six digitized samples and the recommendations for mode and resolution. It concludes that the recommendations do not encompass all the settings of the document to be scanned, and suggest an extension.

Keywords: Textual documents. Scanning. CONARQ. Digital document settings.

LISTA DE IMAGENS

Imagem 1 - Níveis do documento digital	23
Imagem 2 - Exemplo de imagem formada por pixels	26
Imagem 3 - Profundidade de bits (1 Bit).....	26
Imagem 4 - Profundidade de bits (2 Bits)	26
Imagem 5 - Profundidade de bits (4 Bits)	27
Imagem 6 - Profundidade de bits (8 Bits)	27
Imagem 7 - Profundidade de bits (24 Bits)	27
Imagem 8 - Documento datilografado e mimeografado com texto em roxo digitalizado em modo bitonal (B301)	43
Imagem 9 - Documento datilografado e mimeografado com texto em roxo digitalizado em modo tons de cinza (G301)	44
Imagem 10 - Documento datilografado e mimeografado com texto em roxo digitalizado em modo colorido (C301)	44
Imagem 11 - Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência, digitalizado em modo bitonal (B302).....	46
Imagem 12 - Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência, digitalizado em modo tons de cinza (G302)	47
Imagem 13 - Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência, digitalizado em modo colorido (C302).....	48
Imagem 14 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo bitonal (B303)	49
Imagem 15 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo tons de cinza (G303)	50
Imagem 16 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo colorido (C303)	50
Imagem 17 - Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente, digitalizado em modo bitonal (B304)	51
Imagem 18 - Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente, digitalizado em modo tons de cinza (G304)	52
Imagem 19 - Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente, digitalizado em modo colorido (C304).....	53
Imagem 20 - Documento com texto datilografado e digitado em preto, em papel vegetal, digitalizado em modo bitonal (B305).....	54
Imagem 21 - Documento com texto datilografado e digitado em preto, em papel vegetal, digitalizado em modo tons de cinza (G305).....	54
Imagem 22 - Documento com texto datilografado e digitado em preto, em papel vegetal, digitalizado em modo colorido (C305).....	55
Imagem 23 - Documento datilografado em matriz de mimeografo com o texto em roxo, digitalizado em modo bitonal (B306)	56
Imagem 24 - Documento datilografado em matriz de mimeografo com o texto em roxo, digitalizado em modo tons de cinza (G306)	57

Imagem 25 - Documento datilografado em matriz de mimeografo com o texto em roxo, digitalizado em modo colorido (C306)	57
Imagem 26 - Documento datilografado e mimeografado com texto em roxo digitalizado em modo colorido com resolução de 300dpi e 600dpi (C301 e C601) ...	59
Imagem 27 - Documento datilografado e mimeografado com texto em roxo digitalizado em modo colorido com resolução de 300dpi e 600dpi 300dpi e 600dpi (C301 e C601) ampliada	60
Imagem 28 – Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C302 e C602).....	61
Imagem 29 - Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C302 e C602) ampliado	61
Imagem 30 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G303 e G603)	62
Imagem 31 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G303 e G603)ampliados.....	63
Imagem 32 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C303 e C603)	63
Imagem 33 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C303 e C603)ampliados	64
Imagem 34 - Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G304 e G604).....	65
Imagem 35 - Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G304 e G604)ampliados	65
Imagem 36 - Documento com texto datilografado e digitado em preto, em papel vegetal, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G305 e G605)	66
Imagem 37 - Documento com texto datilografado e digitado em preto, em papel vegetal, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G305 e G605) ampliados	67
Imagem 38 - Documento datilografado em matriz de mimeografo com o texto em roxo, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C306 e C606)	68
Imagem 39 - Documento datilografado em matriz de mimeografo com o texto em roxo, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C306 e C606)ampliados	68

LISTA DE QUADROS

Quadro 1 - Representação da profundidade de bits.....	28
Quadro 2 - Resolução de Cor.....	29
Quadro 3 - Características da matriz digital	34
Quadro 4 - Características da derivada de acesso	35
Quadro 5 - Características da derivada de navegação	35
Quadro 6 - Tipos de Scanner	36
Quadro 7 - Relação dos Objetivos específicos e método.....	38
Quadro 8 – Relação de amostras e configurações	40
Quadro 9 - Relação de amostras	41
Quadro 10 - Documentos analisados por modo	74
Quadro 11 - Documentos analisados por resolução	75

SUMÁRIO

1	INTRODUÇÃO	13
1.1	JUSTIFICATIVA	14
1.2	OBJETIVOS	15
1.2.1	Objetivo Geral	15
1.2.2	Objetivos Específicos	16
2	REFERENCIAL TEÓRICO	17
2.1	DOCUMENTO	17
2.1.1	Características físicas do documento	17
2.1.2	Características intelectuais	19
2.2	DOCUMENTO DIGITAL	21
2.2.1	Componentes	21
2.2.2	Níveis de um documento digital	23
2.3	IMAGEM DIGITAL	24
2.3.1	Profundidade de bit	25
2.3.2	Resolução da imagem digital	28
2.3.3	Tipos de imagem digita	30
2.4	DIGITALIZAÇÃO DE UM ACERVO	30
2.5	CONFIGURAÇÃO DO REPRESENTANTE DIGITAL	32
2.5.1	Matriz digital	33
2.5.2	Derivada de acesso	34
2.5.3	Derivada de navegação	35
2.6	ESCOLHA DO SCANNER	36
3	METODOLOGIA	37
3.1	SEGUNDO A NATUREZA	37
3.2	SEGUNDO A ABORDAGEM DO PROBLEMA	37
3.3	SEGUNDO OBJETIVO	38
3.4	SEGUNDO PROCEDIMENTO E COLETA DE DADOS	38
3.5	ANÁLISE DOS DADOS E APRESENTAÇÃO DOS RESULTADOS	39
4	ANÁLISE DOS RESULTADOS	42
4.1	ANÁLISE COM RELAÇÃO AO MODO	42
4.2	ANÁLISE COM RELAÇÃO A RESOLUÇÃO	58
5	CONCLUSÃO	69

REFERÊNCIAS	71
APÊNDICE A - SIGLA DOS DOCUMENTOS ANALISADOS	74

1 INTRODUÇÃO

Atualmente, há várias formas de armazenar, disponibilizar e consultar a informação. Porém, mesmo com toda a tecnologia disponível, percebe-se que o meio mais utilizado ainda é o papel. Provavelmente por ser recorrente no dia a dia, com a informação sendo impressa em jornais, periódicos ou livros.

O papel, porém, tem alguns inconvenientes. Seu grande uso pode danificar o documento, que muitas vezes é único. Com o tempo, o documento passa a apresentar deteriorações, com papel tornando-se amarelado e com manchado; e as informações podem se apagar. Por isso, para preservar a informação de um documento, a digitalização é um meio eficaz. Ela tem como principais vantagens possibilitar que o documento seja acessado e reproduzido diversas vezes, sem danos ao original. O uso da digitalização previne a alterações das propriedades dos originais.

O processo de digitalização consiste em transformar um arquivo físico em digital, através de um equipamento específico para esse fim. O mais utilizado é o *scanner*. Cada documento, porém, tem características que precisam ser observadas antes da digitalização, para que o documento digital gerado seja uma cópia fiel do documento físico. Existem recomendações para a digitalização de documentos. Elas indicam a resolução, modo, profundidade e formato de arquivo para a imagem digital, considerando as características do original. Isto é, o suporte (papel, negativo), seu gênero (textual, textual manuscrito, textual impresso, imagem, mapas, entre outros) e outras características como existência de manchas, cor, papel com deterioração de cor, etc.

No Brasil, o Conselho Nacional de Arquivos (CONARQ), possui recomendações para digitalização de documentos. Esse órgão é vinculado ao Arquivo Nacional do Ministério da Justiça, e tem por objetivo definir políticas para arquivos públicos e privados. (CONSELHO..., 2016).

As recomendações do CONARQ são para uso geral. Isso porque elas não exploram características específicas. No caso dos documento textuais as recomendações não destacam variações dos modos de apresentação. Ele pode ser apresentado com variações de papel (transparente, opaco, com fundo deteriorado) e de tipo de impressão (datilografado e mimeografado). E elas influenciam diretamente o processo de digitalização.

Essa pesquisa teve como foco a digitalização de documentos textuais com características especiais, e sua relação com recomendações do CONARQ. Para esta pesquisa, foram selecionados seis amostras de documentos datilografados com as seguintes variações: mimeografado, em matriz de mimeografo, com o texto escrito com o uso de papel carbono e em papel vegetal. O estudo foi realizado através da digitalização dessas amostras. Através delas, observou-se as configurações de modo de cor e resolução, de acordo com os resultados obtido e o que apresenta na literatura.

1.1 JUSTIFICATIVA

A informação atualmente tem um ciclo de produção muito rápido. E nem tudo o que se produz hoje, é usado ou necessário nesse momento. Mas ninguém garante que não possa ser importante no futuro. Por isso, preservar a informação para as futuras gerações se tornou essencial.

A tecnologia atual permite que possamos registrar e armazenar a informação em diversos formatos além do papel, o suporte físico mais comum. Um desses formatos é o digital. Ele permite gerar diversas cópias do mesmo documento, sem danificar o original.

Para que seja feita a digitalização de um documento, existem recomendações que visam a qualidade do processo. No Brasil, o CONARQ publicou em 2010 recomendações para esse processo. Esse órgão é vinculado ao Arquivo Nacional do Ministério da Justiça, e tem por objetivo definir políticas para arquivos públicos e privados. (CONSELHO..., 2016).

Apesar da publicação do CONARQ, os estudos sobre digitalização de documentos no Brasil são poucos. De acordo com Caregnato e Rocha (2013), isso ocorre devido à falta de centros de preservação digital no Brasil. Segundo ainda eles, a dificuldade de criar esses centros é a necessidade de reunir profissionais tanto da informática quanto da informação interessados no estudo da preservação digital de documentos.

Atualmente, a Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul (FABICO - UFRGS) possui o Centro de Documentação e Acervo Digital de Pesquisa (CEDAP). De acordo com Caregnato e Rocha (2013), ele tem por objetivo o estudo sobre a preservação digital e a

curadoria de dados digitais. Investigar as configurações para textos com características especiais é uma contribuição ao CEDAP. O CEDAP possui, de acordo com Caregnato e Rocha (2013):

[...] equipamentos e serviços para: migração de vídeos a partir de vários suportes; digitalização de documentos em papel, livros, microfilme, imagens radiológicas, etc.; armazenamento, preservação, organização e disponibilização de dados digitais; infraestrutura de rede para seu compartilhamento.

A inauguração do CEDAP e a leitura das recomendações do CONARQ apenas confirmam a falta de estudos sobre digitalização. O material publicado pelo CONARQ é amplo. Os documentos textuais podem ser utilizados como exemplo. As recomendações não destacam variações dos diversos modos de apresentação, citando apenas textos manuscritos e impressos.

Não são contemplados, por exemplo, os datilografados. Apesar de não serem utilizados atualmente, foram muito comuns em décadas passadas. Eles poderiam ser considerados, para efeitos de digitalização, documentos impressos. Mas diferente desse último, ele possui características significativas, como o texto escrito com o uso de papel carbono ou mesmo mimeografados. Outro ponto não contemplado nas recomendações são o papel usado como suporte para os documentos datilografados, que podem ser transparentes, opacos ou mesmo com o fundo de cor que não branca.

A partir da justificativa apresentada, o problema da pesquisa é: Quais as configurações para digitalização de documentos textuais datilografados em suas diferentes apresentações com diferentes características?

1. 2 OBJETIVOS

Nesta seção, serão apresentados os objetivos gerais e específicos.

1.2.1 Objetivo Geral:

Investigar as configurações para digitalização de documentos textuais datilografados em suas diferentes apresentações com diferentes características.

1.2.2 Objetivos Específicos

- a) Analisar as recomendações existentes para digitalização de documentos textuais;
- b) Experimentar as diferentes configurações para a geração de matrizes de documentos textuais;
- c) Analisar as experimentações realizadas, relacionando-as com as recomendações existentes para digitalização de documentos textuais;
- d) Apresentar sugestões para digitalização de documentos textuais datilografados;

2 REFERENCIAL TEÓRICO

Neste capítulo, será apresentado o referencial teórico. O capítulo está dividido em quatro partes: o documento, o documento digital, a imagem digital, a digitalização e a configuração do representante digital.

2.1 DOCUMENTO

O termo documento é muito utilizado na Ciência da Informação para referir a todo e qualquer registro de informação em algum suporte. Siqueira (2015) explica que a origem da palavra documento vem do latim *documentum*, que está relacionado a educação e a instrução. Um documento pode fornecer ambos, além de compartilhar uma informação. Isso ocorre porque uma vez que ela é registrada em um documento, várias pessoas, em diferentes momentos, podem ter acesso.

O Dicionário de Terminologia Arquivística (2014, p. 74) coloca que documento é uma "unidade de registro de informações, qualquer que seja o suporte ou o formato.". Suporte é o material em que as informações são registradas, e formato é um conjunto de características físicas de um documento. (Dicionário de Terminologia Arquivística, 2014). O suporte e o formato influenciam no gênero documental, que é a reunião de espécies de documentos com essas duas características em comum. (Dicionário de Terminologia Arquivística, 2014). Por isso, compreende-se que um documento tem como características principais seu suporte, formato e gênero. São exemplos de gêneros documentais: documentos textuais, documentos iconográficos, cartográfico, entre outros.

Suporte, formato e gênero são características físicas do documento. Isso porque elas influenciam de forma direta de seu tratamento. Os documentos classificados de acordo com as suas características físicas e intelectuais. (GUINCHAT; MENO 1994)

2.1.1 Características físicas do documento

Como características físicas, Guinchat e Menou (1994) citam a natureza, os materiais, a forma de produção, a modalidade de utilização, a periodicidade, as coleções e a forma de publicação. As características físicas influenciam na forma de

tratamento do documento. O suporte, o formato e o gênero são as características da natureza do documento. Isso determina qual é o tipo de informação que eles transmitem. Com relação a natureza, os documentos podem ser divididos em textuais e não textuais.

Por documentos textuais, compreende-se aqueles que apresentam basicamente a informação em formato de texto. O Dicionário de Terminologia Arquivística (2014) coloca esse gênero como composto de manuscritos, datilografados ou impressos. Como exemplo, há livros, periódicos, fichas, documentos administrativos, leis, catálogos, cartas, atas de reuniões, relatórios e demais documentos em que a sua construção seja textual.

E documentos não-textuais são os demais documentos em que a informação é apresentada em outro formato. Ele pode conter texto, porém a sua principal informação é transmitida através de outros elementos. Guinchat e Menou (1994) colocam como sendo os principais documentos não-textuais, o documento iconográfico, o documento sonoro e o documento audiovisual. O documento iconográfico é aquele que contém imagens fixas, impressas ou desenhadas como fotografias e gravuras. O documento sonoro é o que contém registros de sons, como discos de músicas, sons gravados e tudo que possa ser reproduzido de forma sonora. O documento audiovisual é o que contém imagens, fixas ou em movimento, acrescidas de registros sonoros como filmes e fitas de vídeo. (Dicionário de Terminologia Arquivística 2014).

O material e o modo de utilização são relacionados com onde a informação está registrada, ou seja, ao suporte do documento. Um mesmo documento pode ser encontrado em vários suportes. Por exemplo, uma fotografia de uma paisagem (documento iconográfico) pode ser apresentado em papel, em negativo ou em uma mídia digital (arquivo de computador). Ou ainda em um arquivo de computador, formado por bits. E esse mesmo documento (fotografia) transmite a mesma informação (paisagem), porém em materiais diferentes (papel, mídia digital, negativo, bits) e com necessidades de manuseio diferentes. O papel é um suporte que não necessita de outro objeto para dispor a imagem, pode ser manipulado de forma direta. As mídias, como CD's (*Compact Disc*) e DVD's (*Digital Video Disc*) necessitam de um equipamento especial para gerar uma apresentação do documento, como aparelhos reprodutores. O documento em formato digital necessita de um computador, pois ele é capaz de ler e interpretar os *bits*.

A forma de produção de um documento é o modo que ele foi criado. Guinchat e Menou (1994) os dividem em brutos e manufaturados. Os documentos brutos são aqueles encontrados na natureza, como minerais, fósseis e plantas. Eles não têm nenhum tipo de interferência humana em sua produção. Os documentos manufaturados são aqueles produzidos pelo homem. Podem ser feitos de forma artesanal ou industrial. Como exemplos, há os protótipos, obras de arte, relatórios, textos científicos e literários.

A periodicidade e a coleção tratam da produção do documento em relação a volumes e ao tempo de publicação. Periodicidade remete ao periódico, que nada mais é que uma revista ou jornal com intervalo regular. De acordo com Guinchat e Menou (1994), a periodicidade de publicação de um documento pode variar de 24 horas a um ano. Porém, o período é sempre o mesmo, como o nome sugere. A coleção porém, não tem um intervalo regular. A publicação de um volume em uma coleção pode demorar meses ou anos. Porém, os assuntos têm uma relação, e o seu formato é igual.

A forma de publicação distingue os documentos publicados e os não publicados. Documentos publicados são aqueles distribuídos de forma comercial, normalmente por editoras e livrarias. Justamente por isso, Gomes, Mendonça e Souza (2000) denominam esses documentos de literatura branca. Os documentos não publicados são conhecidos também como literatura cinzenta. De acordo com Gomes, Mendonça e Souza (2000), essa expressão vem do inglês *grey literature*, e é utilizada para denominar documentos não publicados de forma convencional. Eles são documentos governamentais, acadêmicos ou da indústria. Sua forma de aquisição normalmente não é por compra e venda, mas muitas vezes por contato direto com quem produziu.

2.1.2 Características intelectuais

Guinchat e Menou (1994) colocam que as características intelectuais podem definir o valor do documento junto ao seu público-alvo. Além disso, influencia diretamente o tratamento da informação. São elas o objetivo, o grau de elaboração, o conteúdo, a origem e os tipos de documento.

O objetivo do documento é a razão pela qual ele foi elaborado. Como por exemplo, servir como prova, apresentar ideias ou auxiliar o preparo de outro

documento. O grau de elaboração de um documento, de acordo com Guinchat e Menou (1994), auxilia na distinção de documentos primários, secundários e terciários. Eles explicam que documentos primários são os documentos originais. Muller (2000) complementa como sendo os documentos produzidos pelo autor de uma pesquisa com sua interferência direta. Exemplos de documentos primários são livros, artigos e relatórios. Os documentos secundários são aqueles que fazem referência aos documentos primários. Auxiliam a classificá-los, e também em utilizar as informações contidas neles. Enciclopédias, dicionários, manuais e revisões de literatura são exemplos de documentos secundários. Documentos terciários são aqueles que guiam o usuário da informação para documentos primários e secundários. Como exemplo as bibliografias, os catálogos e os guias de literatura. (GUINCHAT, MENO, 1994; MULLER, 2000).

O conteúdo, como explica Guinchat e Menou (1994), não trata apenas do assunto. Ele tem relação também com a forma de apresentação, a exaustividade, o nível científico, a acessibilidade, o grau de originalidade e novidade e os dados do documento (números e o tempo que eles tem). Eles colocam que a questão do conteúdo é muito relativa. Isso porque um documento antigo pode estar desatualizado com dados, porém ser um relato importante para determinada época, e um documento recente, com dados atuais, pode não trazer nenhuma informação importante para o assunto abordado ou para a época em que se encontra.

A origem de um documento tem relação direta com a fonte e o autor dele. E esse item influencia no modo que o documento será utilizado. A fonte de um documento pode ser pública ou privada, individual ou coletiva. E o autor pode ser único ou um grupo de pesquisadores, ou até mesmo uma organização, uma empresa. Todos esses dados influenciam de forma direta em como o documento e informação contida nele devem ser tratados. (GUINCHAT, MENO, 1994)

Os tipos de documentos tratam de questões de seleção, armazenamento e tratamento da informação. Guinchat e Menou (1994) colocam que documentos, com relação às suas características intelectuais, podem ser classificados em essenciais e marginais. Os documentos essenciais tratam de assuntos comuns e muito procurados pelo público da unidade de informação em que ele se encontra. Eles devem ser tratados com prioridade. Os documentos marginais são aqueles que podem conter a informação necessária para o público em uma proporção maior ou

menor. Esse tipo de documento necessita de uma análise detalhada para se verificar se deve ou não ficar na unidade de informação.

2.2 DOCUMENTO DIGITAL

Desde que a humanidade começou a produzir informação, o pensamento de preservá-la para a posteridade a acompanha. Muitos foram os métodos empregados para armazenar a informação, desde os tabletes de argila aos livros no formato códex. Atualmente, uma das formas mais comuns de se armazenar a informação é no formato digital.

Utiliza-se a expressão documento digital e documento eletrônico como sinônimos. Porém, isso é incorreto. O Dicionário de Terminologia Arquivística (2014) define que documento eletrônico é aquele que só pode ser acessado através de um equipamento eletrônico. E o documento digital é definido como um documento formado por uma sequência binária, e acessível através de um computador. Como exemplo de documento eletrônico, é possível citar a fita VHS, que é um documento audiovisual que só pode ser acessada por um equipamento de vídeo cassete. E como documento digital, é um arquivo de texto salvo em uma mídia de CD, que só poderá ser acessada em um computador.

O documento digital é então um documento formado por *bits*. Esse tipo de documento, explica Ferreira (2006) pode ter dois modos de origem: física (digitalizados) ou naturalmente digital. O documento digital de origem física é aquele que passou por um processo de digitalização, como uma foto de filme que foi revelado e depois digitalizado. O documento naturalmente digital é aquele já começou de forma digital, como um documento texto criado através do computador.

2.2.1 Componentes

Para que a informação de um documento digital seja acessada, é necessária uma combinação de quatro fatores: *hardware*, *software*, suporte e pessoal. O *hardware*, de acordo com o Conselho Nacional de Arquivologia (2011) é a parte física de um sistema computacional. Ou seja, o computador em si. O *software* é o programa de computador necessário para acessar o documento digital.

Rondinelli (2005) coloca que o suporte de um documento digital é o seu carregador físico. O mesmo documento digital pode ser armazenado e reproduzido por diversos suportes diferentes, sem perder o seu conteúdo.

E o fator pessoal é a pessoa que acessa o documento. Sem esse fator, a informação contida no documento não pode ser vista. Pois é a pessoa que opera o computador.

Com relação a sua composição, Rocha (2014) coloca que o documento digital é formado por quatro pontos. São o conteúdo, a estrutura, a apresentação e a funcionalidade.

O conteúdo nada mais é que a mensagem transmitida através do documento. Pode ser algo essencial, fazendo com que o documento seja importante. Pode ser apenas uma mensagem enviada com um propósito informativo. Ou até mesmo um lembrete. O importante é que tudo que o documento possui e possa transmitir uma mensagem é o seu conteúdo.

A estrutura, de acordo com Rondinelli (2005) trabalha a representação do conteúdo do documento. Como ele é apresentado, distribuído e como isso é mostrado ao usuário do documento. Ela engloba os outros dois componentes do documento, a apresentação e a funcionalidade. Que também podem ser denominadas forma física e forma intelectual, respectivamente.

A apresentação ou forma física são as características de apresentação documento. Como por exemplo, cor e tipo de fonte, idioma, sinalização de anexos, logomarcas, assinaturas e demais características. Elas, em um primeiro momento, não são de extrema significância ao usuário do documento. Porém, qualquer alteração em alguma delas gera um documento diferente.

O documento digital, de acordo com Rocha (2014), necessita seguir um determinado conjunto de regras para ser armazenado e acessado. Esse conjunto é denominado formato de arquivo.

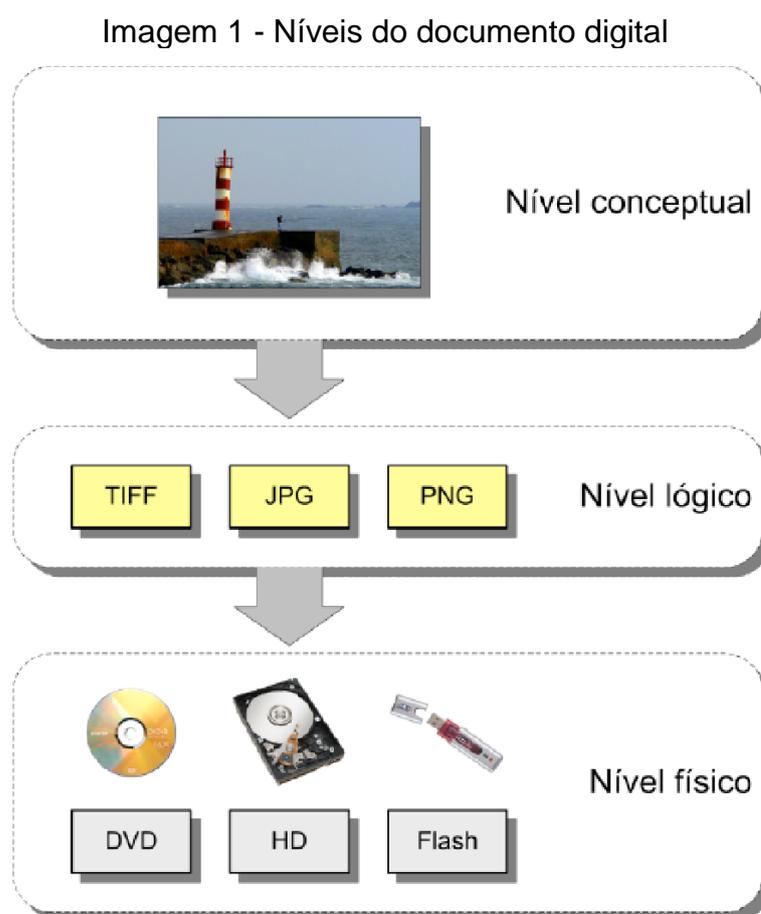
O formato de arquivo, de acordo com o Dicionário de Terminologia Arquivística (2014), são regras para a interpretação dos bits que formam o documento digital. Essas regras, de acordo com o Conselho Nacional de Arquivos (2010) podem gerar quatro tipos de documentos: aberto, fechado, proprietário e padronizado.

O Conselho Nacional de Arquivos (2001) explica que o formato aberto é aquele que as especificações são públicas. Como exemplo, há os arquivos *xml*,

html, *odf* e *rtf*. O formato fechado, ao contrário, é aquele em que as especificações não são divulgadas. O formato proprietário é o que as especificações são definidas por uma organização. Ela tem os direitos do formato de arquivo, e o seu uso pode ser gratuito ou não. E o formato padronizado é aquele que é formado por um órgão de normalização, e são abertos. Como exemplo, o formato *pdf*.

2.2.2 Níveis de um documento digital

É possível determinar que o documento digital é formado por três níveis: o físico, o lógico e o conceitual, conforme mostra a Imagem 1. Ferreira (2006) descreve o nível físico como sendo o local em que os dados do documento se encontram, o suporte físico. Esse nível é representado pelo objeto digital, que de acordo com Ferreira (2006, p.22) "[...] é o que *hardware* é capaz de interpretar."



Fonte: Ferreira (2006, p. 25)

Através da leitura do objeto digital pelo *hardware*, os dados passam a ser interpretados, de acordo com a estrutura do documento. De acordo com Ferreira (2006), essas estruturas são o nível lógico do documento.

A partir do momento que o documento passa para o nível lógico, ele se transforma em algo que o ser humano é capaz de absorver. O documento passa a possuir uma forma conhecida pela mente humana. De acordo com Ferreira (2006), a esse formato formado pela mente humana dá-se o nome de nível conceitual.

De acordo com Ferreira (2006), há ainda um quarto nível, denominado experimental. O nível experimental ocorre a partir do momento em que o ser humano produz a informação, fazendo o caminho inverso ao que foi apresentado. Ou seja, o ser humano absorveu a informação, ou nível conceitual, e a codifica em uma linguagem. Isso pode ser transmitido ou armazenado, através de um processo que transforma a linguagem humana em binária, arquivando-a e um suporte físico. Ferreira (2006) coloca que de forma ideal, o objeto formado pelo nível experimental deverá ser idêntico ao do nível conceitual.

2.3 IMAGEM DIGITAL

Imagem, de acordo com o Dicionário de Terminologia Arquivística (2014, p. 104), é uma "representação gráfica, plástica ou fotográfica de seres, objetos ou fatos.". A imagem digital, por sua vez, trata dessa representação, porém em formato de *bits*. Por se tratar de uma forma de documento digital, ela pode ter sua origem tanto de forma digitalizada como naturalmente digital

Cornell Library (2003a, tradução nossa), explica a imagem digital através da digitalização:

A imagem digital é uma captura instantânea de modo eletrônico, feita a partir de documentos, como fotografias, manuscritos, textos impressos e obras de arte. A imagem digital é amostrado e mapeada como uma grade de pontos ou elementos de imagem (*pixels*).

Uma imagem digital, de acordo com Rocha (2016) pode ser de dois tipos: imagem vetorial e imagem bitmap. A imagem vetorial, de acordo com Martins (2006) é criada em softwares específicos para desenho, como *Corel Draw*, *Illustrator* e *Flash*. Ou seja, ela é uma imagem naturalmente digital. As imagens vetoriais "[...]

são compostas por linhas, ângulos e formas geométricas pré definidas por comandos no computador e por cálculos matemáticos – não só a imagem em si, como também o seu posicionamento em relação à página." (MARTINS, 2016, p. 21)

Uma das características importantes desse tipo de imagem é que ela pode ser reestruturada diversas vezes. Além disso, a resolução é independente da estrutura. Ou seja, pode ser ampliada e reorganizada diversas vezes que a sua resolução não se altera. Isso é importante para casos de artes gráficas. Como por exemplo, uma logomarca para uma empresa, que estampará desde cartões de visita até a fachada do local. (MARTINS, 2006)

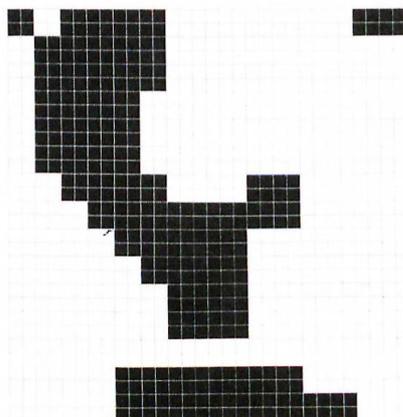
A imagem bitmap, de acordo com Martins (2006, p. 81), é formada por "[...] blocos de dados de *pixels* organizados de determinada maneira.". Cornell Library (2015, tradução nossa), complementa essa explicação. "[...] A imagem digital é mostrada e mapeada como uma grade de pontos ou elementos de imagem (*pixels*).". Esse tipo de imagem pode ser criada digitalmente através de softwares como Photoshop, ou digitalizadas. (ROCHA, 2016)

2.3.1 Profundidade de bit

Martins (2009, p. 16) descreve *pixel* como "[...] a menor unidade gráfica de uma imagem mapeada por *bits* [...]". O *bit* ou *bits*, explica Martins (2006, p. 14) "[...] é a menor unidade de informação usada pelo computador [...]". Cada *bit* assume o valor de 0 ou 1, que representa as cores branco e preto.

A Imagem 2 mostra, por exemplo, uma imagem formada por *pixels*, através de um desenho quadriculado. Cada quadrado representa um *pixel*, e o conjunto deles forma a imagem.

Imagem 2 - Exemplo de imagem formada por pixels



Fonte: Martins (2009, p.15)

Martins (2009) explica que a cor (preto, branco, tons de cinza ou colorido) de cada pixel é representado por um conjunto de *bits*. "Quanto mais *bits* utilizados, mais cores ou tonalidades são obtidas." (p. 14-15). É denominado para a quantidade de *bits* para cada *pixel* profundidade de *bits* de uma imagem.

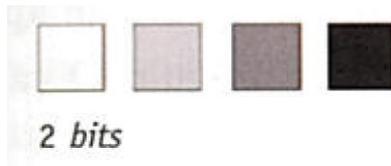
As Imagens 3 a 7 mostram exemplos de números de *bit* por *pixel*. A Imagem 3 mostra um *bit* por *pixel*, que gera duas cores, preto ou branco. Na Imagem 4 mostra 2 *bits* por *pixel*, mostrando quatro tons de cor. Na Imagem 5, há uma representação simplificada de 4 *bits* por *pixel*, mostrando oito dos 16 tons que essa profundidade tem. Na Imagem 6, também traz uma representação simplificada de 8 *bits* por *pixel*. E na Imagem 7, também traz uma representação simplificada de 24 *bits* por *pixel*.

Imagem 3 - Profundidade de *bits* (1 Bit)



Fonte: Martins (2009, p.15)

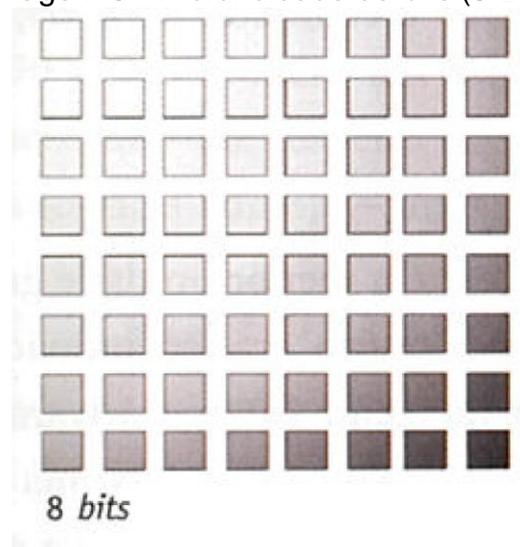
Imagem 4 - Profundidade de *bits* (2 Bits)



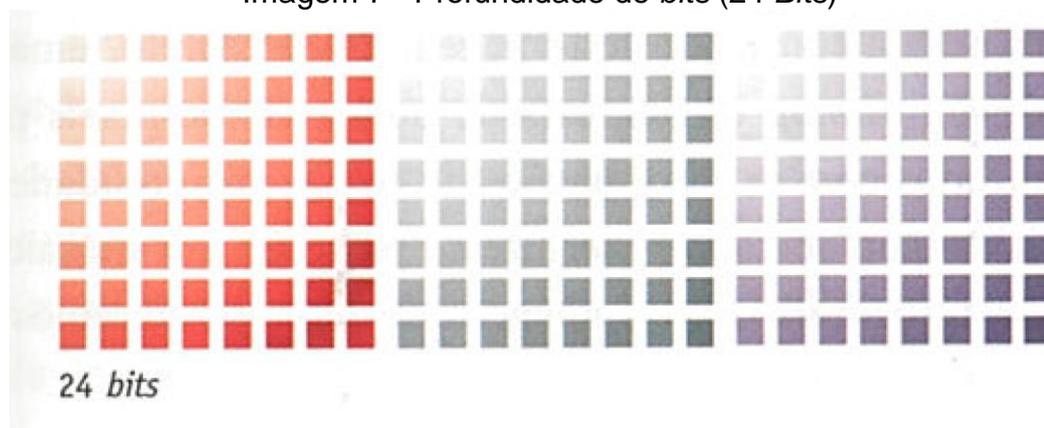
Fonte: Martins (2009, p.15)

Imagem 5 - Profundidade de *bits* (4 Bits)

Fonte: Martins (2009, p.15)

Imagem 6 - Profundidade de *bits* (8 Bits)

Fonte: Martins (2009, p.15)

Imagem 7 - Profundidade de *bits* (24 Bits)

Fonte: Martins (2009, p.15)

Martins (2009), explica que com 2 *bits*, há apenas as cores preto e branco. Para que se possa criar mais cores, é necessário utilizar mais *bits* por *pixel*. Para calcular a quantidade de tons, Martins (2009) coloca que é necessário elevar 2 (possibilidades de cores) ao número de *bits*. O Quadro 01 exemplifica como isso funciona.

Quadro 1 - Representação da profundidade de *bits*

Número de <i>bits</i> por <i>pixel</i>	Possibilidade de cores (2^{bit})
1	2 (preto e branco)
4	16
8	256
16	65.536
24	16 milhões

Fonte: Martins (2005) e Rocha (2016)

2.3.2 Resolução da imagem digital

A profundidade de *bits*, como vista no Quadro 01, pode variar entre dois tons (preto e branco) a 16 milhões de cores, passando por tons de cinza. Por causa disso, Cornell Library (2003b, tradução nossa) explica que a profundidade de *bits* pode ser dividida em três modos: bitonal, escala de cinza ou *grayscale* e modo de cor.

O modo bitonal, explica Cornell Library (2003b, tradução nossa) um *pixel* é formado por um *bit*. Esse *bit* representa 0 ou 1, preto ou branco. Por isso, não há uma gradação de cores entre eles.

A escala de cinza ou *grayscale*, cada *pixel* pode ser formado por dois até oito *bits*, explica Cornell Library (2016). Por causa disso, há uma variação de tons de cinza.

O modo de cor, segundo Cornell Library (2003b, tradução nossa), cada *pixel* pode ser formado por oito até 24 *bits*. Uma imagem de 24 *bits* de profundidade, esses *bits* são divididos em três grupos: oito *bits* para o tom vermelho, oito *bits* para o tom verde e oito *bits* para o tom azul. A combinação desses três tons cria em torno de 16 milhões de cores. Com relação ao modo de cor, Martins (2009) coloca que a cor de 24 *bits* é denominada de cor verdadeira, ou *true color*. Isso porque, com 16 milhões de tons, é possível se aproximar das cores que o olho humano é capaz de capturar.

No Quadro 02, foi elaborada uma comparação entre as os três modos de profundidade de *bit*.

Quadro 2 - Resolução de Cor

Bitonal	Escala de Cinza	Cor
$2^1 = 2$ tons (preto e branco) Cada pixel representa ou preto ou branco 1 <i>bit</i> é utilizado para representar cada pixel Não há gradação de cor	$2^8 = 256$ tons ou $2^{16} = 65.536$ tons Cada pixel representa uma variação de cinza. Pode ter de 8 a 16 <i>bits</i> por pixel. Há gradação de cor.	$2^{16} = 65.536$ tons ou $2^{24} = 16$ milhões Cada pixel representa uma cor Pode ter de 16 a 24 <i>bits</i> por <i>pixel</i> .

Fonte: Adaptado de Conselho Nacional de Arquivos (2010) e Empresa Brasileira de Pesquisa Agropecuária (2006)

A quantidade de pixels em uma imagem digital também tem relação com a resolução da mesma. O Dicionário Brasileiro de Terminologia Arquivística (2005, p. 148) define resolução como "Medida da nitidez de uma imagem, expressa no número de linhas discerníveis por milímetro." Há três tipos de resolução: resolução de imagem, resolução do monitor e resolução da impressão. Martins (2009), porém, esclarece que é comum confundi-las entre si. A resolução de imagem, que comumente é chamada apenas de resolução, é definida por ppi (*pixels per inch* ou *pixels* por polegada). A resolução do monitor, que é chamada de definição, é expressa pela medida de pontos na horizontal versus a medida de pontos na vertical. E a resolução de impressão, é aquela que é definida por dpi (*dots per inch* ou pontos por polegada), que é a quantidade de pontos que a impressora produz ao imprimir um documento. Ambas são importantes para imagem digital, uma vez que as duas primeira trata da resolução na tela da imagem, e a terceira, da impressão da imagem digital.

Martins (2009) coloca que, assim como em um texto, palavras se repetem, em uma imagem, cores, traços ou formas também podem se repetir. Para reduzir essas repetições, a imagem pode passar pelo processo de compactação ou compressão, que faz com que o seu tamanho também diminua.

Cornell Library (2003c, tradução nossa) explica que há dois tipos de compressão: *lossless* e *lossy*. A compressão do tipo *lossless* ou sem perdas, é aquela que é utilizada para "[...] abreviar o código binário sem descartar qualquer

informação, de forma que quando a imagem é descomprimida bit a bit é idêntico ao original." A compressão tipo *lossy* ou com perdas, "[...] utiliza um meio de descartar as informações menos significativo, com base na compreensão da percepção visual." (Cornell Library, 2003c, tradução nossa).

2.3.3 Tipos de imagem digital

A imagem digital pode ser armazenada de vários formatos diferentes. Cornell Library (2003c, tradução nossa) explica que formato de arquivo "[...]consiste nos bits que compõem a imagem e o cabeçalho informações sobre como ler e interpretar o arquivo. [...] variam em termos de resolução, profundidade de bits, recursos de cores e suporte para compressão e metadados." Os formatos mais comuns são o formato TIFF, GIF e JPEG.

A imagem TIFF (*Tagged-Image File Format*), de acordo com Martins (2009) esse formato pode ser utilizado por todos os tipos de *software* de edição, porque é como se fosse um mapa de *bits*. Por causa disso, tem uma elevada resolução. O formato GIF (*Graphics Interchange Format*), Martins (2009) explica, é o mais comum na *internet*. Esse tipo de imagem não comporta cores de 24 e 32 *bits*, somente as 256 cores (ou 8 *bits*), que são as cores constantes no formato RGB. O formato JPEG (*Joint Photographic Expert Group*), de acordo com Martins (2009), foi criado por fotógrafos profissionais, com o objetivo de realizar trocas entre eles. A imagem pode ser compactada, porém haverá perda de qualidade na imagem.

2.4 DIGITALIZAÇÃO DE UM ACERVO

O Conselho Nacional de Arquivos (2010, p. 5-6) coloca digitalização como sendo:

Entendemos a digitalização como 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.

São vários os objetivos para a digitalização de documentos. Entre eles, o Conselho Nacional de Arquivos (2010, p. 6) destaca:

- a) A contribuição para o acesso de forma ampla de documentos por meio de Tecnologias da Informação e Comunicação (TICs);
- b) Permissão de intercâmbio entre acervos por meio de redes de pesquisa informatizadas;
- c) Promoção, difusão e reprodução dos acervos não digitais em formatos diferentes dos originais;
- d) Preservação dos documentos em seus suportes originais.

A Empresa Brasileira de Pesquisa Agropecuária (2006), por sua vez, destaca os seguintes pontos:

- a) Facilidade de acesso aos documentos por outros públicos não usuais;
- b) Pesquisa de textos em formato eletrônico e de forma mais eficiente.

Para que a digitalização ocorra de forma eficaz, ela é dividida em etapas. São elas: a seleção, a conversão, os metadados, o controle de qualidade e o armazenamento.

A seleção, através de critérios pré-estabelecidos, deve indicar quais documentos deverão ser digitalizados. De acordo com a Cornell Library (2003d, tradução nossa), esses critérios são:

- a) Restrições legais (privacidade, copyright e direitos autorais);
- b) Características do documento original (gênero, tipo, formato físico, falhas);
- c) Disponibilidade da documentação (organização);
- d) Uso da digitalização (frequência de uso, benefícios);
- e) Financeiro (recursos).

A conversão, de acordo com Cornell Library (2003e, tradução nossa) envolve as características do documento a ser digitalizado. Elas determinarão as configurações utilizadas na digitalização como resolução, profundidade de bit e níveis de compressão. Essas características são: o tamanho do documento original, detalhamento da imagem, se o documento é colorido ou não e quais cores possui.

Os metadados são informações sobre os dados. Sua função é facilitar a pesquisa e a utilização de um documento em formato digital (EMPRESA..., 2006). De acordo com National Information Standards Organization (2004, tradução nossa) eles podem ser de quatro tipos:

- a) Descritivos: descreve o conteúdo intelectual do documento;
- b) Administrativos: fornecem informações que auxiliam na gestão do documento;
- c) Técnicos: relativo aos aspectos técnicos do documento, como compressão, resolução, formato, entre outros;
- d) Estruturais: relativo a informações da estrutura do documento, possibilitando a navegação.

O controle de qualidade, de acordo com Cornell Library (2003f, tradução nossa) é responsável por garantir que imagem digital tenha suas expectativas atendidas. A Empresa Brasileira de Pesquisa Agropecuária (2006) coloca que alguns dos critérios que devem ser inspecionados são:

- a) Tamanho da imagem
- b) Resolução;
- c) Formato de arquivo;
- d) Modo.

O armazenamento se refere onde estes documentos digitalizados serão mantidos. A Empresa Brasileira de Pesquisa Agropecuária (2006) coloca que o modo mais comum é que o documento mestre ou matriz fique em modo off-line, e que apenas as derivadas de acesso sejam acessíveis de modo online. Esse método visa proteger o documento digital.

2.5 CONFIGURAÇÃO DO REPRESENTANTE DIGITAL

O representante digital de um documento, de acordo com Conselho Nacional de Arquivos (2010), é o resultado da digitalização, ou seja, da conversão de um documento físico em digital. Ele pode ser utilizado para a preservação, acesso e/ou

reprodução. Para cada um desses modos é produzido na digitalização versões diferentes da mesma imagem.

Tanto a Conselho Nacional de Arquivos (2010) quanto o Empresa Brasileira de Pesquisa Agropecuária (2006) recomendam três tipos de imagens, porém com nomenclaturas diferentes. São matriz digital (ou imagem mestra), derivada de acesso (ou imagem de acesso) e a derivada de navegação (ou imagem em miniatura).

2.5.1 Matriz digital

A matriz digital, de acordo com a Empresa Brasileira de Pesquisa Agropecuária (2006), deve ser uma imagem mais fiel possível do documento original. Ou seja, não deve ter alterações como compactação ou edição. Deve possuir alta qualidade, pois a mesma será usada para criar as imagens derivadas (derivada de acesso e de navegação).

Com relação ao formato da imagem, o Conselho Nacional de Arquivos (2010) aponta como sendo o mais utilizado o TIFF, por causa de sua definição de cores elevada. A resolução óptica, a profundidade de bit e o modo de cor variam conforme o tipo de documento original, conforme mostra o Quadro 3.

Quadro 3 - Características da matriz digital

	Tipo de documento	Tipo de reprodução	Formato de documento digital	Resolução mínima, modo de cor
1	Textos impressos, sem ilustração, preto e branco	Bitonal	TIFF sem compressão	300 dpi, escala de 1:1, 4 bits modo bitonal
2	Texto impressos, com ilustrações e preto e branco.	Tons de cinza	TIFF sem compressão	300 dpi, escala 1:1, 4 bits modo bitonal
3	Texto impressos, com ilustração e cor	Cor	TIFF sem compressão	300 dpi, escala 1:1, 24 bits, cor
4	Manuscritos sem a presença de cor	Tons de cinza	TIFF sem compressão	300 dpi, escala 1:1, 8 bits
5	Manuscritos com a presença de cor	Cor	TIFF sem compressão	300 dpi, escala 1:1, 24 bits, cor
6	Fotografias (Preto e Branco e Cor)	Cor	TIFF sem compressão	300 dpi, escala 1:1, 24 bits, cor
7	Negativos fotográficos e diapositivos	Cor	TIFF sem compressão	3000 dpi, 24 bits, cor
8	Documentos cartográficos	Cor	TIFF sem compressão	300 dpi, escala 1:1, 24 bits, cor
9	Plantas	Bitonal	TIFF	600 dpi, 8 bits
10	Microfilmes e Microfichas	Tons de cinza	TIFF sem compressão	300 dpi, 8 bits, modo tons de cinza
11	Gravuras, cartazes e desenhos (Preto e Branco e Cor)	Cor	TIFF sem compressão	300 dpi, escala 1:1, 24 bits, cor

Fonte: Conselho Nacional de Arquivos (2010)

2.5.2 Derivada de acesso

A derivada de acesso, de acordo com a Empresa Brasileira de Pesquisa Agropecuária (2006), deve ser criada a partir da matriz digital. O seu tamanho deve ser menor, para facilitar o carregamento da imagem, que poderá ser acessada através da rede. A qualidade da imagem pode ser inferior a da matriz digital, porém não muito baixa, pois poderá ser utilizada como imagem de impressão ou para *download*.

O Conselho Nacional de Arquivos (2010) recomenda para as derivadas de acesso, o formato JPEG. A resolução deverá ser mais baixa que a matriz digital, uma vez que ela será acessada em rede. Diferente da matriz, a derivada de acesso poderá ter tratamento de imagem, para melhorar a visualização e impressão. A resolução óptica, a profundidade de bit e o modo de cor, assim como a matriz digital, variam conforme o tipo de documento, conforme mostra o Quadro 4.

Quadro 4 - Características da derivada de acesso

Tipo de documento	Tipo de reprodução	Formato de documento digital	Compactação, resolução, modo de cor
Texto	Escala de cinza ou cor	JPEG	20:1, redimensionamentos para 640x480, escala de cinza ou cor
Fotografia	Escala de cinza ou cor	JPEG	7:1 – 10:1 para escalas de cinza; 10:1 - 20:1 para cor Redimensionamento para 640x480, escala de cinza ou cor
Mapas	Escala de cinza ou cor	JPEG	20:1 – dependendo do tamanho do original 1200 pixels para mapas grandes 640x480 para mapas pequenos
Negativos e transparências	Escala de cinza ou cor	JPEG	7:1 – 10:1 para escalas de cinza; 10:1 - 20:1 para cor Redimensionamento para 640x480

Fonte: Conselho Nacional de Arquivos (2010)

2.5.3 Derivada de navegação

A derivada de navegação, de acordo com a Empresa Brasileira de Pesquisa Agropecuária (2006), tem por objetivo auxiliar o usuário, para que ele escolha se quer visualizar a derivada de acesso. Por isso, deve ser bem pequena, menor que ela, para uma rápida visualização em rede.

O Conselho Nacional de Arquivos (2010) recomenda como formato da derivada de navegação os formatos GIF ou JPEG. Porém, o último não é o mais indicado para imagens cujos desenhos e textos com linhas muito finas ou caracteres muito pequenos. A resolução óptica, a profundidade de bit e o modo de cor, assim como os demais representantes digitais, variam conforme o tipo de documento, conforme mostra o Quadro 5.

Quadro 5 - Características da derivada de navegação

Tipo de documento	Tipo de reprodução	Formato de documento digital	Compactação, resolução, modo de cor
Texto	Escala de cinza ou cor	GIF	Compactação típica para formato GIF, resolução: redimensionamento do original para 150-200 pixels em relação a altura
Fotografia	Escala de cinza ou cor	GIF ou JPEG	Compactação típica para formato GIF, redimensionamento do original para 150-200 pixels em relação a altura
Mapas	Escala de	GIF ou JPEG	Compactação típica para formato GIF,

	cinza ou cor		redimensionamento do original para 150-200 pixels em relação a altura.
Negativos e transparências	Escala de cinza ou cor	GIF ou JPEG	Compactação típica para formato GIF, redimensionamento do original para 150-200 pixels em relação a altura

Fonte: Conselho Nacional de Arquivos (2010)

2.6 ESCOLHA DO SCANNER

Outro fator que influencia diretamente a digitalização é a escolha do *scanner*. O equipamento deve ser escolhido de acordo com as características do documento a ser digitalizado, como tamanho, tonalidade e tipo de documento.

Os tipos de *scanners* utilizados para a digitalização, visando a preservação digital são: *scanner* de mesa (*flat bed*), *scanner* planetário, câmera digital, *scanners* para digitalização de negativos e dispositivos fotográficos, *scanners* para digitalização de microformas, *scanners* de produção e alimentação automática (Conselho Nacional de Arquivos, 2010). O Quadro 6 apresenta os tipos de *scanners* e também para quais documentos são recomendados.

Quadro 6 - Tipos de Scanner

Scanner	Documentos recomendados
<i>Scanner</i> de mesa	Documentos planos em folhas simples e ampliações fotográficas em bom estado de conservação.
<i>Scanner</i> planetário	Documentos planos em folhas simples, documentos encadernados com lombada frágil e documentos frágeis.
Câmera digital	Documentos arquivísticas originais, como daguerreótipos, albuminados e ferrótipos.
<i>Scanners</i> para digitalização de negativos e dispositivos fotográficos	Negativos e diapositivos em bom estado de conservação.
<i>Scanners</i> para digitalização de microformas	Diferentes tipos de microformas.
<i>Scanners</i> de produção e alimentação automática	Documentos em excelente estado de conservação.

Fonte: Adaptado de Conselho Nacional de Arquivos (2010) e Empresa Brasileira de Pesquisa Agropecuária (2006).

3 METODOLOGIA

Metodologia, de acordo com Marconi e Lakatos (2010), deve responder as questões de pesquisa tais *como?*, *com o quê?*, *onde?* e *quanto?*. Essas questões são respondidas através da delimitação do tipo de pesquisa.

A pesquisa, de acordo com Gil (2010) é um processo sistemático que responde questões levantadas durante a pesquisa. A pesquisa pode ser classificada conforme a sua natureza, objetivo, procedimento e objeto de coleta. Serão apresentados neste capítulo como a pesquisa será realizada, conforme a classificação.

3.1 SEGUNDO A NATUREZA

De acordo com Silva e Menezes (2005), a pesquisa aplicada é aquela que tem por objetivo gerar um conhecimento que possa ter utilidade prática. Gil (2010) complementa que o conhecimento adquirido nesse método de pesquisa será utilizado em uma situação em específico. Essa pesquisa apresenta um estudo investigativo das configurações para digitalização de documentos textuais datilografados em suas diferentes apresentações e com diferentes características. Ao fazer esse estudo, foram confrontadas as recomendações do CONARQ para a digitalização de documentos textuais. A partir desses resultados, constataram-se quais estavam de acordo com as reais necessidades dos documentos e quais eram equivocados.

3.2 SEGUNDO A ABORDAGEM DO PROBLEMA

A pesquisa qualitativa, de acordo com Silva e Menezes (2005), trata de questões que não podem ser traduzida em números. O ambiente em que ocorre a pesquisa é a fonte dos dados, e o pesquisador, que age como observador, é o instrumento de pesquisa. Isso porque os dados são descritivos, e conhecimento do pesquisador sobre o que está sendo analisado é fundamental para a interpretação do que será observado.

A pesquisa irá tratar da análise de qualidade da digitalização. Não haverá questões quantitativas, tratando-se de uma pesquisa unicamente qualitativa. Para

isso, serão observadas as recomendações sobre digitalização do CONARQ para documentos textuais. Os critérios analisados foram a resolução da imagem e modo de cor. E os dados obtidos através dessa análise foram quais são as configurações mais adequadas para documentos textuais datilografados com diferentes características.

3.3 SEGUNDO OBJETIVO

A pesquisa exploratória, de acordo com Silva e Menezes (2005), envolve a análise bibliográfica e também o estudo de experiências práticas. Como apresentado anteriormente, o objetivo desse estudo foi a investigação das configurações para digitalização de documentos textuais datilografados em suas diferentes apresentações com diferentes características. Para alcançar esse objetivo, a pesquisa foi dividida em quatro objetivos específicos. Cada um deles tem uma função, colaborando para a pesquisa no todo.

O Quadro 7 demonstra de forma ilustrada como cada objetivo específico do trabalho foi desenvolvido durante a pesquisa.

Quadro 7 - Relação dos Objetivos específicos e método

Objetivo	Método
a) Analisar as recomendações existentes para digitalização de documentos textuais;	Estudo teórico das recomendações, com destaque à recomendação do CONARQ.
b) Experimentar configurações para a geração de matrizes de documentos textuais datilografados	Definição das amostras e realização das digitalizações das amostras em várias configurações, comparando os resultados.
c) Analisar as experimentações realizadas com relação as recomendações existentes para digitalização de documentos textuais	Verificação dos resultados da digitalização das amostras considerando recomendações do CONARQ, com foco no modo e resolução.
d) Apresentar sugestões para digitalização de documentos datilografados	Sugestões de como digitalizar esse tipo de documento, considerando os resultados.

Fonte: Elaborado pela autora

3.4 SEGUNDO PROCEDIMENTO E COLETA DE DADOS

O estudo de caso, de acordo com Silva e Menezes (2005), faz um estudo exaustivo de poucos objetivos. Esse tipo de estudo tem como foco apenas uma

unidade: um indivíduo, um pequeno grupo, uma instituição um programa, ou um evento.

Gil (2010, p. 117), coloca que as etapas a serem seguidas em uma pesquisa de estudo de caso são: "Formulação do problema ou das questões de pesquisa, definição das unidades-caso, seleção dos casos, elaboração do protocolo, coleta de dados, análise e interpretação dos dados, redação do relatório."

Essa pesquisa se qualifica com um estudo de caso pois o problema apresentado é focado em um grupo específico de documentos, que são os textuais datilografados. O grupo estudado continha seis amostras de documentos datilografados com diferentes características. Esses documentos pertenciam a um professor universitário gaúcho, e é composto por provas, trabalhos e demais atividades de aula. Essas amostras foram digitalizadas, gerando as matrizes digitais. Partindo do documento gerado, foi analisado o modo de cor e a resolução.

3.5 ANÁLISE DOS DADOS E APRESENTAÇÃO DOS RESULTADOS

Para essa análise foi utilizado o scanner Canon iMager Fórmula DR-G1100, com alimentação automática. A escolha desse *scanner* ocorreu pelo fato da investigação considerar um caso de digitalização de uma grande quantidade de documentos, que não eram encadernados. Optou-se por utilizar a configuração padrão do *scanner*, sem o uso de recursos de processamento de imagem, pois o trabalho teve por foco o estudo de modo e resolução das matrizes. Por isso, o formato escolhido para os documentos digitais foi o TIFF, recomendado para matrizes digitais. A amostra é composta por seis documentos datilografados. Cada um dos documentos possui características especiais, que os diferem uns dos outros.

A análise foi realizada em duas partes: modo e resolução. Para isso, utilizou-se como referência para essa comparação, os valores da matriz digital. Ela tem como objetivo preservar o documento original. Para ambas, foi utilizado o software GIMP2. Ele é utilizado para o processamento de imagens. Possui diversos recursos, como utilizados para tratamento e também para ampliação da imagem, sem alterar as propriedades. (GIMP, 2016). Diferente de um visualizador de imagens comum, o GIMP permite utilizar o formato TIFF sem compressão da imagem. Ele mostra também a resolução (*dpi*) em números, facilitando a comparação de resolução.

A primeira parte da análise foi com relação ao modo. Cada uma das seis amostras foi digitalizada nos modos bitonal, tons de cinza e cor. A resolução utilizada foi a indicada na recomendação do CONARQ, de 300dpi. Foram gerados então 18 documentos digitais, três para cada documento físico. Para facilitar a comparação, cada um dos documentos digitais recebeu um código. A relação de códigos se encontra no Quadro 10 do Apêndice A.

A Quadro 8 uma descrição geral dos documentos e as configurações testadas para cada um deles.

Quadro 8 – Relação de amostras e configurações

Nº da amostra	Descrição do Original	Configurações testadas
1	Documento datilografado e mimeografado com texto em roxo, com marcações manuscritas a lápis em papel amarelado.	Bitonal, 300 dpi; Tons de Cinza, 300 dpi Colorido, 300 dpi
2	Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência	Bitonal, 300 dpi; Tons de Cinza, 300 dpi Colorido, 300 dpi
3	Documento datilografado com o uso de papel carbono, em papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais claro.	Bitonal, 300 dpi; Tons de Cinza, 300 dpi Colorido, 300 dpi
4	Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente.	Bitonal, 300 dpi; Tons de Cinza, 300 dpi Colorido, 300 dpi
5	Documento com texto datilografado e digitado em preto em papel vegetal.	Bitonal, 300 dpi; Tons de Cinza, 300 dpi Colorido, 300 dpi
6	Documento datilografado em matriz de mimeografo com o texto em roxo. Apresenta manchas roxas em torno do texto.	Bitonal, 300 dpi; Tons de Cinza, 300 dpi Colorido, 300 dpi

Fonte: Elaborado pela autora

A análise comparou os três documentos digitais de cada amostra. Observou-se qual deles era mais fiel ao original, escolhendo-o como representante digital do documento físico. Após isso, foram apontadas quais são as configurações de modo indicadas pelo CONARQ de acordo com as características do documento físico. Por fim, comparou-se o modo escolhido de acordo com a análise e o indicado pelas recomendações, comparando-os, e caso necessário, sugerindo opções. Esses resultados são apresentados diretamente no texto.

A segunda análise foi o comparativo de resolução. O indicado nas recomendações do CONARQ (2010) é de 300 dpi. Utilizou-se então para comparação 300dpi e 600dpi, o dobro da resolução indicada. Os documentos então

foram digitalizados com essa configuração de resolução. Para selecionar o modo testado, foi utilizado o resultado da análise anterior. Sendo assim, foram gerados 14 documentos. Para facilitar a comparação, cada um dos documentos digitais recebeu um código. A relação de códigos se encontra no Quadro 11 do Apêndice A.

A Quadro 9 uma descrição geral dos documentos e as configurações testadas para cada um deles.

Quadro 9 - Relação de amostras

Nº da amostra	Descrição do Original	Configurações testadas
1	Documento datilografado e mimeografado com texto em roxo, com marcações manuscritas a lápis em papel amarelado.	Cor, 300 dpi Cor, 600 dpi
2	Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência	Cor, 300 dpi Cor, 600 dpi
3	Documento datilografado com o uso de papel carbono, em papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais claro.	Cor, 300 dpi Cor, 600 dpi Tons de Cinza, 300 dpi Tons de Cinza, 600 dpi
4	Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente.	Tons de Cinza, 300 dpi Tons de Cinza, 600 dpi
5	Documento com texto datilografado e digitado em preto, em papel vegetal	Tons de Cinza, 300 dpi Tons de Cinza, 600 dpi
6	Documento datilografado em matriz de mimeografo com o texto em roxo. Apresenta manchas roxas em torno do texto.	Cor, 300 dpi Cor, 600 dpi

Fonte: Elaborado pela autora

A comparação foi realizada com a utilização de zoom. Para os documentos com resolução de 300dpi foi utilizado zoom de 80%, e para os de 600dpi, de 40%. E uma segunda análise de cada imagem foi feita com zoom de 400% na imagem de 300dpi e zoom 200% para imagem de 600dpi, mostrando detalhes das letras e palavras. Foi utilizado essa proporção para manter o tamanho capturado do documento com o mesmo tamanho em centímetros. Assim como na primeira análise, os resultados são apresentados diretamente no texto.

4 ANÁLISE DOS RESULTADOS

Para a realização dos testes, foram selecionados sete amostras de documentos do gênero textual. Cada documento foi digitalizado em cada modo (bitonal, tons de cinza e cor) e também com resoluções diversas, conforme apresentado nos Quadros 8 e 9 na metodologia. Para efetuar a análise, foi atribuído para cada um deles um código. A relação de códigos por documento encontra-se no Quadro 10 do Apêndice A.

O *scanner* escolhido foi o do tipo produção CANON IMager Fórmula DR-G1100, com alimentador automático. Esse tipo de *scanner* pode ser utilizado para qualquer gênero ou tipo de documento em folha solta, desde que não seja material frágil.

A apresentação dos resultados ocorrerá em duas partes. A primeira será com relação ao modo. E a segunda, com relação à resolução. Serão apresentados um recorte do documento, para fins de comparação. O *software* utilizado é o GIMP. Ele é utilizado para tratamento de imagem.

4.1 ANÁLISE COM RELAÇÃO AO MODO

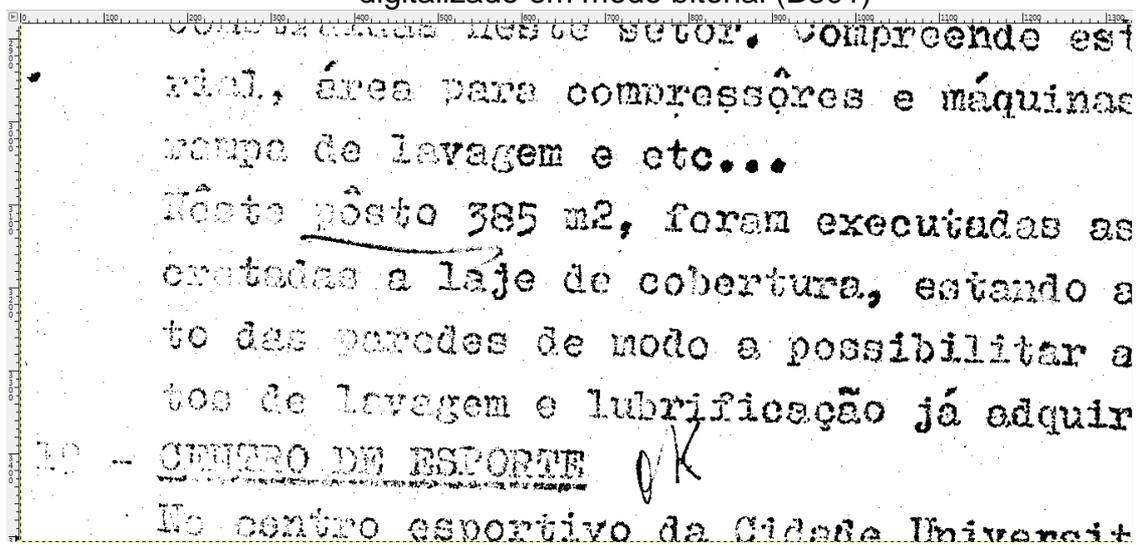
Essa análise tem por objetivo fazer o comparativo do modo (bitonal, tons de cinza e cor) para cada documento da amostra. Através delas, foi verificado quais os modos são mais fiéis ao documento original e comparando com as especificações da literatura, tendo como base as diretrizes do Conselho Nacional de Arquivos (2010). Para essa análise, todas as imagens analisadas estão na resolução de 300 *dpi*, que é a resolução recomendada para matriz digital de documentos textuais. Destacamos que, por serem matrizes, as imagens digitais não foram processadas.

O texto está organizado da seguinte forma: para cada amostra, apresenta as imagens digitalizadas nos modos bitonal, tons de cinza e colorido, destacando características observadas decorrentes da digitalização. A seguir, apresenta uma comparação envolvendo as três imagens, destacando as mais fiéis ao original. Finalizando, analisa o resultado obtido frente às recomendações do Conselho Nacional de Arquivos.

A primeira amostra a ser analisada corresponde a um documento datilografado e mimeografado com texto em roxo. Há marcações manuscritas a

lápiz. O papel é amarelado. As Imagens 8, 9 e 10 apresentam as imagens digitalizadas desse documento nos modos bitonal, tons de cinza e colorido, respectivamente.

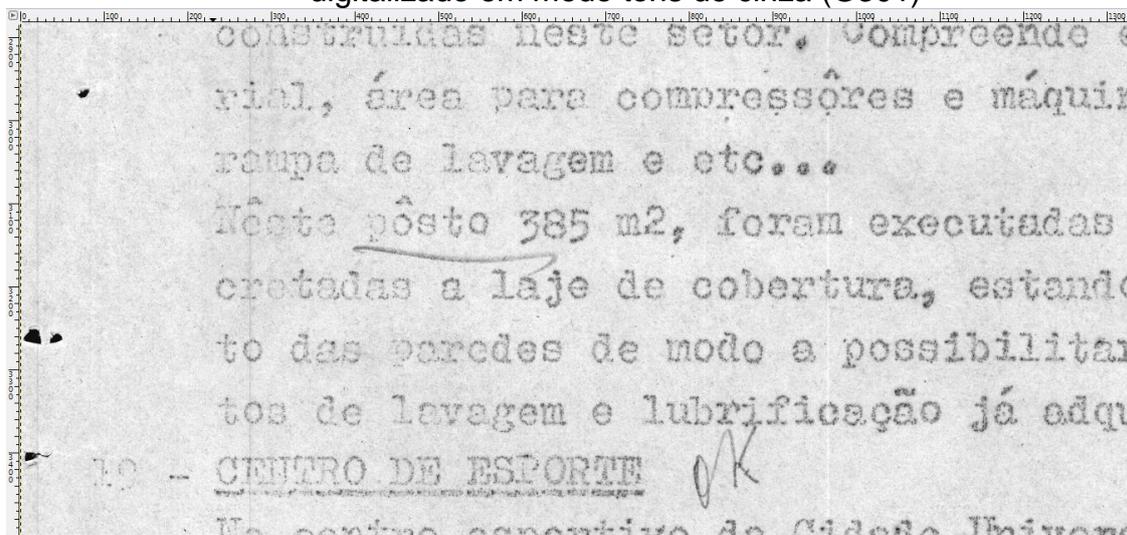
Imagem 8 - Documento datilografado e mimeografado com texto em roxo digitalizado em modo bitonal (B301)



Fonte: Autor

Como mostra a Imagem 8, o documento no modo bitonal (B301) apresenta o fundo branco e o texto em preto. No fundo, retirou as manchas, que são claras, mantendo alguns pontos dessas manchas que são mais escuros, principalmente ao redor do texto. Os pontos marcados a lápis se destacam também em preto. As letras têm os seus contornos com falhas. O texto tem alguns pontos difíceis de compreender, como podemos notar na palavra centro. Isso deve-se ao fato que em muitas palavras a impressão está fraca no original, com muitos pontos sendo assim representado como branco.

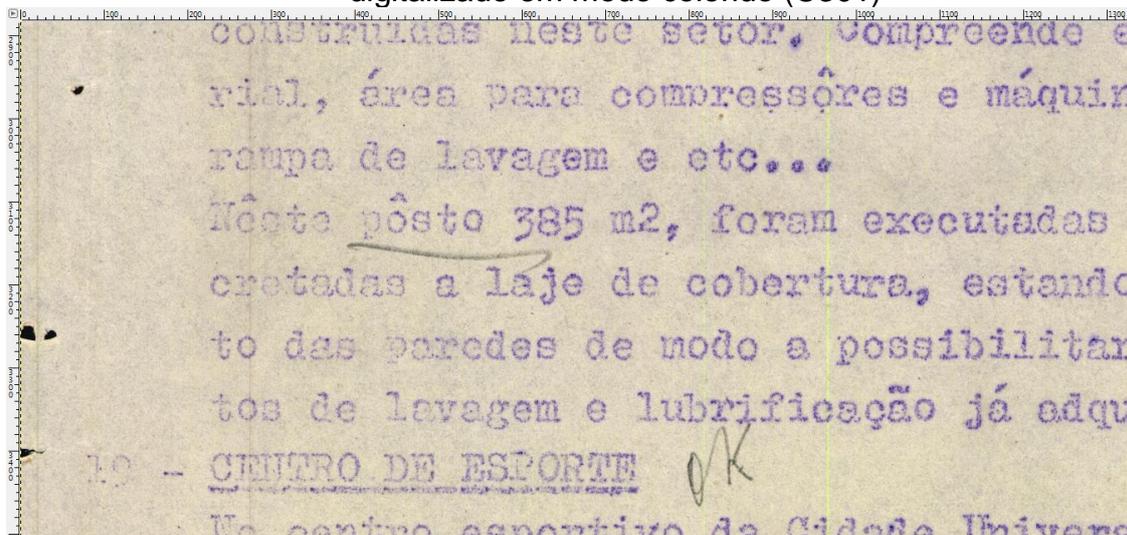
Imagem 9 - Documento datilografado e mimeografado com texto em roxo digitalizado em modo tons de cinza (G301)



Fonte: Autor

Na Imagem 9 temos o documento (G301) no modo tons de cinza. É possível observar que o documento tem o fundo cinza com manchas. O cinza do fundo é acentuado, pois o papel é amarelado. O texto também é cinza, porém mais escuro. O fato do fundo e do texto serem cinza, não traz muito destaque ao texto. Os contornos das letras estão falhos assim como no documento bitonal (B301). Aqui, eles se misturam com o fundo, o que torna a leitura mais difícil. Os pontos a lápis se destacam mais que o texto datilografado.

Imagem 10 - Documento datilografado e mimeografado com texto em roxo digitalizado em modo colorido (C301)



Fonte: Autor

O documento apresentado na Imagem 10 (C301) está no modo colorido. Representa o fundo em um tom amarelado (original) e o texto em roxo. Nota-se

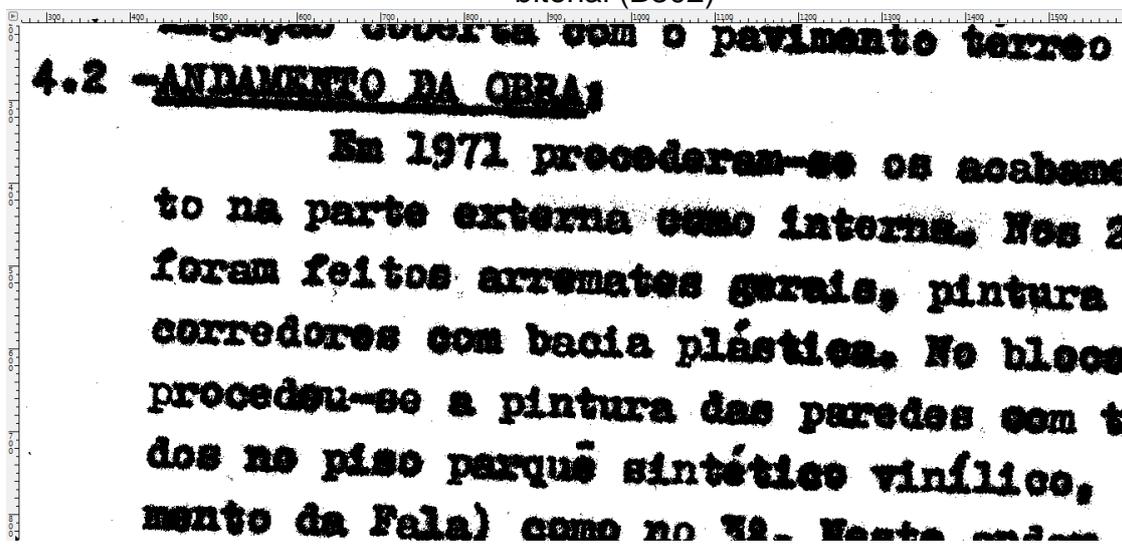
algumas manchas no fundo e no texto. Os pontos marcados a lápis aparecem bem, em tom de cinza amarronzado. O contorno do texto também apresenta falhas, como nos outros dois. Porém, elas são menos perceptíveis, e não prejudicam a leitura do texto.

Comparando as três imagens, a Imagem 10 (cor) representa o documento digital de forma mais fiel. Esse documento não possui imagens, e possui texto impresso datilografado, com fundo amarelado. Entre as recomendações apresentadas no Quadro 03, as que melhor se encaixam são: (1) texto impresso sem ilustração e preto e branco sem manchas; (2) texto impresso sem ilustração e com mancha (considerando o papel amarelado como um tipo de mancha); (3) texto impresso com ilustrações e cor e mancha; (4) texto manuscrito com presença de cor.

Observa-se que nenhuma recomendação representa exatamente o caso, e conclui-se que há necessidade de recomendações mais específicas. As recomendações (1) e (2) poderiam ser utilizadas, visto que o documento original não possui grandes manchas, possui a letra em um único tom (roxo) e não possui ilustrações. Entretanto, pela a análise realizada, observou-se que a cor foi determinante nesse documento, função do papel amarelado e do texto em roxo. Partindo do pressuposto da necessidade de considerar a cor do texto e do fundo, as recomendações mais adequadas seriam as (3) e (4), que sugerem o modo colorido.

A segunda amostra é o documento datilografado, com o uso de carbono, com letras em tom bem escuro. O texto está em preto. O papel é amarelado, quase sem transparência. As Imagens 11, 12 e 13 apresentam as digitalizações em modo bitonal, tons de cinza e colorido, respectivamente.

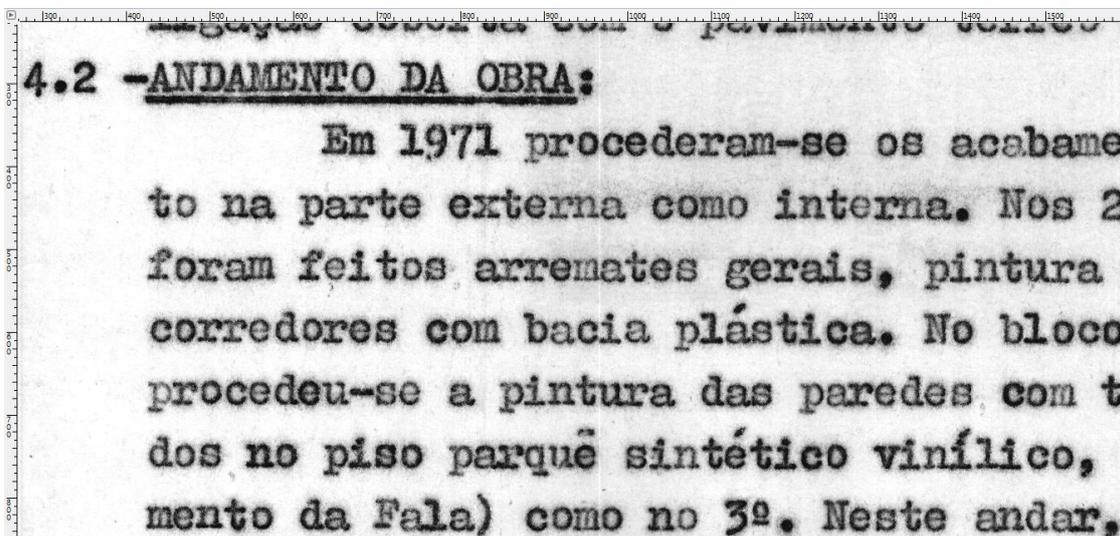
Imagem 11 - Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência, digitalizado em modo bitonal (B302)



Fonte: Autor

A Imagem 11 apresenta o documento (B302) no modo bitonal. O fundo da imagem é branco, e o texto preto. Alguns pontos do texto são difíceis de ler, como em ANDAMENTO. Nesse caso, o contorno das letras está grosso e apresenta falhas, provavelmente em função do uso do papel carbono. Há também pequenas manchas, como a que ocorre entre as palavras externa e como, também provavelmente em decorrência do uso do papel carbono. Essa sombra é mais clara que o texto, mas no momento da digitalização teve suas partes mais escuras interpretadas como preto.

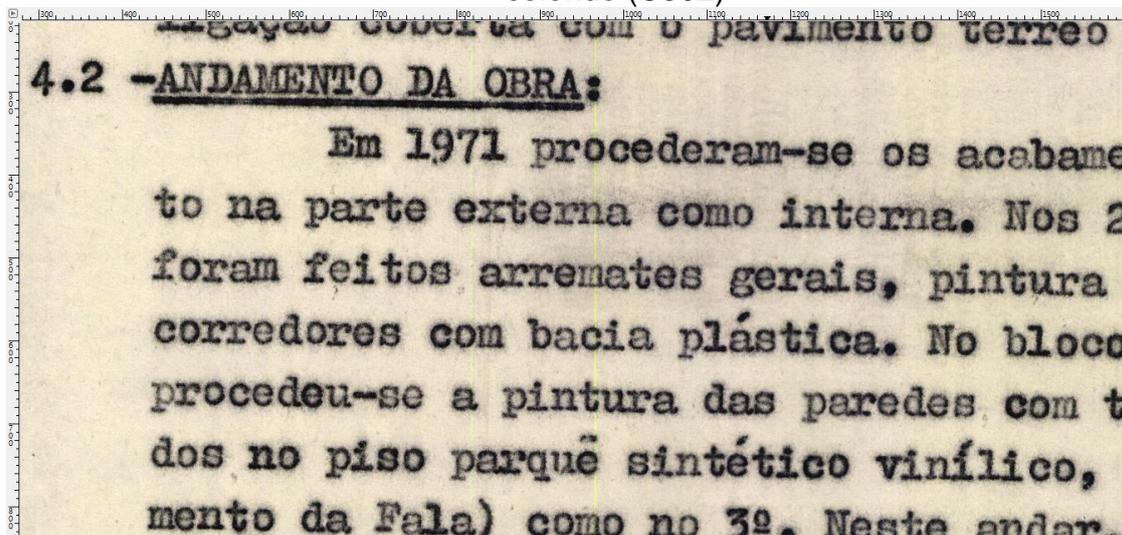
Imagem 12 - Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência, digitalizado em modo tons de cinza (G302)



Fonte: Autor

A Imagem 12 mostra o documento (G302) no modo de tons de cinza. Ele tem o fundo claro, quase branco, com manchas cinza. Essas manchas também aparecem no documento original. O fundo cinza é explicado pelo fato de o papel do documento original ser amarelo. O texto é mais claro que na imagem anterior (Imagem 11). O contorno mais fino, mas apresenta falhas. Aparecem algumas manchas ao redor do texto, como a que ocorre entre as palavras externa e como. Elas, porém, não atrapalham a leitura e o entendimento do documento.

Imagem 13 - Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência, digitalizado em modo colorido (C302)



Fonte: Autor

A digitalização em modo colorido (C302), mostrado na Imagem 13, apresenta o fundo amarelado, como o documento original. Apresenta manchas em cinza em alguns pontos. Essas manchas existem no documento original. O texto está em preto, com falhas no contorno. Em alguns pontos do texto percebe-se uma sombra, no mesmo ponto da Imagem E. Mas está mais fraca.

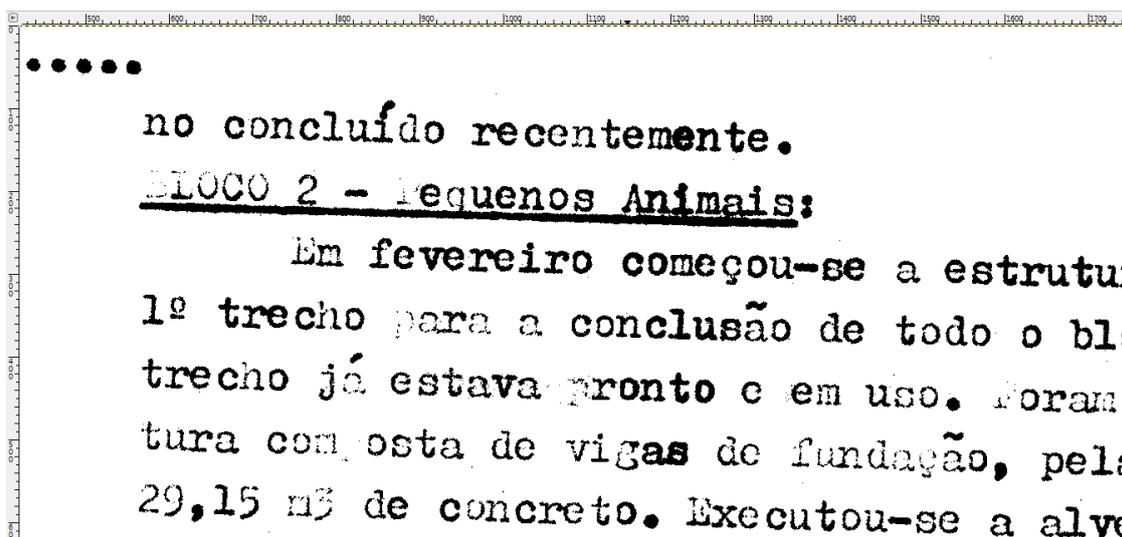
Como esse documento tem o texto em preto, e não contem imagens, as recomendações que mais se encaixam são: (1) texto impresso sem ilustração, preto e branco sem mancha; (2) texto impresso com ilustrações, preto e branco, com mancha.

Observa-se que dentre as recomendações apresentadas, a (1) se encaixaria melhor, em teoria. Na prática, porém, ela não se mostrou suficiente, pois o carbono deixa manchas no fundo e letras com contorno grosso, que são capturadas pelo modo bitonal em preto, como apresentado na Imagem 11 (documento B302). Em modo tons de cinza, como sugerem a recomendação (2) o texto ficou legível, porém as manchas do fundo ainda se destacam. Isso ocorreu por causa da cor de fundo do documento original, que é amarelado, e não branco. Comparando as três imagens, a Imagem 13 (em cor) representa o documento digital de modo mais fiel. Considerando a escolha pelo modo colorido, esse resultado não confere com as recomendações.

A terceira amostra é do documento datilografado em papel carbono. O papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais

claro. E o papel amarelado. As imagens 14, 15 e 16 apresentam as digitalizações em modo bitonal, tons de cinza e colorido, respectivamente.

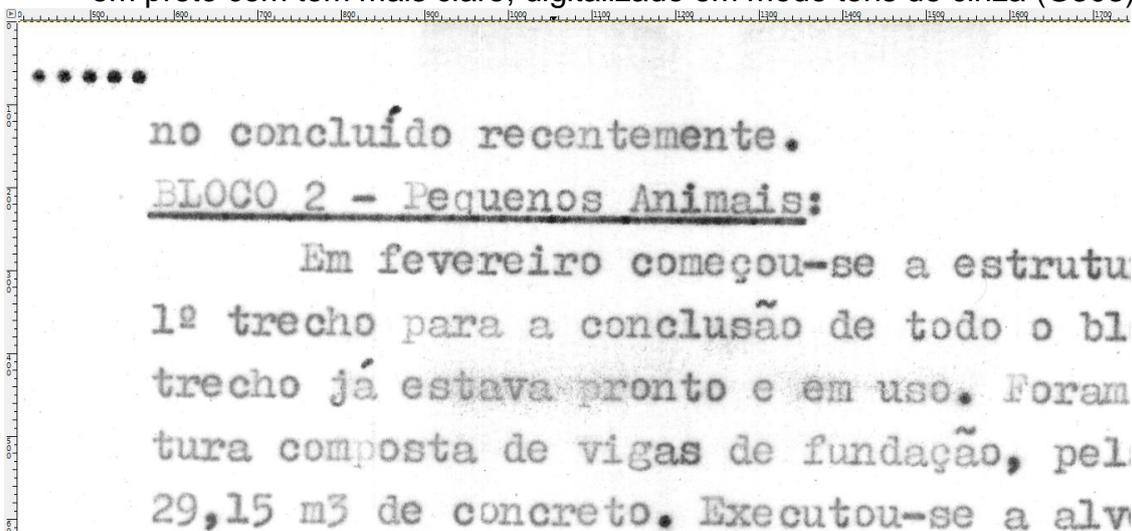
Imagem 14 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo bitonal (B303)



Fonte: Autor

A Imagem 14 mostra o documento (B303) em modo bitonal. O fundo é branco com o texto em preto. Não apresenta manchas. Em alguns pontos do texto, como em **BLOCO**, o contorno está com falhas. Em outros, como **animais**, ele é mais grosso. No documento original, **BLOCO** está mais claro, e **animais** tem pontos mais escuro, como uma sombra, o que explica as falhas. Isso tudo não compromete o entendimento do texto. Também não há manchas em torno do texto na imagem.

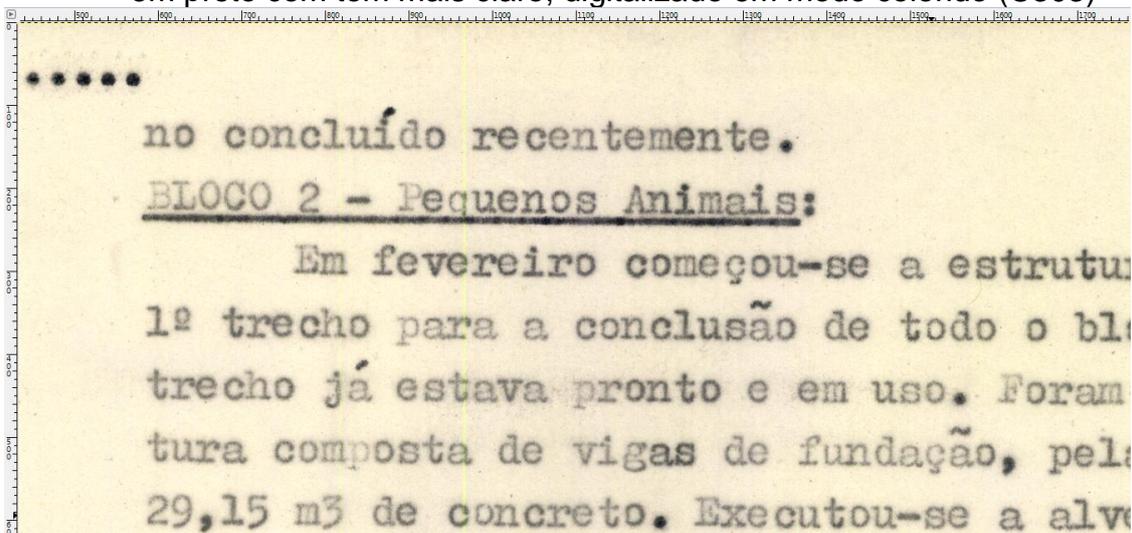
Imagem 15 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo tons de cinza (G303)



Fonte: Autor

A Imagem 15 apresenta o documento (G303) em tons de cinza. O fundo é branco com alguns pontos com manchas. No documento original, o ponto onde está a mancha (palavra estava) está borrado. O texto está em cinza. Em alguns pontos, ele é mais claro e com falhas no contorno. Diferente do da versão em bitonal (Imagem 14, B303), os contornos das letras estão mais padronizados, não há alguns mais largos e fortes do que outros. A leitura do documento não é prejudicada, apesar das falhas e das manchas.

Imagem 16 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo colorido (C303)



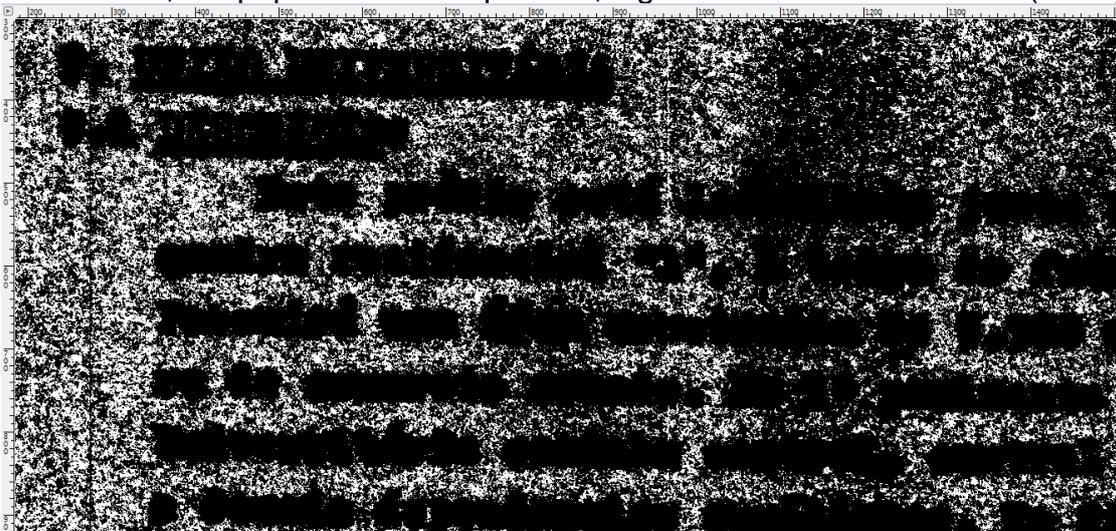
Fonte: Autor

O documento (C303) no modo colorido, mostrado na Imagem 16, apresenta o fundo amarelado com algumas manchas mais claras. O documento original também tem o fundo amarelado, e no ponto da mancha está borrado. O texto está com o contorno limpo, quase sem falhas. Em alguns pontos, o texto é mais claro. No documento original também está assim. Nada disso, porém, interfere na leitura do texto.

Comparando as três imagens, tanto a Imagem 15 (tons de cinza) como a Imagem 16 (cor) seriam suficientes para representar o documento original. Observa-se que as amostras dois e três diferem-se basicamente pelo tom do texto, entretanto nesse caso, considerou-se a imagem em tons de cinza (Imagem 15) suficiente, enquanto que na anterior, apenas a colorida (Imagem 16) foi considerada. Isso demonstra a complexidade em criar recomendações para digitalização, pois pequenas variações podem causar escolhas diferentes. Também mostra a importância de além de seguir recomendações, de realizar experimentações práticas.

A quarta amostra é um documento datilografado com o uso de papel carbono, em papel mais fino e transparente que os das amostras anteriores. O texto é claro. As Imagens 17, 18 e 19 apresentam as digitalizações em modo bitonal, tons de cinza e colorido.

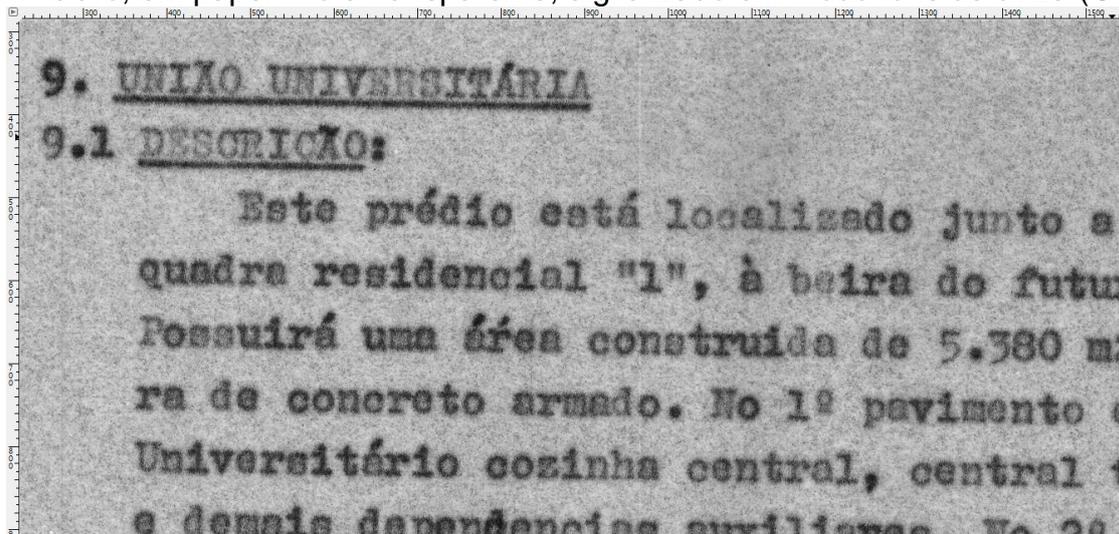
Imagem 17 - Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente, digitalizado em modo bitonal (B304)



Fonte: Autor

O documento em modo bitonal (B304), mostrado na Imagem 17, aparece com o fundo manchado com pontos pretos em todo o documento. Percebe-se algumas manchas mais escuras, formando linhas. É onde deveria estar o texto. Porém, é impossível distinguir qualquer palavra.

Imagem 18 - Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente, digitalizado em modo tons de cinza (G304)

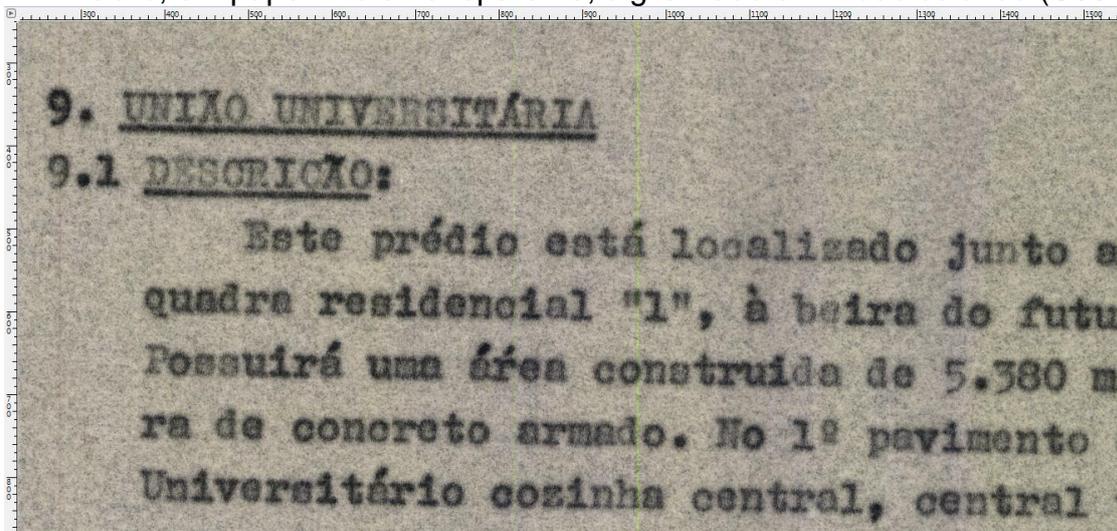


Fonte: Autor

No documento em tons de cinza (G304) mostrado na Imagem 18, o fundo é cinza. A cor de fundo não é uniforme, mas sim manchada e escura se comparado ao original. O papel do documento original é amarelado, o que explica o fundo cinza na digitalização. E por ser mais fino e transparente, ficou mais escuro. Em alguns pontos, como na palavra **LOCALIZANDO**, o texto está mais claro. Isso também ocorre no documento original. Mesmo assim, ainda é possível entender o que está escrito.

O documento em modo colorido (C304) na Imagem 19, é muito similar ao documento em tons de cinza (Imagem 18). O modo colorido captou a nuance amarelado do documento.

Imagem 19 - Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente, digitalizado em modo colorido (C304)



Fonte: Autor

Ao comparar as três imagens (17, 18 e 19), as mais fiéis ao documento original são a 18 e a 19, que representam os documentos em tons de cinza (G304) e colorido (C304), com a preferência para o documento em tons de cinza. Isso porque em modo colorido apresentou o fundo com a cor alterada. Isso provavelmente ocorreu por causa da transparência do papel.

Observando as três análises feitas com o texto datilografado com o uso de papel carbono, nota-se escolhas diferentes. Enquanto que na amostra anterior não houve diferenças significativas entre os documentos em modo tons de cinza e colorido, neste o modo colorido não se demonstrou satisfatório.

A quinta amostra é de um documento com texto datilografado e digitado em preto. O papel desse documento é vegetal, ou seja, transparente. As Imagens 20, 21 e 22 mostram o documento digitalizado em modo bitonal, tons de cinza e colorido respectivamente.

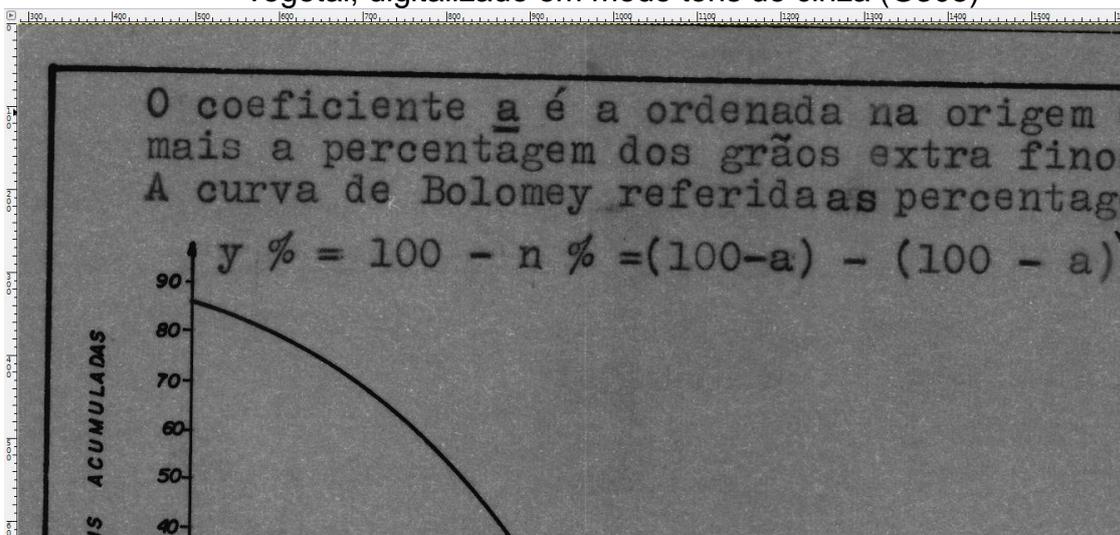
Imagem 20 - Documento com texto datilografado e digitado em preto, em papel vegetal, digitalizado em modo bitonal (B305)



Fonte: Autor

O documento em modo bitonal (B305), representado pela Imagem 20, apresenta o fundo completamente preto. Não é possível identificar nenhum texto escrito. Os únicos pontos claros são duas falhas onde anteriormente havia um grampo. Isso deve ter ocorrido por causa da combinação do papel do documento vegetal, que é transparente, com o modo de cor do documento, que é bitonal.

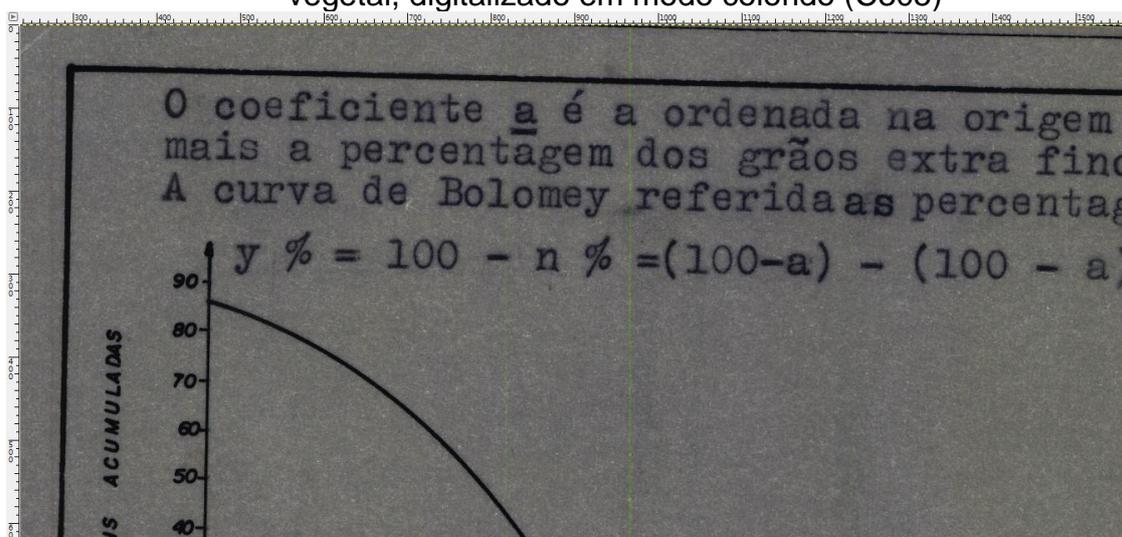
Imagem 21 - Documento com texto datilografado e digitado em preto, em papel vegetal, digitalizado em modo tons de cinza (G305)



Fonte: Autor

O documento em tons de cinza (G305), mostrado na Imagem 21, apresenta o fundo em cinza. A diferença entre a digitalização em tons de cinza e preto é que ele capta diversos tons, dentro da gama de cinza. Isso explica o fundo do documento em cinza. O texto datilografado parece ser mais claro que o texto impresso. Não há borrados nem manchas ao redor do texto. Além disso, é completamente legível.

Imagem 22 - Documento com texto datilografado e digitado em preto, em papel vegetal, digitalizado em modo colorido (C305)



Fonte: Autor

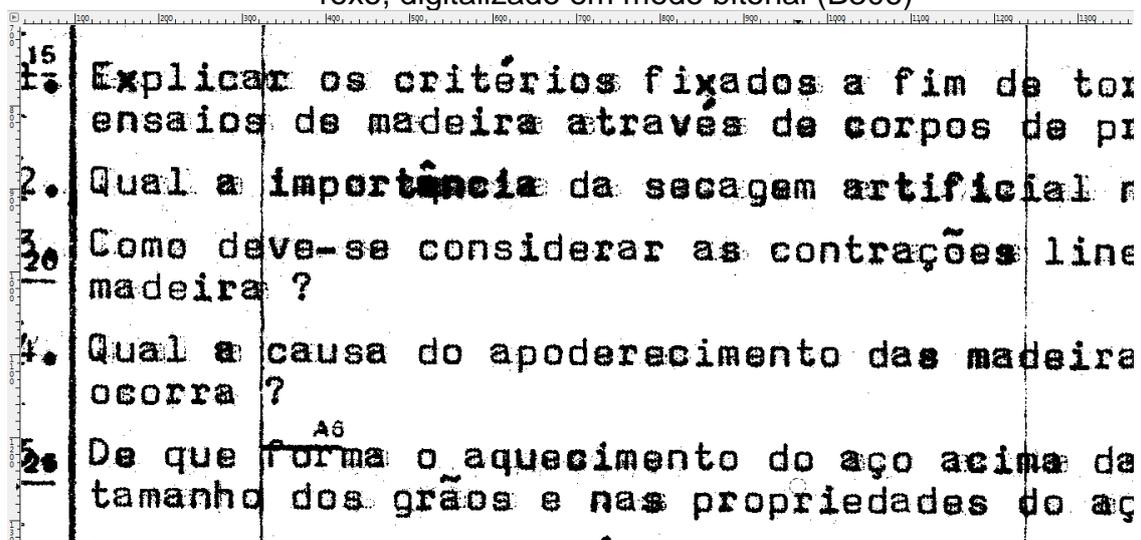
O documento em modo colorido (C305) apresentado na Imagem 22, apresenta o fundo cinza, com alguns pontos com uma leve nuance amarelada. Esse tom não existe no documento original, e pode ser o reflexo da luz do *scanner*. A luz teria refletido no papel, que é transparente, criando esse tom amarelado. O texto, tanto o datilografado quanto o digitado estão iguais ao documento anterior em modo tons de cinza (G305, mostrado na Imagem 22).

Comparando as Imagens 20, 21 e 22, que representam respectivamente os documentos B305, G305 e C305, percebe-se que a imagem P não pode ser considerada, pois não aparece nada. Sendo assim, nesse caso, a comparação fica entre os documentos G305 e C305 (Imagens 21 e 22, respectivamente). Eles, porém, estão praticamente iguais, apenas o segundo (cor) tem o fundo um pouco mais amarelado que o primeiro (tons de cinza). Por isso, o que melhor representa o documento original é a Imagem Q. Entre as recomendações apresentadas, as que se encaixam são: (1) texto impresso com ilustrações e preto e branco; (2) texto

impresso com ilustração e cor. Conforme já visto nas análises anteriores, a transparência favorece a escolha pelo modo de tons de cinza.

A sexta amostra é de um documento datilografado com o texto em roxo. Ele é uma matriz de mimeografo. Em função da sua utilização, apresenta um fundo manchado em roxo. E elas são mais aparentes em torno do texto. As imagens 23, 24 e 25 apresentam os documentos em modo bitonal, tons de cinza e colorido.

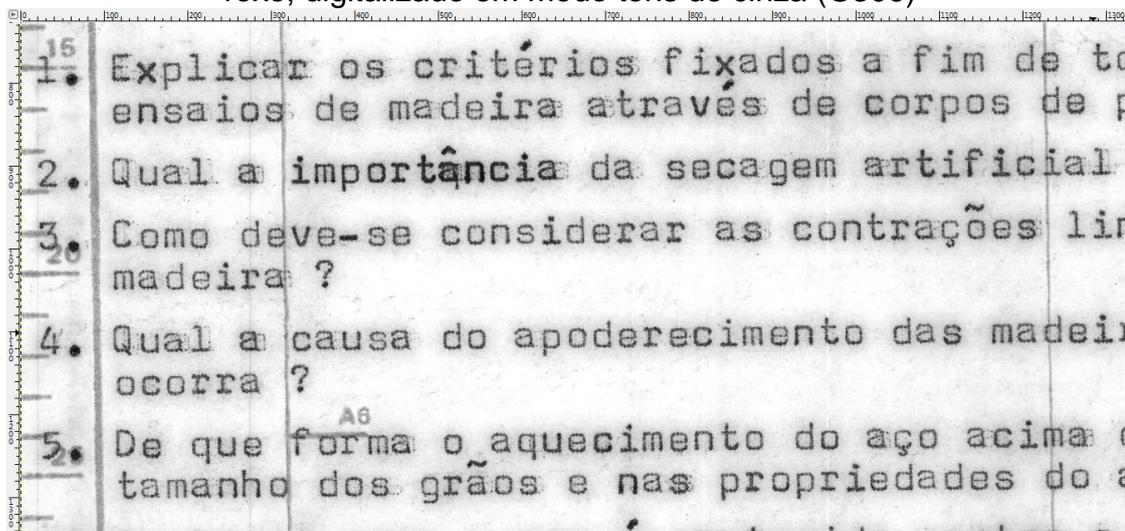
Imagem 23 - Documento datilografado em matriz de mimeografo com o texto em roxo, digitalizado em modo bitonal (B306)



Fonte: Autor

O documento em modo bitonal (B306), mostrado na Imagem 23, apresenta o fundo branco. O texto está em preto, com poucas manchas e com o contorno com falhas. O fundo roxo do documento original aparece em branco. O texto tem falhas, como em qual, e o contorno mais grosso em importância. Essa última tem o contorno mais grosso também no original, além de manchada.

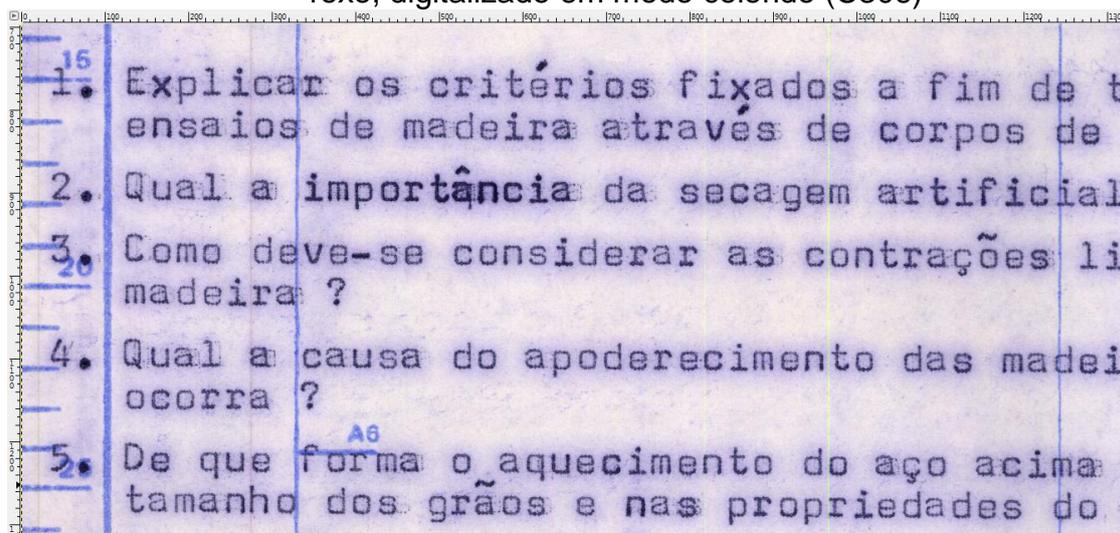
Imagem 24 - Documento datilografado em matriz de mimeografo com o texto em roxo, digitalizado em modo tons de cinza (G306)



Fonte: Autor

O documento (G306) em tons de cinza, mostrado na Imagem 24, apresenta o fundo branco com manchas cinza, principalmente ao redor do texto. O texto está em preto, com algumas falhas, como em qual. Na palavra importância, que foi apontada no documento anterior (B306) com o contorno mais grosso, nesse documento aparece mais escuro, mas com manchas ao redor.

Imagem 25 - Documento datilografado em matriz de mimeografo com o texto em roxo, digitalizado em modo colorido (C306)



Fonte: Autor

O documento colorido (C306), mostrado na Imagem 25 apresenta o fundo claro com manchas roxas, principalmente em volta do texto. No documento original

há manchas roxas, mas elas ficaram mais acentuadas na digitalização. O texto está em roxo e o contorno possui algumas falhas, mas nada que se destaque. A palavra **IMPORTÂNCIA** aparece mais escura que o resto do texto, igual ao documento original.

Fazendo a comparação entre as imagens 23, 24 e 25, nota-se que a última (Imagem 25) apresenta uma cópia mais fiel do documento original. Esse documento apresenta o texto em roxo, e o papel branco com manchas roxas. Nas recomendações analisadas, as que mais se aproximam com as características do original são: (2) texto impresso com ilustração e preto e branco com manchas e (3) texto impresso com ilustrações e cor.

Partindo do pressuposto da necessidade de considerar que o texto e a cor do fundo são coloridos, a recomendação mais adequada seria a (3), que sugerem o modo colorido. Considerando essas escolhas, o resultado obtido pela análise confere com as recomendações.

4.2 ANÁLISE COM RELAÇÃO A RESOLUÇÃO

A segunda análise realizada foi a de resolução. As resoluções comparadas foram a de 300dpi, como indica a recomendação do Conselho Nacional de Arquivologia (2010) para documentos textuais e de 600dpi, que é recomendada para imagens fotográficas.

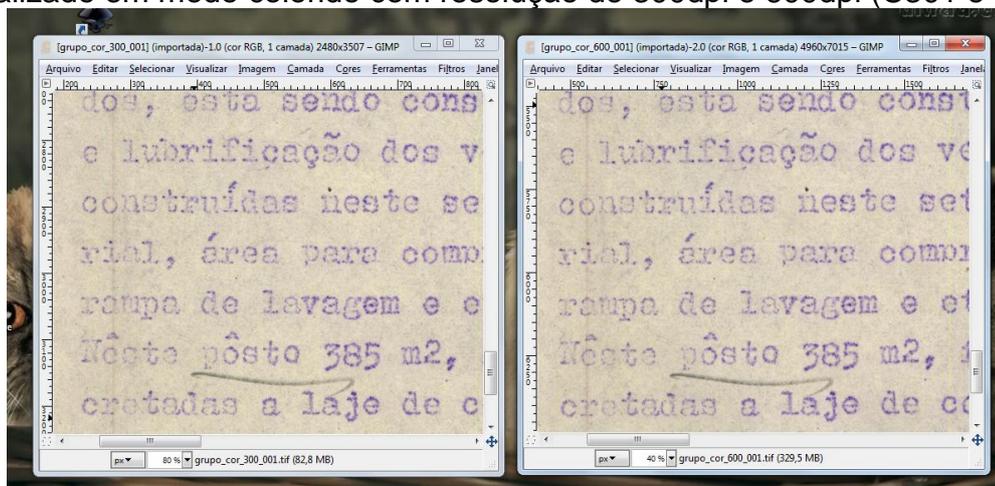
Foram utilizadas os mesmo seis documentos da análise anterior. Mas diferente dessa análise, em que foram comparados os três modos de cor (bitonal, tons de cinza e colorido), foi escolhido apenas um modo de cor. A escolha tomou como base qual dos modos foi apontado como o mais adequado na análise anterior. A relação de códigos por documento encontra-se no Quadro 11 do Apêndice A.

Para realizar a comparação, optou-se em exibir a imagem em dois tamanhos de zoom: 80% e 400% para imagens com 300dpi. Para manter a mesma proporção de dimensão, as imagens em 600dpi foram analisadas com zoom de 40% e 200%. Essa proporção foi mantida para que o tamanho físico das amostras (largura x altura) não fique diferente ao mostrar o texto.

A primeira amostra analisada corresponde a um documento datilografado e mimeografado com texto em roxo. Para essa análise foi escolhido o documento em modo colorido, com as resoluções de 300dpi (C301) e 600dpi (C601).

A Imagem 26 apresenta as imagens em resolução 300dpi e 600dpi, com zoom de 80% e 40%, respectivamente. Observa-se que não há diferenças visíveis entre os dois. O que confirma a recomendação em indicar 300dpi como resolução. O uso de 600dpi praticamente difere-se por exigir maior espaço para armazenamento, que quadruplica. As versões 300dpi e 600dpi exigiram, respectivamente, 25mb e 102Mb de armazenamento.

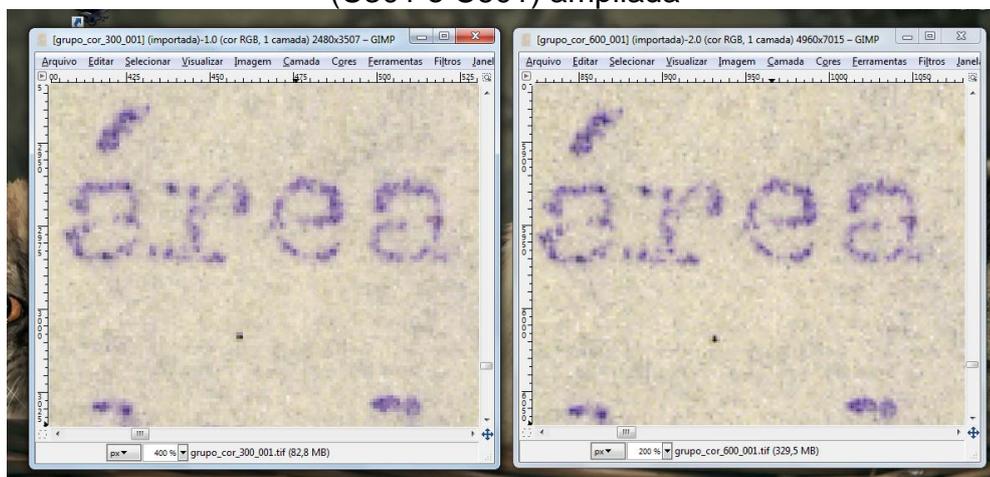
Imagem 26 - Documento datilografado e mimeografado com texto em roxo digitalizado em modo colorido com resolução de 300dpi e 600dpi (C301 e C601)



Fonte: Autor

Para confirmar essa investigação, considerou-se também uma comparação com grande zoom, isto é, analisou-se a imagem de 300dpi com zoom de 400%, e a de 600dpi com zoom de 200%. A comparação pode ser observada na Imagem 27.

Imagem 27 - Documento datilografado e mimeografado com texto em roxo digitalizado em modo colorido com resolução de 300dpi e 600dpi 300dpi e 600dpi (C301 e C601) ampliada



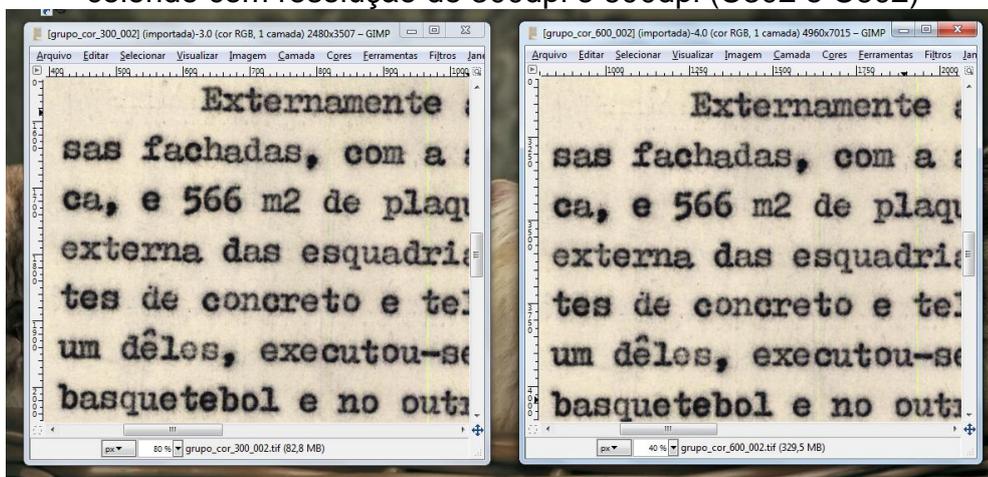
Fonte: Autor

Observando as duas imagens, não há diferenças perceptíveis. As falhas no contorno e no fundo são iguais em ambas. Conclui-se então que não é necessário para esse documento ampliar a resolução.

A segunda amostra é o documento datilografado, com o uso de carbono, com letras em tom bem escuro. O texto está em preto. O papel é amarelado, quase sem transparência. Na análise anterior, modo de cor escolhido para essa amostra foi o colorido. Assim, para essa análise foi escolhido o documento em modo colorido, com as resoluções de 300dpi (C302) e 600dpi (C602).

A Imagem 28 apresenta os documentos em resolução 300dpi (C302) e 600dpi (C602), com zoom de 80% e 40%, respectivamente. Ao comparar as duas, o documento com 300dpi aparenta ter o fundo um pouco mais claro e com as manchas menos acentuadas. Um pixel em 300dpi é representado por quatro pixels em 600dpi, capturando um valor médio de tom equivalente a esses quatro pontos. Esses quatro pontos podem representar detalhes de tons que acabam não sendo considerados na resolução de 300dpi. O uso de 600dpi praticamente difere-se por exigir maior espaço para armazenamento, que quadruplica.

Imagem 28 – Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C302 e C602)



Fonte: Autor

Para confirmar se realmente o fundo está mais claro, ampliou-se o zoom, comparando a imagem de 300dpi com zoom de 400%, e a de 600dpi com zoom de 200%. A comparação pode ser observada na Imagem 29.

Imagem 29 - Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C302 e C602) ampliado



Fonte: Autor

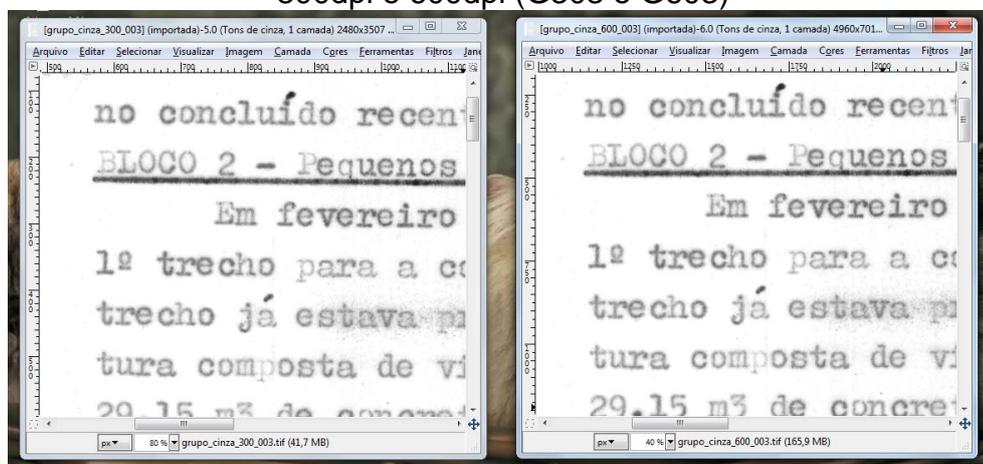
Com o aumento de zoom, confirma-se que o documento com 300dpi realmente tem o fundo mais claro. Além disso, a letra **d** parece com mais pontos (*pixels*) escuros ao redor, na imagem com 300dpi (C302). Nesse caso, o aumento da resolução mostrou diferenças. Comparando ambos, porém, essas diferenças não gera a necessidade de um documento com resolução maior. Principalmente se

considerar o tamanho de armazenamento do arquivo, que no caso de 600dpi é quatro vezes maior que o de 300dpi (102Mb para 600dpi e 25Mb para 300dpi).

A terceira amostra é do documento datilografado em papel carbono. O papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais claro. Esse documento teve duas amostras satisfatórias; A primeira analisada foi a de modo em tons de cinza.

A Imagem 30 apresenta os documentos em modo tons de cinza com resolução 300dpi (G303) e 600dpi (G603), com zoom de 80% e 40%, respectivamente. A comparação entre os documentos com resoluções diferentes mostra que não há diferenças visíveis entre os dois. O que confirma a recomendação em indicar 300dpi como resolução. O uso de 600dpi praticamente difere-se por exigir maior espaço para armazenamento, que quadriplica. As versões 300dpi e 600dpi exigiram, respectivamente, 8Mb e 34Mb de armazenamento.

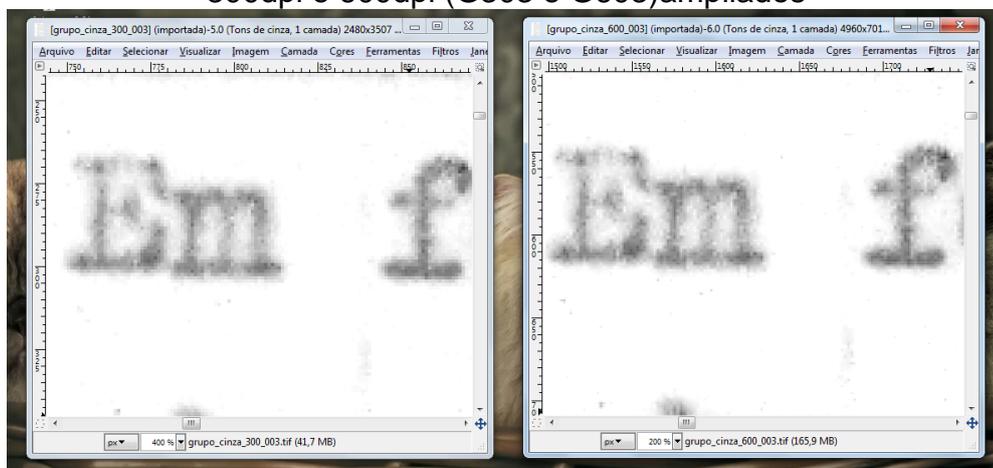
Imagem 30 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G303 e G603)



Fonte: Autor

Para confirmar se realmente não há diferenças, foi realizada uma segunda comparação com grande zoom, isto é, analisou-se a imagem de 300dpi com zoom de 400%, e a de 600dpi com zoom de 200%. A comparação pode ser observada na Imagem F.

Imagem 31 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G303 e G603) ampliados

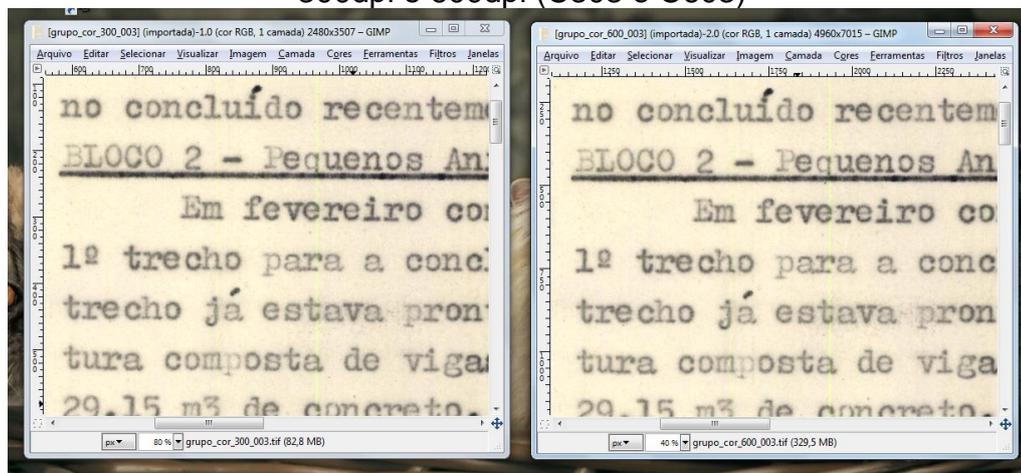


Fonte: Autor

A análise das duas imagens com zoom ampliado confirma que não há diferenças perceptíveis. Conclui-se então que não é necessário para esse documento ampliar a resolução.

A segunda comparação da terceira amostra é em modo colorido. A Imagem 32 apresenta os documentos em resolução 300dpi (C303) e 600dpi (C603), com zoom de 80% e 40%, respectivamente. A comparação entre os documentos com resoluções diferentes mostra que não há alterações visíveis entre os dois.

Imagem 32 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C303 e C603)

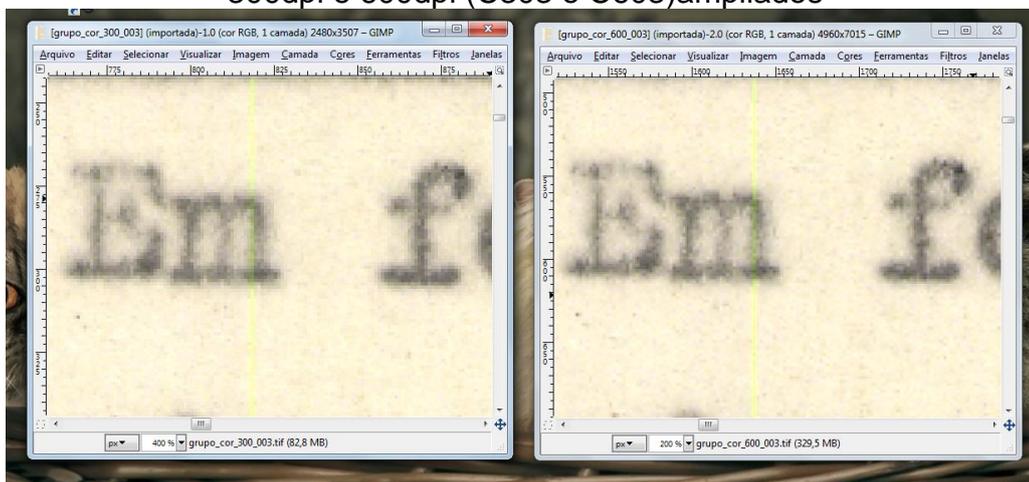


Fonte: Autor

Para confirmar se realmente não há diferenças, foi realizada uma segunda comparação com grande zoom, isto é, analisou-se a imagem de 300dpi com zoom

de 400%, e a de 600dpi com zoom de 200%. A comparação pode ser observada na Imagem 33.

Imagem 33 - Documento datilografado com o uso de papel carbono, com as letras em preto com tom mais claro, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C303 e C603) ampliados



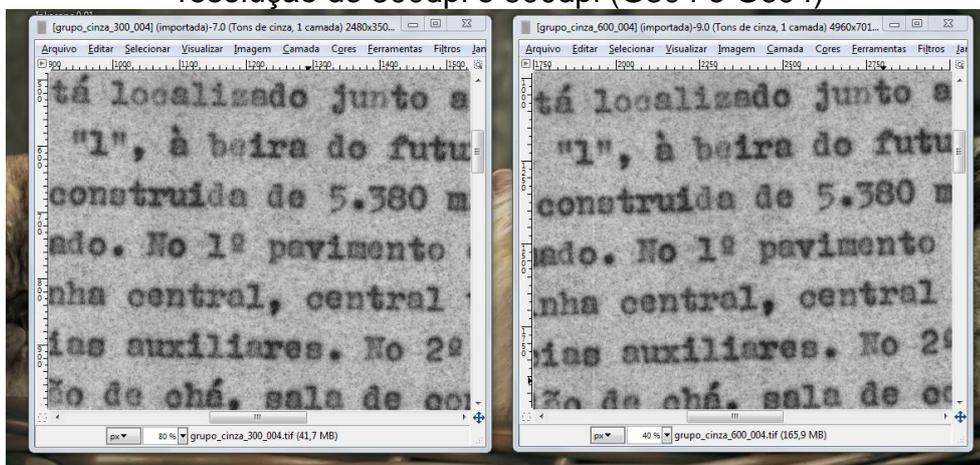
Fonte: Autor

A análise das duas imagens com zoom ampliado confirma que não há diferenças perceptíveis. Conclui-se então que não é necessário para esse documento ampliar a resolução.

A quarta amostra é um documento datilografado com o uso de papel carbono, em papel mais fino e transparente que os das amostras anteriores. O texto é claro. Para essa análise foi escolhido o documento em modo de tons de cinza, com as resoluções de 300dpi (G304) e 600dpi (G604).

A Imagem 34 apresenta os documentos em resolução 300dpi (G304) e 600dpi (G604), com zoom de 80% e 40%, respectivamente. A comparação entre os documentos com resoluções diferentes mostra que não há diferenças visíveis entre os dois. O uso de 600dpi praticamente difere-se por exigir maior espaço para armazenamento, que quadriplica. As versões 300dpi e 600dpi exigiram, respectivamente, 8Mb e 34Mb de armazenamento.

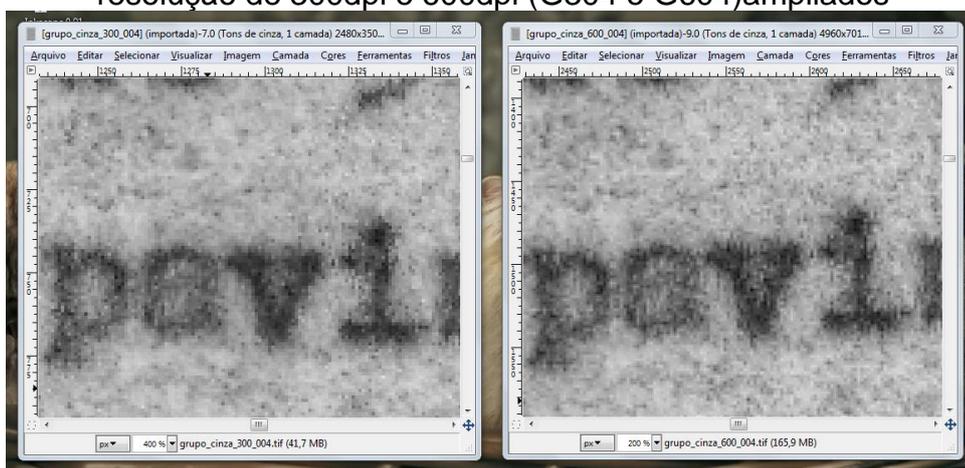
Imagem 34 - Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G304 e G604)



Fonte: Autor

Para confirmar se realmente não há diferenças, foi realizada uma segunda comparação com grande zoom, isto é, analisou-se a imagem de 300dpi com zoom de 400%, e a de 600dpi com zoom de 200%. A comparação pode ser observada na Imagem 35.

Imagem 35 - Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G304 e G604) ampliados



Fonte: Autor

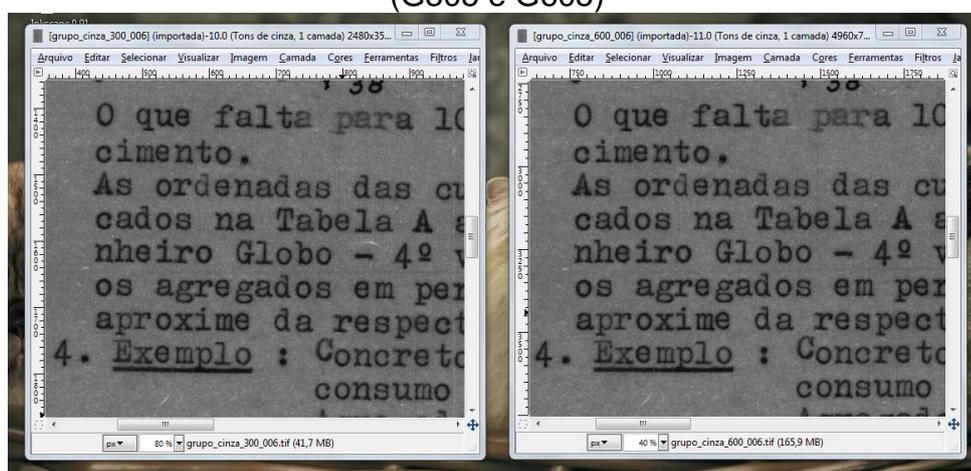
A análise das duas imagens com zoom ampliado confirma que não há diferenças perceptíveis. Conclui-se então que não é necessário para esse documento ampliar a resolução.

A quinta amostra é de um documento com texto datilografado e digitado em preto. O papel desse documento é vegetal, ou seja, transparente. Para essa análise

foi escolhido o documento em modo tons de cinza, com as resoluções de 300dpi (G305) e 600dpi (G605).

A Imagem 36 apresenta os documentos em resolução 300dpi (G305) e 600dpi (G605), com zoom de 80% e 40%, respectivamente. A comparação entre os documentos com resoluções não mostra diferenças entre ambas. O uso de 600dpi praticamente difere-se por exigir maior espaço para armazenamento, que quadriplica. As versões 300dpi e 600dpi exigiram, respectivamente, 8Mb e 34Mb de armazenamento.

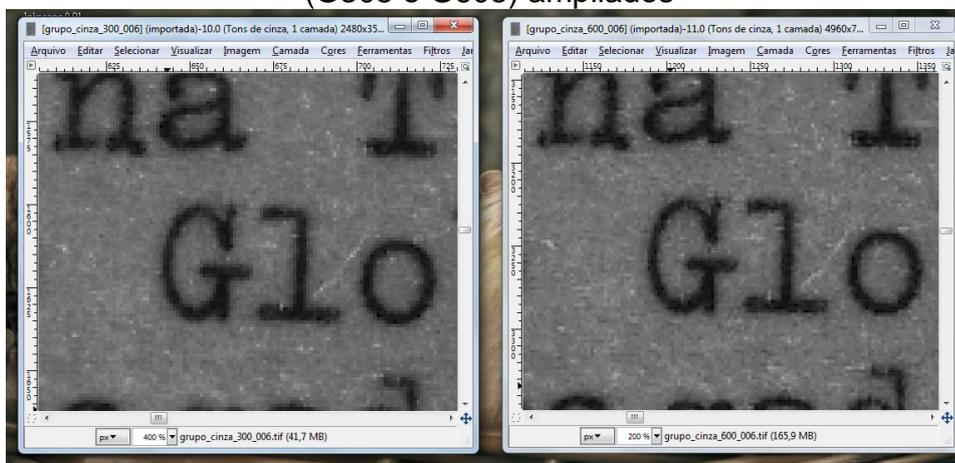
Imagem 36 - Documento com texto datilografado e digitado em preto, em papel vegetal, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G305 e G605)



Fonte: Autor

Para confirmar se realmente não há diferenças, foi realizada uma segunda comparação com grande zoom, isto é, analisou-se a imagem de 300dpi com zoom de 400%, e a de 600dpi com zoom de 200%. A comparação pode ser observada na Imagem 37.

Imagem 37 - Documento com texto datilografado e digitado em preto, em papel vegetal, digitalizado em modo tons de cinza com resolução de 300dpi e 600dpi (G305 e G605) ampliados



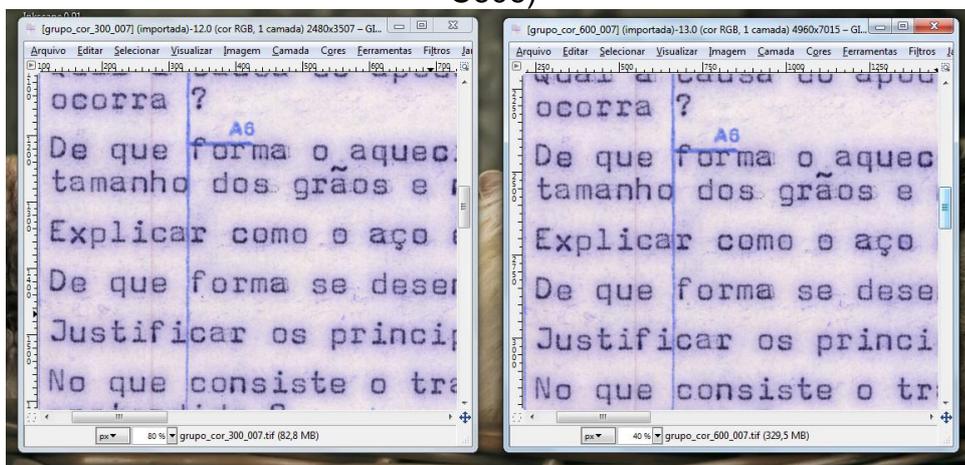
Fonte: Autor

A análise das duas imagens com zoom ampliado confirma que não há diferenças perceptíveis. Conclui-se então que não é necessário para esse documento ampliar a resolução.

A sexta amostra é de um documento datilografado com o texto em roxo. Ele é uma matriz de mimeografo. Em função da sua utilização, apresenta um fundo manchado em roxo. E elas são mais aparentes em torno do texto. Para essa análise foi escolhido o documento em modo colorido, com as resoluções de 300dpi (C306) e 600dpi (C606).

A Imagem 38 apresenta as imagens em resolução 300dpi e 600dpi, com zoom de 80% e 40%, respectivamente. Nota-se que não há diferenças visíveis entre os dois. O que confirma a recomendação em indicar 300dpi como resolução. O uso de 600dpi praticamente difere-se por exigir maior espaço para armazenamento, que quadruplica. As versões 300dpi e 600dpi exigiram, respectivamente, 25mb e 102Mb de armazenamento.

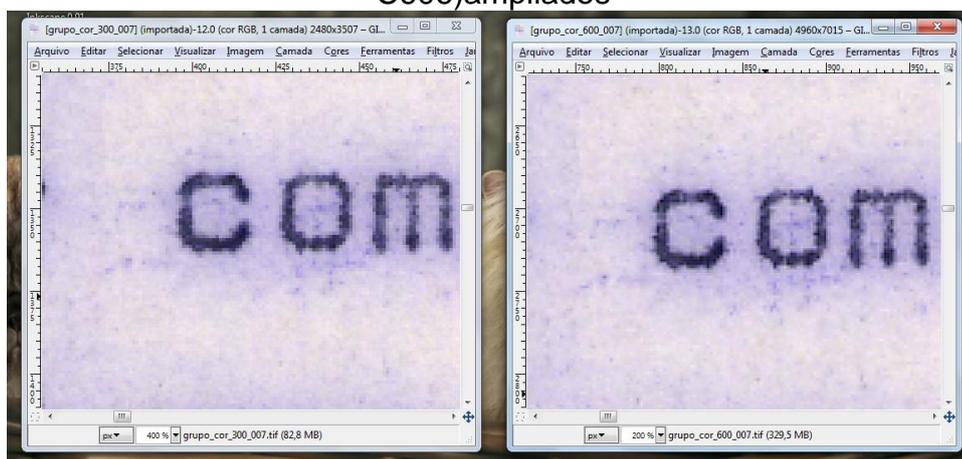
Imagem 38 - Documento datilografado em matriz de mimeografo com o texto em roxo, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C306 e C606)



Fonte: Autor

Para confirmar se realmente não há diferenças, foi realizada uma segunda comparação com grande zoom, isto é, analisou-se a imagem de 300dpi com zoom de 400%, e a de 600dpi com zoom de 200%. A comparação pode ser observada na Imagem 39.

Imagem 39 - Documento datilografado em matriz de mimeografo com o texto em roxo, digitalizado em modo colorido com resolução de 300dpi e 600dpi (C306 e C606)ampliados



Fonte: Autor

A análise das duas imagens com zoom ampliado confirma que não há diferenças perceptíveis. Conclui-se então que não é necessário para esse documento ampliar a resolução.

5 CONCLUSÃO

A digitalização é um processo que cria uma cópia digital de um documento físico. Isso permite preservar a informação de modo eficaz, além de ampliar o seu acesso. Para que essa cópia, ou matriz digital seja realmente eficaz e fiel ao documento original, é necessário seguir uma série de requisitos.

No Brasil, foi publicado pelo CONARQ recomendações para a digitalização de documentos. Elas incluíam as configurações como o formato de arquivo, resolução, modo e profundidade de bits. Porém, constatou-se que elas eram recomendações de aplicação geral, pois considera principalmente as características de gênero, como texto impresso, texto manuscrito, imagens, plantas entre outros. E um documento físico possui outras variáveis, como o suporte, que é o papel e em caso de documentos textuais, o modo de impressão.

O objetivo dessa pesquisa era *investigar as configurações (modo e resolução) para a digitalização de documentos textuais datilografados em suas diferentes apresentações*. Para atingir esse objetivo, era necessário primeiramente alcançar os objetivos específicos.

O primeiro, que se tratava de *analisar as recomendações existentes para digitalização de documentos*, ocorreu através de pesquisa teórica. Foram estudadas as recomendações existentes, tomando como base a do CONARQ.

O segundo objetivo, *experimentar as diferentes configurações para a geração de matrizes de documentos textuais*, foi atingida através de testes com os documentos físicos. Cada amostra foi digitalizada em diferentes modos (bitonal, tons de cinza e cor) e resolução (300 e 600 dpi).

O terceiro objetivo, *analisar as experimentações realizadas, relacionando-as com as recomendações existentes para digitalização de documentos textuais*, é a comparação realizada na análise. Foram comparadas, na primeira parte, o documento original digitalizado em três modos (bitonal, tons de cinza e cor). Percebeu-se a complexidade para a indicação do modo de cor. Ocorreu que documentos com característica de texto parecidas, como os segunda e terceira amostra, obtiveram resultados diferentes. E nesse caso, o motivo foi a tonalidade do texto, que mesmo sendo considerado preto, o da terceira amostra possuía uma tonalidade mais clara. Comparando a terceira e a quarta amostra, que possuem o

texto com tonalidades parecidas, também ocorreram resultados diferentes. Nesse caso, o fator determinante foi o papel, que na quarta amostra era mais fino.

A segunda análise foi a de resolução. Comparou-se entre o indicado pela recomendação do CONARQ, de 300dpi, e 600 dpi por ser o dobro da resolução. Ampliou-se as imagens, comparando tanto com o texto em tamanho para leitura como ampliando-os para mostrar os detalhes. Nessa comparação, confirmou-se então que a resolução indicada está correta. Isso porque não ocorreram diferenças que justificassem a necessidade de ampliação.

O quarto e último objetivo é *apresentar sugestões para a digitalização de documentos textuais datilografados*. Essas sugestões foram realizadas durante a análise de modo, pois foi a que se mostrou necessária. Com relação a resolução, a indicação se mostrou suficiente, não havendo necessidade de sugestões de melhoria.

Com esse estudo, é possível notar que as recomendações do CONARQ são amplas. Elas podem ser utilizadas como base, mas ainda assim requer testes e uma análise do documento a ser digitalizado. Isso porque cada documento, por mais característica em comum que possuem, como o gênero (textual, iconográfico, entre outros), há diferenças que influenciam na digitalização, como o papel e o tipo de texto escrito, não sendo apenas suficiente indicar se é colorido ou preto e branco

Sugere-se que amplie o estudo sobre digitalização, incluindo outros pontos do documento, como o papel e o tipo de escrita, além do impresso e do manuscrito já indicados. Mesmo com essa ampliação nas recomendações, o ideal é não se fixar apenas a elas. Deve-se sempre testar novas configurações, e buscar a cópia ideal é mais fiel possível ao documento original.

REFERÊNCIAS

CAREGNATO, Sônia Elisa; ROCHA, Rafael Port da. Compartilhamento e curadoria de dados de pesquisa: uma proposta para a ciência produzida na Universidade Federal do Rio Grande do Sul. In: **Conferência sobre tecnologia, cultura e memória**, 2013, Pernambuco. Disponível em: <http://www.liber.ufpe.br/ctcm2013/anais/files/6a.CCDP_PCUFRGS.pdf>. Acesso em: 25 out, 2015.

CONSELHO NACIONAL DE ARQUIVOS. **O Conselho**. 2016. Disponível em: <<http://www.conarq.arquivonacional.gov.br/o-conselho-3.html>>. Acessado em: 21 nov. 2016.

_____. **E-Arq Brasil**: modelo de requisitos para sistemas informatizados de gestão arquivística de documentos. Rio de Janeiro: CONARQ, 2011. Disponível em: <<http://www.conarq.arquivonacional.gov.br/publicacoes-ctde/33-modelo-de-requisitos-para-sistemas-informatizados-de-gestao-arquivistica-de-documentos-e-arq-brasil.html>>. Acesso em: 25 out. 2015.

_____. **Recomendações para digitalização de documentos arquivísticos permanentes**. [S.l.]: CONARQ, 2010. Disponível em: <http://www.conarq.arquivonacional.gov.br/images/publicacoes_textos/Recomendacoes_digitalizacao_completa.pdf>. Acesso em: 25 out. 2015.

CORNELL LIBRARY. **Moving Theory into Practice: digital images**. New York: Cornell University, 2003a. Disponível em: <<https://www.library.cornell.edu/preservation/tutorial/intro/intro-01.html>>. Acesso em: 15 out. 2015.

_____. **Moving Theory into Practice: bit deph**. New York: Cornell University, 2003b. Disponível em: <<https://www.library.cornell.edu/preservation/tutorial/intro/intro-04.html>>. Acesso em: 21 nov. 2016.

_____. **Moving Theory into Practice: compression**. New York: Cornell University, 2003c. Disponível em: <<https://www.library.cornell.edu/preservation/tutorial/intro/intro-07.html>>. Acesso em: 21 nov. 2016.

_____. **Moving Theory into Practice: selection**. New York: Cornell University, 2003d. Disponível em: <<https://www.library.cornell.edu/preservation/tutorial/selection/selection-01.html>> Acesso em: 21 nov. 2016.

_____. **Moving Theory into Practice: selection**. New York: Cornell University, 2003e. Disponível em: <<https://www.library.cornell.edu/preservation/tutorial/conversion/conversion-01.html>>. Acesso em: 21 nov. 2016.

_____. **Moving Theory into Practice: quality control**. New York: Cornell University, 2003f. Disponível em: <<https://www.library.cornell.edu/preservation/tutorial/quality/quality-01.html>>. Acesso em: 21 nov. 2016.

DICIONÁRIO BRASILEIRO DE TERMINOLOGIA ARQUIVÍSTICA. Rio de Janeiro: Arquivo Nacional, 2005. Disponível em: <http://www.arquivonacional.gov.br/images/pdf/Dicion_Term_Arquiv.pdf >. Acesso em: 15 out. 2015.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Guia para a digitalização de documentos**: versão 2.0. Brasília, 2006. Disponível em: <<https://www.sct.embrapa.br/goi/manuais/GuiaDigitalizacao.pdf> >. Acesso em: 15 out 2015.

FERREIRA, Miguel. **Introdução à preservação digital**: conceitos, estratégias e actuais consensos. Escola de Engenharia da Universidade do Minho: Portugal, 2006.

GIL, Antônio Carlos. **Como Elaborar projetos de Pesquisa**. 5. ed. São Paulo: Atlas, 2010.

GIMP. **Introdução**. 2016. Disponível em: <https://docs.gimp.org/2.8/pt_BR/introduction.html#introduction-help > Acesso em 21 nov. 2016

GOMES, Sandra Lúcia Rebel; MENDONÇA, Marília Alvarenga Rocha; SOUZA, Clarice Muhlethaler de. Literatura cinzenta. In: CAMPELLO, Bernadete Santos; CENDÓN, Beatriz Valadares; KREMER, Jeannette Marguerite. **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: Ed. UFMG, 2000.

GUINCHAT, Claire; MENOUE, Michael. **Introdução geral às ciências e técnicas da informação e documentação**. 2. ed. Brasília: IBICT, 1994

MARTINS, Nelson. **A imagem digital na editoração**: manipulação, conversão e fechamento de arquivos. Rio de Janeiro: SENAC, 2009.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos da Metodologia Científica**. 7. ed. São Paulo: Atlas, 2010.

MULLER, Suzana Pinheiro Machado. A ciência, o sistema de comunicação científica e a literatura científica. In: **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: Ed. UFMG, 2000.

NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO). **Understanding Metadata**. Bethesda, MD: NISO, 2004. Disponível em: <<http://www.niso.org/publications/press/UnderstandingMetadata.pdf> > Acesso em 21 nov. 2016.

ROCHA, Rafael Port da. **Documento digital: notas de aula**. Porto Alegre, 2014. 69 slides. Apresentação em Power Point.

_____. **Documento digital: imagem**. Porto Alegre, 2016. 54 slides. Apresentação em Power Point.

RONDINELLI, Rosely Curi. **Gerenciamento arquivístico de documentos eletrônicos**. 4. ed. Rio de Janeiro: Ed. FGV, 2005.

SILVA, Edna Lúcia da; MENEZES, Ester Muszkat. **Metodologia da Pesquisa e Elaboração de Dissertação**. 4. ed. Florianópolis: UFSC, 2005.

SIQUEIRA, Jéssica Câmara. Informação e documento: relações simbióticas. **Ponto de Acesso**. Salvador, v.9, n.1, p. 91-110, abr. 2015

APÊNDICE A - SIGLA DOS DOCUMENTOS ANALISADOS

Quadro 10 - Documentos analisados por modo

Documento	Características	Configuração da Imagem digital
B301	Documento datilografado e mimeografado com texto em roxo, com marcações manuscritas a lápis em papel amarelado.	Modo Bitonal, 300dpi, TIFF, 1 Mb
B302	Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência	Modo Bitonal, 300dpi, TIFF, 1 Mb
B303	Documento datilografado com o uso de papel carbono, em papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais claro.	Modo Bitonal, 300dpi, TIFF, 1 Mb
B304	Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente.	Modo Bitonal, 300dpi, TIFF, 1 Mb
B305	Documento com texto datilografado e digitado em preto, em papel vegetal	Modo Bitonal, 300dpi, TIFF, 1 Mb
B306	Documento datilografado em matriz de mimeografo com o texto em roxo. Apresenta manchas roxas em torno do texto.	Modo Bitonal, 300dpi, TIFF, 1 Mb
G301	Documento datilografado e mimeografado com texto em roxo, com marcações manuscritas a lápis em papel amarelado.	Modo Tons de Cinza, 300dpi, TIFF, 8 Mb
G302	Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência	Modo Tons de Cinza, 300dpi, TIFF, 8 Mb
G303	Documento datilografado com o uso de papel carbono, em papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais claro.	Modo Tons de Cinza, 300dpi, TIFF, 8 Mb
G304	Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente.	Modo Tons de Cinza, 300dpi, TIFF, 8 Mb
G305	Documento com texto datilografado e digitado em preto, em papel vegetal	Modo Tons de Cinza, 300dpi, TIFF, 8 Mb
G306	Documento datilografado em matriz de mimeografo com o texto em roxo. Apresenta manchas roxas em torno do texto.	Modo Tons de Cinza, 300dpi, TIFF, 8 Mb
C301	Documento datilografado e mimeografado com texto em roxo, com marcações manuscritas a lápis em papel amarelado.	Modo Colorido, 300dpi, TIFF, 25Mb
C302	Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência	Modo Colorido, 300dpi, TIFF, 25Mb
C303	Documento datilografado com o uso de papel carbono, em papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais claro.	Modo Colorido, 300dpi, TIFF, 25Mb
C304	Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente.	Modo Colorido, 300dpi, TIFF, 25Mb
C305	Documento com texto datilografado e digitado em preto, em papel vegetal	Modo Colorido, 300dpi, TIFF, 25Mb
C306	Documento datilografado em matriz de mimeografo com o texto em roxo. Apresenta manchas roxas em torno do texto.	Modo Colorido, 300dpi, TIFF, 25Mb

Fonte: Elaborado pelo autor.

Quadro 11 - Documentos analisados por resolução

Documento	Características	Configuração da Imagem digital
C301	Documento datilografado e mimeografado com texto em roxo, com marcações manuscritas a lápis em papel amarelado.	Modo Colorido, 300dpi, TIFF, 25Mb
C601	Documento datilografado e mimeografado com texto em roxo, com marcações manuscritas a lápis em papel amarelado.	Modo Colorido, 600dpi, TIFF, 102 Mb
C302	Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência	Modo Colorido, 300dpi, TIFF, 25Mb
C602	Documento datilografado, com o uso de carbono, com letras em tom bem escuro, em papel é amarelado, quase sem transparência	Modo Colorido, 600dpi, TIFF, 102 Mb
G303	Documento datilografado com o uso de papel carbono, em papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais claro.	Modo Tons de Cinza, 300dpi, TIFF, 8 Mb
G603	Documento datilografado com o uso de papel carbono, em papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais claro.	Modo Tons de Cinza, 600dpi, TIFF, 34 Mb
C303	Documento datilografado com o uso de papel carbono, em papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais claro.	Modo Colorido, 300dpi, TIFF, 25Mb
C306	Documento datilografado com o uso de papel carbono, em papel é similar ao da segunda amostra, sendo as letras impressas em preto com tom mais claro	Modo Colorido, 600dpi, TIFF, 102 Mb
G304	Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente.	Modo Tons de Cinza, 300dpi, TIFF, 8 Mb
G604	Documento datilografado com o uso de papel carbono, com o texto claro, em papel fino e transparente.	Modo Tons de Cinza, 600dpi, TIFF, 34 Mb
G305	Documento com texto datilografado e digitado em preto, em papel vegetal	Modo Tons de Cinza, 300dpi, TIFF, 8 Mb
G605	Documento com texto datilografado e digitado em preto, em papel vegetal	Modo Tons de Cinza, 600dpi, TIFF, 34 Mb
C306	Documento datilografado em matriz de mimeografo com o texto em roxo. Apresenta manchas roxas em torno do texto.	Modo Colorido, 300dpi, TIFF, 25Mb
C606	Documento datilografado em matriz de mimeografo com o texto em roxo. Apresenta manchas roxas em torno do texto.	Modo Colorido, 600dpi, TIFF, 102 Mb

Fonte: Elaborado pelo autor.