

**Agnaldo Lopes Martins**

**METADADOS PARA DESCRIÇÃO DE DOCUMENTOS  
REMANESCENTES DE FUNDO ECLESIAÍSTICO**

**Uma proposta de metadados e biblioteca digital  
para os assentos de batismo, casamento e óbito da  
Matriz do Pilar de Ouro Preto**

Dissertação apresentada ao Programa de Pós-graduação em Ciência da Informação da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Mestre em Ciência da Informação.

Área de concentração: Produção, Organização e Utilização da Informação.

Linha de Pesquisa: Organização e Uso da Informação

Orientadora: Profa. Lídia Alvarenga  
Universidade Federal de Minas Gerais

Belo Horizonte  
Escola de Ciência da Informação da UFMG  
2006

Escrivaninhas utilizadas para acomodar os objetos para escrita em documentos, na primeira metade do século XVIII.



A primeira em tartaruga escura, ouro e madrepérola.  
A segunda, de origem italiana, confeccionada em prata.

*“Esclarecida é a sabedoria, e tal que nunca se murcha, e facilmente é vista por aqueles que a amam, e achada pelos que a buscam. Ela se antecipa aos que a cobiçam... aquele que a vigia desde manhã para possuir, não terá trabalho, porque ele a achará sentada à sua porta.”*

Livro da Sabedoria, capítulo III versículo 14-16

## AGRADECIMENTOS

Por mais que busquemos explicações, sejam elas na física na matemática, ou em qualquer outra disciplina, tenhamos sempre em mente que Ele foi o início e nos encontrará no fim. Nossa caminhada depende Dele, e a Ele agradeço.

É difícil lembrar de todas as pessoas que de alguma forma me apoiaram durante esta caminhada. Desde já minhas desculpas pelos nomes que não estarão aqui escritos, saibam que a gratidão não é menor do que para aqueles que aqui estão.

Dedico esta dissertação, a qual marca o fim de uma jornada de dois anos na Escola de Ciência da Informação da UFMG a Míria, minha esposa e ao nosso Thiaguinho, que por uma feliz coincidência nasceu a poucos dias do início desta jornada.

A meu Pai e minha Mãe, não conseguirei retribuir o carinho e a dedicação. Queira Deus eu tenha a mesma paciência de vocês quando cuidar do meu filho.

Eu segui a trilha deixada por várias pessoas que passaram pelo caminho primeiro que eu. Dedico a meus irmãos, pois entendem que uma trilha feita pode deixar tudo mais simples.

Para o sr. Gilson Cheble, pois naquela tarde eu ainda não sabia que me convidaria para fazer o mestrado.

Para a profa. Adalgisa, para quem nunca terei palavras, todas elas juntas, não darão conta de descrever a sua grandeza e elegância.

Para a profa. Lídia, minha orientadora, com quem me identifiquei muito. Sua força no momento mais triste da vida de um ser humano, muito me ensinou.

Agradeço especialmente a Mirian Lott que me orientou em momentos brilhantes em Ouro Preto ou na sala da profa. Adalgisa; ao incansável Júlio o qual facilitou muito meu trabalho de leitura e interpretação do conteúdo dos documentos; aos funcionários da UFMG em especial para Viviane, e dos locais onde visitei em Ouro Preto, principalmente o sr. Carlos (Caju) por ter permitido meu acesso ao acervo original. Ao pessoal do departamento de estatística pelo auxílio durante nossas reuniões e por valorizarem esta pesquisa social. Ao pessoal da Top Training em especial ao Lourival com quem sempre desabafei angústias da pesquisa.

Ao sr. Hiram, por ter me deixado usar os seus computadores até a meia noite em vários dias quando eu tinha quatorze anos (se bem que até hoje eu uso o dele)

A meus alunos deixo o mapa, pois é tudo o que possuo,  
o tesouro vocês encontrarão.

## RESUMO

O presente trabalho propõe metadados para a criação de uma biblioteca digital constituída do acervo de documentos eclesiásticos da Matriz do Pilar de Ouro Preto, Minas Gerais, Brasil, composto de assentos de casamentos, batismos e óbitos, com enfoque para a preservação dos documentos. O crescente aumento do uso do meio digital como forma de disponibilizar informação, combinado com a inexistência de uma metodologia que permitisse a identificação dos metadados para comporem os pontos de acesso aos documentos neste tipo de fundo, indicaram a necessidade da elaboração de um modelo de metadados que contemplasse o maior número de pontos de acesso aos documentos, em caso de sua representação em uma biblioteca digital. A presente pesquisa adotou abordagem experimental, e compreendeu três etapas: visitas ao acervo de documentos; exaustiva leitura dos documentos microfilmados, objetivando identificar metadados; e uso de software de computador para criação de um protótipo de biblioteca digital para este acervo. No processo de identificação dos metadados, que envolveu a análise dos documentos sobre casamentos, batismos e óbitos, trabalhou-se com uma amostragem de documentos, em consonância com os métodos estatísticos, tendo como material os microfilmes presentes na Casa dos Contos de Ouro Preto, órgão subordinado à Escola Superior de Administração Fazendária da Receita Federal, ESAF. Os resultados da investigação incluem: planilhas de metadados para descrição de documentos eclesiásticos ajustados ao modelo Dublin Core; lista de conceitos encontrados nesses tipos de assentos; modelo da base de dados relacional para acomodar a coleta dos dados nos documentos; e protótipo de uma biblioteca digital elaborada em software específico. Como fundamentação teórica e contextual foram incluídos: especificidades do acervo; o contexto de produção e uso dos documentos ao longo dos anos, os pesquisadores e sociedades religiosas; o tratamento digital da informação, incluindo conceitos de metadados, sistemas de recuperação de informação, padrões MARC.e Dublin Core e bibliotecas digitais.

Palavras-chaves: Manuscritos – Metadados – Biblioteca Digital – Fundo eclesiástico – Descrição documental – Matriz do Pilar, Ouro Preto, MG Brasil – Ouro Preto, MG, Brasil.

## **ABSTRACT**

The present dissertation proposes a metadata scheme to create a digital library composed by ecclesiastics documents of the catholic church named Matrix of Pilar (Matriz do Pilar de Ouro Preto), situated in Ouro Preto, Minas Gerais, Brazil. These documents, considered manuscripts, are wedding seats, baptisms and death certifications. The main viewpoint of this work is the documents preservation process. The increasing growth and usage of digital technologies, as a form to make available information, combined with the absence of a methodology that allows a metadata scheme to describe this kind of documents, indicates the need of metadata that considers the biggest number of documents access points for a digital library. The present research adopts an experimental approach in three specific stages: contacts with the documents fund; exhaustive reading of the microfilmed documents to identify metadata; and the use of a computer software to create a prototype of digital library. In the process of metadata identification that includes the analysis of documents about wedding, baptism and death certificate, it is worked with a specific documents sample, according with statistical methods, having as source the microfilms of the House of Tales of Ouro Preto (Casa dos Contos de Ouro Preto). The results of this work includes: metadata scheme for manuscripts description and retrieval, according with the Dublin Core model; list of concepts/subjects found in the documents; model of database related to this data collection; and a prototype of a digital library, using specific software. As theoretical and contextual fundamentals are included: description of documents; document production context; the use of documents through the years; administrative and religious societies; digital treatment of information, including metadata concepts; information retrieval systems; U.S Machine Readable Cataloguing Format, MARC, Dublin Core (simplified MARC) and digital libraries concepts.

### **Key-words**

Manuscripts description - Manuscripts - Metadata – Ecclesiastic Funds - Digital Library – Documents description.

## LISTA DE ABREVIATURAS E SIGLAS

AACR2	- Anglo American Cataloging Rules
CECO	- Centro de Estudos do Ciclo do Ouro
DC	- Dublin Core
DVD	- Digital Vídeo Disk
EAD	- Encoded Archival Description)
ESAF	- Escola Superior de Administração Fazendária
ETDMS	- Eletronic Thesis and Dissertation Metadata Standard
FGDC	- Federal Geographic Data Committee
HTML	- Hipertext Markup Language
MARC	- Machine Readable Catalogue
OAI	- Open Archives Initiative
RFC	- Request For Comments
SI	- Sistemas de Informação
SRI	- Sistemas de Recuperação da Informação
TEIH	- Text Encoding Initiative Headers
UFMG	- Universidade Federal de Minas Gerais
W3C	- Word Wide Web Consortium

## LISTA DE FIGURAS

<b>Figura 2.1</b>	Mapa da região de Vila Rica	14
<b>Figura 3.1</b>	Local onde os documentos estão ficam estocados antes de serem tratados pelo técnicos em preservação	18
<b>Figura 3.2</b>	Estantes onde os documentos ficam alocados em Ouro Preto	21
<b>Figura 3.3</b>	Foto de um livro para assentos de batismo do século XVIII	22
<b>Figura 3.4</b>	Livro com capa de papelão e bordas em couro	23
<b>Figura 3.5</b>	Capa revestida em tecido	24
<b>Figura 3.6</b>	Capa com acabamento em couro	24
<b>Figura 3.7</b>	Capa em detalhes imitando mármore	25
<b>Figura 3.8</b>	Livro de 1780 com marcações feitas por um usuário que utilizou caneta esferográfica caracterizando um ato de vandalismo	27
<b>Figura 3.9</b>	Reprodução de um assento de óbito de 1715	28
<b>Figura 4.1</b>	Memex descrito por Vannevar Bush	43
<b>Figura 5.1</b>	Um assento de casamento de 1821 – Rolo de microfilme 107	54
<b>Figura 6.1</b>	Arquitetura final da Biblioteca Digital proposta	71
<b>Figura 6.2</b>	Interface para digitação dos assentos de casamento	78
<b>Figura 6.3</b>	Trecho do código em HTML para metadados de casamento	80
<b>Figura 6.4</b>	Tela principal para utilização pelo bibliotecário	81
<b>Figura 6.5</b>	Criando uma nova coleção	81
<b>Figura 6.6</b>	Importando os arquivos HTML para dentro da coleção criada	82
<b>Figura 6.7</b>	Processo da criação da biblioteca finalizado	82
<b>Figura 6.8</b>	Busca por documentos com a palavra chave Francisco	84
<b>Figura 6.9</b>	Resultado da busca realizada pelo Greenstone	84
<b>Figura 6.10</b>	Exibição de todos os campos no documento original	85
<b>Figura 6.11</b>	Exibição do documento digitalizado	85

## LISTA DE TABELAS E GRÁFICOS

<b>Tabela 1.1</b>	Comparação entre a condição ideal sugerida pelo comitê e a condição observada no local	22
<b>Tabela 4.1</b>	Características dos metadados	37
<b>Tabela 4.2</b>	Conjunto padrão de metadados proposto pelo Dublin Core	40
<b>Tabela 5.1</b>	Intervalo de documentos utilizados para criação da amostra piloto	52
<b>Tabela 5.2</b>	Quantidade final de documentos utilizados	52
<b>Tabela 5.3</b>	Metadados finais no Dublin Core: Planilha para casamento	58
<b>Tabela 5.4</b>	Metadados finais no Dublin Core: Planilha de batizados	61
<b>Tabela 5.5</b>	Metadados finais no Dublin Core: Planilha de óbitos	63
<b>Tabela 5.6</b>	Resumo dos elementos e qualificadores utilizados ou criados	64
<b>Tabela 5.7</b>	Resumo das tabelas de apoio no preenchimento dos metadados	65
<b>Tabela 5.8</b>	Significado de alguns termos menos comuns	66
<b>Tabela 6.1</b>	Metadados finais de casamento representados no MySQL: tabela de batismo	73
<b>Tabela 6.2</b>	Metadados finais de casamento representados no MySQL: tabela de casamento	74
<b>Tabela 6.3</b>	Metadados finais de casamento representados no MySQL: tabela de óbito	75
<b>Tabela 6.4</b>	Outras tabelas que constituem o sistema proposto	76
<b>Tabela 7.1</b>	Quantidade de casamentos realizados ordenados por padre que realizou e separados por condição social	87
<b>Tabela 7.2</b>	Palavra chave digitada pelo usuário e outras que seriam consideradas pelo sistema, verificando a semelhança sonora	92
<b>Gráfico 5.1</b>	Gráfico das distribuições dos campos Igreja e Cidade entre 1800 e 1850 nos assentos de casamento	53
<b>Gráfico 7.1</b>	Número de óbitos separados por estado civil e profissão	89

## SUMÁRIO

1.	INTRODUÇÃO.....	1
1.1	Objetivo geral.....	5
1.2	Objetivos específicos.....	5
1.3	Estrutura da dissertação.....	5
2.	CONTEXTUALIZAÇÃO HISTÓRICA E CONTEMPORÂNEA.....	8
2.1	O mundo dos navios de madeira.....	11
2.2	O surgimento de Vila Rica.....	13
2.3	As Igrejas e as irmandades.....	15
2.4	As cerimônias que originaram os documentos.....	16
3.	O ACERVO, DOCUMENTOS E USOS.....	18
3.1	Aspectos de preservação.....	18
3.2	A grafia nos documentos.....	27
3.3	Usos dos assentos.....	29
4.	O TRATAMENTO DIGITAL DA INFORMAÇÃO.....	31
4.1	Sistemas de Informações (SI) e Sistemas de Recuperação da Informação (SRI).....	33
4.2	Metadados e os SRI.....	35
4.3	O padrão MARC.....	38
4.4	O padrão Dublin Core.....	39
4.5	Bibliotecas digitais.....	42
5.	IDENTIFICAÇÃO DOS METADADOS.....	48
5.1	Aspectos metodológicos.....	48
5.1.1	Metodologia para identificação dos metadados.....	49
5.2	Aspectos interferentes durante a leitura paleográfica.....	54
5.3	Metadados segundo o Concílio de Trento.....	56
5.4	Resultados: Planilhas finais de metadados identificados.....	57
5.4.1	Resumo dos elementos e qualificadores utilizados e acrescentados.....	64
6.	A BIBLIOTECA DIGITAL: Métodos e resultados.....	69
6.1	Escolha do software para gestão da biblioteca digital.....	69
6.2	Proposta do banco de dados.....	72
6.3	Proposta da interface para digitação das planilhas.....	77
6.4	Exportação para a linguagem HTML.....	78
6.5	Criação da coleção no Greenstone.....	80
6.6	Recuperação a Informação pelo Greenstone.....	83
6.7	Aspectos sobre a visualização do texto completo.....	86
6.8	Outras formas de se visualizar a informação a partir dos metadados propostos.....	86
7.	CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS.....	89
	Referências.....	94
	Anexos.....	98

## 1. INTRODUÇÃO

O homem sempre buscou aumentar suas capacidades intelectuais e físicas, através de idéias e artefatos construídos por ele, externando dessa forma reflexões e pensamentos antes comunicados oralmente. Os primeiros desenhos em cavernas já retratavam esta necessidade de preservar e divulgar o conhecimento, apreendendo-o por meio de um artefato externo ao seu intelecto. Foi assim desde o princípio, e no mundo contemporâneo as tecnologias interferem positivamente nos processos de comunicação de informações. Por exemplo o telescópio que aumentou nossa capacidade de visão, a fotografia que nos auxilia a memorizar imagens, e o telefone que hoje elimina distâncias no processo de comunicação.

Mas dentre todas as invenções criadas pelo engenho humano, o computador talvez seja o instrumento tecnológico mais importante. Através dele o homem viu sua capacidade de acumular e processar informações aumentada sobremaneira, funções antes realizadas apenas nos cérebros dos emissores e receptores de mensagens. Esta máquina que possui funcionamento baseado na matemática dos números binários, tornou-se ferramenta indispensável em inúmeras atividades. Embora não determine a evolução histórica e a transformação social, a tecnologia (ou a falta dela) incorpora a capacidade de transformação das sociedades, bem como os usos que a ela decide dar ao seu potencial tecnológico (CASTELLS, 2003, p.44).

Com a invenção da Internet nos anos 70, amparada pela disseminação dos computadores em nível mundial, gigantescos acervos eletrônicos passaram a ser transmitidos pelo mundo e suprem de informações usuários e pesquisadores, a uma velocidade jamais imaginada. Ao apoiar e disponibilizar a produção científica humana, os computadores estão, neste momento, realizando a grande tarefa para a qual foram projetados. Podemos concluir que a quantidade de informações que trafega pela rede mundial pode ser utilizada como um indicador da capacidade de influenciar o futuro das sociedades. Assim, a questão fundamental a ser refletida pelas organizações de uma maneira geral (universidades, centros de pesquisa, laboratórios, empresas, governo, entre outros) é, sem dúvida, a capacidade de se inserir como um dos agentes ativos no uso desta nova tecnologia, como forma de disseminar o conhecimento.

Apropriando-se da tecnologia da informação, como aliada do historiador e dos profissionais da área de informação, e focalizando o lado do pesquisador, principalmente aquele envolvido com atividades que remetem a documentos de difícil acesso, é certo que seu trabalho tem sido sobremaneira facilitado. Isto fica evidente no tempo e no espaço necessários à realização de seus trabalhos de campo. Até pouco tempo, para se realizar uma pesquisa de ponta era necessário que os pesquisadores recorressem diretamente às fontes primárias, tarefa muitas vezes dispendiosa e, em alguns casos, impossível de ser realizada. Nos dias atuais boa parte das pesquisas podem ser realizadas sem que o pesquisador saia do seu local original de trabalho.

Evidenciando a pesquisa em história, o historiador muitas vezes utiliza-se dos fatos históricos, que são consultados em livros, periódicos, revistas científicas, arquivos, monumentos, bibliotecas digitais, acervos *on-line*, jornais e vários outros meios. Várias são as instituições ligadas a preservação de acervos que são normalmente utilizados por historiadores. Museus, arquivos públicos e privados, bibliotecas, universidades, órgãos do governo e diversos outros, são exemplos dessas instituições que mantêm acervos como fontes de informações para pesquisadores. A história dos povos e suas instituições está inscrita em objetos e documentos criados por nossos antepassados. Nesse sentido trabalhar pela sobrevivência dos acervos históricos talvez seja a chave para garantir nosso compromisso com o futuro. A tarefa de identificação, organização, descrição e transferência da informação dos documentos custodiados nos diferentes arquivos do país é gigantesca, até mesmo aterradora diante da escassa parcela levada a cabo até hoje (BELLOTTO, 1998).

Uma das grandes preocupações se deve ao fato de que a preservação desses documentos conta com o dilema entre guardá-los em local seguro, tendo como uma das conseqüências às dificuldades de acesso, ou expô-los ao público, o que sem dúvida traz enormes problemas de degradação contínua. O uso da tecnologia da informação surge como uma solução moderna que permitirá organizar, armazenar e disponibilizar o acesso aos acervos. Não se trata de um artigo de luxo mas, antes de tudo, de uma necessidade, não só sob o ponto de vista da preservação mas, sobretudo, de disponibilização.

Pode-se enumerar diversos motivos que levam as instituições a buscarem na tecnologia da informação uma aliada na preservação e disponibilização de seus acervos de documentos. Alguns dos principais são:

- Preservar o acervo original evitando o manuseio direto;
- Facilitar a consulta ao acervo de documentos já microfilmados ou digitalizados;
- Facilitar a criação de relações entre os documentos através do cruzamento de informações;
- Permitir o acesso simultâneo em várias regiões distintas com o uso das redes de computadores;
- Possibilitar uma melhora na qualidade de imagens através de recursos como brilho e contraste;
- Implementar métodos de busca de pesquisa alternativas como textual, palavras chaves, cruzamento de dados, entre outros;
- Permitir o uso de hipertexto, tornando as pesquisas mais dinâmicas e amigáveis.

Ao se analisar estas fontes e dispor-as eletronicamente, estamos simplificando o trabalho do pesquisador, que antes de recorrer às fontes primárias, pode utilizar uma base de dados documental como parte de sua atividade que é o “fazer história”.

Os documentos que serviram de objeto para esta pesquisa compõem o Acervo da Matriz do Pilar de Ouro Preto. A documentação eclesiástica presente nesta Paróquia abrange um acervo completo e preservado, resultado da atuação e empenho desta instituição, constituindo um dos arquivos mais importantes do Brasil, em se tratando de informações sobre a população do Brasil Colônia e Império. Vale ressaltar que a documentação paroquial estudada cobrirá apenas os registros de batismo, casamento e óbitos, porém, são muitos os documentos de irmandades como recibos, cartas, bilhetes, dentre outros existentes no acervo.

A pesquisa está limitada a estes documentos por dois motivos: primeiro, o curto espaço de tempo e a falta de recursos necessários a realização de uma pesquisa de maior envergadura. Segundo, porque mesmo os registros que serão tratados (assentos de batismo, casamento e óbitos) carecem

de maiores estudos pela história. Uma vez descritos em uma base digital poderão se tornar mais facilmente acessíveis, facilitando o trabalho dos pesquisadores.

Convém, já nesta introdução, destacar a importância da Igreja Católica detentora de ricos acervos de fatos históricos registrados em certidões de batismo, casamento e óbitos, os quais tanto ajudam a decifrar a sociedade de épocas passadas, como auxiliam uma possível interpretação do momento atual. Decorre deste fato a grande importância da preservação dos arquivos da Igreja que, durante muito tempo da história mundial, assumiu o papel de controle que mais tarde seria repassado ao Estado, tendo sido por muitos anos a única instituição responsável pelo recenseamento da população.

A existência deste acervo ocorreu porque a Igreja Católica desempenhava um papel fundamental na sociedade que se desenvolvia ao seu redor. “Sempre ao surgir um arrumado de simples aglomerados de casinhas, logo se cuidava da igrejinha, em geral, construída pelo próprio descobridor que se encarregava da construção dessas capelas iniciais” (SALLES, 1999, p49).

No período Brasil colônia não haviam escolas, hospitais ou estradas. Tudo o que a Coroa Portuguesa queria era extrair ao máximo as riquezas dessa terra. Por este motivo, como uma forma de construir e manter novos templos, Portugal estimulava a formação de irmandades religiosas. Essas, durante várias décadas, promoveram a construção de matrizes e capelas em todo o território das Minas Gerais. Uma delas, a Matriz do Pilar de Ouro Preto, de talha erudita e que cobre toda a nave e capela mor, representa autenticamente a pompa barroca. Sua construção foi arrematada por João Francisco de Oliveira em 1720, passados apenas pouco mais de vinte anos do descobrimento do caminho que levaria de São Paulo à região de Vila Rica em Minas.

Sua importância é retratada por diversos historiadores. Era lá que os governadores tomavam posse quando chegavam a Minas. Sua inauguração ocorreu no ano de 1733, ensejando acontecimento singular e raro na vida de Ouro Preto, durante o primeiro quarto do século XVIII. Para os historiadores de nosso tempo, toda a rica documentação presente no acervo desta Igreja representa uma inesgotável fonte de informações que permanece, em pleno século XXI, carente de pesquisas de maior envergadura.

Trata-se de um desafio para nossa sociedade, uma vez que as informações contidas neste e em outros acervos eclesiásticos desperta interesse de outros países e religiões, como a igreja dos Mormons também conhecida como a Igreja dos Santos dos Últimos Dias, que busca retratar a genealogia das famílias batizando os mortos, como forma de aliviar suas dores (MORMON, 1995). Segundo esta igreja, as pessoas mortas podem receber as ordenanças, o batismo por exemplo, mas não podem fazer isso por elas mesmas. Assim, conhecendo seus antepassados, as pessoas desta religião podem solicitar o batismo em qualquer capela da comunidade Mormon. Por isso, conhecer o acervo de documentos da igreja Católica, é uma importante fonte de informação para essa igreja.

### **1.1 Objetivo geral**

Identificar os metadados que respondam às necessidades de recuperação da informação na parte da documentação já microfilmada que compreende os assentos de batismo, casamentos e óbitos do acervo eclesiástico da Matriz do Pilar de Ouro Preto, os quais respondam às necessidades de recuperação da informação, na criação de uma biblioteca digital.

### **1.2 Objetivos específicos**

- Definir os metadados necessários à representação dos documentos;
- Organiza-los segundo o padrão Dublin Core para documentos manuscritos;
- Projetar um protótipo de biblioteca digital de textos completos que facilite o acesso às informações contidas nos documentos, utilizando software específico, preservando dessa forma o acervo primário e microfilmado.

### **1.3 Estrutura da dissertação**

O capítulo 2 trata da necessidade de preservação do conhecimento produzido pelo homem, como fator de sobrevivência social, partindo do conceito de capital cultural proposto por Bourdieu.

Apresenta a biblioteca de Alexandria, que pretendeu manter sob suas paredes nada mais do que todo o conhecimento humano que existisse sobre a terra, traçando um paralelo entre esta e a biblioteca do Congresso Americano, uma das maiores do nosso tempo. Explora o momento histórico mundial, brasileiro e mineiro, ambiente sob o qual o caminho que levava de São Paulo à Vila Rica foi descoberto, uma revisitação necessária, pois, o acesso ao acervo sem o conhecimento do contexto no qual ele está inserido, significaria pouco para a recomposição da história.

O capítulo 3 intitulado “O acervo, documentos e usos”, apresenta as características do acervo primário focalizado nesta pesquisa, que dará base à sua descrição para ser convertido ao meio digital. Foi realizada uma pesquisa documental para compreender melhor o suporte utilizado naquela época. Um fato interessante foi percebido: na época não existia a intenção de utilizar instrumentos (papel, capa, caneta) que durassem por tantos anos, mas vale destacar que se passaram quase trezentos anos e a maior parte dos documentos estão em perfeitas condições. Hoje não se imagina utilizar suportes que possam durar tanto tempo, mesmo que pelo processo de cópia, a informação possa se perpetuar.

Acervos como este constituem uma parcela significativa da memória de um país ou de um povo, e salvaguardam os direitos legais do Estado, das instituições e das pessoas. São eles que, através dos seus documentos, preservam os testemunhos da trajetória da evolução dos estados, e das suas comunidades, registrando elementos sobre a sua formação, seu desenvolvimento, ou sua decadência. A preservação resgata para a sociedade um dos maiores bens que ela possui: sua cultura, uma moeda forte, por vezes esquecida pelas sociedades modernas, imediatistas.

O capítulo 4 apresenta aspectos sobre o tratamento digital da informação. Foram identificados os processos básicos para organização e uso da informação no meio digital, e as etapas deste tratamento. São apresentadas as principais características dos sistemas de informação e seus módulos. Também neste capítulo procurou-se levantar as características dos metadados, os quais estão em posição de destaque no que diz respeito ao planejamento de sistemas para recuperação da informação. Caracteriza-se o padrão Dublin Core e as Bibliotecas Digitais.

O capítulo 5 denominado “Identificação dos metadados”, trata da metodologia e a planilha final com os resultados dos metadados identificados. Os documentos eclesiásticos tratados nesta dissertação foram criados por padres da Igreja Católica. Por sua vez, estes deveriam seguir os padrões definidos pelo Concílio de Trento<sup>1</sup> no momento de efetivar o registro seja de casamento, de batismo ou de óbito. Porém, e por diversos motivos, vários desses assentos possuem outros elementos diferentes dos definidos pela Igreja Católica. Além das informações comuns encontradas nos assentos, os detalhes tornaram-se uma rica fonte de estudos, pois a partir deles várias pesquisas podem ser desencadeadas. Este capítulo apresenta em detalhes os campos dos metadados identificados e discutidos com pesquisadores da área de história.

O capítulo 6 intitulado “A biblioteca digital”, apresenta os passos seguidos para criação e o protótipo da Biblioteca Digital de documentos Eclesiásticos a partir dos metadados identificados. Foi elaborado um protótipo que pretende cobrir os casamentos entre os anos de 1800 a 1850, os batismos e óbitos entre os anos de 1800 e 1830.

O capítulo 7 são as considerações finais apresentado as contribuições, limitações e os trabalhos futuros.

---

<sup>1</sup> O Concílio de Trento, realizado de 1545 a 1563, foi o 19º concílio ecumênico, convocado pelo Papa Paulo III para assegurar a unidade de fé e a disciplina eclesiástica. A sua convocação surge no contexto da reação da Igreja Católica à divisão que se vive na Europa do século XVI quanto à apreciação da Reforma Protestante. O Concílio de Trento foi o mais longo da história da Igreja: é chamado Concílio da Contra-Reforma. Emitiu numerosos decretos disciplinares. Este Concílio teve especial importância para os pesquisadores de genealogia devido a uma de suas resoluções, esta determinava que toda criança, para ser batizada na igreja católica deveria possuir um nome cristão e um sobrenome de família, desta forma, as famílias que ainda não o possuíam foram obrigadas a assumir o termo que os identifica, o uso de sobrenomes familiares foi então implantado definitivamente.

## 2. CONTEXTUALIZAÇÃO HISTÓRICA E CONTEMPORÂNEA

Bourdieu foi quem delineou o conceito de capital a ser utilizado quando referimo-nos a uma análise social, relacionando-se a cultura aos valores econômicos. Trata-se de uma metáfora utilizada para se referir a um tipo de poder que pode conduzir os indivíduos de uma sociedade a um nível socioeconômico mais elevado. No campo das ciências sociais o conceito formulado por Bourdieu permite ser utilizado em diversos campos como a educação, a arte a literatura e a cultura (BOURDIEU, 2003).

Lahire (2003), utilizando a teoria de Bourdieu, destaca que a privação do acesso a um determinado bem cultural constitui uma carência, uma deficiência ou uma injustiça inaceitável. Para Bourdieu o capital cultural é pessoal e leva tempo, sendo um trabalho do sujeito sobre si mesmo, tornando-se parte integrante da pessoa, destacando também que o legado de bens culturais transmitidos entre as gerações, pertencem realmente aos que detêm os meios para dele se apropriarem, ou seja, os bens culturais enquanto simbólicos só podem ser apreendidos por aqueles que possuem os meios para decifrá-los. Isso pressupõe instrumentos para sua apropriação.

Apesar dos esforços de diferentes instituições, o acesso a objetos culturais por pessoas com menores condições sociais, continua a ser um fator de desigualdade, uma vez que o acesso, segundo Bourdieu, continuará sendo apenas para aqueles que detêm o código de apropriação. O autor evidencia o uso da informática, como um fator diferenciador entre aqueles que utilizam a grande rede para realizar operações, e aqueles que ainda permanecem distantes desta realidade chamada sociedade da informação.

Na sociedade da informação, sistemas de disseminação da informação e utilização se tornaram de fundamental importância. Apesar dos conhecidos problemas da falta de infra-estrutura para uma real inclusão-digital, a cada ano mais e mais pessoas passam a ter acesso a computadores e outros recursos eletrônicos.

A estrutura tecnológica atual permite que as informações em repositórios como as bibliotecas digitais possam ser compartilhadas entre usuários, que também podem colaborar escrevendo artigos, resolvendo problemas e processando informações (CHEN, 1998). Em termos de colaboração, vários domínios do conhecimento podem se beneficiar do potencial das bibliotecas digitais, incluindo projetos de engenharia, pesquisas científicas, ensino a distância e diversos outros, criando um ambiente de memória social.

Mas, a busca pelo acúmulo de conhecimento, como base para um poder simbólico e potencialmente pronto para se tornar capital econômico, tem seu início há muito tempo atrás e se manifestou de diversas formas, dentre elas cultural, científico e tecnológico dos povos. Um bom exemplo foi a Biblioteca de Alexandria.

Alexandria foi fundada por Alexandre, o Grande, em 332 a.C. e pretendia, com todos os méritos, ser a capital intelectual mundial. Um centro cosmopolita, com praças, mercados, judeus, persas, árabes, romanos e tantas outras nações. Sua posição entre rotas navais, fluviais e terrestres de três continentes, e sua efervescência, fez com que ela fosse uma cidade de destaque em sua época, uma Paris ou Nova York do nosso tempo.

O museu de Alexandria pode ser considerado a primeira universidade do mundo, uma vez que servia como um grande e verdadeiro instituto de investigação da história. A biblioteca de Alexandria não era apenas um depósito de rolos de papiro e de livros, mas fonte de inspiração para homens, filósofos, e todas as áreas da ciência, que ali podiam desbravar o mundo do conhecimento e das emoções. Sua história, antiga, de uma biblioteca que durou mais de mil anos, não pode ser resgatada de maneira clara, não por falta de informação sobre ela, mas pelo mesmo mau que assola nossa sociedade, excesso de informações que muitas vezes se contradizem. Nem mesmo a data e os motivos de sua destruição são conhecidos, acredita-se que seja por volta do século VII d.C, feito realizado pelos árabes ou por Teodósio que teria mobilizado uma multidão cristã que se encarregaria de sua demolição, pois a Igreja Católica acreditava que ali estava um depósito de maldades do paganismo<sup>2</sup> e do ateísmo<sup>3</sup> (CÂNFORA, 1989).

---

<sup>2</sup> Paganismo – religião onde existem vários Deuses.

<sup>3</sup> Falta de crença em Deus.

Observando Alexandria, conhecemos um povo que fez da conservação do passado o centro de sua cultura, deixando feitos até hoje admirados, como a reconstrução do pensamento de Aristóteles, a tradução da Bíblia para o Grego e sua divulgação no ocidente, recuperando e preparando todo conhecimento dos poetas gregos para que fosse, ainda hoje, a base de muitos esforços intelectuais.

A biblioteca havia sido idealizada por Demétrio, um talentoso filósofo que pretendia com a empreitada, tornar-se rival cultural de Atenas. Como plenipotenciário da biblioteca, a qual por vezes era alvo da visita do rei, ele perguntava – Quantos rolos temos? Demétrio como o guardião de todo conhecimento da terra, bibliotecário supremo, atualizava os números. Cálculos haviam sido feitos, e os números eram grandiosos até mesmo para aquela época. Chegaram a uma conclusão que seriam necessários 500 mil rolos para que todo o conhecimento escrito ao redor do mundo pudesse estar em Alexandria.

Para que os propósitos da biblioteca fossem atingidos, Ptolomeu havia escrito uma carta a todos os governantes da terra. Na carta pedia-lhes que não hesitassem em enviar para Alexandria tudo o que fosse escrito sobre tudo o que existisse nas terras sobre as quais comandassem. Ordenou também que tudo o que existisse nos navios que passassem por Alexandria deveria ser copiado, ficando os originais em Alexandria e sendo as cópias devolvidas. Em certos momentos, Demétrio fazia uma exposição ao Rei:

Em obediência à tua ordem de acrescentar às coleções da biblioteca, para completá-la com os livros que ainda faltam, e de restaurar adequadamente os defeituosos, dediquei grande cuidado, e agora faço-te um relatório:...

... Demétrio ao grande Rei  
(CANFORA, 1989, p.23)

Ser o bibliotecário de Alexandria era atingir o Olimpo, era conduzir todo o conhecimento do mundo, viver cercado por milhares de manuscritos sobre todas as áreas do conhecimento. Assim como em qualquer biblioteca, Alexandria também sofre a ação da perda do acervo por alguns incidentes. Dentre eles, boa parte se queimou em um incêndio que pôs a perder parte dos 700 mil

rolos que chegou a possuir. Acontecimentos como este, de algum modo, impediram que ela continuasse sendo visitada por vários cientistas da época. A biblioteca também foi alvo de falsários; conta a lenda que uma das tarefas dos bibliotecários era de distinguir obras falsas das autênticas.

Toda essa história só pode ser contada porque documentos foram preservados. Nunca, por exemplo, os arqueólogos acharam qualquer indício físico da existência do museu de Alexandria, nem ruínas, nem qualquer tipo de objeto. Apenas através das descrições de escritores antigos é possível reconstruir esta e outras histórias. H. G. Wells em seu livro “Uma Breve História do Mundo”, considera que a fundação do museu e da Biblioteca de Alexandria representa o verdadeiro início da História Moderna (FLOWER, 1999).

Em nossos dias o ápice do desejo de controlar o capital cultural pode ser percebido na biblioteca do Congresso Americano, inaugurada em 1802, sendo a maior do mundo, com seus mais de 14 milhões de livros distribuídos por 858 quilômetros de prateleiras, apresentando o poder das nações que possuem conhecimento e acesso a ele. Em seu ambicioso projeto, a Biblioteca do Congresso pretende digitalizar seu acervo e todas as publicações que venham surgir no futuro. O acesso, que hoje é feito apenas pessoalmente, e a um livro de cada vez, poderia ser feito remotamente e por qualquer cidadão do mundo.

Sob o olhar de Bourdieu a questão do acesso perpetuará, pois o código para decifrar estes documentos está nas mãos de uma pequena parcela da sociedade.

No próximo item apresenta-se o contexto histórico mundial da época onde os documentos considerados foram criados.

## **2.1 O mundo dos navios de madeira**

Este trabalho não estaria completo se não se reportasse ao momento histórico mundial e particularmente da colônia brasileira e da região de Vila Rica, local no qual os documentos tratados nesta dissertação tiveram origem. Obviamente que mesmo sem nos aprofundarmos nos

requintes históricos daquela época, pode-se sentir a atmosfera dos navios de madeira que aportavam em nossa costa marítima, trazendo tendências marcantes e influenciando nossa sociedade.

O final do século XVII foi um momento de grandes transformações no mundo. Alguns historiadores como John E. Wills, relatam que o mundo moderno começou nesta época. Era um tempo de enorme poder. Pedro, o Grande, começava um reinado que iria transformar a Rússia e a Europa; Kangxi governava o vasto império Manchu a partir da Cidade Proibida em Pequim. Carlos II, rei da Espanha, extraía poder das minas de prata na América do Sul e do México. Ao lado deles, Isaac Newton fazia brilhantes descobertas na ciência, Locke instigando-nos a pensar sensível e cuidadosamente sobre os enigmas mais profundos da realidade, conhecimento e direito político (WILLS, 2001).

O mundo islâmico se estendia de Pequim até a Nigéria, mas só alcançava o novo mundo quando algum desafortunado muçulmano africano sobrevivía à travessia para se tornar escravo de algum senhor na América.

Neste mundo reinavam as florestas e alguns poucos povoados, pastagens que hoje se tornaram cidades. Era um mundo silencioso, de doenças infecciosas, de morte por enforcamento em praça pública. Quase todos viviam influenciados pelos seus ancestrais, a mudança social era sempre muito lenta. Aliás, as grandes mudanças por vezes eram apenas para resgatar alguma velha tradição que já estivesse sendo esquecida. (WILLS, 2001, p. 25)

Mesmo assim, alguns intelectuais europeus sentiam a possibilidade de mudanças. Eles se apoiavam em algumas observações que podiam fazer sobre tudo que estava ao redor: o surgimento da ciência, o crescimento de algumas cidades, políticas para promover o crescimento econômico e uma variedade imensa de publicações.

O mundo ainda não estava geograficamente desenhado como conhecemos hoje, algumas regiões como a Austrália, as costas do Pacífico e a Antártida sempre eram deixadas em branco ou apenas esboçadas. Os europeus que cruzavam o mundo, entre eles os portugueses, estavam construindo as civilizações no mundo novo, deparando-se com uma quantidade de culturas ainda nem sonhadas.

Neste contexto, a colonização do Brasil impôs a Portugal a tarefa de construir uma sociedade a partir do nada, pois nada havia de lucrativo em uma terra de longas distâncias e imensas florestas.

No final do século XVII, próximo da descoberta do caminho que levaria de São Paulo à região de Ouro Preto, a comunicação entre o mundo dependia inteiramente de entregadores de cartas, os quais podiam levar anos para chegarem aos destinos.

Segundo a pesquisadora Vera Alice Silva, ao longo do século XVIII a coroa portuguesa experimentou diversas fórmulas para organização do governo na colônia brasileira, na verdade, sempre pretendendo elaborar um modelo mais eficaz de controlar as atividades que resultassem em lucros. A fórmula básica partia do princípio de que a ocupação, exploração e defesa do território deveriam se autofinanciar e, além disso, permitir o máximo de lucro para a coroa. Para isso Portugal tomava alguns cuidados, como não deixar autoridades máximas em seus cargos por muito tempo e não permitir que os governadores tivessem interferências nos assuntos econômicos. Neste período, mais precisamente, na segunda metade do século XVIII, o governo do Brasil, agora centralizado, comprou as últimas capitânicas donatárias. Nesta época, até mesmo a Igreja era sustentada pelos colonos sendo que, o próprio Rei se beneficiava do pagamento do dízimo. Ficava assim estabelecido um padrão de controle que se auto-sustentava e garantia o envio de lucros para Portugal (SILVA, 1999).

## **2.2 O surgimento de Vila Rica**

O surgimento de Vila Rica não poderia ter ocorrido em outro momento. Como sobredito, o mundo passava por transformações históricas que o levaria às mais importantes descobertas. Vila Rica, atual Ouro Preto, nasceu da aventura de Antônio Dias de Oliveira, no ano de 1696, que bravamente estabeleceu o caminho que levaria de São Paulo ao Pico do Itacolomi, que determinava o ponto certo onde a primeira expedição havia dito encontrar ouro em abundância, reluzindo sob as águas de qualquer riacho. Segundo Marcos Ferreira, a comarca de Ouro Preto era a menos extensa da província mineira, mas mantendo sua importância até quase o final do império, e compreendia hoje a Ouro Preto e Mariana (ANDRADE, 1997).

De Vila Rica dependiam dois grandes distritos, Nossa Senhora do Pilar e Antônio Dias. A figura 2.1 mostra como a cidade se organizava naquela época.

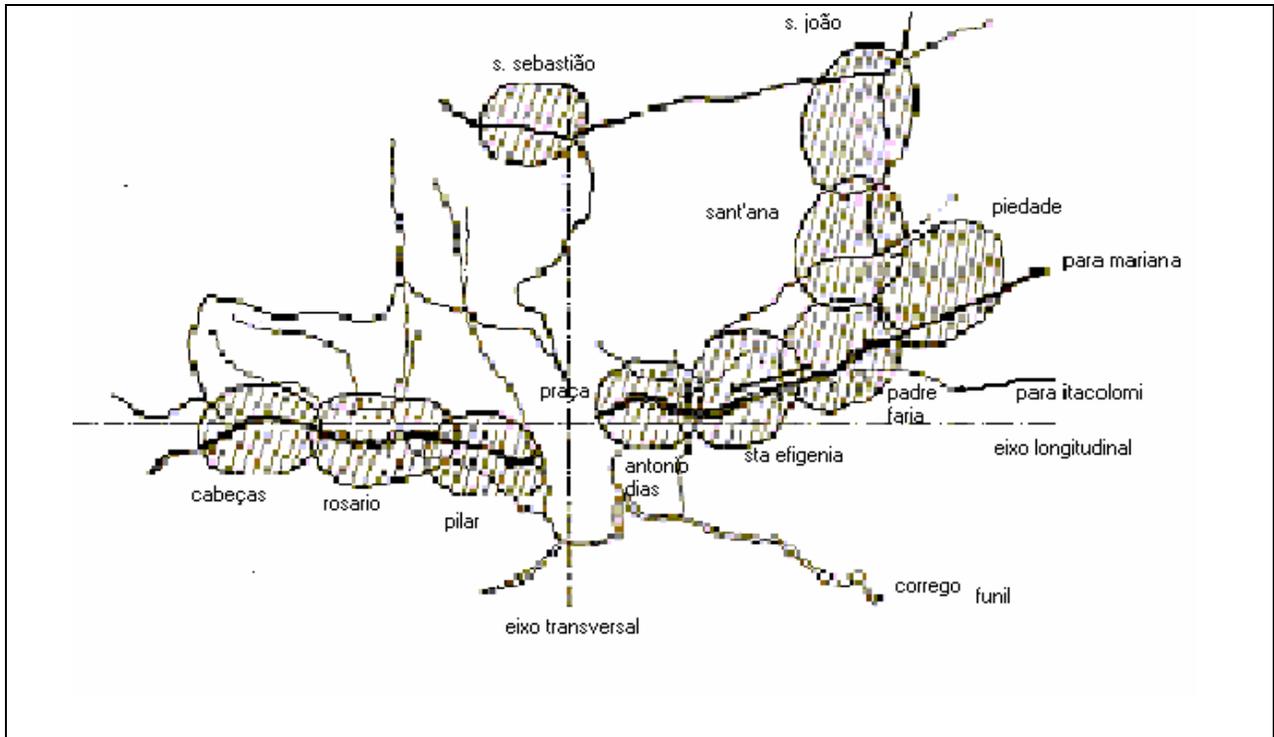


Figura 2.1 – Mapa da região de Vila Rica

Fonte: VASCONCELLOS, Sylvio de. *Vila Rica*. São Paulo: Editora Perspectiva, 1977, p.78.

No auge da mineração, a população havia atingido aproximadamente vinte mil habitantes. Mas com o fim do Ouro na região, a cidade começou a perder seus encantos e várias casas estavam semi-abandonadas.

Caio Prado Jr. em “A formação do Brasil contemporâneo” (PRADO JUNIOR, 1963) investigou a sociedade brasileira contemporânea a partir da sua formação colonial e salientou, o que também era percebido em Vila Rica, que o sistema instalado pela metrópole era caótico e irracional.

O temor de contrabando de ouro, somando-se as precárias condições de fiscalização foram, segundo a pesquisadora Cláudia Maria, um dos maiores empecilhos para que as vias de comunicação das terras centrais com as regiões litorâneas se desenvolvessem. Apenas no início do século XIX isso começou a ocorrer (CHAVES, 1996).

Hoje, revivendo o Ciclo do Ouro, podemos ver no interior perdido do Brasil por entre serras e pedras com imensas matas virgens, a única estrada para o litoral rigorosamente policiada por forças militares que impediam o acesso a estrangeiros. Passado o momento de glória pelo ouro, ficou a cultura, toda música, pintura, poesia e escultura. Era gente a quem tinha sido negada qualquer forma de saber, mas que respirava arte. Uma escultura poderosa, um barroco renovado e uma poesia romântica que já rompia com as tradições portuguesas, tomando seu próprio destino. Criava-se um pensamento político, impulsionado pelo profundo anseio de independência que até hoje molda a sensibilidade nacional.

Assim se fez e se refaz Ouro Preto, berço da nossa história. Dona de um rico acervo histórico, de documentos, escritos com mãos firmes, com tinta grossa, que perdura até hoje e nos permite realizar uma releitura do passado.

### **2.3 As Igrejas e as irmandades**

Como uma forma de transferir para o povo o custeio da construção das Igrejas, a coroa portuguesa incentivou desde o início a criação de irmandades. Com isso, as igrejas em Minas não foram construídas como as do litoral, as quais foram erguidas com recursos das grandes ordens religiosas européias.

As irmandades construíram altares laterais nas igrejas, fato que, após conflitos, algumas abandonavam a matriz e criavam seus próprios templos que seriam as capelas filiais.

Já no final do século XVII, os arraiais nasciam do trabalho dos aventureiros que recebiam um pedaço de terra para exploração do ouro. Surgiam então as capelas, que com o passar dos anos eram aumentadas e ganhavam estrutura de igreja.

Segundo Marcos Magalhães, a organização administrativa fez com que a Igreja concentrasse em suas mãos a fiscalização das irmandades, aprovando compromissos, acompanhando o cumprimento dos instituidores, além de verificar as finanças através dos livros de receita e

despesa (AGUIAR, 2003). Naquele tempo, segundo Ramon Fernandes Grossi, o mundo era o espaço do sagrado, pois todos os habitantes estavam em contato com o plano divino. O mundo terreno e o mundo divino se constituíam em um único espaço onde só a fé cristã era considerada verdadeira. A Igreja só enxergava o sagrado por ela estabelecido e normatizado, numa vida mágica para os fieis.

Os moradores das Minas não seguiam o modelo de vivência religiosa defendido pela Igreja de inspiração tridentina. Podia-se perceber a existência de uma cultura religiosa singular, expressada no cotidiano...A religião, enquanto um conjunto de regras estabelecidas no seio da instituição eclesiástica, convivia com a religiosidade, esta enquanto um número diversificado de comportamentos e ações que emergiam do processo diário de construção do viver em Minas” (GROSSI, 1996, p.90)

Os documentos existentes na Matriz do Pilar desempenharam um papel importante no controle imposto pela Coroa Portuguesa. Era através desta informação que o Rei tomava medidas com o objetivo de manter o controle sobre a vida da população. Informações sobre quantidade de crianças, mortos, ricos, etc, eram totalmente conhecidos por Portugal.

#### **2.4 As cerimônias que originaram os documentos**

No início do século XVI surgiu na Europa um movimento chamado Contra Reforma. Neste movimento teve início a necessidade de se controlar cada membro da Igreja Católica. O tema foi levado ao Concílio de Trento, em 1545, onde foi debatido durante longo tempo. Naquele momento estava por ficar estabelecida uma resolução que mudaria para sempre a vida dos católicos, resolução esta que ainda hoje somos submetidos. Ficou decretado que o Cura<sup>4</sup> passaria a ser responsável pelo registro de cada batismo e de cada matrimônio celebrado em sua paróquia. Já neste momento a fórmula a ser utilizada foi minuciosamente estabelecida, formato este que tornou-se único em todo o mundo. Logo a Igreja se deu conta de que seria necessário também registrar as mortes, principalmente para evitar os abusos, como a bigamia. Esta obrigatoriedade foi estabelecida pelo Papa Paulo V em 1614. Obviamente que após os registros tornarem-se rotineiros pelo mundo, em algumas dioceses, como da Europa e mesmo do Brasil, eles foram aprimorados e enriquecidos com outras informações. Por exemplo, no caso da França, assinalava-

---

<sup>4</sup> O pároco; o vigário.

se a idade de cada cônjuge ao casar. No Brasil colonial, o registro de casamento de famílias proprietárias mereciam um registro especial de página inteira (MARCÍLIO, 2004).

Esta breve contextualização histórica é de vital importância para qualquer implementação de bibliotecas digitais que busquem levar um acervo existente no mundo real para o mundo virtual, uma vez que longe do seu contexto original, o percurso historiográfico pode se tornar fundamental para o sucesso deste tipo de projeto. Conforme exposto pode-se perceber claramente a riqueza das informações contidas na invejável história da cidade de Ouro Preto e mais especialmente a Matriz do Pilar, pois nos livros de batismo, casamento e óbito estavam inscritos toda a população, sejam ricos, pobres, nobres, brancos, negros, índios, homens, mulheres, todos sem exceção. Assim todos passaram a estar registrados em documentos que, após trezentos anos, temos em nossas mãos para uma importante releitura da sociedade daquela época.

### 3. O ACERVO, DOCUMENTOS E USOS

#### 3.1 Aspectos de preservação

Como não poderia deixar de ser, num país como o Brasil, os entraves com a preservação deste tipo de acervo são ainda mais problemáticos do que em países desenvolvidos. A falta de pessoal qualificado e recursos financeiros é algo recorrente em várias instituições. Além de outros motivos, este é um dos que contribuem para que boa parte dos registros já não existam mais e, encontrar séries completas como a presente no acervo da Matriz do Pilar, é uma oportunidade rara. Esta constatação já justifica a criação de bibliotecas digitais que podem, até certa medida, aliviar a manipulação dos acervos históricos preservando o documento. Essa é uma preocupação de diversas entidades como arquivos e museus que hoje, mais do que nunca, utilizam o formato digital para apresentarem suas coleções, sejam em *CD-Rom* ou via Internet.

O que percebemos atualmente é que, mesmo o acervo da Matriz do Pilar, que se constitui como um dos mais ricos do Brasil, é armazenado de forma inadequada (figura 3.1) em local impróprio e ainda sujeito à ação de vândalos.



Figura 3.1 – Local onde os documentos ficam estocados antes de serem tratados pelos técnicos da preservação

O acesso aos originais está cada dia mais restrito por diversos fatores, que vão desde a falta de pessoal para o atendimento aos visitantes, até o uso inadequado pelos próprios usuários. Antes de

ser implantado tal acesso restrito, não era incomum a presença de pessoas mal intencionadas que roubaram ou danificaram parte das coleções.

Porém, e apesar de todos os problemas a serem enfrentados, a renovação dos estudos utilizando as fontes paroquiais continuam a dar frutos para a história do Brasil. Prova disso são os inúmeros trabalhos apresentados na revista *Vária*<sup>5</sup>, que representou um desafio nos estudos da história mineira e brasileira. Os vários estudos realizados e em andamento, ora utilizam o acervo original, ora os documentos micro-filmados e por vezes o banco de dados instalado em um microcomputador no Departamento de História da UFMG, onde estão transcritos os resultados de leituras realizadas nos microfilmes.

Para destacar a importância deste acervo, estão apresentados alguns artigos que foram escritos por pesquisadores de diversas instituições, tanto do Brasil quanto do exterior, os quais tiveram como suporte a documentação da Paróquia: Donald Ramos da Universidade de Cleveland nos Estados Unidos estudou “O lugar do batismo e compadrio na sociedade de Vila Rica durante o século do ouro”; Adalgisa Arantes do Nascimento e Renato Franco (FAFICH) realizaram estudos sobre “O significado religioso do batismo”; Maria Luiza Marcílio da USP estudou “Os registros paroquiais e a história do Brasil”; Douglas Cole e Tarcísio Botelho estudaram o “Batismo de crianças legítimas e naturais na Paróquia do Pilar”; Vera Alice Silva da UFMG estudou os “Aspectos da função política das elites na sociedade colonial brasileira”; Maria de Fátima Gouvêa da UFF estudou “Os poderes de Vila Rica”; Liana Maria Reis da PUC de Minas Gerais estudou “Os escravos, armas e política de desarmamento na capitania setecentista”.

Estas são apenas algumas das pesquisas que contaram com informações retiradas do acervo, tanto diretamente na fonte primária, quanto pelo acesso aos registros digitados em sistema computadorizado.

No início da década de 70, sob direção de Rui Mourão, então diretor executivo da Fundação de Arte de Ouro Preto, coordenação de Myriam Andrade Ribeiro e com a participação de Suely

---

<sup>5</sup> Revista do departamento de História da UFMG. A edição 31 foi reservada para publicação de artigos que utilizaram de alguma maneira o acervo da Matriz do Pilar de Ouro Preto. Uma nova edição com este propósito está prevista para 2008.

Perucci, foi iniciado o primeiro trabalho de imunização, organização e catalogação dos arquivos das igrejas da cidade de Ouro Preto, dentre elas a Matriz do Pilar. Posteriormente, sob a coordenação do Museu da Inconfidência, Suely Perucci retomou o trabalho colocando à disposição dos pesquisadores um catálogo que tornou evidente a riqueza da documentação existente.

Em seguida, nos idos de 1973, o Centro de Estudos do Ciclo do Ouro (CECO), órgão de pesquisa subordinado à Escola Superior de Administração Fazendária (ESAF), instalado em 1973 no Museu Casa dos Contos em Ouro Preto, realizou a microfilmagem do acervo Casa dos Contos. O projeto inicial foi ampliado com a implantação do laboratório de microfilmagem e a constituição de equipe especializada. Além da documentação de caráter econômico-financeiro, outros acervos referentes Minas do período colonial foram se integrando à documentação original. Através de parcerias, o CECO se encarregou da classificação e microfilmagem de arquivos históricos, colocados à disposição dos pesquisadores. Portanto, além de subsidiar a pesquisa histórica, através de amplo trabalho de inventariação de fundos documentais, o CECO tem contribuído para a preservação de nossas fontes primárias. Foram classificados, microfilmados e inventariados os seguintes documentos:

- Arquivo Criminal do Fórum de Ouro Preto;
- Arquivo Judiciário do Fórum de Ouro Preto;
- Arquivo do Ministério da Fazenda em Minas Gerais;
- Arquivo da Câmara Municipal de Ouro Preto;
- Arquivo da Paróquia de Nossa Sr<sup>a</sup> da Conceição de Ouro Preto ;
- Arquivo da Paróquia de Nossa Sr<sup>a</sup> do Pilar de Ouro Preto.

A documentação presente no acervo da Matriz do Pilar data do período a partir de 1700 até nossos dias, uma vez que a igreja continua realizando casamentos e batismos. Trata-se de uma rica fonte de estudos, como pode-se perceber no catálogo que serve para indicar a referência exata de qualquer material disponível no acervo<sup>6</sup>.

---

<sup>6</sup> Este catálogo foi criado em 1971 pela Casa dos Contos, e ainda hoje é o principal guia de acesso para localização de documentos no acervo original.

Os livros que contem os assentos estão atualmente embalados e armazenados cronologicamente como mostra a figura 3.2. Quando da necessidade de leitura de determinado volume, o visitante pode contar com o catálogo, o qual aponta para o local onde está o volume, de acordo com o período a ser pesquisado.

Apesar de já terem passado pelo processo de preservação, com a eliminação de fungos e climatização, o armazenamento não cumpre todas as regras que precisariam ser aplicadas à preservação deste tipo de acervo, um dos principais motivos é a falta de recursos, como sobredito.



Figura 3.2 –Após o processo de limpeza e higienização, os livros são acondicionados em uma estante, cronologicamente.

Obviamente a tarefa de preservação é árdua e movimenta elevadas cifras. Segundo o trabalho realizado pelo Comitê Técnico de Obras Raras apresentado pela Universidade Federal do Rio de Janeiro (UFRJ) em 1993, a conservação de acervos deve ser realizada por dois motivos:

- Para uma conservação preventiva: que atua na preservação do acervo com o objetivo de prevenir danos. Inclui o monitoramento, higienização, planejamento de desastres, monitoramento das condições ambientais, dentre outros.
- Para conservação corretiva: servindo para remediar a deteriorização física ou química, e deve contar com profissionais qualificados, se caracterizando por ser um processo de altos custos.

O trabalho da UFRJ destaca ainda a necessidade de um planejamento da preservação, prevenindo o acervo da melhor forma possível dos agentes causadores de danos aos documentos, como insetos e a própria ação do tempo. Segundo este comitê, os fatores que mais fragilizam ou danificam a maior parte dos acervos são: temperatura, umidade (medida por instrumento conhecido como higrômetro), iluminação, poluição, agentes biodeterioradores e a ação do homem muitas vezes com atos de vandalismos ou descuido. A figura 3.3 mostra um livro com as folhas dobradas.



Figura 3.3 – Um livro com registros de batismos do século XVIII com as folhas dobradas decorrentes de manuseio inadequado

A seguir temos a situação do acervo da Matriz do Pilar observadas a partir de visitas ao local, comparadas com as normas ideais de preservação. Destacam-se os principais fatores a serem observados, a situação ideal e a condição atual.

**Tabela 1.1 – Comparação entre a condição ideal sugerida pelo comitê e a condição observada no local**

<b>Fator</b>	<b>Ideal</b>	<b>Condição atual do acervo</b>
Temperatura	19° a 23°	Ambiente – Não monitorada
Umidade relativa	50% a 60%	Ambiente – Não monitorada
Iluminação	Não incidir diretamente, as janelas devem possuir filtros	Não existe filtro, as grandes janelas ficam abertas durante o dia
Poluição atmosférica	Utilização de filtros no ambiente	Não existem filtros
Insetos	Eliminar a presença de roedores de superfície, baratas, traças, cupins	Controle simples
Ação do homem	Crítérios para manusear a documentação	São observados os critérios mínimos para a manipulação

Existem diversos procedimentos que podem aumentar a vida útil de um acervo documental, como não manusear os documentos com as mãos sujas; não manter plantas aquáticas nos locais próximos ao acervo; evitar a infiltração de goteiras; em dias muito úmidos (que são muito

comuns em Ouro Preto) não abrir portas ou janelas; não fumar próximo ao acervo; não utilizar fitas adesivas nos documentos; não retirar o livro da estante pela sua borda; utilizar carrinhos especiais para o transporte entre locais, dentre outros<sup>7</sup>. Obviamente que é do conhecimento e desejo de todos que lidam com a preservação deste acervo a necessidade de um melhor manejo, entretanto, a falta de recursos ou a falta de vontade política, algo muito comum nas instituições públicas brasileiras, impede que um trabalho melhor seja realizado.

No que diz respeito aos documentos presentes no acervo, um detalhe que chama a atenção é a pompa das capas utilizadas. Pelas visitas realizadas pode-se observar que aparentemente existiria uma variação ao longo do tempo, mas, em alguns casos, a capa possui o luxo que merece, como no caso de livros utilizados pelas irmandades os quais podem ser encontrados confeccionados em veludo. Um exemplo é o livro de compromisso da Irmandade de Rosário dos Pretos o qual foi trabalhado em veludo azul com um brasão em prata bem ao centro. Alias, o veludo era um tecido nobre naquela época, fato que contribuiu para que os livros mais importantes o utilizassem como acabamento da capa. Uma das capas utilizadas para acondicionar as folhas dos registros era confeccionada em papelão com as bordas em couro, ela pode ser encontrada, por exemplo, nos documentos de óbito entre 1818 e 1854. Apesar do tempo decorrido, aproximadamente 200 anos, ela ainda está bem conservada (figura 3.4).



Figura 3.4 – Livro com capa de papelão com bordas em couro

---

<sup>7</sup> Estas regras foram retiradas do Manual de Conservação de Acervos Bibliográficos da UFRJ e pode ser encontrado na Web pelo endereço: <http://www.forum.ufrj.br/biblioteca/artigo.html>.

Uma outra capa encontrada possui uma espécie de tecido o que, aparentemente, facilitou o ataque de traças (figura 3.5), suas condições de preservação são mínimas. A sensação é que a qualquer momento ela pode se desmanchar.



Figura 3.5 – Capa revestida em tecido

A terceira capa encerrava um livro onde estão assentos de casamentos entre 1804 e 1833. Uma capa totalmente em couro, com ricos detalhes em baixo relevo, mas também em estado bastante avançado de decomposição (figura 3.6).



Figura 3.6 – Capa com acabamento em couro

A quarta capa também é de papelão assim como a primeira, porém menos denso. Datava assentos entre 1834 a 1865 registrando os casamentos celebrados.

A última capa analisada foi toda feita em couro, mas com detalhes em papel que imitava a textura do mármore. Trata-se de um livro de óbitos de 1827 a 1876, bastante preservado (figura 3.7).



Figura 3.7 – Capa com detalhes imitando o mármore

Uma conclusão que pode-se chegar é que, de uma maneira geral, a capa escolhida variava de acordo com a utilização diária do livro, por exemplo, a capa em couro seria utilizada quando o livro fosse manuseado com mais frequência, pois o couro iria evitar um desgaste maior pelo uso. Capas em veludo seriam utilizadas em livros que teriam um manuseio menor e fossem mais importantes, como por exemplo os missais e os livros com as regras sociais das irmandades, conhecidos como livro de compromisso.

Ao analisar o papel na época, percebe-se que até final do século XVIII, como era de se esperar, apresenta-se bem mais artesanal do que o encontrado a partir do século XIX. Com isso, alguns registros feitos a caneta em papel mais fino, transpassa a folha em ambos os lados. Isso traz ao processo de microfilmagem ou digitalização um problema adicional, pois é necessário utilizar técnicas que não permitam que a imagem da parte de trás da folha, interfira no processo. Além disso, o papel do século XVIII é mais fibroso e as vezes de cor esverdeada. O papel do século XIX já aparece com pautas, como o papel utilizado hoje. Um detalhe interessante é que em algumas folhas, como no livro 509, existe uma figura em alto relevo onde está escrito a palavra *Almasso* e um símbolo que lembra um brasão.

Há de se observar, como lembra a pesquisadora Mirian Lott <sup>8</sup> que as folhas vinham avulsas e, apenas eram encadernadas após serem preenchidas.

---

<sup>8</sup> Mirian Lott é aluna do curso de doutorado em história da UFMG. Atualmente trabalha com os assentos de casamentos realizados na segunda metade do século XIX, tendo orientado algumas visitas à Matriz do Pilar durante esta pesquisa.

Obviamente que o papel utilizado, principalmente nos primeiros anos, vinha diretamente da Coroa Portuguesa. De acordo com Arnaldo Campos em “A Arte do livro”, o papel não sofreria modificações desde a sua invenção no século XII até meados do século XVIII onde permaneceria espesso e escuro, dependente de trapos brancos e de água límpida, na proporção de dois mil litros por quilo de papel a ser produzido. O primeiro papel de melhor qualidade foi criado por John Baskerville e o primeiro livro utilizando este papel foi editado em 1757. O preço do papel nesta época representava 50% do preço final do livro (CAMPOS, 1994).

Uma particularidade que pode ser observada em alguns livros é o preenchimento do documento com tinta a base de minério, elemento em abundância na região de Ouro Preto. A utilização deste tipo de tinta no preenchimento de alguns livros, faz com que o restaurador tenha a necessidade de redobrar os seus cuidados no momento da limpeza dos originais. Isto se deve ao fato de que esta tinta fica na superfície da folha de papel, fazendo com que, ao se tentar limpar o documento, a escrita seja danificada.

### **A ação de vândalos**

A ação de vândalos é um dos principais fatores contra a preservação, neste trabalho advogada como processo fundamental. Os atos de vandalismos não se restringem apenas à pessoas despreparadas, mas também pode ser percebido pela ação de pessoas mal intencionadas ou que buscam vantagem própria. Por isso os atos de vandalismo podem ser divididos em duas classes. Na primeira, temos o vandalismo por pessoas inocentes no que diz respeito a reconhecer a necessidade eminente da preservação. São facilmente encontrados riscos a caneta, cortes com estiletes, marcações a lápis e diversos outros (figura 3.8).



Figura 3.8 – Livro de 1780 com marcações feitas por um usuário que utilizou uma caneta esferográfica para rabiscar a lateral do livro.

Outra classe é composta por pessoas interessadas em vantagem própria. Em diversos documentos pode-se ver o suposto roubo de registros importantes, que vão em seguida parar nas mãos de colecionadores particulares.

### **3.2 A grafia nos documentos**

Um dos maiores desafios nos processos de análise tal como a desenvolvida no âmbito desta dissertação, é a interpretação da grafia utilizada em determinados momentos da história. Hoje, analisando documentos de diversas épocas, podemos perceber uma freqüente escrita que resulta em complexidade na leitura. Vale lembrar que poucas eram as pessoas que sabiam escrever nos séculos XVIII e XIX.

Para realizar a leitura dos documentos daquela época é necessário que se conheça a leitura paleográfica, segundo o dicionário Silveira Bueno esta é a arte de decifrar escritos antigos.

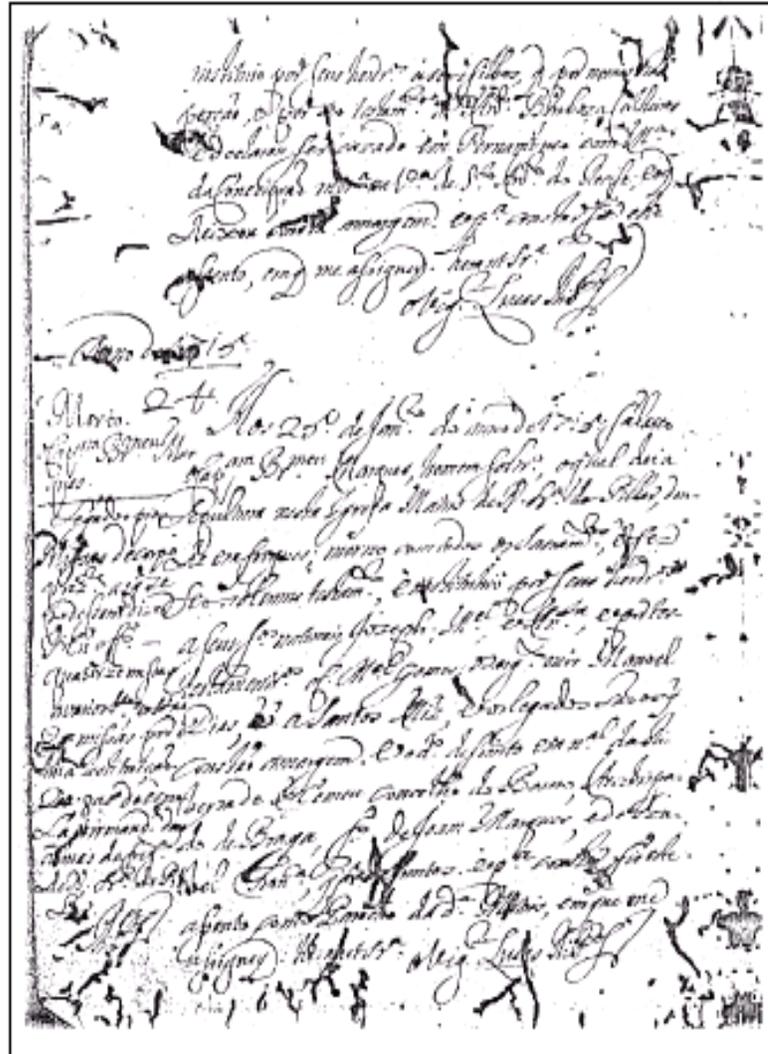


Figura 3.9 – Reprodução de um assento de óbito de 1715, presente na coleção da Matriz do Pilar, apresentando as dificuldades da leitura paleográfica.

Como a imprensa só chegou no Brasil em 1808, a maior parte da documentação existentes nas Igrejas do Brasil Colônia está manuscrita.

O conhecimento de paleografia utilizada neste período torna-se importante para efeitos da leitura e identificação do conteúdo do qual trata o documento, que será necessário para uma transcrição para um banco de dados, criação de índices, catálogos, dentre outros. Basicamente o paleógrafo precisa de duas habilidades: primeiro saber transcrever cada letra para sua correspondente a qual está mais familiarizado. Segundo, saber identificar as abreviações, algo muito comum naquela época. O segredo pode estar também em se conhecer o estilo que estava sendo utilizado na época

ou em determinado tipo de documento. As dificuldades e técnicas que podem ser utilizadas para a leitura paleográfica serão explicitadas no capítulo 5, que trata da identificação dos metadados nos documentos manuscritos.

### 3.3 Usos dos assentos

O acervo presente na Matriz do Pilar de Ouro Preto, nas palavras de Donald Ramos, da Universidade de Cleveland, representa um recurso histórico da maior importância, pois abre uma janela para a história de Vila Rica e, sem dúvida, toda a zona da mineração daquela região, permitindo que imaginemos um mapa social cheio de teias, ligando a comunidade de livres e escravos.

Os estudos históricos já realizados a partir deste acervo são amplos e originaram artigos, dissertações e teses sobre os mais variados assuntos, tanto de natureza individual dos habitantes deste período quanto social, de cunho religioso, econômico, político, cultural, artístico e demográfico. Isto é possível graças aos infindáveis cruzamentos que podem ser realizados, a partir dos dados registrados sobre batismo, casamentos e óbitos, como já citadas. Ressalta a historiadora Adalgisa Arantes Campos que análises mais profundas só foram possíveis graças ao advento do computador pessoal.

Não restritos apenas a estudos históricos, acervos com estes da Igreja Católica também são importantes para outras igrejas, como a dos Mormons já citados no primeiro capítulo. A Igreja dos Santos dos Últimos Dias possui em um dos seus princípios básicos a idéia de que todas as pessoas, vivas ou já falecidas, podem ser batizadas e assim receberem suas ordenanças<sup>9</sup>. Segundo eles, o Pai Celestial preparou um meio para que os mortos desfrutassem das bênçãos do evangelho.

Enquanto o corpo de Jesus jazia no túmulo, depois de Sua crucificação, Seu espírito foi ao mundo espiritual. Ali, Ele pregou o evangelho ao espírito dos justos.

(I Pedro 3: 18-20; 4-6)

---

<sup>9</sup> Colocação de ordens eclesiásticas.

Segundo os Mormons, esta obra missionária continua ainda hoje. Pessoas que estão no mundo espiritual podem exercer fé e aceitar a mensagem do evangelho, mas não podem receber as ordenanças do evangelho por si mesmas. O senhor nos orientou para que realizássemos essas ordenanças por elas (MORMONS, 1998).

Por este motivo conhecer as pessoas que já falecerem se torna de fundamental importância para a uma família Mormon. Esta Igreja criou um sistema sem igual, com alcance mundial onde qualquer pessoa (membro da igreja ou não) pode consultar sua árvore genealógica e recuperar nomes de pessoas que serão submetidas ao batismo em uma de suas capelas. Através de pesquisa de campo, pessoas ligadas a esta Igreja percorrem todos os anos centenas de Igrejas católicas ao redor do mundo, à procura de documentos sobre casamentos e batismos que podem revelar os ancestrais de suas famílias. Trata-se de um sistema, de proporções continentais, totalmente interligado, com o intuito de disseminar não apenas a informação contida nestes assentos antigos, como em alguns casos, com o próprio documento já sendo visualizado na sua forma digital.

#### **4. O TRATAMENTO DIGITAL DA INFORMAÇÃO**

Ao tratarmos os documentos presentes em acervos permanentes, nossa preocupação passa por aspectos como a organização do acervo, preservação e o uso efetivo da informação. Segundo Buckland, a vinculação do termo informação a um objeto que contém uma informação reporta-se a um dos conceitos de documento. O autor trabalha o conceito de informação como coisa, designando o documento que porta alguma informação. Os documentos de um acervo podem ser vistos como coisas informativas e desencadeadoras de processos de produção de conhecimento (BUCKLAND, 1991).

O tratamento da informação é um campo de conhecimento específico dentro da ciência da informação, que se destina a organizar a informação, de maneira que possa ser recuperada e disseminada no futuro, quer seja no ambiente convencional ou digital, sabendo-se estar este sujeito às várias etapas da organização da informação. Segundo Langridge (1977) a organização do conhecimento é a expressão mais abrangente para designar a função das bibliotecas, significando não apenas sua habilidade em identificar itens, como também de mostrar os assuntos disponíveis em suas coleções.

Lancaster (1993) coloca a organização e o controle no ponto central dos serviços a serem realizados por um sistema de informação, que dentre suas funções ressalta-se a interface entre uma população em particular e um universo de informação. De uma maneira geral a organização e o tratamento da informação, também conhecidos como processos de representação do conhecimento, dividem-se em duas etapas; a descrição física, onde se pretende identificar os elementos físicos e/ou extrínsecos de um documento como autor, local, data, etc.; e a representação temática, onde se busca identificar os conceitos que traduzem a essência do conteúdo. Desta atividade irão resultar representações que se apresentam como sínteses e facilitam a manipulação e acesso por sistemas e usuários.

O grau de tratamento ao qual os documentos serão submetidos irá depender, basicamente, da finalidade do sistema. A criação de sistemas digitais de recuperação da informação não pode

negligenciar todo o conhecimento técnico adquirido ao longo dos anos no contexto tradicional da informação (não automático). Descartar estes recursos revela ignorância a respeito das funções dos sistemas tradicionais (DIAS, 2001).

Segundo Dias, sistemas de informações que fossem voltados para a satisfação do usuário deveriam definir a sua clientela e, em função desta definição, os objetivos do sistema seriam estabelecidos. Este sistema estaria completo se contasse com uma ferramenta de recuperação da informação que, da mesma forma, poderia apenas descrever os recursos existentes no sistema, da forma mais conveniente para o usuário.

Sem dúvida alguma, a Internet é um sistema de comunicação eficiente e que facilita enormemente o acesso a sistemas de informação e de recuperação da informação. Porém, ao tornar-se hegemônico, o contexto digital necessitará de tratamento com o apóio de instrumentos já bem assimilados pelos processos de representação tradicionais. Definir os recursos a serem disponibilizados, as ferramentas para tratamento da informação e os objetivos do sistema a ser implementado são tarefas essenciais.

Torna-se importante destacar que com as novas tecnologias novos trabalhos e possibilidades do tratamento da informação surgiram requerendo a atuação direta de diferentes profissionais.

Vickery (1986, apud ALVARENGA 2001, p.2) já preconizava que, com o advento do computador, a tarefa de tratamento da informação extrapolaria o campo da biblioteconomia, e passaria a envolver profissionais de outras áreas ligados a atividades como a estruturação de registros em base de dados, a elaboração de programas de computador, a idealização de estruturas semânticas da linguagem, a representação do conhecimento com inteligência artificial, e um maior conhecimento sobre o modelo de pensamento humano.

Méndez (2002) destaca que não são poucos os desafios a enfrentar por quem se propõe a realizar pesquisas que tratem, por um lado, dos aspectos relacionados aos problemas de recuperação da informação, como padrões e métodos para estruturar registros do conhecimento e, por outro, de

questões conceituais que a eles dizem respeito, no sentido, tanto de evitar imprecisões terminológicas ou descontextualizações, quanto de definir categorias em suas relações recíprocas.

#### **4.1 Sistemas de Informações (SI) e Sistemas de Recuperação da Informação (SRI)**

As primeiras preocupações com a recuperação da informação já foram identificadas no século II a.C. quando Galeno publicou o *De libris propriis liber*, um catálogo de suas próprias obras, visando distinguir as de sua autoria daquelas escritas por outros autores e a ele indevidamente atribuídas (ARAÚJO, 1994).

Segundo ALLEN (1996, p. 5 apud MOREIRA 2002, p.50) um sistema de informação é um sistema conectado e relacionado de entidades (incluindo um ou mais dispositivos de informação) que fornecem acesso para um ou mais corpos de conhecimento e age como um mecanismo através do qual indivíduos podem informar outras pessoas ou se tornarem informados. Moreira destaca que os Sistemas de Informação têm por objetivo repassar informações aos usuários tornando-se assim Sistemas de Comunicação. Quando este meio de comunicação é um dispositivo computacional, pode-se dizer que trata-se de um sistema de informação computacional ou baseado em meios eletrônicos (MOREIRA, 2002). Estas definições se aproximam da definição de um sistema para disponibilizar a informação, uma vez que o objetivo gira em torno de recuperar a informação para o uso.

Nesta visão, Lancaster (1994) coloca a recuperação da informação como a atividade mais importante dos centros de informação. Recuperar a informação é um processo que faz parte de um Sistema de Informação. Estes seriam centrais onde a informação estaria disponível aos usuários.

Os SRI são mecanismos projetados especificamente para proporcionar a recuperação da informação. Araújo (1994) apresenta várias características presentes nos SRIs: 1) As mensagens podem ser transientes ou registradas; 2) a audiência pode ser individual ou em grupo; 3) podem ser centralizadas ou dispersas; 4) os tipos de usuários devem ser caracterizados pela natureza dos

serviços prestados; 5) a informação fornecida pode ser apenas um referência ou o texto completo; 6) a entrega pode ser oral, impressa ou eletrônica.

Dentre as possíveis subdivisões de um sistema de recuperação da informação pode-se destacar:

*Subsistema de entrada:* preparam os documentos para serem introduzidos no SRI. É a parte do sistema responsável pela seleção intelectual dos documentos que irão compor o sistema, e provê condições para execução dos mecanismos do subsistema de saída. Este processo pode ser realizado através de ferramentas como catálogos e resenhas e será composto das etapas de: *Descrição:* é a representação dos registros que se tornaram os substitutos dos documentos. Trata-se de uma parte importante, pois será sobre esta descrição que o processo efetivo de recuperação irá atuar; *Representação:* onde os aspectos intrínsecos do documento são identificados, tais como sua tematicidade e atinência. Este trabalho também é conhecido como indexação; *Organização:* trata da representação dos documentos em seus aspectos físicos, estruturais e lógicos.

*Subsistema de saída:* Em geral boa parte dos trabalhos de negociação de questões sobre buscas e recuperação são tratadas neste módulo do SRI (Araújo, 2004). Trata-se da interface do usuário com o sistema, compreendendo o problema de informação do usuário, a comunicação entre usuário e sistema e as estratégias de busca. A análise das questões para criação de um SRI se dividem em duas partes: Na primeira buscamos uma compreensão dos problemas do usuário. Na segunda temos uma clarificação, a especificação de vários aspectos da pergunta, e envolve também uma intensa troca de informações entre o usuário e o especialista na busca. Neste processo o especialista também pode tornar claro para o usuário quais são as potencialidades e limitações do sistema. Segundo Araújo, existe pouca literatura sobre este módulo, uma vez que é grande a complexidade e o número de variáveis envolvidas. Saracevic (1983) identifica cinco elementos básicos em uma situação de busca por informação em um SRI: 1) o problema 2) o estado interno de conhecimento 3) a intenção 4) a questão 5) o pedido. Segundo Rees e Saracevic (1963) o problema de informação de um usuário não implica necessariamente em uma necessidade de informação e, uma vez que o problema seja transformado em pergunta, esta não necessariamente precisa coincidir com a necessidade real. O curioso é que quanto mais perguntas maiores as dificuldades em se expressa-las. *Estratégias de busca e recuperação:* O maior

objetivo deste subsistema é tentar coincidir os termos utilizados na busca pelo usuário com os termos propostos pelo sistema. Em geral a estratégia será montada utilizando operadores booleanos E, OU, NÃO de George Boole<sup>10</sup>.

Na tentativa de avaliar o desempenho de sistemas de recuperação de informações Araújo (1994), destaca que um dos grandes problemas é que o grau de relevância não é uma propriedade dicotômica, mas ocorre em graus. Além disso, os SRI precisam trabalhar com uma certa imprecisão, principalmente por parte do usuário, ambigüidades, indeterminismos e estes aspectos não são cobertos pela lógica de Boole.

Uma das etapas da organização da informação para posterior recuperação, é a identificação dos metadados que compõem os pontos de acesso nos SRI. Estes por sua vez podem ser intrínsecos ou extrínsecos aos documentos e serão considerados no item a seguir.

## **4.2 Metadados e os SRI**

Por ser fundamentalmente um sistema para recuperação da informação, as bibliotecas digitais podem possuir grandes quantidades de dados que precisam ser manipulados de uma maneira eficiente, para que seu uso seja considerado satisfatório pelo usuário. Parte deste objetivo pode ser alcançado observando-se a qualidade dos metadados presentes na biblioteca digital, os quais irão servir como pontos de acesso à coleção. Uma descrição eficiente tende a facilitar o acesso.

Segundo Dempsey e Herry os metadados são dados associados aos objetos e permitem aos usuários conhecer as suas características, suportando uma grande gama de operações, tanto pelo usuário humano quanto pela máquina. (DEMPSEY, 1997)

Uma definição mais geral pode ser obtida em Baca:

Dados sobre Dados - metadados representam dados sobre objetos com o objetivo de descrever, administrar, acrescentar

---

<sup>10</sup> George Boole foi o matemático que primeiro estabeleceu a lógica baseada nos operadores E, OU, NÃO, utilizada em todos os equipamentos computacionais nos dias atuais.

funcionalidades técnicas, estabelecer requerimentos legais e de preservação. BACA (1998, p.37)

Segundo Borgman (2000) os metadados atendem a uma grande gama de usuários e servem para vários propósitos. Obviamente que podem haver divergências entre os padrões de metadados mais gerais e os mais específicos. Para o público leigo pode não ser óbvio em um primeiro momento, mas os metadados podem causar um impacto muito grande no desenvolvimento efetivo de um sistema de informação, principalmente no momento da recuperação da informação.

Na atualidade várias são as instituições que fazem uso de padrões de metadados para descrever objetos. Exemplos de profissionais envolvidos em planejar os metadados são museólogos, bibliotecários, arquivistas, administradores de bancos de dados, e vários outros. Na verdade, qualquer pessoa que esteja envolvida em processo que pretenda descrever ou identificar pontos de acesso a grupos de objetos (coleções) estará em algum momento identificando os metadados necessários. As funções de um grupo de metadados podem ser divididas em: administrativa; descritiva; de preservação; técnica e uso (BACA, 1998). Todos eles são importantes no desenvolvimento de sistemas de informações digitais, sendo de extrema importância para a recuperação das informações em bibliotecas digitais. Apesar de parecer tarefa simples, a criação e o gerenciamento de metadados pode se tornar uma tarefa complexa, envolvendo processos manuais e automáticos, evoluindo de acordo com a vida do objeto. As fases da vida de qualquer objeto compreendem etapas de criação; organização; recuperação; utilização e preservação.

Na literatura podem ser encontrados diferentes padrões de metadados tais como o Dublin Core (DC) que por ser muito flexível é utilizado em diversos tipos de coleções. Este padrão foi adotado pelo W3C, instituição responsável pelo gerenciamento da Internet e que conta com profissionais bibliotecários, analistas, museólogos entre outros. Sua principal finalidade é descrever os vários recursos disponíveis na Internet facilitando a procura por informações neste ambiente. O DC se constitui como um excerto do MARC o qual permite o compartilhamento de recursos bibliográficos reduzindo a duplicidade de trabalho.

A utilização de padrões em um ambiente livre como a Internet tem se tornado uma atividade fundamental para o sucesso de qualquer trabalho ou pesquisa que tenha como objetivo utilizar

este suporte. Padrões facilitam a integração e o compartilhamento de informações entre diferentes sistemas, os quais podem possuir diferentes propósitos. O uso de padrões estabelecidos por instituições nacionais ou internacionais implica em um compromisso em aceitar terminologias e regras estabelecidas. Dempsey e Herry agrupam os metadados em três grupos possíveis de acordo com cinco características:

*Índices de texto completo:* os quais podem ser encontrados em máquinas de busca na web, onde todo o texto pode representar pontos de acesso.

*Formatos genéricos:* que são estruturados segundo uma norma padrão. Neste grupo encontra-se o Dublin Core.

*Formatos mais complexos:* utilizados em domínios específicos como o MARC.

Estes três grupos possuem características que irão diferenciá-los. Eles são agrupados segundo o ambiente de uso, a função, a forma de criação, o uso final, a forma de recuperação e o status (DEMPSEY, 1997).

Segundo BACA (1998) os metadados também possuem características listadas na tabela a seguir:

**Tabela 4.1 – Características dos metadados**

<b>Atributo</b>	<b>Características</b>	<b>Exemplos</b>
Origem dos metadados	Gerado de acordo com o momento da criação do objeto ou posteriormente pelo processo de digitalização, etc.	Nomes dos arquivos Estruturas de diretórios Tipos de compressão
Método de criação	Gerados automaticamente ou manualmente	Índices automáticos Observação de logs
Natureza dos metadados	Define se os metadados foram criados por um especialista ou por um leigo	Metatags em um documento web
Status	Define se são estáticos ou dinâmicos	Resolução das imagens Direitos autorais
Estrutura	Qual o padrão utilizado para criação dos metadados	MARC Objeto relacional
Semântica	Vocabulário utilizado	AACR2 HTML Texto livre
Nível	Se irá descrever itens isolados ou toda a coleção	MARC Índice próprio

Os metadados são importantes para aumentar a acessibilidade, uma vez que o documento digitalizado pode continuar com problemas para acesso. Segundo Baca os metadados permitem ainda expandir o uso; separar os verbos; tratar os aspectos legais e de preservação (BACA, 1998).

Hayes e Becker (1970, apud Borgman, 2000) também atribuem como qualidade dos metadados o fato de permitirem a coleta de informações sobre aquisições, descartes e armazenamento.

No meio digital, os metadados também podem facilitar a criação de funcionalidades difíceis de se conseguir no mundo real, como o estabelecimento de links entre documentos diferentes, por exemplo, proporcionando a exibição de documentos com versões diferentes. Os metadados também podem facilitar o controle através de estatísticas não apenas sobre qual material foi acessado, mas também sobre como foi acessado, além de referências sobre quem acessou, informações estas que podem ser tratadas tanto no individual quando na coleção por inteiro. Os metadados também podem favorecer um melhor agrupamento de documentos com características semelhantes, realizando o que Borgman descreveu como sendo uma organização intelectual, baseada em comodidades e relações entre os metadados, por exemplo: agrupamento por datas, por versão, por autor, por formato, por interesse e diversos outros.

### **4.3 O padrão MARC**

O padrão MARC (Machine Readable Cataloguing - USMARC, 1996) foi criado nos Estados Unidos com objetivo de facilitar o intercâmbio de informações entre bibliotecas. A estrutura dos registros USMARC segue padrões nacionais e internacionais, dentre eles os padrões ANSI Z39.2 (Formato de Intercâmbio de Informações) e ISO 2709 (Formato para Troca de Informações).

A maior parte dos elementos encontrados no registro USMARC na verdade são definidos fora dele, como as Regras de Catalogação Anglo Americana (AACR2).

Segundo o MARC e conseqüentemente o Dublin Core, baseados na AACR2, um importante detalhe no tratamento do acervo objeto desta dissertação, é o fato desses formatos trabalharem

especificamente com os manuscritos que, segundo as normas, constituem-se de documentos feitos a mão ou datilografados, tais como livros, dissertações, teses e catálogos.

O MARC trata como sendo a principal fonte de informação em um manuscrito o próprio manuscrito, dando preferência aos textos que façam parte do próprio para em seguida considerar outras informações pré-textuais como capa, legenda, cabeçalho, e também informações contextuais. Recomenda-se ainda que toda a coleção seja tratada como fonte principal. Para documentos manuscritos vários são os campos que podem ser representados. Dentre eles destacam-se: título; títulos equivalentes; indicação de responsabilidade; data; descrição física; acessibilidade e direitos literários; estilo de escrita; nota ou observações.

Caso o documento manuscrito não possua um título, deve-se utilizar um título criado que indique a sua natureza, utilizando uma palavra apropriada que o descreva seguida do lugar ou ocasião em que ocorreu. As indicações de responsabilidades deverão ser preenchidas caso sejam consideradas como relevantes. Deve-se dar indicações de responsabilidades que não estejam claras na área do título, relatando pessoas ou instituições que estejam relacionadas com o documento. A data poderá fazer parte do título ou ser colocada separadamente. Deverá compreender o ano, o mês e o dia em que ocorreu. O campo de notas poderá ser utilizado para descrever o tipo de documento, com registros de dados suficientes para identificar a pessoa, podendo compreender um breve resumo do conteúdo do manuscrito.

De uma forma genérica, modelos antigos como MARC, oferecem uma visão muito complexa e rígida dos objetos descritos, não possibilitando a criação “ad-hoc” de metadados de forma a atender a propósitos de comunidades específicas ou a necessidades ocasionadas pela própria evolução de conceitos. Qualquer alteração no padrão MARC (para inclusão de um novo campo, por exemplo) exige um consenso de sua comunidade internacional (BARRETO, 1999).

#### **4.4 O padrão Dublin Core**

Ressalta-se que o Dublin Core é um padrão originado do MARC, mais flexível, e incluindo dois níveis: simples e qualificado. O primeiro é composto por quinze elementos padrões que cobrem

aspectos como autoria, formatos, datas e diversos outros. O segundo conta com três elementos adicionais que irão qualificar o recurso sendo eles a audiência, a proveniência e os direitos autorais. Cada um dos elementos pode ser opcional e pode ser repetido. Muitos deles também sofreram uma espécie de qualificação e refinamentos através da inclusão de subníveis. A seguir é relacionado o conjunto de elementos definidos pelo DC (1999).

**Tabela 4.2 – Conjunto padrão de metadados proposto pelo Dublin Core**

Título	Um título do recurso
Criador	Entidade principal responsável pela elaboração do conteúdo do recurso
Assunto	Assunto referente ao conteúdo
Descrição	Uma descrição do recurso
Editor	A instituição responsável pela difusão do recurso
Contribuinte	Uma entidade responsável pela contribuição ao conteúdo do recurso
Data	Data associada com um evento no ciclo de vida do recurso
Tipo	A natureza ou gênero do conteúdo
Formato	Manifestação física ou digital do recurso
Identificação	Identificação não ambígua do recurso dentro de um contexto
Fonte	Uma referência para um outro recurso do qual o presente é derivado
Idioma	Idioma do conteúdo intelectual do recurso
Relação	Uma referência a um outro recurso que se relaciona com este
Cobertura	A extensão ou cobertura espaço-temporal do conteúdo
Direitos	Informações sobre direitos autorais do recurso

Com estes elementos o W3C estabeleceu quatro objetivos básicos para o DC, são eles:

- Simplicidade de criação e manutenção: Permitir a especialistas e não especialistas criar de forma simples maneiras de se recuperar documentos no mundo web.
- Semântica fácil de ser entendida: os metadados escolhidos deverão facilitar a compreensão sobre “o que” aqueles metadados estão representando e que sejam, na medida do possível, metadados universalmente aceitos.
- Escopo internacional: Apesar de ser originalmente desenvolvido em Inglês, existe considerações sobre sua utilização como um padrão multilinguas.
- Extensibilidade: O DC prevê extensões para atender a diversos tipos de objetos e recursos, expandindo seus elementos básicos de forma a atender objetivos específicos.

Segundo Borgman, o padrão Dublin Core de metadados é um dos mais utilizados por servir a diferentes tipos de usuários, sendo o resultado de um esforço em se criar um padrão que servisse tanto a documentos quanto a objetos reais. Esta pesquisa procurou se adequar ao padrão DC principalmente pelo fato dele ser adotado por inúmeras instituições de padrão internacional e em agências governamentais.

O grupo W3C estabeleceu alguns princípios que devem ser observados na utilização do DC para descrever objetos, são eles:

- *One-To-One*<sup>11</sup>: Em geral a proposta de metadados deve descrever cada elemento de forma individual. Por exemplo, uma imagem JPG (um formato de arquivo para armazenamento de imagens) da Mona Lisa deve ser representada individualmente e não deve tentar descrever também a imagem original no Museu do Louvre.
- *Dumb-Down*: Um usuário poderá ignorar qualquer qualificador e utilizar um outro não qualificado. Isto poderá gerar uma perda de especificidade, mas continuará sendo uma forma correta de se definir o objeto em si.
- *Appropriate values*: Qualquer elemento do DC poderá variar de acordo com o projeto ou o contexto utilizado, mas não podemos prever como os sistemas irão interpretar os metadados encontrados uma vez que a estratégia de busca pela informação partirá da cabeça de cada usuário.

Atualmente o padrão Dublin Core não está restrito apenas aos elementos intrínsecos do documento. Dados extrínsecos como direitos de uso, propriedade intelectual, custo, entre outros, também podem ser descritos pelo formato proposto pelo DC.

Do ponto de vista administrativo, o DC pode ser expandido e passar a fornecer funcionalidades para a gerência do acervo digital, com informações sobre aquisição, direitos de reprodução,

---

<sup>11</sup> Os termos em inglês foram mantidos por serem termos técnicos mais conhecidos do que suas traduções.

requerimentos de segurança e acesso, controle de versões, entre outros. Apesar de ser simples o padrão fornece uma solução adequada para um grande problema: A necessidade de se criar uma descrição simples e genérica que possa ser utilizada para descrever uma grande gama de objetos.

O Open Archives Initiative (OAI) desenvolveu o OAI- Protocol for metadata harvesting (OAI-PMH) que tem como função desenvolver padrões uniformes para o tratamento de metadados aumentando a performance do modelo de comunicação atualmente utilizado na web. A iniciativa pretende criar um conjunto de metadados para transporte de dados entre repositórios, resolvendo assim problemas de interoperabilidade. A base da iniciativa é também o padrão Dublin Core e irá dividir as instituições em dois grupos que são Provedores de dados e Provedores de serviços.

Uma vez identificados os metadados que irão representar uma classe de objetos do mundo real ou digital, é necessário estabelecer a linguagem que será utilizada para representá-los quando implementados em um sistema. Dentre as várias opções disponíveis destacam-se os padrões HTML e XML por se tratarem de padrões internacionalmente aceitos e serem utilizados por diversas instituições que buscam interoperação entre sistemas.

#### **4.5 Bibliotecas digitais**

Em 1945 Vannevar Bush estabeleceu a primeira idéia do que seria um sistema de informação, para disseminar e gerenciar o vasto conhecimento científico sendo produzido durante aquela década, no período pós-guerra. Bush descreveu um equipamento tecnológico em que o indivíduo iria armazenar toda a sua coleção de livros e registros, os quais seriam recuperados pelo sistema de uma forma rápida e eficiente.

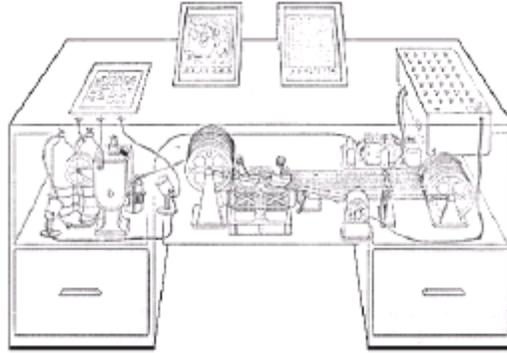


Figura 4.1 - Memex descrito por Vannevar Bush

Cinquenta anos depois o mundo se depara com a tecnologia necessária para que seu sonho se realize, através de modernas infra-estruturas, bibliotecas digitais, navegação por hipertexto e várias outras técnicas e aparatos tecnológicos.

Segundo Alvarenga,

“Com o advento e intensificação do uso da Internet, a chamada biblioteca digital tornou-se realidade. Em constante evolução, esse novo e potente meio de comunicação vem possibilitar, finalmente, a concretização do sonho da biblioteca sem paredes, da grande biblioteca universal, preconizado por pensadores, dentre os quais podem ser citados os enciclopedistas franceses e Jorge Luís Borges, que um dia sonharam com a possibilidade da existência de uma biblioteca que pudesse abarcar a totalidade dos conhecimentos produzidos, em todos os recantos do mundo. A concretização, enfim, de uma situação dantes utópica e recorrente em inúmeros momentos da evolução do pensamento científico e filosófico da humanidade”. (ALVARENGA, 2001 p.3)

Marchiori (1997) destaca que a concepção de bibliotecas digitais apresenta-se como uma possível quebra no paradigma de tratamento e disseminação de informações representado pelos recursos, atividades e serviços da "biblioteca tradicional". A transição da mídia impressa para a mídia digital afirma-se como passo decisivo para a criação de uma nova estrutura de bibliotecas onde a informação poderá ser acessada remotamente e compartilhada de forma quase instantânea, a custos relativamente baixos.

Lancaster percebe as bibliotecas digitais como uma oportunidade para os profissionais da informação combinarem suas atividades altamente especializadas com as exigidas pela realidade tecnológica (LANCASTER, 1994).

Segundo Levacov a principal diferença é que “onde” o documento reside torna-se menos importante. O conceito de lugar passa a não ser mais significativo nem para o usuário nem para o bibliotecário. Não se pode deixar de imaginar que no futuro o conteúdo das bibliotecas serão menos tangíveis, o que contraria enormemente a tendência das bibliotecas convencionais. O acesso remoto passa a ser a principal característica e seu produto é frequentemente intangível (LEVACOV, 2001). O conceito de “tempo” também se altera. Levy chama isso de “velocidade qualitativa e espaço-tempo mutantes” (LEVY, 1996).

Entretanto, segundo Chen (1998), as bibliotecas convencionais permaneceram como principal local para acesso ao conhecimento, e as bibliotecas digitais irão expandir as capacidades das bibliotecas tradicionais, uma vez que também irão permitir a recuperação de textos completos.

Então a questão que se levanta é: porque precisamos de bibliotecas digitais uma vez que as bibliotecas tradicionais são sistemas importantes e com funcionamento adequado, capaz de atender muitas das expectativas dos usuários? Sem dúvida esta não é uma resposta simples, porém, um fator importante leva cada vez mais ao desenvolvimento de pesquisas com bibliotecas digitais – a quantidade e a velocidade com que a informação é produzida e deve ser acessada no mundo moderno.

A principal função de qualquer biblioteca é disseminar conhecimento para seus usuários, através de objetos de diferentes formatos tais como livros, filmes, vídeos, mídia para computador como disquetes, CD, DVD e vários outros tipos como mapas, revistas, entre outros. O escopo original das bibliotecas, que é o de preservar a memória social, também é uma importante função nos dias atuais.

As principais características de uma biblioteca digital também foram sumarizadas pela Association of Research Libraries (DRABENSTOTT, 1994):

- Não é um entidade única, ligando recursos de várias bibliotecas;
- Estas ligações devem ser transparentes para os usuários;

- Seu principal objetivo é o acesso universal;
- Não são limitadas a documentos mas também a artefatos digitais os quais possuem as seguintes características: Identificação única; lista de transações realizadas (log); propriedades como direitos de acesso e métodos para acesso; metadados; tipo de conteúdo (música, vídeo, imagem, etc); assinatura digital (opcional).

Como as bibliotecas digitais geralmente são sistemas complexos e que muitas vezes se interoperam, é importante o uso de padrões e modelos que tenham como objetivo formalizar o ambiente criado, o que poderá facilitar o entendimento e a descrição de como o sistema irá funcionar em sua plenitude.

Para Cunha (1994) em uma biblioteca digital encontramos uma ou várias das características abaixo:

- Acesso remoto do usuário via computador conectado a rede;
- Utilização simultânea de um documento por mais de um usuário;
- Acesso não apenas a referência, mas também ao texto completo;
- Utilização de forma que a biblioteca local não necessite ser a dona da coleção;
- Utilização de vários tipos de suporte (vídeo, som, texto);
- Sistemas auxiliares na recuperação da informação.

Durante a elaboração de bibliotecas digitais o primeiro passo a ser dado é a construção ou automação dos catálogos sendo, em seguida, disponibilizado o texto completo da obra os quais podem ser obtidos por meio de metadados.

Para sistemas de recuperação da informação pode-se ter três modelos clássicos que são: vetorial, booleano e probabilístico (ATKINS, 1999). No campo das bibliotecas digitais, isso ainda não está totalmente maduro uma vez que trata-se de um campo em plena efervescência. Mesmo assim, observa-se o surgimento de alguns modelos, dentre eles o de Wang que propôs um sistema híbrido de aproximação para bibliotecas digitais contendo um banco de dados e uma interface para o usuário, suportando conteúdo multimídia.

Kalinichenki (2000) propôs o modelo Canônico onde as bibliotecas digitais são consideradas uma coleção de documentos web.

Os documentos presentes em uma biblioteca digital podem ser basicamente em: texto, imagens, sons e vídeo. Dependendo do tipo de sistema implementado um dos formatos pode ser mais conveniente que outro. Em muitos casos, os documentos de texto são convertidos em imagem pelo processo de digitalização, muitas vezes a única forma de disponibilizar o conteúdo de documentos impressos diretamente na tela de um computador.

Existem dezenas de softwares que podem ser utilizados no processo de digitalização dos documentos que irão pertencer à biblioteca digital. Independentemente daquele utilizado, alguns aspectos importantes devem ser observados durante o processo: deve-se fazer uma varredura nos documentos para verificar aspectos como sujeira, grampos, colas os quais devem ser tratados antes da realização da digitalização; deve-se observar se o documento original não está inclinado o que pode trazer distorções no momento da digitalização; estabelecer a melhor qualidade possível ajustando brilho e contraste, nível de nitidez e resolução de saída; corrigir manchas, eliminar pontos, definir limites do documento digitalizado.

Uma das etapas do processo de criação e disponibilização de uma biblioteca digital, seja em CDROM ou na web é a aquisição e configuração do software que irá ser utilizado para esta tarefa. Existem atualmente vários softwares para gerenciamento de bibliotecas digitais, vários deles atendem ao padrão OPEN SOURCE, ou seja, possuem uso irrestrito e não exigem a aquisição ou o pagamento de royalties pelo uso. Bons exemplos são: Phronesis criado pela universidade de Monterrey no México; E-prints da universidade de Southampton; o D-Space criado pelo MIT e a HP Labs; e o Greenstone criado pela universidade Waikato de N. Zeland.

Por ser uma iniciativa apoiada pela UNESCO, o software escolhido para criar o protótipo da biblioteca digital proposta foi o Greenstone. As características listadas abaixo justificam o seu uso pois atendem a quase totalidade dos requisitos deste projeto:

- Multiplataforma (UNIX/ Windows);
- Interfaces em diversas línguas;

- Busca em texto completo, navegação por hierarquia e busca por metadados;
- Suporta figuras, sons, textos;
- Suporta grandes coleções;
- Possuem funções administrativas como login de usuário e estatísticas de uso;
- Pode ser disponibilizado na web ou via CD-ROM

O panorama atual de implementação de bibliotecas digitais é amplo e resultou em diversos sistemas que podem ser acessados via Internet, tais como:

**Projeto Gutenberg:** que coloca a disposição do público uma enorme gama de textos eletrônicos de domínio público (<http://www.promo.net/pg>)

**National Digital Library:** abrange a produção de documentos não textuais, dentre eles a biblioteca do vaticano com 100 mil manuscritos e 150 mil gravuras.  
(<http://rs6.loc.gov/ammem/ndlpedu>)

**NASA Technical Report Server:** que disponibiliza o acesso a milhões de documentos técnicos em diversas bibliotecas digitais. (<http://techreports.larc.nasa.gov/cgi-bin/ntrs>)

**Biblioteca Digital da PUCRS:** Uma biblioteca digital que possibilita o acesso a documentos em diversos formatos. (<http://www.cglobal.pucrs.br/bibdigital/bib>)

**SciELO Scientific Electronic Library Online:** coleção virtual de artigos de revistas científicas brasileiras disponíveis na Internet. (<http://www.scielo.br>)

## 5. IDENTIFICAÇÃO DOS METADADOS

### 5.1 Aspectos metodológicos

Essa pesquisa se insere em um campo interdisciplinar, apoiando-se no contexto da informação histórica em que, a partir dos aspectos descritos pela arquivologia e pela ciência da computação, pretende-se, sob a luz da ciência da informação, identificar fatores pertinentes ao acesso a documentos microfilmados ou em meio digital, realizando a identificação dos metadados necessários a sua representação.

Com isso torna-se imprescindível uma investigação e discussão com os campos científicos, tanto da história, da arquivologia, da computação e da ciência da informação. Ressalta-se o método qualitativo para o processo de pesquisa intensamente contextualizado, a metodologia exploratória com viés experimental, com a criação de um protótipo de biblioteca digital.

Comparada a pesquisa disciplinar, a qual trata a realidade apenas de um único nível, a interdisciplinaridade prevê a transferência de métodos de uma disciplina para outra, o que proporciona uma maneira nova de conhecer. A metodologia tratada neste capítulo, diz respeito ao processo de identificação dos metadados, ficando o processo de criação do protótipo da biblioteca digital, exposto no capítulo 6.

Como dito anteriormente, na primeira parte do trabalho foram realizados contatos com vistas a obter a participação de vários responsáveis pelo acervo primário. Entre elas a Casa Setecentista do Pilar, com sede no município de Ouro Preto, onde contamos com a colaboração do atual diretor da instituição, Carlos Aparecido Oliveira<sup>12</sup>, que mantém a guarda da documentação da Matriz do Pilar. Pretendia-se com este contato dispor de acesso às fontes primárias, uma vez que seria importante e enriquecedor conhecer não apenas as fontes secundárias (microfilmadas), como também a documentação original.

---

<sup>12</sup> Nosso agradecimento ao sr. Carlos, por ter gentilmente facilitado nosso contato com os documentos originais, durante as visitas ao acervo.

O principal objetivo desta etapa foi estabelecer um contato físico com o local onde os documentos foram criados e estão armazenados. Tratou-se de um momento raro, uma vez que a visita foi assistida pela pesquisadora Adalgisa Arantes<sup>13</sup> do departamento de história da UFMG. Concluídas as primeiras visitas, iniciou-se a etapa de definição dos objetos de análise. Percebeu-se que a documentação existente é ampla e cobre documentos tanto de batismos, casamentos e óbitos, como também documentos de irmandades, recibos, partituras musicais, contratos, testamentos e diversos outros.

Por já ser utilizado em diversas pesquisas, e também por uma questão do curto espaço de tempo disponível para fins desta dissertação, a documentação que teve seus campos de metadados identificados compõem os seguintes intervalos:

<b>Coleção</b>	<b>Primeiro Assento tratado</b>	<b>Último Assento tratado</b>
Assentos de casamento	1800	1850
Assentos de batismos	1800	1830
Assentos de óbito	1800	1830

Trata-se de um período rico e de transição da história brasileira. Este período, além de ser de grande interesse para os pesquisadores, também facilitou a pesquisa, uma vez que a leitura paleográfica desta época é mais simples se comparada àquela realizada em documentos do século XVIII. Outro fator que contribuiu é o fato de que os documentos desta época estão melhor conservados, o que resultou em microfilmes de melhor qualidade final, facilitando a leitura.

### **5.1.1 Metodologia para identificação dos metadados**

Estabelecidos o acervo e períodos que seriam tratados, uma pergunta que veio a mente é: como identificar metadados em documentos manuscritos os quais surgiram em uma época de pouca

---

<sup>13</sup> A profa. Adalgisa Arantes Campos é profunda conhecedora e estudiosa dos assuntos relacionados a sociedade Ouropretana, sobretudo no campo das artes.

padronização, uma vez que as determinações da igreja católica para os dados que precisavam constar nos assentos, nem sempre eram seguidas pelos párocos. Em meio a milhares de assentos realizados, quantos documentos precisariam ser lidos para que tivéssemos a menor margem de erros na identificação dos metadados? Esta não é uma resposta simples, e poderá ter pontos de vista diferentes em áreas diferentes do conhecimento.

Optou-se por realizar uma análise estatística para uma primeira aproximação de quais metadados poderiam ser identificados nos assentos, uma vez que ela traria uma metodologia que poderia evitar a consulta em todos os assentos o que poderia se tornar um trabalho inviável, principalmente porque vários assentos não estão em bom estado de conservação.

Foi realizado um contato com o Departamento de Estatística da UFMG, com o objetivo de se identificar uma metodologia que permitisse a coleta de uma amostra, com tamanho suficiente para identificar os metadados e que resultasse em uma margem de erro pequena.

Conforme orientação do Departamento de Estatística, este trabalho foi dividido em duas etapas: na primeira etapa foram identificados todos os metadados presentes em 50 assentos de forma aleatória, mas que cobrisse o intervalo de tempo de 50 anos com a qual iríamos trabalhar. O objetivo desta amostra piloto foi o de conhecer um pouco sobre o comportamento dos metadados para possibilitar uma amostragem melhor e mais elaborada em seguida.

Nesta amostra percebeu-se que a variação poderia ocorrer em função de diversos fatores, tais como: a época em que foi criado; o assentador que fez o registro; as pessoas envolvidas no assento; as leis da igreja vigentes no momento, dentre outros. Como esta amostragem visava o estudo de vários metadados, sabe-se que o período interfere no aparecimento ou não de cada um deles no assento, não sendo conhecido se os metadados tem a mesma tendência de aparecimento e permanência ao longo do tempo. Fez-se necessário garantir amostras periódicas.

Para um teste piloto a proposta da estatística foi denominada amostragem sistemática. A tabela 5.1 apresenta um resumo da coleta realizada. Quando dois ou mais assentos caíssem no mesmo dia, uma escolha aleatória deveria ser feita.

**Tabela 5.1 – Intervalo de documentos utilizados para criação da amostra sistemática**

<b>Tipo de assento</b>	<b>Tamanho da amostra</b>	<b>Primeira amostra</b>	<b>Intervalo</b>
Batismo	56	Julho de 1800	11 meses
Casamento	24	Abril de 1801	2 anos e 1 mês
Óbito	40	Janeiro de 1800	1 ano e 3 meses

De posse desta amostra piloto, o Departamento de Estatística estabeleceu a quantidade de documentos que deveria ter os metadados identificados, permitindo, segundo a estatística, uma margem de erro de 5%.

**Tabela 5.2 – Quantidade final de documentos utilizados**

<b>Tipo de assento</b>	<b>Quantidade de assentos a ser consultada</b>
Batismos	145 assentos (para 50 anos)
Casamentos	50 assentos (para 30 anos)
Óbitos	50 assentos (para 30 anos)

Com o início das leituras de alguns documentos, percebeu-se que alguns metadados poderiam surgir ou desaparecer com o passar dos anos. Apesar disso, considerou-se como sendo um metadado qualquer campo que aparecesse pelo menos uma vez nas séries observadas, pois os assentos foram tratados em suas particularidades, e não como dossiês do ponto de vista arquivístico.

Exemplifica-se pelo gráfico 5.1 a existência ou não do metadado *Igreja e Cidade* nos assentos de casamento, na primeira metade do século XIX. Vê-se claramente que o nome da igreja foi muito utilizado nos início do século até aproximadamente 1840. Em seguida este metadado não foi mais citado nos assentos, voltando a ser citado novamente próximo à metade do século. Já o campo Cidade era pouco utilizado nos assentos do início do século, foi muito utilizado entre os anos de 1825 e 1835 e novamente desapareceu até o final da primeira metade do século.

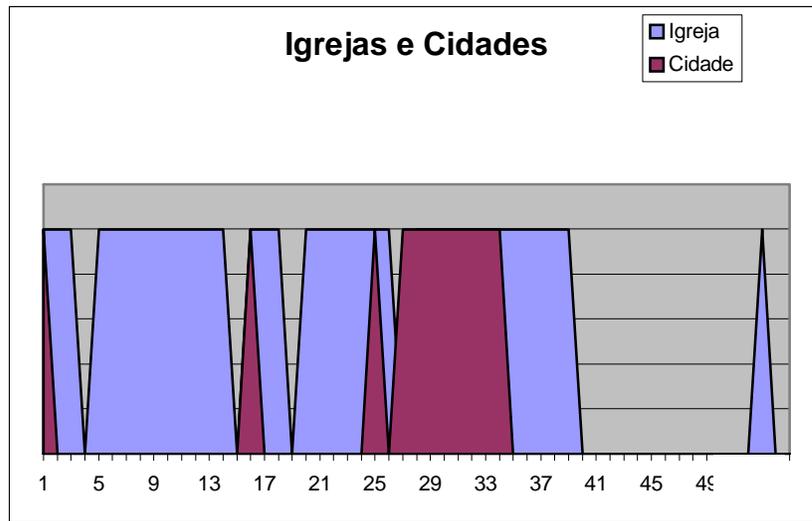
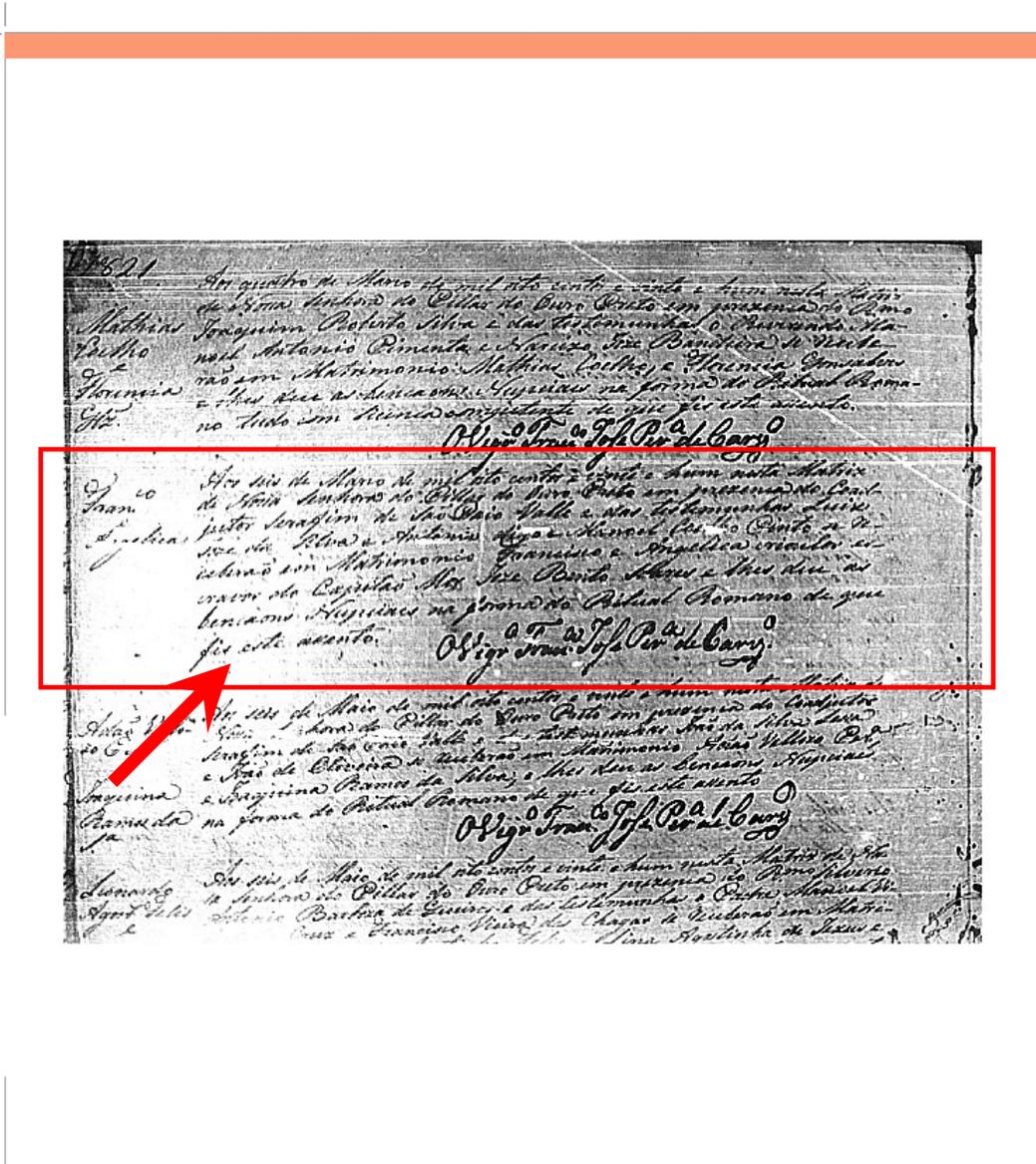


Gráfico 5.1 - Gráfico da distribuição dos campos Igreja e Cidade entre 1800 e 1850 nos assentos de casamento

De posse dos documentos em formato de microfilme, o primeiro trabalho foi de transcrevê-los seguindo as quantidades propostas pelo departamento de estatística. A título de exemplo, a figura 5.2 apresenta um documento utilizado, em formato digital, e sua respectiva transcrição.



ECI/UFMG - Documento em formato digital

Figura 5.2 – Um assento de casamento do ano de 1821 - Rolo de microfilme 107

“Aos Seis de Mayo de mil oitocentos e vinte e hum nesta matriz de Nossa Senhora do Pillar de Ouro Preto em presença do coadjutor Serafim de Sãopaio Valle e das testemunhas Luiz José da Silva e Antonio digo Manoel Coelho Pinto se receberam em Matrimonio Francisco e Angelica crioulos escravos do Capitão Mor José Bento Álvares e lhes dei as bencaons nupciais na forma do Ritual Romano que fiz esse assento.”

Vigário Francisco Jose Pereira da Cruz

A partir do documento transcrito, a identificação dos metadados tem início. No caso do assento apresentado na figura 5.1, os seguintes campos de metadados foram identificados:

- Data de realização da cerimônia: 06 de mayo de 1821
- Nome do templo: Matriz de Nossa Senhora do Pillar
- Padre que celebrou: Serafim de Sãopaio
- Nome da primeira testemunha: Luiz José da Silva
- Nome da segunda testemunha: Manoel Coelho Pinto
- Nome do noivo: Francisco
- Condição social do noivo: Escravo
- Nome da noiva: Ângela
- Condição social da noiva: Escrava
- Nome do dono dos dois, pois eram escravos: José Álvares
- Profissão do dono: Capitão mor
- Nome do padre que fez o assento: Francisco José Pereira da Cruz

Este mesmo processo foi realizado para todos os tipos de assentos e com a frequência sugerida pelo departamento de estatística da UFMG, resultando na tabela de metadados final, a qual foi apresentada ao departamento de história da UFMG. O protótipo de biblioteca digital proposta no capítulo 6 apresenta maiores detalhes sobre os documentos escolhidos para compor os testes preliminares de recuperação da informação na biblioteca digital proposta.

## **5.2 Aspectos interferentes durante a leitura paleográfica**

Durante a leitura paleográfica para identificação dos metadados, percebeu-se que vários aspectos podem influenciar numa pior ou melhor identificação. Nomeou-se como “fonte de ruído”, sendo qualquer tipo de interferência que poderá influenciar no processo de leitura paleográfica culminando em erros na transcrição.

O ruído poderá influenciar tanto no processo de identificação dos metadados potenciais nos microfilmes, quanto no processo de leitura dos documentos para preenchimento dos conteúdos

dos metadados, o que agrava ainda mais o problema, resultando em transcrições que não são exatamente o que está escrito no documento.

A leitura de documentos antigos se apresenta como um grande desafio. Além de outros fatores, em geral as condições de armazenamento não estão de acordo com padrões propostos pela comunidade internacional, o que acarreta documentos deteriorados tornando a leitura de alguns extremamente complexa. Além disso, a própria ação do tempo torna o conteúdo de vários documentos praticamente inacessível.

Vários destes fatores também serão reproduzidos no microfilme e no documento digital, uma vez que estes processos pretendem reproduzir fielmente, as características dos documentos originais. Algumas técnicas podem ser utilizadas para melhorar a qualidade da representação secundária, mas em alguns casos, pode até piorar a leitura. Não são raros os casos em que o usuário precisa recorrer ao documento original para dirimir dúvidas sobre a fonte secundária.

Outra fonte de ruído é o equipamento de leitura que será utilizado. Estes aparelhos possuem um alto preço o que torna inviável para muitas instituições, a utilização de equipamentos novos. Equipamentos de leitura de microfilmes mais antigos contam com problemas principalmente de foco, por ser manual, e na qualidade da lâmpada que eles utilizam, a qual poderá estar com sua vida útil além do previsto. Isso pode acarretar problemas de interpretação, principalmente no cabeçalho e no rodapé dos documentos, uma vez que a qualidade ficará melhor no centro do documento do que nas extremidades. Isso faz necessário realizar a leitura em salas com pouca iluminação, na tentativa de reverter o incômodo.

Com isso, torna-se importante um campo de metadados administrativos no qual o leitor poderá se relatar os detalhes observados durante a leitura do documento. Este campo será utilizado para que ele relate dúvidas e decisões que foram tomadas durante o processo de leitura de um determinado documento. Ainda neste sentido, torna-se importante que os leitores possuam um manual para ser consultado diante de uma dúvida, como por exemplo: nomes das freguesias da época, abreviaturas utilizadas, profissões, nomes de irmandades, nomes de igrejas, entre outros. Desta forma, dúvidas durante a leitura poderão ser resolvidas acessando este guia. Este poderá ser

criado dinamicamente durante a leitura e servirá também como um histórico das observações que foram feitas pelos leitores.

Alguns assentos podem estar fora de ordem o que, no caso de documentos digitalizados, e que contenham um índice por data pode se tornar de fácil acesso, mas no caso de microfimes, este trabalho pode se tornar desgastante com as constantes idas e vindas de forma manual, necessárias a se localizar os períodos a serem lidos.

O que se propõe é que a leitura tenha uma etapa posterior de confirmação. Esta confirmação poderá ser realizada de maneira sistemática por alguém com maiores conhecimentos sobre a historicidade da época, dirimindo dúvidas durante a leitura, e tornando o trabalho em pares, cujos resultados serão mais consistentes.

### **5.3 Metadados segundo o Concílio de Trento**

A primeira pista que se pode considerar a respeito dos metadados que serão utilizados para recuperação da informação no acervo da Matriz do Pilar, já foram previstos no ano de 1545 pelo Concílio de Trento. Naquele momento ficava estabelecida uma resolução que mudaria para sempre a vida dos Católicos, resolução esta que ainda hoje somos submetidos. Ficou decretado que a Cura passaria a ser responsável pelo registro de cada batismo e de cada matrimônio celebrado em sua paróquia, formato este que tornou-se único em todo o mundo. (MARCÍLIO, 2004).

Os dados sobre a cerimônia do batismo que deveriam ser redigidos obrigatoriamente eram: data do batismo, nome do batizando, nome do pai caso tivesse, nome da mãe, local de origem dos pais ou responsáveis e o nome completo de pelo menos um padrinho. Pode-se observar que com o passar dos anos algumas informações foram incorporadas aos registros, como a condição do batizando, isto é, se era forro, livre ou escravo. Por sua vez, o registro de casamento deveria constar: data do casamento, nome de cada conjugue e filiação, residência, naturalidade, nomes dos padrinhos e residência e naturalidade dos padrinhos. Para o óbito as informações que

deveriam estar presentes eram: data do falecimento, nome do morto, estado civil e o nome dos pais no caso de solteiro.

Obviamente que se estivéssemos considerando o mundo contemporâneo este seria um trabalho bastante simplificado, uma vez que o uso de padrões e de formulários pré-impressos para coleta de informações é algo freqüente.

O diferencial deste tipo de coleta de metadados é que a não observação das normas pela maioria dos párocos, fez com que praticamente todos os documentos precisassem ser tratados como exceção, uma vez que por um motivo ou por outro, muitos assentos não estavam de acordo com as normas estabelecidas pela Igreja. Isto não só agigantou a demanda por trabalho metódico, como também passou a ser a principal contribuição desta dissertação, pois estas informações adicionais anotadas pelos padres tornaram-se uma rica fonte para a pesquisa em história, como também a outros campos do conhecimento.

Depois de lidos e transcritos todos os documentos, observando as quantidades propostas pela estatística, os seguintes campos de metadados foram identificados em cada tipo de assentos e estão relatados a seguir.

#### **5.4 Resultados: Planilhas finais de metadados identificados**

Como o objetivo de ser exaustivo no processo de identificação dos metadados, o simples aparecimento de uma ocorrência nos documentos, já se tornou motivo de sua permanência na planilha final de metadados. Os fatores que levaram a decisão de mantê-los foi, em primeiro lugar, o grande poder computacional para manipulação de registros nos computadores atuais, o que no passado foi um fator limitador no armazenamento das informações. Em segundo lugar, o corte de algum metadado poderia levar a falta de informações em projetos futuros e que ainda não teriam sido considerados nas pesquisas atuais.

Abaixo são apresentadas as planilhas finais de casamento, batismo e óbito, acrescidas dos metadados administrativos necessários a representação dos assentos na biblioteca digital,

conforme o padrão Dublin Core e também observando os campos sugeridos pelo MARC para descrição de manuscritos.

**Tabela 5.3 – Metadados finais no Dublin Core:  
Planilha para casamento**

<b>Elemento/Qualificador</b>	<b>Descrição dos elementos e qualificadores</b>
<b>DC.Title</b>	Título do documento, contendo o nome dos noivos.
<b>DC.Source</b>	Fonte utilizada para gerar esta descrição. No caso, considerou-se os documentos originais e os microfilmados como sendo as duas fontes possíveis. (Tabela específica 1) <sup>14</sup>
<b>DC.Format</b>	Formato físico do documento. (Tabela específica 2)
<b>DC.Creator</b>	Instituição que criou o documento original. Foi considerada como <i>creator</i> a Matriz do Pilar de Ouro Preto em todos os assentos
<b>DC.Language</b>	Idioma utilizado nos documentos: Português
<b>DC.Identifier</b>	Código de identificação único no sistema
<b>DC.Date</b>	Data da cerimônia
DC.Date.Microfilmagem	Data de microfilmagem
DC.Date.Digitalização	Data de digitalização do documento
<b>DC.Coverage</b>	Nome do local onde o casamento aconteceu. (Tabela específica 3)
DC.Coverage.Freguesia	Nome da freguesia. (Tabela específica 4)
DC.Coverage.Localidade	Nome da localidade, podendo ser uma cidade, etc.
DC.Coverage.Templo	Nome do templo. (Tabela específica 5)
DC.Coverage.Arcebispo	Arcebispo ao qual o local pertence. (Tabela específica 6)
DC.Coverage.Horario	Horário em que o casamento aconteceu
<b>DC.Contributor</b>	Nome do padre que realizou a cerimônia. (Tabela específica 7)
DC.Contributor.Assentador	Nome de quem fez o assento no original. (Tabela específica 7)
DC.Contributor.Provisao	Nome de quem recebeu a provisão. (Tabela específica 7)
<b>DC.Subject.Noivo</b>	Nome do noivo
DC.Subject.Noivo.Idade	Idade do noivo
DC. Subject.Noivo.Cor	Cor do noivo. (Tabela específica 8)
DC. Subject.Noivo.CondicaoSocial	Condição social do noivo. (Tabela específica 9)
DC. Subject.Noivo.Legitimidade	Legitimidade do noivo. (Tabela específica 10)
DC. Subject.Noivo.NomeDoDono	Nome do dono, caso o noivo seja escravo
DC. Subject.Noivo.Profissao	Profissão do noivo. (Tabela específica 11)
DC. Subject.Noivo.Pai	Nome do pai do noivo
DC. Subject.Noivo.Pai.Profissao	Apresenta a profissão do pai. (Tabela específica

<sup>14</sup> Vários campos precisam ser preenchidos a partir de tabelas específicas. Um resumo do conteúdo de cada tabela pode ser visto no quadro 5.7.

	11)
DC. Subject.Noivo.Mae	Nome da mãe do noivo
DC. Subject.Noivo.País	País de origem do noivo. (Tabela específica 13)
DC. Subject.Noivo.Localidade	Local onde o noivo mora, descrito pelo assentador.
DC. Subject.Noivo.FreguesiaOrigem	Freguesia de origem do noivo. (Tabela específica 4)
DC. Subject.Noivo.FreguesiaBatismo	Freguesia onde o noivo foi batizado. (Tabela específica 4)
DC. Subject.Noivo.TemploDeBatismo	Templo onde o noivo foi batizado. (Tabela específica 5)
DC. Subject.Noivo.IgrejaParoquiano	Igreja onde o noivo é paroquiano. (Tabela específica 5)
DC. Subject.Noivo.ExMulher	Nome da ex-mulher caso seja um novo casamento
<b>DC. Subject.Noiva</b>	Nome da Noiva
DC. Subject.Noiva.Idade	Idade da noiva
DC. Subject.Noiva.Cor	Cor da noiva. (Tabela específica 8)
DC. Subject.Noiva.CondicaoSocial	Condição social da noiva. (Tabela específica 9)
DC. Subject.Noiva.Legitimidade	Legitimidade da noiva. (Tabela específica 10)
DC. Subject.Noiva.NomeDoDono	Nome do dono caso a noiva seja escrava
DC. Subject.Noiva.Pai	Nome do pai da noiva.
DC. Subject.Noiva.Pai.Profissao	Apresenta a profissão do pai. (Tabela específica 11)
DC. Subject.Noiva.Mae	Nome da mãe da noiva.
DC. Subject.Noiva.País	País de origem da noiva. (Tabela específica 13)
DC. Subject.Noiva.Localidade	Local descritivo onde a noiva mora
DC. Subject.Noiva.FreguesiaOrigem	Freguesia de origem da noiva. (Tabela específica 4)
DC. Subject.Noiva.FreguesiaBatismo	Freguesia onde a noiva foi batizada. (Tabela específica 4)
DC. Subject.Noiva.TemploDeBatismo	Templo onde a noiva foi batizada. (Tabela específica 5)
DC. Subject.Noiva.IgrejaParoquiano	Igreja onde a noiva é paroquiana. (Tabela específica 5)
DC. Subject.Noiva.ExMarido	Nome do ex-marido caso seja um novo casamento
<b>DC. Subject.Testemunha1</b>	Nome da primeira testemunha
DC. Subject.Testemunha1.Sexo	Sexo da testemunha
DC. Subject.Testemunha1.Residencia	Local de residência da testemunha
DC. Subject.Testemunha1.Profissao	Profissão da testemunha. (Tabela específica 11)
DC. Subject.Testemunha1.CondicaoSocial	Condição social da testemunha. (Tabela específica 9)
DC. Subject.Testemunha1.Cor	Cor da testemunha (Tabela específica 8)
DC. Subject.Testemunha1.NomeDoDono	Nome do dono da testemunha
<b>DC. Subject.Testemunha2</b>	Nome da primeira testemunha
DC. Subject.Testemunha2.Sexo	Sexo da testemunha.
DC. Subject.Testemunha2.Residencia	Local de residência da testemunha
DC. Subject.Testemunha2.Profissao	Profissão da testemunha. (Tabela específica 11)
DC. Subject.Testemunha2.CondicaoSocial	Condição social da testemunha. (Tabela específica 9)
DC. Subject.Testemunha2.Cor	Cor da testemunha. (Tabela específica 8)
DC. Subject.Testemunha2.NomeDoDono	Nome do dono da testemunha
<b>DC. Relation</b>	Nome do documento digital ao qual se reporta

DC.Relation.Original	Documento original com número e página do livro
DC.Relation.Microfilme	Documento microfilmado com número do rolo e volume
<b>DC. Impedimentos</b> (Elemento adicionado ao DC)	Impedimentos para a realização do casamento
<b>DC.Description</b>	Descrição geral do documento tratando dos aspectos físicos
DC.Description.Decisoões	Decisões tomadas sobre a leitura e interpretação feita do conteúdo lido.

**Tabela 5.4 – Metadados finais no Dublin Core:  
Planilha para batizados**

<b>Elemento/Qualificadores</b>	<b>Descrição dos elementos e qualificadores</b>
<b>DC.Title</b>	Título do documento, contendo o nome do batizando.
<b>DC.Source</b>	Fonte utilizada para gerar esta descrição. No caso, considerou-se os documentos originais e os microfilmados como sendo as duas fontes possíveis. (Tabela específica 1)
<b>DC.Format</b>	Formato físico do documento. (Tabela específica 2)
<b>DC.Creator</b>	Instituição que criou o documento original. No caso consideramos como <i>creator</i> a Matriz do Pilar de Ouro Preto.
<b>DC.Language</b>	Idioma utilizado nos documentos: Português
<b>DC.Identifier</b>	Código de identificação único no sistema
<b>DC.Date</b>	Data do batismo
DC.Date.Microfilmagem	Data de microfilmagem
DC.Date.Digitalização	Data de digitalização do documento
<b>DC.Coverage</b>	Nome do local onde o batismo aconteceu. (Tabela específica 3)
DC.Coverage.Templo	Nome do templo. (Tabela específica 5)
DC.Coverage.Freguesia	Nome da freguesia. (Tabela específica 4)
DC.Coverage.Modos	Modo como foi realizado o batismo. metadado para marcar S/N no caso de ter sido batizado <i>In-extremis</i>
<b>DC.Contributor</b>	Nome do padre que realizou a cerimônia. (Tabela específica 7)
<b>DC.Subject.Batizando</b>	Nome do batizando
DC.Subject.Batizando.Idade	Idade presumida do batizando (Tabela específica 20)
DC.Subject.Batizando.Legitimidade	Legitimidade do batizando. (Tabela específica 10)
DC.Subject.Batizando.Sexo	Sexo do batizando
DC.Subject.Batizando.ExpostoPara	Se a criança foi exposta, registra para quem ele foi exposta.
DC.Subject.Batizando.DataExposto	Data em que foi exposto
DC.Subject.Batizando.Alforriado	Indica se o batizando foi alforriado em pia batismal
DC.Subject.Batizando.Pai	Nome do pai
DC.Subject.Batizando.Pai.Viuvo	Campo que indica se o pai da criança era viúvo.
DC.Subject.Batizando.Pai.Cor	Cor do pai do batizando. (Tabela específica 8)
DC.Subject.Batizando.Pai.Condicao	Condição social do Pai. (Tabela específica 9)
DC.Subject.Batizando.Pai.Profissao	Profissão do pai do batizando. (Tabela específica 11)
DC.Subject.Batizando.Mae	Nome da mãe do batizando
DC.Subject.Batizando.Mae.Cor	Cor da mãe do batizando (Tabela específica 8)
DC.Subject.Batizando.Mae.Viuva	Campo que indica se a mãe era viúva
DC.Subject.Batizando.Mae.Condicao	Condição social da mãe. (Tabela específica 9)
DC.Subject.Batizando.FreguesiaDosPais	Freguesia de origem dos pais do batizando. (Tabela específica 4)
DC.Subject.Batizando.SantosOleos	Campo que apresenta se o batizando recebeu os

	santos óleos
<b>DC.Subject.Padrinho</b>	Nome do padrinho
DC.Subject.Padrinho.Localidade	Local de residência do padrinho
DC.Subject.Padrinho.Freguesia	Freguesia do padrinho. (Tabela específica 4)
DC.Subject.Padrinho.Casado	Indica se o padrinho era casado
DC.Subject.Padrinho.Profissao	Profissão do padrinho. (Tabela específica 11)
DC.Subject.Padrinho.Procuracao	Indica se o padrinho era padrinho por procuração
<b>DC.Subject.Madrinha</b>	Nome da madrinha
DC.Subject.Madrinha.Casada	Indica se o madrinha era casada
DC.Subject.Madrinha.Freguesia	Freguesia da marinha. (Tabela específica 4)
DC.Subject.Madrinha.Localidade	Local de residência da madrinha
DC.Subject.Madrinha.Procuracao	Indica se a madrinha era madrinha por procuração
<b>DC.Relation</b>	Nome do documento digital ao qual se reporta
DC.Relation.Original	Documento original com página e número do livro
DC.Relation.Microfilme	Documento microfilmado com rolo e volume
<b>DC.Description</b>	Descrição geral do documento tratando dos aspectos físicos
DC.Description.Decisooes	Decisões tomadas sobre a leitura e interpretação feita do conteúdo lido.

**Tabela 5.5 – Metadados finais no Dublin Core:  
Planilha para óbitos**

<b>Elemento/Qualificadores</b>	<b>Descrição dos elementos e qualificadores</b>
<b>DC.Title</b>	Título do documento, contendo o nome do falecido e a data de falecimento.
<b>DC.Source</b>	Fonte utilizada para gerar esta descrição. No caso, considerou-se os documentos originais e os microfilmados como sendo as duas fontes possíveis. (Tabela específica 1)
<b>DC.Format</b>	Formato físico do documento. (Tabela específica 2)
<b>DC.Creator</b>	Instituição que criou o documento original. No caso consideramos como creator a Matriz do Pilar de Ouro Preto.
<b>DC.Language</b>	Idioma utilizado nos documentos
<b>DC.Identifier</b>	Identificador único no sistema
<b>DC.Date</b>	Data do óbito
DC.Date.Microfilmagem	Data de microfilmagem
DC.Date.Digitalização	Data de digitalização do documento
<b>DC.Coverage</b>	Nome do local de enterro (Tabela específica 21)
DC.Coverage.Freguesia	Freguesia (Tabela específica 4)
<b>DC.Contributor</b>	Nome do padre que realizou a cerimônia. (Tabela específica 7)
DC.Contributor.Assentador	Nome de quem que fez o assento. (Tabela específica 7)
<b>DC.Subject.Falecido</b>	Nome do falecido ou falecida
DC.Subject.Falecido.CausaDoObito	Motivo que levou ao falecimento
DC.Subject.Falecido.Idade	Idade presumida do falecido(a)
DC.Subject.Falecido.Cor	Cor do falecido(a). (Tabela específica 8)
DC.Subject.Falecido.Sexo	Sexo do falecido(a)
DC.Subject.Falecido.Mae	Nome da mãe do falecido(a)
DC.Subject.Falecido.Pai	Nome do pai do falecido(a)
DC.Subject.Falecido.Legitimidade	Condição de legitimidade do falecido. (Tabela específica 10)
DC.Subject.Falecido.EstadoCivil	Estado civil do falecido. (Tabela específica 22)
DC.Subject.Falecido.Conjugue	Nome do conjugue
DC.Subject.Falecido.CondiçãoSocial	Sua condição social quando faleceu. (Tabela específica 9)
DC.Subject.Falecido.Profissao	Profissão do falecido. (Tabela específica 11)
DC.Subject.Falecido.LocalDeResidencia	Local de residência. Campo textual pois depende muito dos detalhes que o assentador quis passar.
DC.Subject.Falecido.Irmandade	Irmandade que acompanhou. (Tabela específica 12)
DC.Subject.Falecido.NomeDoDono	Em caso do falecido ser escravo, conterà o nome do dono.
DC.Subject.Falecido.ProfissaoDoDono	Profissão do dono. (Tabela específica 11)
DC.Subject.Falecido.TumbaouCaixa	Campo para escolher se o falecido utilizou tumba ou caixão. (Tabela específica 19)
DC.Subject.Falecido.Mortalha	Tipo de mortalha utilizada, escolhido em uma lista de opções. (Tabela específica 16)
DC.Subject.Falecido.Missa	Tipo de missa que foi celebrada. (Tabela

	específica 18)
DC.Subject.Falecido.Sacramentos	Campo para escolher os sacramentos que o falecido recebeu. (Tabela específica 17)
DC.Subject.Falecido.Oficio	Campo de marcar (SIM/NÃO) se houve ofício.
DC.Subject.Falecido.Testamento	Campo de marcar (SIM/NÃO) indicando se deixou testamento
DC.Subject.Falecido.InExtremis	Campo de marcar (SIM/NÃO) indicando se foi batizado inextremis.
DC.Subject.Falecido.Bocal	Campo de marcar (SIM/NÃO) indicando se o falecido era boçal.
DC.Subject.Falecido.Acompanhado	Campo de marcar (SIM/NÃO) indicando se foi acompanhado
<b>DC.Relation</b>	Nome do documento digital ao qual se reporta
DC.Relation.Original	Documento original com número da página e livro
DC.Relation.Microfilme	Documento microfilmado com rolo e volume
<b>DC.Description</b>	Descrição geral do documento tratando dos aspectos físicos
DC.Description.Decisoões	Decisoões tomadas sobre a leitura e interpretação feita do conteúdo lido.

#### 5.4.1 Resumo dos elementos e qualificadores utilizados e acrescentados

Como dito anteriormente, o Dublin Core é composto de 15 elementos que podem receber qualificadores de maneira a tornar a descrição mais específica para a coleção em questão. Além da utilização de qualificadores, em alguns casos até mesmo novos elementos podem ser criados, de maneira a atender da melhor forma possível a descrição de um documento. A tabela 5.6 mostra um resumo dos elementos e qualificadores utilizados, acrescentados ou não utilizados durante a elaboração dos metadados para os assentos.

**Tabela 5.6 – Resumo dos elementos e qualificadores utilizados ou criados**

	<b>Casamento</b>	<b>Batismo</b>	<b>Óbito</b>
Elementos utilizados	12	12	12
Elementos não utilizados	3	3	3
Elementos criados	1	0	0
Qualificadores criados	58	36	33

Vários foram os campos de metadados que deverão ser preenchidos a partir de uma tabela auxiliar. Nestas tabelas, muitas delas cadastradas previamente, estão à lista de opções que o usuário deverá escolher no preenchimento do metadado em questão. Abaixo tem-se a lista de todas as tabelas, o tipo de conteúdo e um resumo das opções nelas contidas.

**Tabela 5.7 – Resumo das tabelas de apoio no preenchimento dos metadados**

<b>Número da tabela</b>	<b>Conteúdo</b>	<b>Opções para seleção</b>
1	Fontes	- Documento original - Documento microfilmado
2	Formato	- Documento digital - Documento microfilmado
3	Locais	- No templo - Em casa - Outros
4	Freguesias	- Listagem das freguesias (cadastro prévio)
5	Templos	- Listagem dos templos (cadastro prévio)
6	Arcebispados	- Listagem dos arcebispados (cadastro prévio)
7	Párocos	- Listagem dos párocos (cadastro prévio)
8	Cor	- Branco - Preto - Negro - Outros - NC (Não consta)
9	Condição social	- Livre - Forro - Escravo - Quartado - Aldeiado - NC
10	Condição de legitimidade	- Legítimo - Natural - Exposto - Não se aplica
11	Profissão	- Listagem das profissões da época (cadastro prévio)
12	Irmandades	- Listagem das irmandades (cadastro prévio)
13	Países	- Listagem de países (cadastro prévio)
14	Acompanhamento	- Indica se foi acompanhado pelo pároco
16	Mortalha	- Nome das mortalhas (cadastro prévio)
17	Sacramentos	- Todos - Parcial - Nenhum

		NC
18	Tipos de missa	Nome da missa celebrada (cadastro prévio)
19	TumbaCaixão	- Nome da tumba utilizada ou opção para escolher caixão. (cadastro prévio)
20	Idade	- Inocente - Adulto - NC
21	Local de enterro	- Listagem dos locais de enterro (Cadastro prévio)
22	Estado civil	- Casado - Solteiro - Viúvo - NC

**Tabela 5.8 – Significado de alguns termos menos comuns**

<b>Termo</b>	<b>Significado</b>
Assentador	Pároco que fez o assento no livro.
Provisão	O casamento era função do pároco, mas este podia licenciar o capelão ou o coadjutor (auxiliar) para realizá-lo.
Alforriado	O escravo podia receber ou comprar sua carta de alforria, tornando-se liberto. Este, portanto era alforriado.
Boçal	Escravo africano recém-chegado da África. Não conhecia a língua portuguesa e portanto só podia receber os sacramentos da igreja católica, a começar pelo batismo, com um tradutor presente, pois durante a cerimônia ele precisava responder a algumas perguntas, como por exemplo: Crê em um único Deus? Resposta: Sim. Como as respostas eram sempre uma afirmação ou uma negativa, era possível que o escravo fosse minimamente "orientado" para respondê-las, sem passar por uma catequização da maneira como havia sido prescrita pela legislação católica.
Santos óleos	Os sacramentos da Igreja católica são ministrados a partir de falas pré-determinadas, gestos e ações. Uma delas era ungir a testa da pessoa com óleo sacramentado. Havia um para o batismo, outro para o crisma e outro para a extrema-unção. Somente a partir da unção com os santos óleos o sacramento estava completo.
Procuração	Há casos em que os padrinhos convidados para o batismo ou testemunhas do casamento não podiam estar presentes na cerimônia. Por isso outra pessoa se apresentava por procuração do padrinho ou testemunha "oficial".
Tumba caixão	Nome da tumba em que o falecido seria colocado, ou caixão conforme o caso.
Mortalha	Se o defunto era irmão da Ordem Terceira de São Francisco de Assis ou

	do Carmo, podia pedir para ser enterrado com a mortalha daquela irmandade, ou seja, com a vestimenta própria daqueles irmãos.
Acompanhamento	Quando o padre acompanhava o cortejo e enterro.
Impedimentos	Havia vários motivos que impediam um casamento, como por exemplo parentes com entre os noivos. Entretanto, os noivos podiam ser dispensados de vários impedimentos. O casamento se realizava, mas esta dispensa é citada no assento de casamento: "após terem sido dispensados do impedimento de ..."
Legitimidade	Se a criança é fruto de pai e mãe casados de acordo com o ritual católico, a criança é legítima. Caso ela seja fruto de união consensual (concubinada) ou filho de mãe solteira, é natural.
Condição social	Numa sociedade escravista, a condição social de cada um é bem determinada. Livres, escravos, forros. A hierarquia social não se fazia somente a partir da condição econômica, mas de seu lugar na sociedade. Esta caracterização se aproximava da sociedade portuguesa nobiliárquica, onde o nascimento definia o papel social de cada um.
In-extremis	O batismo deveria ser ministrado o mais rapidamente nas crianças, para evitar que morressem sem recebê-lo. Para isso as Constituições Primeiras orientavam que qualquer pessoa podia ministrá-lo em caso de perigo de morte. Como a mortalidade infantil era alta, muitas crianças eram batizadas in-extremis, ou seja, com urgência. Posteriormente, sobrevivendo, deveriam ir à igreja para confirmar o sacramento, inclusive se colocando os santos óleos.
Data exposto	Data em que a criança foi exposta, ou seja, abandonada, na maioria das vezes na porta de alguma família ou pessoa que pudesse criá-la.
Sacramento	Indicação de quais os sacramentos que a pessoa recebeu. Em geral o padre destacava se era todos, parcial ou nenhum. Na hora da morte, o padre ministrava a extrema-unção. Este ritual era formado na realidade por três sacramentos: o arrependimento e confissão dos pecados, a eucaristia e a unção com óleo bento. Dependendo da gravidade do caso, um ou mais dos sacramentos podiam deixar de ser realizados, por exemplo: a pessoa em coma não tem como confessar ou receber a eucaristia. Portanto o enfermo recebia a extrema unção de maneira parcial.
Testamento	Se o falecido deixou testamento.

Como podemos perceber nas planilhas finais apresentadas, são vários os metadados que podem servir como ponto de acesso neste tipo de documento, com isso, encontrar um padrão que atendesse todas as particularidade da recuperação da informação, quer seja na arquivologia ou na biblioteconomia não foi viável. Por vezes o excesso e a particularidade nos manuscritos criados pelos padres, fez com que qualquer que fosse o padrão de metadados escolhidos para representar

os documentos, fosse ampliado enormemente. A escolha pelo Dublin Core se deve principalmente, por se tratar de um modelo internacionalmente aceito e que representa bem os documentos digitais, onde seu uso é recomendado.

## **6. A BIBLIOTECA DIGITAL: Métodos e resultados**

Na primeira etapa desta dissertação, que contemplou o planejamento e seleção dos documentos e uma posterior identificação dos metadados, já se idealizou a criação de uma biblioteca digital que pudesse tornar mais fácil o acesso às informações contidas nos assentos.

O desenvolvimento de um protótipo de biblioteca digital finaliza esta dissertação, apresentando o resultado final a partir dos metadados propostos. Protótipo, segundo o dicionário Aurélio, é algo produzido para ser o primeiro tipo ou exemplar, servindo como um modelo preliminar. A criação de um protótipo pode ser uma etapa importante no desenvolvimento de softwares, uma vez que a partir deles os usuários poderão testar o sistema, permitindo uma avaliação antes da implementação definitiva. Desta forma, a massa de documentos utilizada servirá apenas para uma análise preliminar do funcionamento do sistema, uma vez que a coleção de documentos da Matriz do Pilar ainda não se encontra digitalizada.

As etapas identificadas neste processo foram:

1. Escolha do software para gestão da biblioteca digital
2. Proposta do banco de dados
3. Proposta da interface para digitação das planilhas de metadados
4. Prototipação do módulo administrativo
5. Criação da coleção no software escolhido para biblioteca digital
6. Criação do módulo para exportação dos registros do banco de dados para a biblioteca digital
7. Testes de funcionalidade tanto do módulo administrativo quanto da recuperação da informação na biblioteca digital

### **6.1 Escolha do software para gestão da biblioteca digital**

Dentre os diversos softwares que poderiam ser utilizados na implementação da biblioteca digital, optou-se por uma avaliação daqueles que atendessem aos padrões de software livre, que são:

- A liberdade de executar o programa, para qualquer propósito;
- A liberdade de estudar como o programa funciona, e adaptá-lo para as suas necessidades, permitindo o acesso ao código fonte.
- A liberdade de redistribuir cópias de modo que você possa ajudar ao seu próximo;
- A liberdade de aperfeiçoar o programa, e liberar os seus aperfeiçoamentos, de modo que toda a comunidade se beneficie.

Após a avaliação de algumas alternativas para o software de biblioteca digital, a escolha recaiu sobre o aplicativo Greenstone, produzido pela Zealand Digital Library Project em conjunto com a University of Waikato, sendo indicado e distribuído pela UNESCO e pela ONG Human Info, e que se tornou um dos softwares mais utilizados no mundo, em se tratando de disponibilização de documentos digitais. O Greenstone está disponível em três versões e pode ser executado em vários sistemas operacionais tanto da família Windows, quanto Unix/Linux. Isto faz dele uma ótima plataforma em se tratando de portabilidade entre ambientes. Existem atualmente três versões do Greenstone, sendo que a principal diferença é o ambiente onde ele irá executar, são elas:

- Greenstone Local: Executado em versão *desktop*;
- Greenstone Rede: Para execução e acesso via rede de computadores, inclusive, via Internet;
- Greenstone código fonte: O qual inclui os códigos que podem ser utilizados na geração de CDRom.

Apesar de ser um software muito completo, e atender plenamente vários dos requisitos para criação de uma biblioteca digital, ele ainda se mantém fácil de ser utilizado, com uma boa velocidade na recuperação da informação e permitindo, através de *Plugins*<sup>15</sup>, o acesso a uma série de tipos de documentos digitais tais como DOC, XLS, PDF, JPG, GIF.

---

<sup>15</sup> Pequenos programas que acrescentam funcionalidades aos sistemas onde eles são instalados.

De uma maneira geral, o trabalho do Greenstone se resume a criar uma base de documentos em XML a partir dos documentos para ele importados em outros formatos. Desta maneira o usuário pode realizar buscas em todas as partes de um documento, inclusive, utilizando operadores booleanos. Manipulando o Greenstone podemos configurar como o software deverá disponibilizar o acesso aos documentos pelo usuário final. Estes ajustes são muito fáceis de serem realizados, dispensando qualquer tipo de programação técnica.

Ainda nesta etapa, e após a escolha do Greenstone, foi definida a arquitetura geral de funcionamento da biblioteca digital. A figura 6.1 apresenta esta arquitetura:

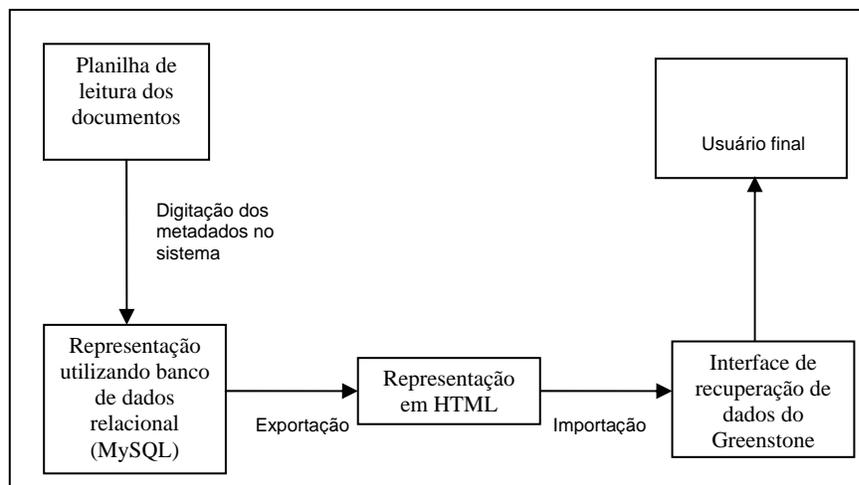


Figura 6.1 – Arquitetura final da biblioteca digital proposta

O planejamento da construção do sistema seguiu as etapas propostas pela metodologia da “análise estruturada” (YOURDON, 1990) a qual normatiza um processo para construção de software para aplicações interativas que cobre as etapas de: análise, planejamento, diagrama de entidades e relacionamento, diagrama de fluxo de dados, codificação e testes. Questões sobre a preparação e digitalização dos documentos não foram cobertos nesta proposta de biblioteca digital, uma vez que este não era um dos objetivos desta dissertação. Entretanto, na criação da biblioteca digital em sua versão final, a digitalização e todos os seus aspectos precisam ser analisados.

Apesar disso, com a criação do protótipo foi possível avaliar, com um grau aceitável de confiança, tanto o hardware quanto o software necessário, uma vez que a Casa dos Contos de Ouro Preto disponibilizou alguns documentos em formato digital para testes. Com estes documentos inseridos no sistema foi possível testar a recuperação da informação a partir de palavras chaves diretamente no Greenstone, testes realizados pelo próprio mestrando.

Obviamente que a tecnologia digital requer gestão permanente, para evitar os problemas trazidos pela obsolescência, procedimentos inadequados e riscos de perdas de dados. Uma das atividades contempladas no projeto final deverá tratar da implantação de uma política de gestão permanente dos acervos digitais, que visa determinar responsabilidades sobre as operações de conservação do acervo, bem como um planejamento estratégico de longo prazo para a tecnologia no arquivo.

## 6.2 Proposta do banco de dados

Uma vez definidos tanto o software para recuperação dos documentos, quanto a arquitetura geral da aplicação que seria construída, a próxima etapa foi a criação do banco de dados que irá comportar a digitação dos assentos. O software de banco de dados escolhido nesta pesquisa foi o MySQL por comportar, sem maiores riscos, a quantidade de dados tratados e por não possuir necessidade de aquisição de licença pelo uso. A seguir estão apresentadas as três principais tabelas criadas para receber a digitação das planilhas de batismo, casamento e óbito.

**Tabela 6.1 – Metadados finais representados no banco de dados MySQL:**

**Tabela de batismo**

Field	Type	Default	Extra
DCTitle	varchar(200)		
DCSource	int(11)	0	
DCFormat	int(11)	0	
DCCreator	varchar(200)	Matriz do Pilar de Ouro Preto	
DCLanguage	int(11)	0	
DCIdentifier	int(11)		auto_increment
DCDate	date	0000-00-00	
DCDateMicrofilmagem	date	0000-00-00	
DCDateDigitalizacao	date	0000-00-00	
DCCoverage	int(11)	0	
DCCoverageFreguesia	int(11)	0	
DCCoverageTemplo	int(11)	0	
DCContributor	int(11)	0	
DCSubjectBatizando	varchar(200)		

DCSubjectBatizandoldade	int(11)	0	
DCSubjectBatizandoLegitimidade	int(11)	0	
DCSubjectBatizandoSexo	int(11)	0	
DCSubjectBatizandoExpostoPara	varchar(200)		
DCSubjectBatizandoDataExposto	date	0000-00-00	
DCSubjectBatizandoMaeCor	int(11)	0	
DCSubjectBatizandoMaeCondicao	int(11)	0	
DCSubjectBatizandoAlforriado	int(11)	0	
DCSubjectBatizandoPai	varchar(200)		
DCSubjectBatizandoPaiViuvo	int(11)	0	
DCSubjectBatizandoPaiCor	int(11)	0	
DCSubjectBatizandoPaiCondicao	int(11)	0	
DCSubjectBatizandoPaiProfissao	int(11)	0	
DCSubjectBatizandoMae	varchar(200)		
DCSubjectBatizandoMaeViuva	int(11)	0	
DCSubjectBatizandoFreguesiaDosPais	int(11)	0	
DCSubjectBatizandoSantosOleos	int(11)	0	
DCSubjectPadrinho	varchar(200)		
DCSubjectPadrinhoLocalidade	varchar(200)		
DCSubjectPadrinhoFreguesia	int(11)	0	
DCSubjectPadrinhoCasado	int(11)	0	
DCSubjectPadrinhoProfissao	int(11)	0	
DCSubjectPadrinhoProcuracao	int(11)	0	
DCSubjectMadrinha	varchar(200)		
DCSubjectMadrinhaCasada	int(11)	0	
DCSubjectMadrinhaFreguesia	int(11)	0	
DCSubjectMadrinhaProcuracao	int(11)	0	
DCRelation	varchar(200)		
DCRelationOriginal	varchar(200)		
DCRelationMicrofilme	varchar(200)		
DCDescription	text		
DCDescriptionDecisoes	text		

**Tabela 6.2 – Metadados finais representados no banco de dados MySQL:  
Tabela de casamento**

Field	Type	Default
DCTitle	varchar(200)	
DCSource	int(11)	0
DCFormat	int(11)	0
DCCreator	varchar(200)	Matriz do Pilar de Ouro Preto
DCLanguage	int(11)	0
DCIdentifier	int(11)	
DCDate	date	0000-00-00
DCDateMicrofilmagem	date	0000-00-00
DCDateDigitalizacao	date	0000-00-00
DCCoverage	int(11)	0
DCCoverageFreguesia	int(11)	0
DCCoverageLocalidade	int(11)	0
DCCoverageTemplo	int(11)	0
DCCoverageArcebisado	int(11)	0
DCCoverageHorario	varchar(20)	
DCContributor	int(11)	0
DCContributorAssentador	int(11)	0
DCContributorProvisao	int(11)	0
DCSubjectNoivo	varchar(200)	
DCSubjectNoivoldade	tinyint(4)	0
DCSubjectNoivoCor	int(11)	0
DCSubjectNoivoCondicaoSocial	int(11)	0
DCSubjectNoivoLegitimidade	int(11)	0
DCSubjectNoivoNomeDoDono	varchar(200)	
DCSubjectNoivoProfissao	int(11)	0
DCSubjectNoivopai	varchar(200)	
DCSubjectNoivoPaiProfissao	int(11)	0
DCSubjectNoivomae	varchar(200)	
DCSubjectNoivoPais	int(11)	0
DCSubjectNoivoLocalidade	varchar(200)	
DCSubjectNoivoFreguesiaOrigem	int(11)	0
DCSubjectNoivoFreguesiaBatismo	int(11)	0
DCSubjectNoivoTemploBatismo	int(11)	0
DCSubjectNoivolgrejaParoquiano	int(11)	0
DCSubjectNoivoExMulher	varchar(200)	
DCSubjectNoiva	varchar(200)	
DCSubjectNoivaIdade	int(11)	0
DCSubjectNoivaCor	int(11)	0
DCSubjectNoivaCondicaoSocial	int(11)	0
DCSubjectNoivaLegitimidade	int(11)	0
DCSubjectNoivaNomeDoDono	varchar(200)	
DCSubjectNoivaPai	varchar(200)	
DCSubjectNoivaPaiProfissao	int(11)	0
DCSubjectNoivaMae	varchar(200)	
DCSubjectNoivaPais	int(11)	0
DCSubjectNoivaLocalidade	varchar(200)	
DCSubjectNoivaFreguesiaOrigem	int(11)	0
DCSubjectNoivaFreguesiaBatismo	int(11)	0
DCSubjectNoivaTemploBatismo	int(11)	0

DCSubjectNoivaIgrejaParoquiana	int(11)	0
DCSubjectNoivaExMarido	varchar(200)	
DCSubjectTestemunha1	varchar(200)	
DCSubjectTestemunha1Sexo	int(11)	0
DCSubjectTestemunha1Residencia	varchar(200)	
DCSubjectTestemunha1Profissao	int(11)	0
DCSubjectTestemunha1CondicaoSocial	int(11)	0
DCSubjectTestemunha1Cor	int(11)	0
DCSubjectTestemunha1NomeDoDono	varchar(200)	
DCSubjectTestemunha2	varchar(200)	
DCSubjectTestemunha2Sexo	int(11)	0
DCSubjectTestemunha2Residencia	varchar(200)	
DCSubjectTestemunha2Profissao	int(11)	0
DCSubjectTestemunha2CondicaoSocial	int(11)	0
DCSubjectTestemunha2Cor	int(11)	0
DCSubjectTestemunha2NomeDoDono	varchar(200)	
DCRelation	varchar(200)	
DCRelationOriginal	varchar(200)	
DCRelationMicrofilme	varchar(200)	
DCImpedimentos	varchar(200)	
DCDescription	text	
DCDescriptionDecisoies	text	

**Tabela 6.3 - Metadados finais representados no banco de dados MySQL:  
Tabela de óbito**

Field	Type	Default
DCTitle	varchar(200)	
DCSource	int(11)	0
DCFormat	int(11)	0
DCCreator	varchar(200)	Matriz do Pilar de Ouro Preto
DCLanguage	int(11)	0
DCIdentifier	int(11)	0
DCDate	date	0000-00-00
DCDateMicrofilmagem	date	0000-00-00
DCDateDigitalizacao	date	0000-00-00
DCCoverage	int(11)	0
DCCoverageFreguesia	int(11)	0
DCContributor	int(11)	0
DCContributorCelebrante	int(11)	0
DCContributorAssentador	int(11)	0
DCSubjectFalecido	varchar(200)	
DCSubjectFalecidoCausadoObito	varchar(200)	
DCSubjectFalecidoIdade	int(11)	0
DCSubjectFalecidoCor	int(11)	0
DCSubjectFalecidoSexo	int(11)	0
DCSubjectFalecidoMae	varchar(200)	
DCSubjectFalecidoPai	varchar(200)	
DCSubjectFalecidoLegitimidade	int(11)	0
DCSubjectFalecidoEstadoCivil	int(11)	0
DCSubjectFalecidoConjuge	varchar(200)	

DCSubjectFalecidoCondicaoSocial	int(11)	0
DCSubjectFalecidoProfissao	int(11)	0
DCSubjectFalecidoLocalDeResidencia	varchar(200)	
DCSubjectFalecidoIrmandade	int(11)	0
DCSubjectFalecidoNomeDoDono	varchar(200)	
DCSubjectFalecidoProfissaoDono	int(11)	0
DCSubjectFalecidoTumbaOuCaixao	int(11)	0
DCSubjectFalecidoMortalha	int(11)	0
DCSubjectFalecidoMissa	int(11)	0
DCSubjectFalecidoSacramentos	int(11)	0
DCSubjectFalecidoOficio	int(11)	0
DCRelation	varchar(200)	
DCRelationOriginal	varchar(200)	
DCRelationMicrofilme	varchar(200)	
DCDescription	text	
DCDescriptionDecisoies	text	

Alem das tabelas apresentadas acima, a metodologia de representação através de bancos de dados relacionais, como é o caso do MySQL, necessita de tabelas adicionais as quais irão conter os dados periféricos em relação aos dados principais que auxiliaram no momento do preenchimento dos metadados, são elas:

**Tabela 6.4 – Outras tabelas do sistema proposto**

<b>Tipo</b>	<b>Conteúdo</b>
Acompanhamento	Lista dos tipos de acompanhamento
Arcebispos	Lista dos arcebispos
Celebrantes	Lista dos celebrantes
Condição social	Lista das condições sociais
Cor	Lista de cor
Freguesias	Lista de freguesias
Igrejas	Lista de Templos
Irmandades	Lista das irmandades
Legitimidade	Tipos de legitimidade
Local de enterro	Locais de enterro
Localidade	Localidades
Mortalha	Tipos de Mortalha
Países	Países
Profissões	Profissões
Sacramentos	Tipos de sacramentos recebidos
Tipos de missa	Missas encomendadas
Tumba caixão	Nome da tumba ou caixão

### 6.3 Proposta da interface para digitação das planilhas

Uma vez criado o banco de dados, tem-se a etapa de elaboração do que chamamos de “módulo administrativo”. Este módulo estará disponível para que os operadores digitem os dados resultantes das leituras realizadas. Este processo é conhecido como transcrição dos documentos e contempla o preenchimento de planilhas como a mostrada no anexo IVa, IVb e IVc. Esta planilha é formada por todos os campos dos metadados identificados e seus dados irão alimentar o banco de dados elaborado.

As funcionalidades do módulo administrativo compõem o núcleo do sistema de informação, local onde podem ser realizados a inclusão de documentos, consultas ao acervo e geração do arquivo em HTML que será utilizado pelo Greenstone.

A partir da planilha preenchida manualmente durante a leitura paleográfica, o usuário pode utilizar o módulo administrativo para digitação dos dados diretamente nas tabelas do banco de dados. A figura 6.2 apresenta a *interface* criada para digitação dos assentos de casamentos. Para cada registro gerado são gravadas cópias digitais que ficam guardadas em repositórios (pastas em memória secundária) com os documentos digitalizados com o texto completo.

Como previsto pelo DublinCore, cada registro conta com campos descritivos e administrativo, preparados para receber dados sobre possíveis problemas ou mesmo rastrear o histórico resultante das leituras paleográficas. Os campos “DCDescription e DCDescriptionDecisões” deverão conter qualquer anotação sobre o processo de leitura ou sobre a qualidade do documento transcrito que o operador julgue ser interessante para o usuário que realizará o acesso posterior ao sistema utilizando o Greenstone.

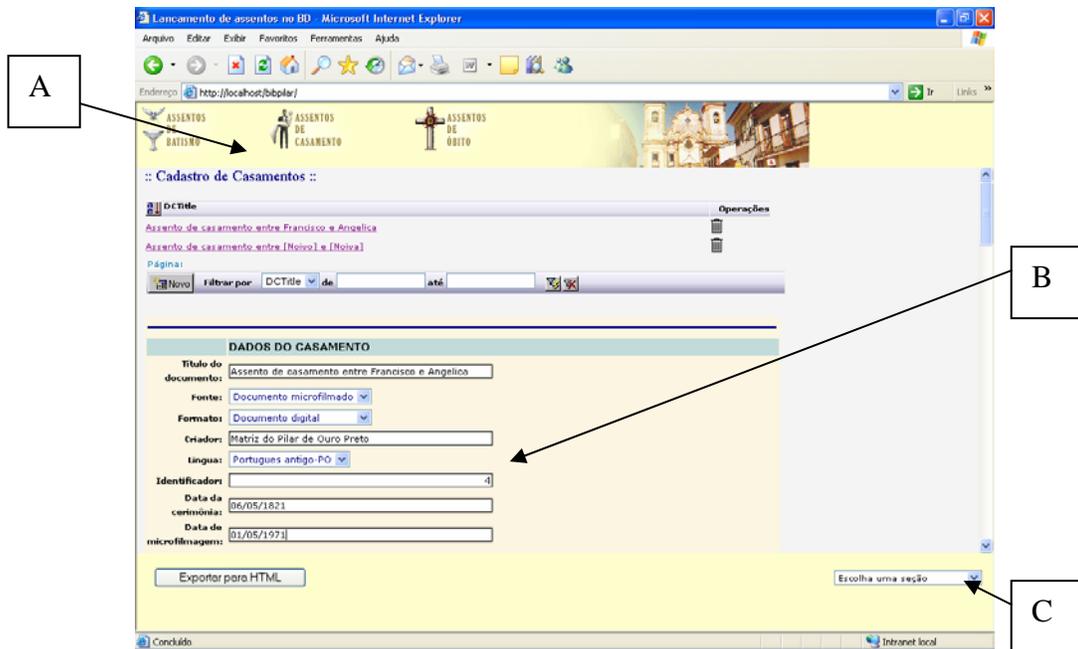


Figura 6.2 – Interface para digitação dos assentos de casamentos

A principal preocupação na criação desta interface foi o de facilitar o processo de utilização. Na parte superior (A), o operador escolhe o tipo de assento a ser digitado no sistema, podendo ser de batismo, casamento ou óbito. Na parte central da tela (B) o sistema apresenta a planilha de metadados conforme a escolha do operador. Os demais dados das tabelas de apoio o usuário, conforme a tabela 6.4, poderão ser acessadas através do *menu* localizado na parte inferior direita da tela (C). A interface para digitação de batismos e óbitos segue o mesmo raciocínio. Esta padronização permite uma curva de tempo para aprendizado menos acentuada.

#### 6.4 Exportação para a linguagem HTML

Após a digitação dos dados no sistema via módulo administrativo, através da interface apresentada no item anterior, os dados já estarão armazenados no banco de dados MySQL. Entretanto, o software Greenstone não suporta a recuperação de dados diretamente neste banco de dados. Além disso, como a proposta é de se criar um mecanismo que possibilite a recuperação da informação no ambiente WEB, optou-se por migrar os dados do MySQL para o padrão HTML, uma vez que este é um dos padrões mais utilizados para intercâmbio de dados entre equipamentos na Internet.

HTML (Hypertext Markup Language) é a “língua franca” para publicar hipertexto na Web. Esta linguagem tornou-se bastante difundida, tanto do ponto de vista acadêmico quanto comercial, devido a sua simplicidade. O HTML é derivado da linguagem SGML (Standard Generalized Markup Language) sendo um padrão internacional (ISO 8879) publicado em 1986 e regula o uso de marcações descritivas mescladas ao documento. O SGML também fornece um método padrão para nomear as estruturas de um texto, definindo modelos hierárquicos para cada tipo de documento produzido. Uma vantagem do HTML em relação a outros padrões como o PDF e PS, se refere ao tamanho final do arquivo que, usando HTML será bem menor, uma vez que ele utiliza textos simples para representar tanto o conteúdo quanto a formação final que o documento deverá ter.

A linguagem HTML funciona através de marcadores estruturando os documentos para que possam ser lidos por navegadores web. Ela possui semântica simples, além de não diferenciar entre maiúsculas ou minúsculas. Foi criado um mecanismo automático que permite que todos os registros no banco de dados seja exportado para HTML para depois ser importado pelo Greenstone o qual realiza a consulta através da linguagem XML.

No HTML cada elemento é demarcado por uma TAG inicial e final também conhecida como rótulo. O rótulo inicial terá o seguinte formato <nome\_da\_TAG>, e o rótulo final terá o mesmo nome do inicial, acrescido de uma barra </nome\_da\_TAG>. Por exemplo:

```
<bold> Assento de casamento </bold>
```

Esta marcação diz ao navegador para colocar o texto “Assento de casamento” em negrito. Algumas tags permitem o uso de atributos, redefinindo propriedades, como por exemplo, tamanhos, posicionamentos e funcionalidades.

```
<p align=”center”>José Angola</p>
```

Neste caso a TAG “p” irá gerar um novo parágrafo. Entretanto, temos o atributo “align” o qual permite centralizar o texto a ser exibido pelo navegador, neste parágrafo apenas.

Parte de uma representação em HTML contendo os metadados de um assento de casamento pode ser vista na figura 6.3.

```

<html>
<head>
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1">
</head>

<body bgcolor="#FFFFFF" text="#000000">
<b>Data de realiza&ccedil;&atilde;o da cerim&ocirc;nia: </b>06 de mayo de 1821<br>
<b>Nome do templo: </b>Matriz de Nossa Senhora do Pillar<br>
<b>Padre que celebrou:</b> Serafim de S&atilde;opaio<br>
<b>Nome da primeira testemunha:</b> Luiz Jos&eacute; da Silva<br>
<b>Nome da segunda testemunha:</b> Manoel Coelho Pinto<br>
<b>Nome do noivo</b>: Francisco<br>
<b>Condi&ccedil;&atilde;o social do noivo: </b>Escravo<br>
<b>Nome da noiva:</b> &Aacute;ngela<br>
<b>Condi&ccedil;&atilde;o social da noiva:</b> Escrava<br>
<b>Nome do dono dos dois, pois eram escravos: </b>Jos&eacute; &Aacute;lvaes<br>
<b>Profiss&atilde;o do dono:</b> capit&atilde;o mor<br>
<b>Nome do padre que fez o assento: </b>Francisco Jos&eacute; Pereira da Cruz<br>
</body>
</html>

```

Figura 6.3 – Trecho do código em HTML para metadados de casamento

## 6.5 Criação da coleção no Greenstone

Com todos os registros exportados para arquivos HTML, pode-se agora criar a biblioteca no módulo *Greenstone Librarian Interface*, a qual permite o acesso pelos usuários responsáveis pela gestão da biblioteca. A figura 6.4 apresenta a interface do bibliotecário.

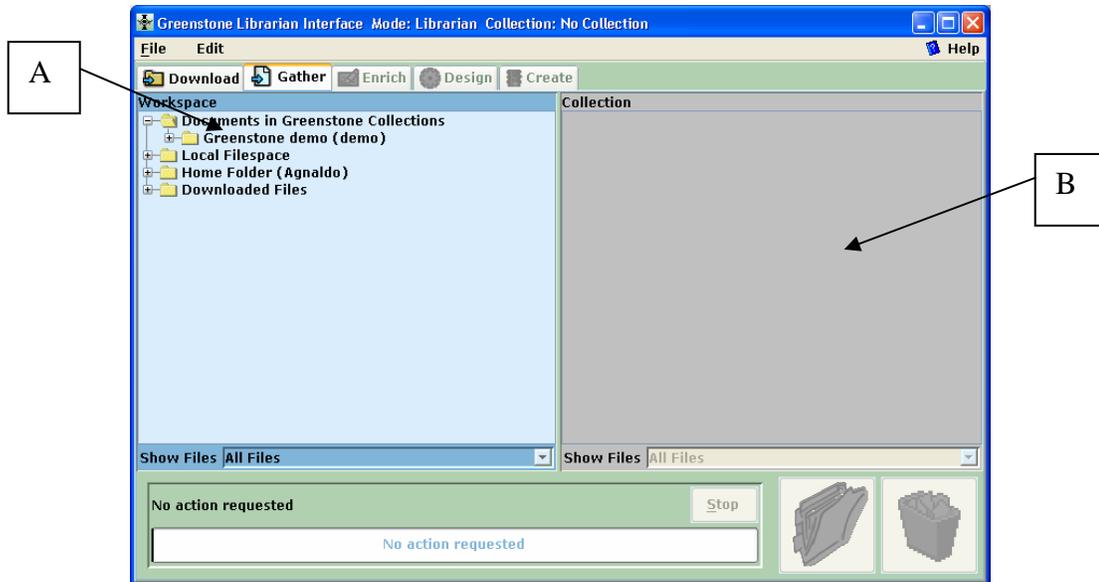


Figura 6.4 – Tela principal de utilização pelo bibliotecário.

Na parte esquerda da tela (A) o operador tem acesso aos principais módulos do Greenstone. Nesta figura podemos ver a existência de uma biblioteca de demonstração (demo). No lado direito aparecem todos os arquivos que compõem a biblioteca selecionada (neste caso ela está vazia).

Para criar uma nova biblioteca o operador deverá clicar no menu File...New. Em seguida uma tela para informações gerais sobre a nova biblioteca irá aparecer.

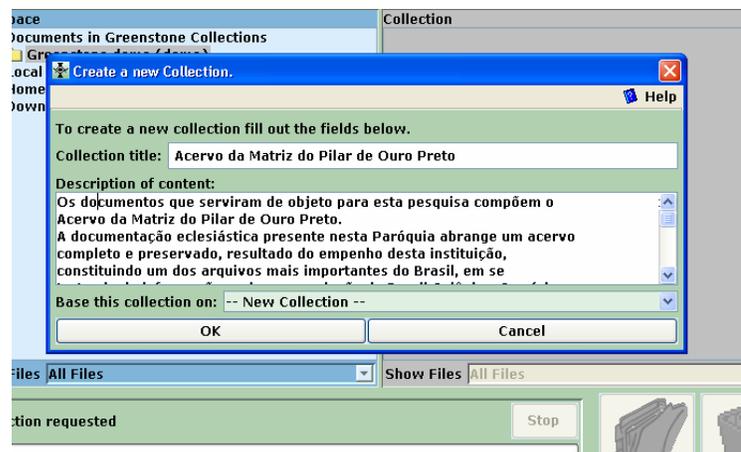


Figura 6.5 – Criando uma nova coleção.

Após a criação da nova biblioteca, falta apenas a importação dos arquivos que serão utilizados. Isso pode ser feito de maneira simples, clicando e arrastando os documentos com os metadados, a

partir do seu local de origem no sistema (A) para dentro da nova coleção (B). A figura 6.6 ilustra este processo.

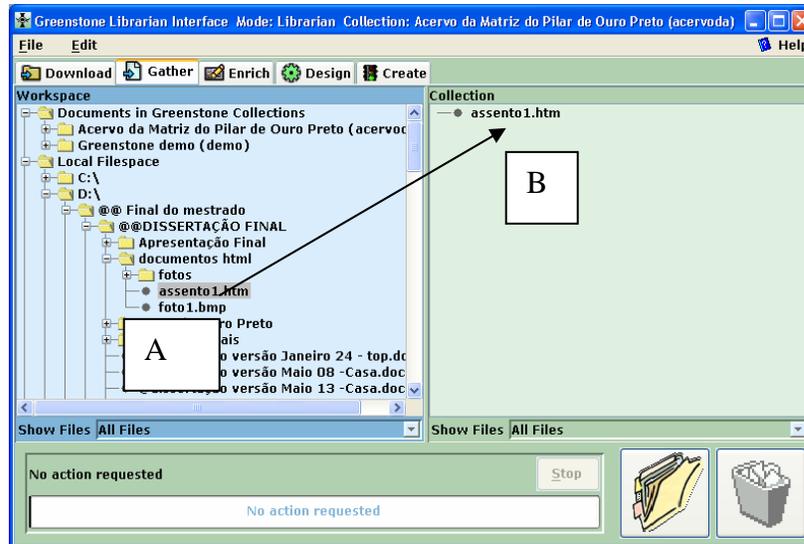


Figura 6.6 – Importando os arquivos HTML para dentro da coleção criada.

Para finalizar a criação da biblioteca digital, basta o operador repetir este mesmo processo com todos os documentos digitais que compõem a coleção. Isso também vale para os documentos digitalizados os quais contem os assentos em texto completo. Clicando sobre o botão *Create*, e aguardando alguns segundos a coleção já estará pronta para acesso pelos usuários.

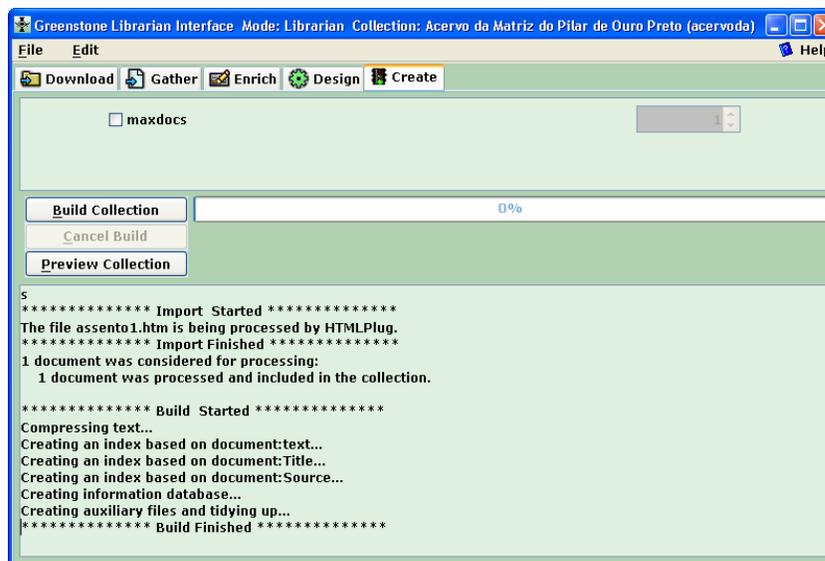


Figura 6.7 – Processo de criação da biblioteca finalizado.

## 6.6 Recuperação a Informação pelo Greenstone

Flexibilizar o acesso à informação é uma das principais funções de uma biblioteca digital. O propósito da arquitetura de informação é representar as informações ricas e variadas de uma biblioteca. No caso deste protótipo, o objetivo principal é facilitar a recuperação de documentos digitais do acervo Eclesiástico da Matriz do Pilar. Para que essas operações sejam realizadas com sucesso, as informações na biblioteca digital devem estar organizadas efetivamente e disponíveis via mecanismos como a busca por palavra chave.

O Greenstone pode manter diversas coleções separadas individualmente as quais podem crescer com o passar do tempo com a adição de novos documentos. Para encontrar a informação o usuário poderá contar com mecanismos como buscas por palavra chave específicas, as quais podem estar no texto de um documento ou em uma de suas seções e sub-seções. A consulta também pode ser realizada por título, através de um *click* sobre ela a partir de uma listagem; e também sendo possível consultar as coleções através de temas, como em uma organização típica de uma prateleira em uma biblioteca convencional.

Como sobredito, os documentos disponibilizados pelo Greenstone podem estar em vários formatos, por exemplo Texto puro, HTML, WORD ou PDF. A consulta aos metadados identificados, os quais servirão como guias que identificam a localização do documento de texto completo, se dará através de metadados armazenados no formato XML.

Para localizar um documento, o consulente pode utilizar-se de um texto-livre o qual será comparado com todos os campos de metadados do documento. Dependendo da configuração disponível no Greenstone o consulente também pode realizar uma busca nos campos de observação.

Neste protótipo a tela do Greenstone foi mantida como na original. Ela permite que se realizem buscas utilizando palavras chaves em todo o corpo do documento. A figura 6.8 mostra a realização de uma busca pela palavra chave “Francisco”, a qual deverá ser digitada no campo

apontado pela letra (A). Em seguida o usuário clica sobre o botão *Begin Search* para que a recuperação tenha início.

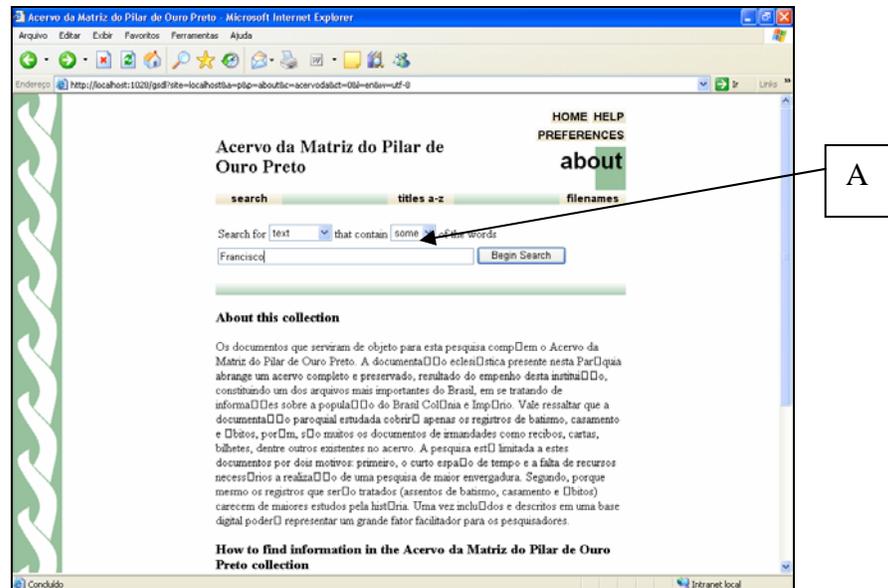


Figura 6.8 – Busca por documentos com a palavra chave Francisco.

Neste momento o Greenstone procura em toda a sua base documental, por registros que contenham esta palavra chave. Como resposta ele apresenta uma tela como mostrada na figura 6.9.

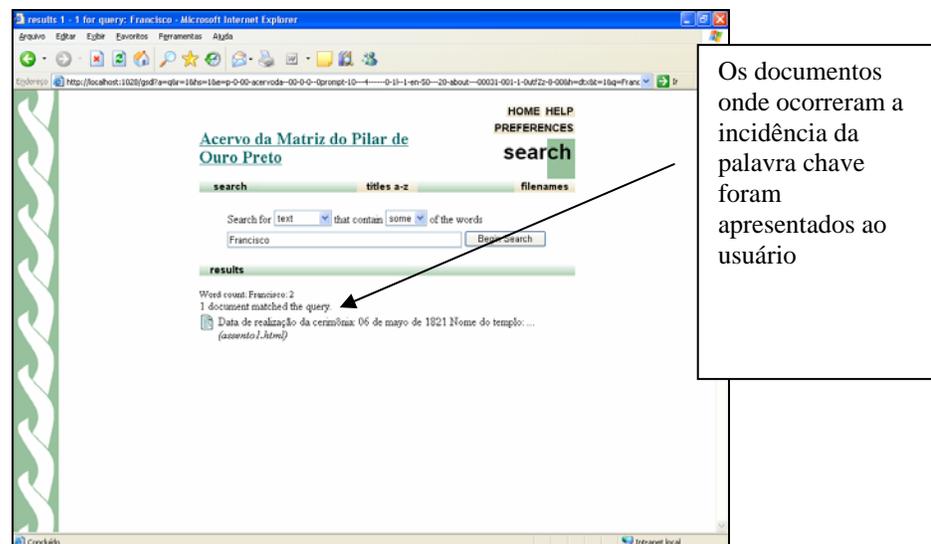


Figura 6.9 – Resultado da busca realizada pelo Greenstone

A figura 6.10 apresenta a listagem de todos os documentos contendo a palavra chave digitada pelo usuário. Este documento possui um link que levará o usuário a uma tela que irá exibir o documento digitalizado.

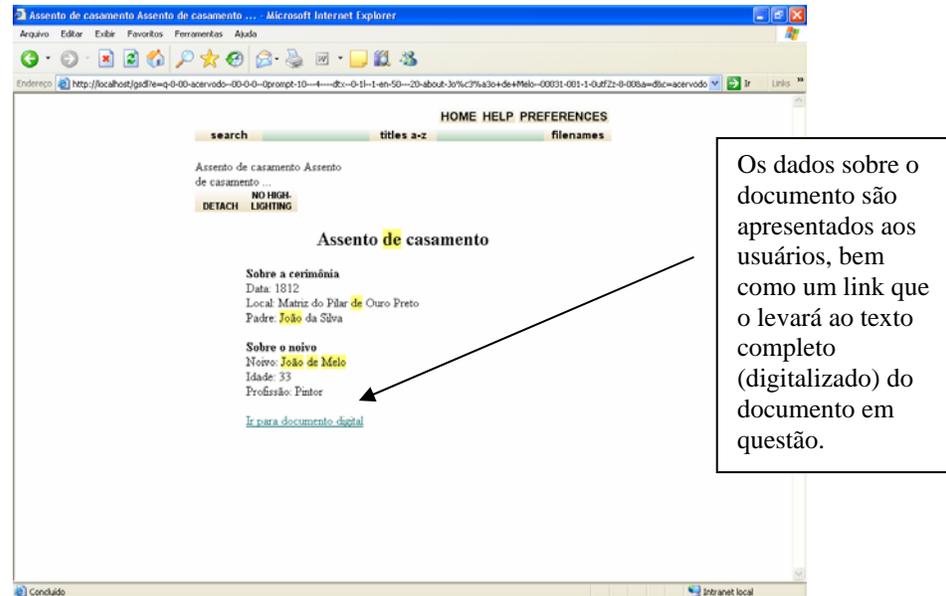


Figura 6.10 Exibição de todos os campos do documento original.

A escolha e o acionamento de um dos links em destaque permite que o usuário visualize o documento em sua forma digitalizada.

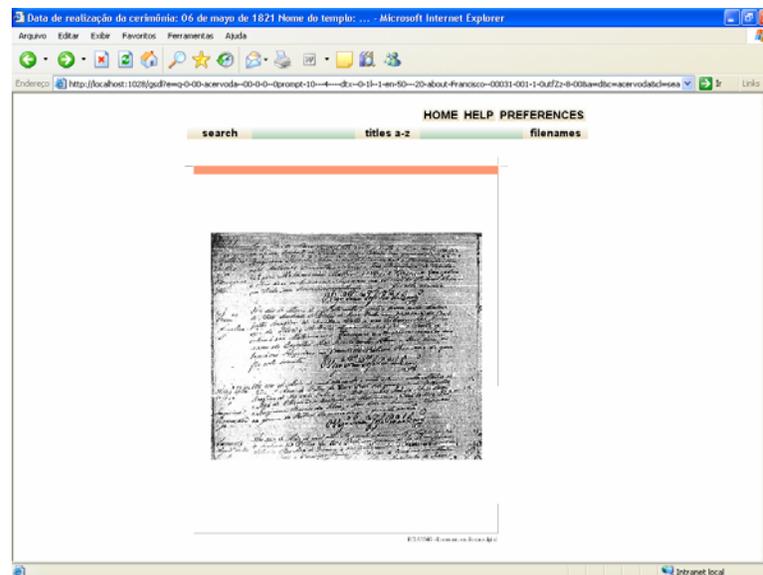


Figura 6.11 – Documento digitalizado com texto completo.

Esta recuperação só foi possível graças a todo o levantamento dos metadados realizados no item 5 desta dissertação. Estes descritores, conhecidos como metadados, incluíram todas as partes dos documentos que fossem relevantes no processo de busca.

### **6.7 Aspectos sobre a visualização do texto completo**

Em se tratando de tecnologia digital e sua disponibilização via web, a digitalização dos acervos é uma etapa importante pois a maior parte das vantagens proporcionadas pelas bibliotecas digitais se deve ao fato de que, o documento digitalizado, ou seja, acessível pelo meio virtual, irá permitir que vários dos requisitos para bibliotecas digitais sejam atendidos, dentre eles, os citados por Cunha: Acesso remoto do usuário via computador conectado a rede; utilização simultânea de um documento por mais de um usuário; acesso não apenas a referência, mas também ao texto completo; utilização de forma que a biblioteca local não necessite ser a dona da coleção; utilização de sistemas auxiliares na recuperação da informação Cunha (1999).

Trata-se de um momento fundamental na elaboração do sistema como um todo. As informações devem ser digitadas de forma precisa observando os padrões que foram definidos, uma vez que este processo será realizado por diversos operadores os quais não devem tomar decisões de forma individual quer seja no preenchimento da planilha quer seja durante a digitação no sistema informatizado. Qualquer caminho tomado deverá ser discutido com os demais membros da equipe, e as decisões devem ser informadas ao usuário quando da elaboração do mecanismo de recuperação biblioteca digital.

### **6.8 Outras formas de se visualizar a informação a partir dos metadados propostos**

Uma importante contribuição dos metadados é na elaboração de gráficos e cruzamento de informações.

Contagem de NumeroDoVol	Condicao					Total geral
	NomeDoCelebrante	Escravo	Forro	Livre	N/C	
Agostinho Ribeiro Cabellos				3		3
Alexandre da Silva Vas				2		2
Antonio Alvares Ferreira Rodrigues				1		1
Antonio Alves Nogueira			1			1
Antonio Correia de Faria		1	4	4	1	10
Antonio de Almeyda Vellozo				1		1
Antonio de Pinna		1		1		2
Antonio de Souza Braga				2		2
Antonio Dias Nicolao Barreto					1	1
Antonio Gonçaves de Moraes e Castro		1				1
Antonio Jose Ribeiro Bhering				1		1
Antonio Metella Velho				1		1
Antonio Ribeiro de Azevedo				4		4
Belchior Pinheiro de Oliveira				1		1
Bernardo Hygino Dias Coelho				10		10
Bruno Joze de Souza e Castro				1		1
Diogo Martins Lassa				1		1
Emmerenciano Maximino de Azeredo Coutinho				4		4
Felecianno Joze Dias				1		1
Feliciano Ferreira de Carvalho		1		21		22
Felipe de Almeyda				1		1
Felix Antonio Lisboa				1		1
Francisco da Sylva e Almeida		5	5	18	1	29
Francisco de Almeida Pinto		7	1	29	1	38

Tabela 7.1 – Quantidade de casamentos realizados ordenados por padre que realizou e separados pela condição social.

A tabela 7.1 apresenta uma listagem que conta a quantidade de casamentos realizados por cada padre a partir de 1712. Pela listagem podemos ver que o padre que mais fez casamentos foi o pároco Francisco de Almeida Pinto que realizou durante sua permanência na Igreja Matriz do Pilar um total de 38 casamentos. Esta, por exemplo, poderia ser uma pista que desencadearia uma pesquisa de maior envergadura na busca pela influência que ele teria na vida dos habitantes de Ouro Preto.

Esta listagem foi possível porque o grupo de estudos do banco de dados da Matriz do Pilar de Ouro Preto, possui um sistema criado no aplicativo *Access*, com vários documentos transcritos dos microfimes.

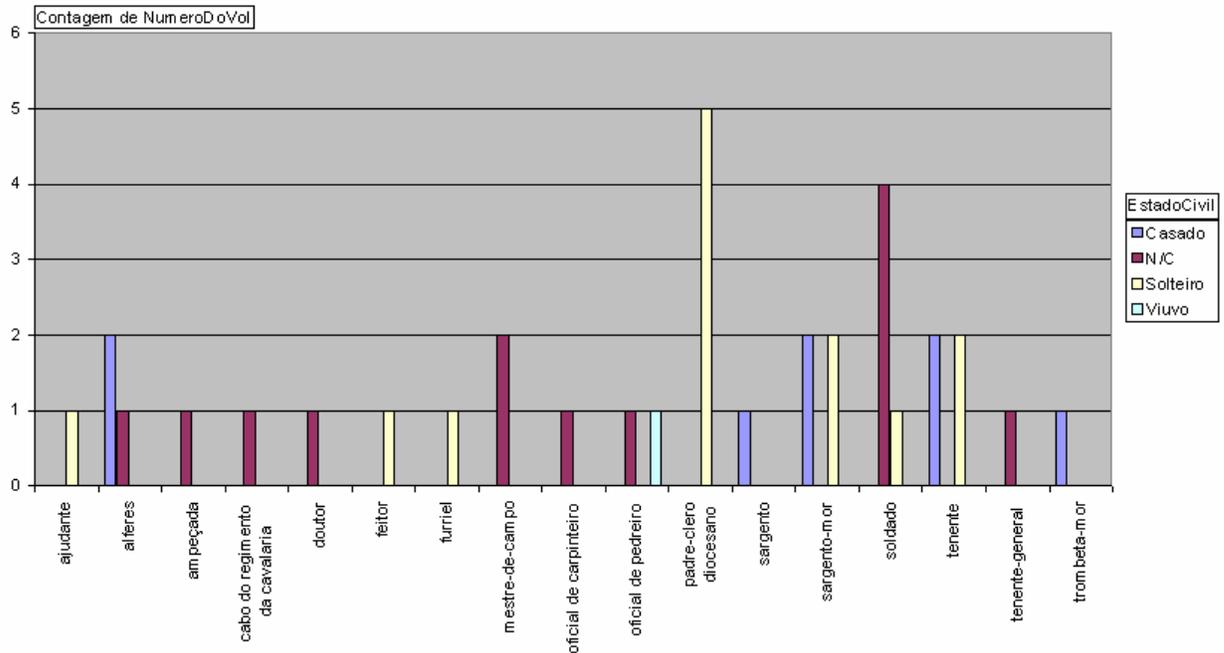


Gráfico 7.1 – Número de óbitos separados por estado civil e profissão.

O gráfico 7.1 traz algumas informações interessantes, por exemplo, a maior parte dos óbitos neste período foram de padres, seguido de soldados. Dentre os soldados na maioria dos assentos não constava se ele era casado ou não. Entretanto, boa parte deles eram solteiros. Ou seja, porque não colocavam se o soldado morto era casado? Essas e outras questões só são possíveis graças aos infundáveis cruzamentos de dados possíveis neste tipo de análise. Elas se tornam possíveis a partir da separação consistente em metadados os quais permitem esta forma de recuperação da informação multidimensional.

## 7. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

O caráter exploratório da presente pesquisa, aliada às novas tecnologias do tratamento e processamento de informações, possibilitou a problematização do potencial informativo do acervo em questão, assim como a proposta de soluções para a descrição e organização desse tipo de documento, soluções essas calcadas em bases teóricas e produtivas. O levantamento de metadados, visando-se à recuperação ágil de importantes informações históricas, a partir da análise multidimensional dos assentamentos de registros civis da entidade religiosa católica, a Matriz do Pilar de Ouro Preto, MG, veio sem dúvida contribuir para o aperfeiçoamento do acesso direto ao conteúdo de documentos otimizando a pesquisa em várias campos de conhecimento, ressaltado-se dentre estes a história, o direito, a economia, a sociologia.

A análise detalhada e identificação dos elementos informativos, como etapa que antecede a criação de um biblioteca digital, deu origem aos metadados que servirão para descrever os assentamentos, tal como demonstrado no aplicativo criado também no âmbito desta dissertação, possibilitando uma recuperação ampla de seus conteúdos. No que tange a esse levantamento ressaltam-se os seguintes comentários: A dificuldade da leitura de manuscritos antigos; A cobertura o mais exaustiva possível dos elementos de descrição; O atendimento às comunidades de interesse; As dificuldades de digitalização dos originais devido ao seu estado.

O objetivo principal deste trabalho foi o de identificação dos metadados, processo básico para que se obtenha sucesso com a recuperação da informação. Quando elaboramos os metadados e os representamos através de uma biblioteca digital estamos colaborando para que o acesso seja facilitado.

No que se refere às atuais bibliotecas digitais, a ciência da informação oferece tecnologias para a melhor organização dos dados a serem disponibilizados, de forma a dar maior segurança, e facilitar o manuseio, bem como permitir que sejam criadas inter-relações entre os dados, permitindo aos pesquisadores um elevado grau de sucesso quando da necessidade de uma determinada informação.

Percebeu-se que uma importante contribuição da CI para a pesquisa histórica é a explanação dos planos que podem ser seguidos, padrões e métodos que podem ser utilizados para tornar a pesquisa menos árdua para o pesquisador. Ela pode contribuir desde a problematização em detalhes, identificando fraquezas, sugerindo a utilização das tecnologias, como as bibliotecas digitais, tanto na elucidação de conceitos quanto na própria elaboração do projeto, fazendo jus a sua denominação de área social aplicada. Um acompanhamento durante a elaboração dos metadados, da criação da biblioteca digital e principalmente na digitação dos assentos da biblioteca digital é de fundamental importância, podendo trazer variações significativas na qualidade final do sistema que será criado.

Acreditamos que esta dissertação também proporcione o levantamento de recursos para disponibilização deste tipo de acervo, uma vez que várias etapas e requisitos foram identificados e nomeados, servindo de referência não apenas para a documentação da Igreja Matriz do Pilar como também de outras Igrejas que possuem este tipo de documentação. O percurso agora está delineado e menos sujeito a surpresas.

Ao se tentar propor uma planilha de metadados, uma das limitações identificadas foi o pouco tempo que dispusemos para visitar outras instituições que guardam este tipo de documento (eclesiástico), uma vez que existem outras igrejas que podem possuir este tipo de documentação, ou mesmo órgãos públicos como o Arquivo Público Mineiro. A coleta de amostras nestas instituições poderia ter sido considerada, apesar que, em princípio não deveria trazer grandes novidades em termos de metadados, uma vez que as definições básicas já estavam propostas pela igreja católica e valeriam para todas as igrejas do Arcebispado. Mesmo assim, essa constatação poderia ser de importância para uma planilha que pretendesse ser universalmente aceita.

A elaboração do protótipo, como apresentado no capítulo 6, foi de certa forma facilitada pelo uso do Greenstone o que comprovou a viabilidade técnica em se implementar o sistema. Entretanto, a elaboração e disponibilização de uma biblioteca digital como esta, tem como requisito o planejamento e execução de algumas etapas prévias e posteriores que precisam ser levadas em consideração neste tipo de sistema. Uma destas etapas é o estudo da viabilidade econômica. A

implementação de qualquer sistema passa pelo uso de mão de obra de pessoal especializado e isso deve ser levado em consideração no custo do sistema. Outro custo a ser considerado é a aquisição dos equipamentos, seja para digitalização do acervo, para digitação em um sistema informatizado e para acesso pelos usuários. Além destes custos, existem gastos com pessoal especializado em história, os quais devem estar aptos a dirimir dúvidas sobre as informações contidas no sistema no que diz respeito ao conteúdo, provendo uma consultoria histórica aos usuários. Do ponto de vista gerencial e tecnológico também seriam necessárias pessoas envolvidas, pois existiriam manutenções no sistema digital, com operações de backup, análise de logs, resolução de problemas de acesso como erros em logins, cadastro de novos usuários, controles de acesso, entre outros.

O Greenstone se mostrou eficiente neste protótipo de biblioteca digital, mas outros softwares poderiam ser avaliados, principalmente quando o sistema definitivo estiver sendo tratado. A representação do acervo no Greenstone se mostrou fácil de ser realizada, mas as limitações na recuperação da informação pode deixar o sistema ineficaz. Isso pode acontecer porque, basicamente, a consulta se baseia em palavras chaves e, em vista disso, devemos destacar que a escrita portuguesa daquela época não era como conhecemos hoje. Vários nomes de pessoas e lugares não eram escritos como atualmente. Desta forma, o usuário deveria conhecer as várias maneiras de se escrever um nome para conseguir recuperar a informação no Greenstone. Isso pode ser um dificultador para a localização de um documento digital na biblioteca.

A quantidade de documentos utilizados para o protótipo foi o mínimo necessário para apresentar sua funcionalidade. Entretanto, uma quantidade maior de documentos deverá ser utilizada para realização de testes de maior envergadura. Outros usuários também poderão ser selecionados para testar tanto a interface de recuperação dos documentos, quanto validar a informação recuperada.

Estudar aspectos sobre digitalização de documentos não foi objetivo desta pesquisa, porém, trata-se de uma etapa importante na criação do sistema de biblioteca digital para a Matriz do Pilar (caso ela ocorra), pois o estado de boa parte dos documentos irá requerer um completo domínio de técnicas de digitalização e equipamentos modernos, uma vez que a ação do tempo foi cruel com boa parte deles.

Aspectos legais também deverão ser levados em conta. Atualmente este acervo está sob a guarda da própria Matriz do Pilar e, disponibilizar o acesso deve ser algo resolvido em conjunto com esta instituição e os órgãos do governo que determinam as leis para este tipo de arquivo.

Para estudos futuros sugere-se a análise e implementação de um sistema que realizasse a recuperação por palavra chave, mas que considerasse também a semelhança fonética entre as palavras. Isso iria aumentar o número de documentos retornados na busca. A tabela 7.1 apresenta alguns nomes e as recuperações que seriam feitas.

<b>Palavra chave digitada pelo usuário</b>	<b>Documentos que seriam recuperados por semelhança fonética conteriam também:</b>
José	Jose, Joze, Jozé
Tomaz	Thomaz, Tomaz, Thomas, Tomas
Anastácio	Anastazio, Anasthacio
Bartholomeu	Bartholomeo, Bartolomeo, Bartolomeu

Tabela 7.1 – Palavras chaves do usuário e outras que seriam consideradas pelo sistema, verificando a semelhança fonética.

Como trabalho futuro sugere-se também a pesquisa e utilização de outros padrões de metadados, como o EAD (Encoded Archival Description), o qual permite de forma eficiente tanto arquivos quanto manuscritos, padrão que foi proposto pela Sociedade Americana dos Arquivistas (SAA).

No que diz respeito ao acesso, ao colocar este acervo na forma digital estaremos ampliando de forma inimaginável as possibilidades de realização de pesquisas com informações da época tratada. Trabalhos neste sentido vêm ocorrendo todos os dias em países de primeiro mundo. Hoje, estudar a história da Grécia, por exemplo, pode ser feito com um grau elevado de credibilidade, sem que o pesquisador tenha que recorrer diretamente às fontes primárias o que, em muitos casos, inviabiliza totalmente a pesquisa, seja por falta de apóio financeiro, ou por indisponibilidade dessas fontes.

Futuramente o sistema também deverá levar em consideração o seu funcionamento em vários tipos de infra-estrutura de informática, seja em CDROM, DVD, acessado via internet, considerando os principais sistemas operacionais como Windows/Linux e Unix.

Apesar das limitações e dos trabalhos futuros a serem desenvolvidos, o percurso de pesquisa e elaboração da planilha de metadados para toda a coleção de documentos da Matriz do Pilar de Ouro Preto, com a posterior construção do protótipo de biblioteca digital, mostrou-se como uma importante ferramenta de avaliação do sistema que poderia ser implementado de forma definitiva. Com seu desenvolvimento envolvendo instituições financiadoras, estaríamos contribuindo não apenas com a preservação da memória do estado de Minas Gerais, como também ampliando seu acesso.

## Referências

- AGUIAR, Marcos Magalhães de. *Estado e Igreja na capitania de Minas gerais: Notas sobre mecanismos de controle da vida associativa*. Revista Varia História, Belo Horizonte, n.2, p.80-114, 1997.
- ALVARENGA, Lídia. *A Teoria do Conceito Revisitada em Conexão com Ontologias e Metadados no Contexto das Bibliotecas Tradicionais e Digitais*. DataGramZero-Revista de Ciência da Informação, v.2 , n.6, 2001.
- ANDRADE, Marcos Ferreira de. *Rebeliões escravas na comarca de Ouro Preto*. Revista Varia História, Belo Horizonte, n.17, p.237-257, 1997.
- ARAÚJO, Vânia M. R. H. de. *Sistemas de recuperação da informação: nova abordagem teórico-conceitual*. 1994. 240 f. (Doutorado em Educação) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1994.
- ATKINS, D. L., Ball, T., Bruns, G. and Mawl, K. C. *A domin-specific language for form-based services*. IEEE Transactions on software engineering, 1999.
- BACA, Murtha. *Introduction to metadata: pathways to digital information*. Los Angeles: Getty Information Institute, 1998. 41p.
- BARRETO, Cássia Maria. *Modelo de metadados para a descrição de documentos eletrônicos na web*. Instituto Militar de Engenharia, 1999.
- BELLOTTO, Heloisa Liberalli. *Arquivos permanentes: tratamento documental*. São Paulo: T. A. Queiroz, 1991. 198p.
- BORGMAN, Christine L. *From Gutenberg to the global information infrastructure: access to information in the networked world*. Cambridge: MIT Press, 2000.
- BOURDIEU, Pierre. *Contrafogos 2: por um movimento social e europeu*. Rio de Janeiro: Jorge Zahar, 2001, 115p.
- \_\_\_\_\_. *A economia das trocas simbólicas*. 5. ed. In: Miceli, S. (org). São Paulo: Perspectiva, 2001, 361p.
- BUCKLAND, Michael. *Information is think*. Journal of the American Society of Information Science, Washington, n.5, v.42, p.351-360,1991.
- BULLOCK, Aliso. *Preservation of digital information; issues and current status*. Disponível em: <<http://www.nlc-bnc.ca/publications/1/pl-259-e.html>>. Acesso em 25 mar. 2005.

- CÂNFORA, Luciano. *A biblioteca desaparecida: historias da biblioteca de Alexandria*. São Paulo: Cia. das Letras, 1989, 195p.
- CAMPOS, Arnaldo. *A Arte do livro em A Magia do papel*. Porto Alegre: Riocell/ Marprom, 1994.
- CASTELLI., Meghini, D., Pagano,P. *A Foundations of a Multidimensional Query Language for Digital Libraries*. Lectures Notes in Computer Science, 2002.
- CASTELLS, Manuel. *A sociedade em rede*. 7. ed. São Paulo: Paz e Terra, 2003. 698 p.
- CHAVES, Cláudia Maria das Graças. *Especulação e monopólio no comércio mineiro colonial: um estudo sobre mercados pré-capitalistas*. Revista Varia História, Belo Horizonte, n.16, p.181-185, 1996.
- CHEN, Su-Shing. *Digital Libraries – The life cycle of information*. Comumbia: BE Publisher, 1998.
- COLE, Douglas, BOTELHO, Tarcísio R. *Filhos de Deus: Batismos de crianças legítimas e naturais na Paróquia de Nossa Senhora do Pilar de Ouro Preto*. Revista Varia História, Belo Horizonte, n.31, p.69, 2004.
- CUNHA, Murilo Bastos da. *As tecnologias de informação e a integração das bibliotecas brasileiras*. Ciência da Informação, Brasília, v.23, n.2, 1994.
- DEMPSEY, L., & Heery, R. *A review of metadata: a survey of content resource description formats*. Retrieved, v.1, 1997.
- DIAS, Eduardo José Wense. *Contexto digital e tratamento da informação*. DataGramZero - Revista de Ciência da Informação, v.2, n.5, 2001.
- DRABENSTOTT, Karen Markey; Vizine-Goetz, Diane; Ocl. *Using subject headings for online retrieval: theory, practice, and potential*. San Diego: Academic Press, 1994.
- FLOWER, Derek Adie. *Biblioteca de Alexandria: as histórias da maior biblioteca da antigüidade*. São Paulo: Nova Alexandria, 2002, 215p.
- GROSSI, Ramon Fernandes. *Buscando a salvação da alma: um estudo sobre medo e morte nas minas da segunda metade dos setecentos*. Revista Varia História, Belo Horizonte, n.16, p.78, 1996.
- KALINICHENKO, L. A. Briukhov, D.O., Skvortsov, N.A. *Infraestrutura of the subject mediating environment aiming at semantic interoperability of heterogeneous digital library collections*. In Proceedings of the Second Russian Scientific Conference on Digital Libraries, v.09, n.01, 2003.

- LANCASTER, F. *Ameaça ou Oportunidade? O futuro dos serviços de biblioteca a luz das inovações tecnológicas*. Revistas da escola de biblioteconomia da UFMG, Belo Horizonte, v.23, n.1, 1994.
- LAHIRE, Bernard. *Crenças coletivas e desigualdades culturais*. Educação & Sociedade, São Paulo, v.34, n.84, p. 983-995, 2003.
- LESK, Michael. *Practical digital libraries: books, bytes, and bucks*. Kaufmann, 1997.
- LEVACOV, Marília. *Bibliotecas virtuais (r)evolução? Disponível em: <http://www.ibict.br/cionline/artigos/2629702.htm>. Acesso em: 01 jul. 2006*
- LEVY, Pierre. *O que é virtual?* São Paulo: Ed. 34, 1996.
- MARCHIORI, Patricia Zeni. “Ciberteca” ou biblioteca virtual: uma perspectiva de gerenciamento de recursos de informação. *Ciência da Informação*, Brasília, v.26, n.2, p.115-124, 1997.
- MARCÍLIO, Maria Luiza. *Os registros paroquiais e a história do Brasil*. . Revista Varia História, Belo Horizonte, n.31, p.13, 2004.
- MARTINS, Roberto A. *O cientista face à informação Arquivística: um exemplo da história da ciência e suas lições*. São Paulo: Universidade Estadual de Campinas, 2000.
- MÉNDEZ, Rodrigues Eva. *Metadados y recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales*. Gijón: Trea, 2002. 429 p.
- MIRANDA, Antonio; SIMEÃO, Elmira. *Ciência da informação: teoria e metodologia de uma área em expansão*. Brasília: Thesaurus, 2003. 212 p.
- MORMON, O livro de. *O Outro testamento de Jesus Cristo*. Nova York: Intellectual Reerve, 1995.
- PRADO JUNIOR, Caio. *A formação do Brasil contemporâneo*. 7. ed. São Paulo: Brasiliense, 1963. 390p.
- POZO, David Patrício Viscara. *Modelagem e Prototipação de uma Biblioteca Digital usando a abordagem 5S*. 2004. 88f. Dissertação (Mestrado em Ciência da Computação da UFMG) - Universidade Federal de Minas Gerais, 2004.
- REES A.,SARACEVIC, T. *Conceptual analysis of questions in information retrieval systems*. *Proceedings of the annual meeting of the American Document Institute*, 1963, p-175-177.
- SALLES, Fritz Teixeira de. *Vila Rica do Pilar*: Belo Horizonte: ed. Itatiaia, 1999.
- SARACEVIC, T. *On a method for studying the structure and nature of request in information retrieval*. *Proceedings of the ASIS*, v20, p.22-55, 1983.

SILVA, Vera Alice. *O Sustento financeiro da administração colonial*. Revista Varia História, Belo Horizonte, n.21, p.209-221, 2000.

Thomaz, Katia de Padua; Andrade, Maria Eugênia Albino; Universidade Federal De Minas Gerais. *A preservação de documentos eletrônicos de caráter arquivístico novos desafios, velhos problemas*. 2004. Tese Doutorado (Pós-graduação em Ciência da Informação) - Universidade Federal de Minas Gerais, 2004.

THOMAZ, K. Santos, Vilma M dos. *Metadados para o gerenciamento eletrônico de documentos de caráter arquivístico GED/A: estudo comparativo de modelos e formulação de uma proposta preliminar*. Datagramazero – Revista de ciência da Informação v4, n4. 2003.

YOURDON, Edward. *Análise estruturada moderna*. Rio de Janeiro: 1990, 836p.

WILLS, John E. *1688 O Início da era Moderna*. Rio de Janeiro: Campus. 2001.

## Anexos

### ANEXO I – Acontecimentos históricos no período escolhido

1801	Tratado de Badajós (a Espanha renuncia à posse dos Sete Povos das Missões, e Portugal confirma o direito espanhol à Colônia do Sacramento).
1806	Napoleão Bonaparte decreta o Bloqueio Continental contra a Inglaterra.
1807	Recusando-se a aderir ao Bloqueio Continental, Portugal é invadido por tropas franco-espanholas. Em novembro desse ano, a família real abandona o território português, transferindo a sede do reino para o Brasil.
1808	D. João chega ao Brasil. Pressionado pela Inglaterra, assina o decreto da abertura dos portos, rompendo com o monopólio do comércio colonial.
1810	Portugal e Inglaterra assinam um tratado de comércio, que fixa em 15% a taxa alfandegária ( <i>ad valorem</i> ) sobre produtos ingleses vendidos para o Brasil. Os demais países pagavam 24%, e Portugal 16%.
1815	Elevação do Brasil à categoria de Reino Unido à Portugal e Algarves.
1817	Explode a Revolução Pernambucana, tendo como ideal a proclamação da república e a elaboração de uma Constituição liberal. Os revoltosos ocupam o poder, por pouco tempo. A revolução foi violentamente reprimida
1820	Eclode em Portugal a Revolução do Porto, liderada pela burguesia lusitana. Os revoltosos exigem a volta de D. João VI ao país
1821	D. João VI é obrigado a deixar o Brasil, depois de tê-lo governado por mais de 12 anos. Em seu lugar fica o príncipe regente D. Pedro.
1822	No dia 7 de setembro, D. Pedro proclama a independência do Brasil. Inicia-se o período monárquico, que durou 67 anos.

### Brasil Império

1823	Instala-se, no dia 3 de maio, a Assembléia Constituinte encarregada de elaborar a primeira Constituição do Brasil. Criando um sistema eleitoral baseado na renda em farinha de mandioca, o anteprojeto ficou conhecido como a Constituição da Mandioca. Em novembro, D. Pedro I dissolve a Assembléia.
1824	É outorgada no dia 25 de março, por decreto imperial, a primeira Constituição do país. No dia 25 de julho, explode a Confederação do Equador.
1825	Os revolucionários cisplatinos assumem o controle militar da província Cisplatina.
1826	Morre, em Portugal, D. João VI. Aberta a sucessão do trono, D. Pedro I torna-se o legítimo pretendente, mas abdica seu direito em favor de sua filha D. Maria da Glória.
1828	É assinado um tratado de paz entre Brasil e Argentina, pelo qual ambos os países aceitavam a fundação de República Oriental do Uruguai.
1831	D. Pedro I abdica o trono brasileiro em favor de seu filho Pedro de Alcântara. A Regência Trina Provisória assume o poder até junho desse ano. A partir dessa data, o governo é transferido para a Regência Trina Permanente. O período regencial vai até 1840.
1832	O ministro da Justiça, padre Feijó, renuncia ao cargo.
1834	D. Pedro I morre em Portugal. É aprovado o Ato Adicional, introduzindo modificações na Constituição do império.
1835	Início da Regência Una do padre Feijó. No Pará, explode a Cabanagem; no sul, a Farroupilha.
1837	O Regente Feijó renuncia ao cargo. Tem início a Regência Una de Araújo Lima. Na

	Bahia, explode a Sabinada.
1838	A Balaiada tem início, no Maranhão.
1840	Termina o Período Regencial, com a decretação da maioria de D. Pedro II. Início do Segundo Reinado. É aprovada a lei interpretativa do Ato Adicional, limitando a autonomia das províncias.
1842	Revolta dos liberais em São Paulo e Minas Gerais.
1847	É criado o cargo de presidente do Conselho de Ministros. Introdução dos primeiros imigrantes na fazenda de café Ibicaba, em São Paulo.
1848	Explode a Revolução Praieira, a última grande revolta liberal do império.
1850	É extinto o tráfico de escravos no Brasil.

**Anexo II – Questionários**

Nome	Formação
Área de pesquisa	Entidade

Objetivo: Conhecer como o pesquisador em história percebe o uso da tecnologia em suas pesquisas.

- 1) Como você se sente utilizando recursos tecnológicos para realização de suas pesquisas
  
- 2) Qual a importância que você percebe ao utilizar a tecnologia na realização de suas pesquisas
  
- 3) Relate sucintamente alguma experiência onde a pesquisa que você desenvolve (ou desenvolveu) contou com o uso de recursos tecnológicos
  
- 4) Como você vê o uso da tecnologia no futuro de suas pesquisas
  
- 5) Quais as limitações que você percebe que inviabilizam o uso de tecnologia nas suas pesquisas

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
**ESCOLA DE CIÊNCIA DA INFORMAÇÃO**  
**Proposta de metadados para disponibilização**  
**digital do acervo de documentos da Matriz do Pilar de Ouro Preto.**

ID	Metadado	Definição	Aceito? ___S ___N	Comentários
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				

Participantes: \_\_\_\_\_, Data: \_\_\_\_\_

Tipo: (B)atismo (C)asamento (Ó)bitos



## Anexo IVa – Planilha para os leitores dos assentos – Modelo: Casamentos

<b>Data do casamento:</b>		<b>Local:</b>	
<b>Freguesia:</b>		<b>Localidade:</b>	
<b>Templo:</b>		<b>Arcebispado:</b>	<b>Horário</b>
<b>Padre:</b>		<b>Assentador:</b>	<b>Provisão:</b>
<b>Noivo:</b>		<b>Idade:</b>	
<b>Cor:</b>		<b>Condição social:</b>	
<b>Legitimidade:</b>		<b>Nome do dono:</b>	
<b>Profissão:</b>		<b>Nome do pai:</b>	
<b>Profissão do pai:</b>		<b>Nome da mãe:</b>	
<b>Pais de origem:</b>		<b>Localidade onde mora:</b>	
<b>Freguesia de origem:</b>		<b>Freguesia de batismo:</b>	
<b>Templo de batismo:</b>		<b>Igreja onde é paroquiano:</b>	<b>Nome da ex-mulher</b>
<b>Noiva:</b>		<b>Idade:</b>	
<b>Cor:</b>		<b>Condição social:</b>	
<b>Legitimidade:</b>		<b>Nome do dono:</b>	<b>Nome do pai:</b>
<b>Profissão do pai:</b>		<b>Nome da mãe:</b>	
<b>Pais de origem:</b>		<b>Localidade onde mora:</b>	
<b>Freguesia de origem:</b>		<b>Freguesia de batismo:</b>	
<b>Templo de batismo:</b>		<b>Igreja onde é paroquiana:</b>	<b>Nome do ex-marido</b>
<b>Nome da testemunha1:</b>	<b>Sexo:</b>	<b>Cor:</b>	<b>Condição social:</b>
<b>Residência:</b>	<b>Profissão:</b>		<b>Nome do dono:</b>
<b>Nome da testemunha2:</b>	<b>Sexo:</b>	<b>Cor:</b>	<b>Condição social:</b>
<b>Residência:</b>	<b>Profissão:</b>		<b>Nome do dono:</b>
<b>Impedimentos:</b>		<b>Arquivo digital:</b>	<b>Referência do doc. Original (livro/pág):</b>
<b>Observações do assento:</b>	<b>Observações de leitura (dúvidas, conclusões, situações da leitura que acha interessante relatar)</b>		<b>Ref. Do microfilme (Rolo/Volume):</b>

## Anexo IVb – Planilha para os leitores dos registros – Batismos

<b>Data do batismo:</b>		<b>Local:</b>	
<b>Templo:</b>		<b>Freguesia:</b>	<b>In-Extremis? S N</b>
<b>Celebrante</b>		<b>Alforriado? S N</b>	
<b>Nome do batizando</b>		<b>Idade:</b>	<b>Legitimidade:</b>
<b>Sexo:</b>	<b>Exposto para:</b>	<b>Exposto em:</b>	<b>Mãe</b>
<b>Cor da mãe:</b>	<b>Mãe viúva? S N</b>	<b>Condição da mãe:</b>	
<b>Nome do pai:</b>		<b>Pai viúvo? S N</b>	
<b>Cor do Pai:</b>		<b>Condição social do Pai:</b>	
<b>Profissão do Pai:</b>		<b>Freguesia dos pais:</b>	
<b>Recebeu os santos óleos? S N</b>			
<b>Padrinho:</b>		<b>Localidade:</b>	
<b>Freguesia:</b>		<b>Estado civil:</b>	
<b>Profissão:</b>		<b>Padrinho por procuração? S N</b>	
<b>Madrinha:</b>		<b>Estado civil:</b>	
<b>Freguesia:</b>		<b>Localidade:</b>	
<b>Madrinha por procuração? S N</b>			
<b>Arquivo digital:</b>		<b>Referência do doc. Original (livro/pág):</b>	<b>Ref. Do microfilme (Rolo/Volume):</b>
<b>Observações do assento:</b>		<b>Observações de leitura (dúvidas, conclusões, situações da leitura que acha interessante relatar)</b>	

### Anexo IVc – Planilha para os leitores dos registros - Óbitos

<b>Data do batismo:</b>		<b>Local:</b>	
<b>Freguesia:</b>		<b>Celebrante:</b>	
<b>Assentador:</b>			
<b>Nome do falecido(a):</b>		<b>Causa da morte:</b>	
<b>Idade:</b>	<b>Cor:</b>	<b>Sexo:</b>	
<b>Nome do Pai:</b>		<b>Nome da mãe:</b>	
<b>Legitimidade:</b>		<b>Estado civil:</b>	
<b>Nome do conjugue:</b>		<b>Condição social:</b>	
<b>Profissão:</b>		<b>Local de residência:</b>	
<b>Irmadade:</b>		<b>Nome do dono:</b>	
<b>Tumba/caixão:</b>		<b>Mortalha:</b>	
<b>Missa:</b>		<b>Sacramentos:</b>	
<b>Ofício: S N</b>		<b>Testamento: S N</b>	
<b>In-extremis: S N</b>		<b>Boçal: S N</b>	
<b>Acompanhamento:</b>			
<b>Arquivo digital:</b>		<b>Referência do doc. Original (livro/pág):</b>	<b>Ref. Do microfilme (Rolo/Volume):</b>
<b>Observações do assento:</b>		<b>Observações de leitura (dúvidas, conclusões, situações da leitura que acha interessante relatar)</b>	









## Anexo VI – Rolos de microfilme utilizados durante a leitura paleográfica

Os microfilmes que compõem a documentação da Matriz do Pilar, estão atualmente sob a guarda da casa dos contos em Ouro Preto. O acesso a eles é bastante restrito, podendo cada pesquisador ter acesso a apenas um rolo de microfilme por vez (caso precise retirá-lo da casa dos contos). O acesso aos rolos também podem ocorrer na casa dos contos a qual possui estrutura para que os pesquisadores leiam o conteúdo dos microfilmes.

A seguir destacamos os períodos e número de rolos que correspondem à documentação do período escolhido (1800-1850).

### Batismos

Época	Volume	Rolo	Microfilme
1803	0501	029	0379-0393
1821	0502	029	0394-0405
1786 a 1865	0497	104	0324-0604
1834 a 1852	0498	104	0605-0762
1789 a 1808	0494	028	0511-0980
1772 a 1838	0495	104	0035-0323
1781 a 1889	0496	028	0981-1041
1781 a 1889	0496	029	0001-0083

### Casamentos

1753 a 1804	0507	105	0065-0307
1804 a 1833	0508	105	0308-0452
1834 a 1865	0509	105	0453-0597

### Óbitos

1786 a 1888	1854	106	0558-0692
1802 a 1823	1855	106	0693-0773
1802 a 1823	1855	107	0001-0162
1824 a 1844	1856	107	0163-0316
1844 a 1860	1857	107	0317-0425

Com isso conseguimos identificar todos os rolos que serão necessários para nosso trabalho:

Batismos : 029 – 028 – 104  
 Casamentos : 105  
 Óbitos : 106 – 107