

ESTUDO DE PADRONIZAÇÃO DE METADADOS PARA PRESERVAÇÃO DA MEMÓRIA DOCUMENTAL DA BIBLIOTECA DIGITAL DO INPE

Resumo: O uso correto de bibliotecas digitais para disseminação da informação bem como a preservação da memória, são temas atualmente discutidos na área da ciência da informação. Este artigo apresenta um estudo de padronização de metadados usados na Biblioteca Digital do Instituto Nacional de Pesquisas Espaciais (INPE), para preservação da memória documental a fim de melhor descrever os documentos, contextualizando-os no ambiente do INPE como um todo.

Palavras-chave: biblioteca digital, memória documental, produção científica, padronização de metadados.

Abstract: The correct use of digital libraries for information dissemination as well as the memory preservation, is subjects currently argued area of the information science. This paper presents a study of metadados standardization used in the National Institute of the Space Research (INPE) Digital Library, for preservation of the documentary memory in order better to describe documents, in the context the INPE environment INPE as a whole.

Key-words: digital library, documental memory, scientific production, metadata standardization.

1. Introdução

O avanço tecnológico e seu impacto na sociedade fizeram com que serviços de informação reavaliassem a forma de disseminação da informação atendendo às necessidades de seus usuários, particularmente à comunidade científica. Além da disseminação da informação, há outro fator de interesse na preservação dos documentos produzidos, sejam eles de natureza científica, técnica, administrativa ou de divulgação. Desta forma, a área da ciência da informação vem abordando temas como a preservação da memória e o desenvolvimento de bibliotecas digitais.

De acordo com Sayão (1996), quando um pesquisador, diante de um microcomputador ligado a um banco de dados, que pode estar em qualquer parte do mundo, vasculha suas estantes eletrônicas à procura de informações que definam, completem ou estabeleçam as fronteiras do seu trabalho de pesquisa, ele repete o mesmo gesto de quem mergulha na memória de seu grupo para reconstruir as lembranças comuns e, dessa forma, manter íntegra a sua comunidade.

A preservação da memória é vista com interesse pela comunidade brasileira e abordada como fundamento essencial para preservação cultural e intelectual, seja de uma ou de um grupo de pessoas, uma instituição ou de uma região territorial. Por sua característica e valor, a memória pode ser vista como patrimônio intelectual, a qual merece ser preservada, registrada e disseminada como uma herança do conhecimento.

De acordo com Delvizio (2004), a Constituição Brasileira de 1998 refere-se ao patrimônio cultural como compreensão de duas classificações: a primeira referindo-se ao patrimônio de valor imaterial ou intangível, cujas características relacionam-se aos modos de criar, fazer e viver, encerrando na compreensão de seu significado subjetivo a “memória cultural” (religião, culinária, vestuário, costume, etc); a segunda refere-se aos bens de valor

tangível, que correspondem ao produto concreto das ações dos seres humanos, englobando objetos, artefatos e construções.

Para Granato (2003), os bens culturais são o produto e o testemunho das diferentes tradições e realizações intelectuais do passado e constituem um elemento essencial da personalidade dos povos. Reconhecendo essa importância, faz-se mister, segundo o autor, transmitir da melhor maneira possível esse patrimônio cultural às gerações futuras.

Nesse sentido, uma alternativa para preservar e disseminar a memória, ou seja, o “patrimônio cultural e científico tangível” produzido por um indivíduo ou um grupo particular, está na criação e no desenvolvimento de bibliotecas digitais que atendam à toda essa produção correlacionando todo contexto existente entre os diversos tipos de documentos.

A biblioteca digital é necessária ao desenvolvimento da pesquisa, modificando o modo que o pesquisador dissemina a informação de sua produção, tanto para os pares quanto para o público em geral. Além disso, as facilidades de disseminação e as características da biblioteca digital constituem uma justificativa concreta para o registro da produção científica, hoje no País, que se encontra em constante evolução.

Desta forma, a biblioteca digital tem se destacado por sua facilidade na disseminação, no acesso à informação, na divulgação da produção institucional e na agilização da transferência da informação. Muitas iniciativas já foram desenvolvidas para a disseminação de teses e dissertações e para publicações científicas e arquivísticas.

Uma destas iniciativas é a do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), que coordena o projeto da Biblioteca Digital de Teses e Dissertações (BDTD), buscando integrar os sistemas de informação de teses e dissertações existentes nas Instituições de Ensino Superior (IES) e os institutos de pesquisa brasileiros, bem como estimular o registro e a publicação de teses e dissertações em meio eletrônico.

Segundo o IBICT (2005), a BDTD adota um modelo distribuído utilizando-se das tecnologias de arquivos abertos. As IES e os institutos de pesquisa são provedores de dados e o IBICT opera nesta rede como provedor de serviços e agregador, coletando metadados de teses e dissertações destes provedores de dados, promovendo serviços de informação sobre esses metadados e expondo-os para serem coletados por outros provedores de serviços.

Costa (2004) afirma que, para uma biblioteca digital, é necessário combinar um sistema de representação da informação altamente padronizado (os metadados) com um sistema de armazenamento dos textos completos. Através da *web* os computadores se complementam no ato da pesquisa e fornecem ao usuário uma resposta completa.

A partir deste contexto, e, considerando a utilização correta da biblioteca digital para preservação da memória técnico-científica, este artigo tem como proposta apresentar um estudo de padronização dos metadados usados na Biblioteca Digital da Memória Técnico-científica do Instituto Nacional de Pesquisas Espaciais (INPE), para materiais como clippings, artefatos tridimensionais e materiais gráficos, a fim de melhor definir os campos de descrição de cada tipo de documento, facilitar a recuperação e a troca de informação e,

de modo especial, tornar possível a descrição de diferentes documentos, contextualizando-os no ambiente do INPE como um todo.

2. Padronização para Bibliotecas Digitais

No desenvolvimento de uma biblioteca digital, a padronização é o processo essencial para caracterização e preservação dos documentos, a fim de garantir que qualquer pessoa ou qualquer máquina possa ter acesso à visualização e leitura do documento requerido. Segundo Lorist e Meer (2001), os padrões garantem a interoperabilidade, o consenso interdisciplinar em conceitos, técnicas e procedimentos. Para preservação de documentos digitais, os autores destacam alguns padrões a serem definidos:

- Padrões para servir como modelo de referência (arquitetura), descrevendo a funcionalidade e o comportamento de uma Biblioteca Digital, procedimentos e conceitos.
- Padrões para preservar o formato do documento digital.
- Padrões de metadados para preservar o acesso ao conteúdo, descrevendo o contexto técnico, procedência e semântica, permitindo a interpretação futura do documento.
- Padrões para interoperabilidade.

De acordo com Chaves (2002), entende-se por padrões um conjunto de regras ou especificações para projeto ou operação de um dispositivo computacional. Através da adoção de padrões, segundo o autor, é possível separar conteúdos próprios de cada domínio do conhecimento, facilitar a busca por tecnologias específicas e organizar as informações de forma estruturada, tanto na Web como nas bibliotecas digitais. Os documentos disponíveis nas bibliotecas digitais devem estar codificados em formatos e linguagens de marcação padronizados, de modo a facilitar o intercâmbio, o acesso e a manutenção das informações.

Para a padronização da arquitetura da biblioteca, deve-se viabilizar a troca de informações e de serviços, considerando, segundo Chaves (2002), os relacionamentos entre materiais digitais, versões, direitos e permissões dos autores, redes e sistemas de computadores e os componentes.

A identificação dos padrões para formato do documento deve ser definida, pensando sempre na integridade do documento original. Segundo Lorist e Meer (2001), padrões para conteúdo deveriam retardar o processo de envelhecimento da recuperação física e semântica do documento que está sendo preservado. Formatos como *Portable Document Format* (PDF) e o *Extensive Markup Language* (XML) são freqüentemente utilizados para documento em bibliotecas digitais.

A interoperabilidade é um dos focos importantes a ser discutido, pois compartilha informações e registros entre bibliotecas digitais com diferentes arquiteturas, formatos e tecnologias. De acordo com Marcondes e Sayão (2002), a possibilidade que os acervos digitais possam ser consultados simultaneamente, sem que um usuário acesse cada site individualmente é chamada de *interoperabilidade*. Existem diversos protocolos para a interoperabilidade, entre os quais destacam-se o *Open Archives Initiative* (OAI), o *Dublin Metadata Core Element Set* (Dublin Core) e o MARC. .

A padronização de metadados é um também tópico a ser definido, visto que uma descrição detalhada do documento, de acordo com suas características, auxilia na contextualização do mesmo dentro de seu universo institucional. Arellano (2004) ressalta afirmando que é preciso chamar a atenção para a importância de informar o contexto do objeto digital a ser registrado (e preservado) para que os futuros usuários possam entender o ambiente tecnológico no qual ele foi criado.

Segundo Chaves (2002), o conceito de metadados quando aplicado ao contexto de bibliotecas, digitais ou tradicionais, normalmente refere-se a informações que produzem uma (geralmente breve) categorização do objeto de informações nas coleções de uma biblioteca, armazenam os conteúdos dos catálogos digitais como em bibliotecas tradicionais, e são usadas para ajudar os usuários a acessarem objetos de informação de interesse.

Para Lorist e Meer (2001) a definição de padrões necessita de uma descrição do objeto digital ou de várias partes que compõem um único objeto. Assim, os metadados deveriam incluir a informação contextual que é necessária para gerenciar, recuperar e interpretar a informação eletrônica. Para os autores, há diversos esquemas de metadados que devem ser padronizados, cada um com um objetivo específico e domínio de aplicação.

A padronização de metadados pode ser definida através da utilização de códigos e normas que melhor descrevam o documento e seu contexto. Normas e códigos como da Associação Brasileira de Normas Técnicas (ABNT), Z39.50, Norma Geral Internacional de Descrição Arquivística (ISAD) e Código de Catalogação Anglo-Americano (AACR2) podem ser utilizados em conjunto, garantindo a eficácia na descrição do metadado definido.

Considerando a importância da padronização para bibliotecas digitais, o item a seguir apresenta algumas características do INPE, como sua Biblioteca Digital está estruturada e o estudo para padronização de metadados para o acervo documental do Instituto.

3. A Memória do INPE

O INPE tem como principal finalidade "promover e executar estudos, pesquisas científicas, desenvolvimento tecnológico e capacitação de recursos humanos, nos campos da Ciência Espacial e da Atmosfera, das Aplicações Espaciais, da Meteorologia e, da Engenharia e Tecnologia Espacial, bem como em domínios correlatos da política definida pelo Ministério da Ciência e Tecnologia (MCT)". Também tem como missão a implementação de uma política industrial envolvendo a transferência de tecnologia ao setor produtivo e o fomento da capacitação da indústria espacial.

Como resultado desta finalidade, o INPE é reconhecido internacionalmente pelos trabalhos desenvolvidos em suas áreas de competência. Segundo Miranda (2002), o conhecimento acumulado pelo Instituto ao longo de sua história constitui-se no grande diferencial. O efetivo investimento na formação de profissionais, doutores e mestres, em níveis compatíveis com a formação encontrada em países de primeiro mundo, possibilitou ao INPE atingir um alto grau de maturidade em suas pesquisas.

Assim, ao longo dos 44 anos de existência da pesquisa espacial no Brasil, o INPE gerou inúmeros documentos, tais como relatórios técnicos, manuais, notas técnicas, artigos de periódicos e trabalhos apresentados em eventos internacionais e nacionais, capítulos de livros e livros, teses e dissertações que somam mais de 16.000 publicações, as quais encontram-se publicadas em múltiplos meios de disseminação. O conjunto destas publicações forma a Memória Técnico-Científica do INPE.

Para a montagem e a disponibilização dos acervos eletrônicos da Biblioteca Digital, o INPE utiliza desde 1995, o software *URLibService*. Este software permite, em particular, a disponibilização de texto completo através do protocolo http e garante a persistência de vínculos entre documentos depositados em acervos distintos.

As características do *URLibService* fizeram dele uma plataforma adequada para receber a Memória Técnico-Científica e Documental do Instituto. O trabalho consolida-se com a disponibilização do acervo da Biblioteca Digital do INPE à comunidade nacional e internacional, provendo mais um mecanismo de difusão da informação.

3.1 O *URLibService*

O *URLibService* é um software criado para servir na montagem e manutenção de uma biblioteca digital com acervos distribuídos seguindo o paradigma de repositórios uniformes para uma biblioteca (Banon e Banon, 2005). Na Biblioteca Digital, cada repositório contém o nome do acervo local que hospeda a obra original. Por sua vez, os repositórios criados num determinado acervo local podem ser copiados, sem atropelo, para qualquer outro acervo local. Desta forma, mesmo fazendo cópias de um repositório para outros acervos locais, sabe-se qual é o acervo local que possui o original (Banon, Ribeiro e Banon, 2004).

A definição dos padrões para Biblioteca Digital, faz com que o *URLibService* possa ser, segundo Banon e Banon (2005), utilizado para:

- Armazenar os documentos em qualquer formato de arquivo (HTML, PDF, DOC, etc);
- Achar os documentos na Web através de palavras chaves;
- Emitir relatórios de pesquisa, tabelas, sumários e índices por autores;
- Facilitar o download dos documentos;
- Compartilhar o acervo local com outros;
- Instalar uma copia de um documento proveniente de um outro acervo local;
- Controlar as versões dos documentos;
- Localizar na Web a versão original de um documento;
- Registrar os repositórios em base de dados independentes;
- Armazenar e exibir as estatísticas de acesso de cada documento;
- Permitir vínculos relativos entre documentos, tornando o acesso persistente;
- Garantir a robusteza dos vínculos em relação a mudanças de estrutura do acervo;
- Permitir vínculos sensíveis à preferência de idioma do usuário;
- Identificar automaticamente os documentos em construção;
- Exibir os metadados dos documentos no formato BibTeX ou Refer (EndNote);
- Disponibilizar os metadados dos documentos através do protocolo OAI (Open Archives Initiative) para a coleta de metadados;

Segundo os autores, o *URLibService* permite a construção de hipertextos estáveis cujas partes podem ser copiadas ou mudadas de um site da Biblioteca Digital para outro sem a necessidade de modificar os vínculos ou de recorrer a um serviço global de nome. Assim, um dos destaques do *URLibService* é oferecer uma solução interessante para o problema dos **vínculos de citações**. Qualquer que seja o local físico onde se encontra um documento no acervo da Biblioteca Digital seu acesso é persistente. Outro fator importante é o sistema de revisão on-line e edição automática para anais de eventos, que vem sendo utilizado com sucesso nos últimos anos. O *URLibService*, inclui em seus padrões de formato, o *Dublin Core*, *BibTeX*, *Refer*, *xrefer*, além do formato *mtd-br* para disponibilização de suas teses e dissertações na Biblioteca Digital de Teses e Dissertações do IBICT e de padrão bibliográfico próprio da Instituição denominado *BibINPE*.

Desta forma, destacam-se a versatilidade e multifuncionalidade do software em diversas atuações, confirmando seu papel de gerenciador do conhecimento técnico-científico.

Atualmente, na Biblioteca Digital, é registrada toda a publicação técnico-científica do INPE: teses, dissertações, livros, capítulos de livros, artigos apresentados em eventos e publicados em revistas, além de materiais audiovisuais, relatórios, miscelâneas, patentes, programas de computador e clippings. O total de documentos disponíveis soma 16.517, sendo que 4.590 destas referências possuem vínculo para texto completo. O Quadro 1 apresenta os números de documentos inseridos na biblioteca digital até setembro de 2005, de acordo com o tipo de publicação:

Tipo de Documento	Total
Artigos apresentados em eventos	7.571
Artigos publicados em revistas	3.923
Capítulos de Livros	379
Clippings	339
E-print	239
Livros	57
Livros editados	38
Manuais técnicos	48
Material audiovisual	22
Material didático	142
Miscelânea	901
Notas Técnicas	375
Patentes	7
Programas de computador	29
Relatórios	1059
Teses e Dissertações	1.388

Quadro 1 – Número de registros disponíveis na *URLib* até setembro de 2005.

Fonte: INPE. SID (2005)

A Biblioteca Digital está sendo preparada para a inclusão do registro de novos tipos de documentos como materiais gráficos e objetos tridimensionais, além da inclusão de novos campos para documentos existentes, visto que esta já possui uma relação de campos definidos os quais podem ser utilizados a partir do estudo de padronização dos metadados, como mostra o Anexo A.

A padronização de metadados para os documentos técnico-científicos estão consolidados em normas que descrevem estes documentos de acordo com sua finalidade. Porém, além dos documentos técnico-científicos, a Biblioteca Digital possui canais para inclusão de documentos como: clippings (*Newspaper article*) e fotografias (*Image*) que já possuem uma definição prévia para estes metadados.

Para iniciar o trabalho de registro, depósito de materiais gráficos e objetos tridimensionais além da revisão dos registros de clippings e fotografias, foi realizado um estudo para padronização de metadados, levando a inclusão de novos campos em formulário existente ou até a criação de novos formulários, de acordo com a produção documental do Instituto. Parte deste estudo é descrito no item a seguir.

4. O Estudo

O registro de documentos como clippings, materiais gráficos, audiovisuais e tridimensionais necessitam de códigos e normas que descrevam o documento em si (como o Código de Catalogação Anglo-Americano – AACR2 e a ABNT-NBR 6023), e também de normas que descrevem detalhadamente o documento e o contextualize dentro de seu espaço e proposta de criação. Por isso, realizou-se um estudo baseado na Norma Geral Internacional de Descrição Arquivística (ISAD), que é uma norma que define diretrizes para a preparação de descrições arquivísticas cujo principal objetivo é identificar e explicar o contexto e o conteúdo de documentos de arquivo e promover o acesso aos mesmos.

De acordo com o Conselho Internacional de Arquivos (2000) as regras do ISAD para descrição arquivística podem ser aplicadas independentemente da forma ou do suporte dos documentos. O Conselho ainda ressalta que o ISAD pode ser utilizado em combinação com manuais que expõem regras de descrição para documentos, possibilitando, assim, uma adequada descrição de documentos especiais.

O estudo para padronização dos metadados foi baseado na norma ISAD comparada com os capítulos referentes à materiais gráficos, objetos tridimensionais e publicações seriadas do AACR2, que é um código para descrição de documentos para catalogação. Deste estudo, foi desenvolvido um quadro comparativo das áreas do AACR2, dos campos descritos pelo ISAD e dos campos pré-definidos na Biblioteca Digital para descrição de clippings e fotografias.

Para documentos de materiais gráficos e objetos tridimensionais, o quadro apresentou apenas o estudo comparativo do ISAD com o AACR2 a fim de definir os campos na Biblioteca Digital para inclusão destes documentos.

Como exemplo deste estudo é apresentado no Quadro 2 a análise dos campos para fotografias. O quadro compara o capítulo do AACR2 referente a materiais gráficos (que contempla a descrição de fotografias), o campos do ISAD e a descrição da Biblioteca Digital (*Image*):

Quadro 2 Parte 1 – Estudo comparativo dos campos de descrição para materiais gráficos.

Itens	AACR2 – Materiais Gráficos	ISAD	Biblioteca Digital – URLib
Código de identificação		<u>Código de referência do material</u> (código do país / detentor / referência local / número de controle ou identificador único)	Nome do repositório
Título	Título principal. Se um item não tiver um título, forneça, assim como a uma coleção de itens gráficos, usando um título pelo qual é conhecido ou que indique a sua natureza.	Título formal ou título atribuído (que vai depender da hierarquia – fundo, série, sub-série, dossiê ou item documental)	Título
Data		Data da unidade de descrição. Registre uma data única ou datas-limite.	Data de criação ou atualiz. do documento
Níveis		Fundo / Série / Subsérie / Dossiê/processo / Item documental	Tipo
Designação geral do material	Dado opcional: indica o tipo de material (ilustração, diafilme, diapositivo...) Não se registram estes elementos para um item gráfico não publicado e nem para uma coleção não publicada destes itens, mesmo que a coleção inclua itens publicados.		Tipo secundário
Área de publicação, distribuição, etc			
Descrição física / Dimensão e suporte	Extensão do item (número de unidades físicas de um item gráfico em números arábicos) acompanhado da designação específica do material. a) Outros detalhes físicos (som, cor,...) b) Dimensões	<u>Dimensão e suporte</u> : Dimensão física ou lógica e suporte da unidade de descrição (algarismos arábicos). Ex: 134 rolos de microfilme, 35mm	Formato
Indicação de Responsabilidade / Nome do Produtor	<u>Indicação de responsabilidade</u> : pessoas ou entidades que aparecem creditadas como participantes maiores na criação ou produção de um item gráfico, como diretores, produtores, artistas, desenhistas.	<u>Nome do produtor</u> : organização ou pessoa responsável pela produção, acumulação e manutenção dos documentos	Autor(es) Crédito

Continua

Quadro 2 Parte 2 – Estudo comparativo dos campos de descrição para materiais gráficos.

Itens	AACR2 – Materiais Gráficos	ISAD	Biblioteca Digital – URLib
Biografia	Área de notas: Edição e história . Resumo	<u>Biografia</u> : História administrativa ou dados biográficos do produtor do documento para contextualizar o acervo e torná-lo mais compreensível.	-Projeto -Área (do conhecimento) -Endereço de e-mail
História Arquivística		Informação sobre a história da unidade de descrição que seja significativa para sua autenticidade, integridade e interpretação (transferências de propriedade, custódia).	
Procedência	Notas: Proprietários anteriores	<u>Procedência</u> : Origem imediata de aquisição ou transferência da unidade de descrição.	Base
Âmbito e Conteúdo	Notas: Natureza ou forma artística Resumo	<u>Âmbito e conteúdo</u> : sumário do âmbito (período de tempo e localização geográfica) e conteúdo (como formas de documentos, assuntos, processos administrativos). Descrição do fundo ou série	Tema Resumo Palavras-chave
Avaliação, eliminação e temporalidade		<u>Informações</u> sobre a avaliação, seleção e eliminação ocorridas ou planejadas	
Incorporações		<u>Acréscimos</u> previstos à unidade de descrição.	
Sistema de arranjo		<u>Estrutura</u> interna, ordem e/ou sistema de arranjo da unidade de descrição.	
Condições de acessos		<u>Estatuto</u> legal ou outros regulamentos que restrinjam ou afetem o acesso à unidade de descrição.	Permissão de leitura
Condições de reprodução		Indicar quaisquer restrições quanto à reprodução da unidade de descrição.	
Idioma		<u>Idioma</u> , escrita ou sistema de símbolos utilizados na unidade de descrição.	Idioma
Características físicas e requisitos técnicos		<u>Condições</u> físicas importantes, como requisitos de preservação, que afetem o uso da unidade de descrição.	

Continua.

Quadro 2 Parte 3 – Estudo comparativo dos campos de descrição para materiais gráficos.

Itens	AACR2 – Materiais Gráficos	ISAD	Biblioteca Digital – URLib
Instrumentos de pesquisa	Notas: Material adicional (materiais gráficos) Material adicional (artefatos tridimensionais ou reália)	Instrumentos de pesquisa que a retentora ou o produtor possa dispor, que forneça informação relativa ao contexto e conteúdo.	
Existência e localização dos originais	Notas: Proprietários Nota relativa ao original	Existência e localização dos originais: Existência, localização, disponibilidade e/ou destruição dos originais	Detentor da cópia
Existência de localização de cópias		Localização de cópias disponíveis da unidade de descrição.	
Unidades de descrição relacionadas		Unidades de descrição existentes na mesma instituição ou qualquer outro lugar que sejam relacionadas por proveniência.	
Notas de Publicação		Referência ou informação sobre publicações que sejam sobre ou baseadas no uso, estudo ou análise da unidade de descrição.	
Notas	Informações que não possam ser incluídas em outras áreas / Público a que se destina Exemplar que está sendo descrito	Informações que não possam ser incluídas em qualquer outra área. Indicar se o título foi formal ou atribuído	Notas
Nota do arquivista		Como a descrição foi preparada e por quem e suas revisões.	Usuário avançado
Regras ou convenções		Regras ou convenções internacionais, nacionais e/ou locais seguidas na preparação da descrição.	
Data da descrição		Data em que a descrição foi preparada e/ou revisada	
URL			URL
Grupo			Grupo
Afiliação			Afiliação
Data de acesso			Data de acesso
Nome do arquivo			Nome do arquivos
Senha			Senha

Conclusão. Fontes: Ribeiro (2002), Conselho Internacional de Arquivos (2000) e INPE. SDI (2005).

A partir da análise do quadro, pôde-se perceber que os campos de inclusão das informações dentro da Biblioteca Digital estavam incompletos, não atendendo todas as necessidades de descrição dos documentos, de acordo com a característica e o objetivo de contextualização no processo de produção intelectual e histórico do INPE. Sendo assim, foi realizada uma avaliação de cada campo de descrição e um estudo sobre a definição de cada tipo de documento, sua natureza, suporte e forma de divulgação.

A partir desta análise, foi criada, para documentos de clippings e fotografias uma lista de sugestões para inclusão de novos campos e a descrição para novos tipos na Biblioteca Digital, ou seja, materiais gráficos e objetos tridimensionais.

Como exemplo, o Quadro 3 apresenta os campos sugeridos para materiais fotográficos, que após o estudo comparativo, deveriam ser separadas de materiais gráficos (devido à sua característica e funcionalidade).

Proposta para Fotografias
1. Separar metadados distintos para fotos e materiais gráficos (apesar da foto ser considerada um material gráfico). Isto facilita na descrição específica da foto e auxilia na busca do material
2. Código de identificação
3. Descrição física (dimensão e suporte)
4. Cor
5. Utilizar o campo “Resumo” para descrição biográfica.
6. Procedência
7. Âmbito e conteúdo: sumário do âmbito (como o período de tempo, localização geográfica) e conteúdo (como formas de documentos, assuntos, processos administrativos)
8. Características físicas e condições de acesso (dados de conservação da foto e requisitos de preservação que afetem o uso)
9. Instrumentos de pesquisa: instrumentos de pesquisa que a instituição retentora ou o produtor do documento possam dispor, que forneçam a informação relativa ao contexto e conteúdo
10. Existência e localização de originais
11. Existência e localização de cópias
12. Unidades de descrição relacionadas (outros documentos e tipos de documentos existentes na instituição ou Qualquer outro lugar que sejam relacionadas por proveniência)
13. Retirar o campo Idioma do formulário
14. Notas de publicação (referência ou informação sobre publicações que sejam sobre ou baseadas no uso, estudo ou análise da foto).
15. Resolução da foto (em DPI)
16. Orientação da foto (retrato ou paisagem)

Quadro 3 – Proposta de inclusão de campos para categoria Fotografias

A sugestão de inclusão dos novos campos para clippings já foi avaliada pelo grupo de estudo de padronização e de desenvolvimento da Biblioteca Digital. Após esta análise, foram efetuadas as novas modificações para esta categoria, facilitando o trabalho de inclusão e permitindo a contextualização das notícias divulgadas na imprensa nos projetos e trabalhos de pesquisa desenvolvidas pelo Instituto. Vale ressaltar que alguns dos novos campos definidos para Clippings estão direcionados para apresentação de indicadores desta projeção das atividades do INPE na mídia.

5. Considerações Finais

A padronização de bibliotecas digitais é essencial para preservação digital de documentos. Na padronização de metadados, faz-se necessário estudar e identificar normas e códigos que possam servir para descrição de campos de acordo com a necessidade da instituição e os objetivos de criação do documento. A partir da definição dos campos a serem incluídos consegue-se não somente descrever o documento, mas também transformá-lo num elemento de construção do conhecimento e da história da instituição.

Vale ressaltar que o estudo do ISAD para inclusão dos campos de descrição para documentos referentes à Memória Documental foi de fundamental valor para o aprofundamento e melhor visibilidade da natureza e descrição de documentos com características arquivísticas.

A próxima etapa do projeto deste trabalho é a reestruturação dos novos campos para fotografias e a inclusão dos documentos de materiais gráficos e objetos tridimensionais, além da revisão de todos os registros já inseridos na Biblioteca Digital para Clippings.

Além disso, sentiu-se a necessidade de criar um vocabulário controlado para documentos como clippings (utiliza de uma linguagem jornalística); e materiais gráficos e objetos tridimensionais (que possuem, ao mesmo tempo, uma linguagem visual e subjetiva). Este vocabulário utilizar-se-á do Thesaurus da NASA, que é utilizado para os descritores de documentos técnico-científicos, além de um vocabulário específico, identificando novos descritores a partir dos artigos na mídia e da indexação de imagens. Desta forma o INPE contará com uma terminologia específica e padronizada, construindo assim sua taxonomia.

Destaca-se também a preservação do documento físico, uma vez que o acesso a toda publicação da instituição será feita de maneira digital. Para esta preservação, foi criado um grupo de preservação e conservação que está sendo capacitado para exercer as atividades de limpeza, acondicionamento e manutenção do documento e do ambiente.

Um outro fator que merece ser ressaltado é a constituição de uma equipe multidisciplinar para a realização deste trabalho, contando com o apoio e colaboração de vários profissionais de diversas áreas do conhecimento. Sem esta equipe, este trabalho não seria eficiente em seus processos e nem eficaz em seus resultados.

Espera-se, com este trabalho, iniciar a preservação da memória documental do INPE, bem como seu acesso a toda comunidade. A utilização da Biblioteca Digital para este trabalho é essencial para que toda produção intelectual do INPE seja acessível a toda comunidade nacional e internacional.

Referências bibliográficas

ARELLANO, M.A. Preservação de documentos digitais. **Ciência da Informação**, v.33, n.2, p.15-27, maio/ago. 2004.

BANON, .G.J.F.; BANON, L.C. O que é a URLib?. Disponível em:
<<http://iris.sid.inpe.br:1905>>. Acesso em 07 de julho 2005.

BANON, G. J. F.; RIBEIRO, M. L.; BANON, L. G. Preservação digital da memória técnico-científica do INPE. In: SIMPÓSIO INTERNACIONAL DE BIBLIOTECAS DIGITAIS, 2., 2004, Campinas. **Anais eletrônicos...** Disponível em: <<http://libdigi.unicamp.br/document/?view=8264>>

CHAVES, M.S. **Padrões em bibliotecas digitais**. 2002. Dissertação (Mestrado em Ciência da Computação) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2002.

CONSELHO INTERNACIONAL DE ARQUIVOS. **ISAD(G)**: norma geral internacional de descrição arquivística. 2.ed. Rio de Janeiro: Arquivo Nacional, 2000. 119p.

COSTA, M.V.S.B.; et al. Estratégias de implantação da biblioteca digital de teses e dissertações da UNIRIO. In: Simpósio Nacional de Bibliotecas Universitárias, 12., 2004, Natal. **Anais...** Natal: Universidade Federal do Rio Grande do Norte, 2004.

DELVIZIO, J.B.U. Patrimônio tangível e memória: a arquitetura moderna na cidade de Corumbá MS como vetor do desenvolvimento local. Disponível em: <http://www.pantanal2002.ucdb.br/eixos/eixo04/e4_20.pdf>. Acesso em 07 de julho 2004.

GRANATO, M. **Restauração de instrumentos científicos históricos**. 2003. Tese (Doutorado em Engenharia metalúrgica de materiais) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2003.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Sobre a BDTD. Disponível em: <<http://bdtb.ibict.br/bdtb/utilitarios/sobre/sobre.jsp>>. Acesso em 12 de julho 2005.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. SERVIÇO DE INFORMAÇÃO E DOCUMENTAÇÃO. Memória técnico científica do INPE. Disponível em: <<http://iris.sid.inpe.br:1905/col/sid.inpe.br/banon/2001/04.03.15.36.19/doc/mirror.cgi>>. Acesso em 04 de julho 2005.

LORIST, H.H.J.; MEER, K.V.D. Standards for digital libraries and archives: digital longevity. 2001. Disponível em: <http://www.betade.tudelft.nl/publications/LoristMeer_ICEIS2001.pdf>. Acesso em: 21 de setembro 2005.

MIRANDA, L.C. Gerando tecnologia de reconhecimento mundial. **Techno**, 2002.

RIBEIRO, A. M. C. M. **Catálogo de recursos bibliográficos pelo AACR2**. 2.ed. Brasília: A.M.C.M. Ribeiro, 2002.

SAYÃO, F. Bases de dados: a metáfora da memória científica. **Ciência da Informação**, Brasília, v.25, n.3, p.314-318, set./dez. 1996.

SHIRI, A. Digital library research: current developments and trends. **Library Review**, v.52, n.5, p.198-202, 2003.

Anexo A – Lista de Campos – Biblioteca Digital

Nome de Campo	Abreviação
Accession Number	accessi
Acervo Hospedeiro	hos
Afiliação	af
Alternate Journal	alternatej
Alternate Title	alternatet
Ano	y
Ano de Acesso	accessy
Area	ar
Area Principal	firstar
Arquivo Alvo	ta
Assignee	as
Autor Secundário	secondarya
Autor(es)	au
Banca	comm
Base	ba
Calculador	comp
Call Number	cal
Cartógrafo(s)	car
Chave de Citação	cita
Chave Secundária	secondaryk
Cidade	city
Crédito	cr
Curso	cour
Custom 1	cu
Data	date
Data da Última Atualização	lastupdated
Data de Acesso	accessd
Data do Número	issued
Data Secundária	secondaryd
Detentor da Cópia	cop
Divulgação	di
Editor(es)	edito
Editor(es) da Série	seriee
Editora (Publisher)	publisher,
Edição	editi
Edição Anterior	pre
Edição Posterior	nex
Endereço da Editora	publishera
Endereço de e-Mail	e-m
Endereço(s) de e-Mail do(s) Autor(es)	el
Escala	sc
Estágio do Documento	do
Fonte	so
Forma de Publicação	how
Formato	fo
Grupo de Leitores	reade
Grupo de Usuários	userg
Grupo(s)	g
Idioma	lan
Índice	ind
Instituição	ins

Nome de Campo	Abreviação
ISBN	isbn,
ISBN/ISSN	isbn/
ISSN	issn
Jornal	new
Linhagem	li
Local da Publicação	pl
Localização do Evento	conferencel
Magazine	mag
Mês	mo
Nome do Evento	conferencen
Nota	mar
Notas	no
Número	issue,
Número da Patente	pat
Número de Arquivos	numberoff
Número de Páginas	numberofp
Número de Volumes	numberofv
Número do Relatório	reportn
Organização	o
Orientador(es)	sup
Pais	coun
Palavras-Chave	k
Permissão de Leitura	readp
Presidente da Banca	firstco
Primeira Afiliação	firstaf
Primeira Base	firstb
Primeira Fonte	firstso
Primeiro Autor	firstau
Primeiro Cartógrafo	firstca
Primeiro Editor	firste
Primeiro Editor da Série	firstse
Primeiro Grupo	firstg
Primeiro orientador	firstsu
Primeiro Programador	firstp
Primeiro Repórter	firstr
Primeiro Tradutor	firstt
Producer	prod
Programador(es)	prog
Projeto	proj
Pub. Source	pubs
Páginas	pag
Repositório	repos
Repositório de Banco de Dados	data
Repositório do Metadado	metadatar
Repositórios Filhos	ch
Repositórios Pais	par
Repórter(es)	reporte
Resumo	ab
Revista	j
Rótulo	lab
Site	sit

Nome de Campo	Abreviação
Subsidiary Author	sub
Tamanho	siz
Tema	them
Tertiary Author	tertiarya
Tertiary Title	tertiaryti
Tertiary Type	tertiaryty
Tipo	type,
Tipo da Referência	ref
Tipo da Tese	thes
Tipo de Suporte	typeofm
Tipo de Trabalho	typeofw
Tipo do Conteúdo	cont
Tipo Secundário	secondaryty
Tradutor(es)	tr
Título	ti
Título da Série	seriest
Título do Livro	bo
Título em Inglês	en
Título Secundário	secondaryti
Última Atualização	lastupdate,
Última Atualização do Metadado	metadatal
Unidade	aca
Universidade	un
URL	ur
Usuário Avançado	usern
Versão	ve
Visibilidade	vi
Volume	vo

**ESTUDO DE PADRONIZAÇÃO DE METADADOS PARA PRESERVAÇÃO DA
MEMÓRIA DOCUMENTAL DA BIBLIOTECA DIGITAL DO INPE**

Barbedo, Simone Angélica Del-Ducca

Ribeiro, Marciana Leite

Instituto Nacional de Pesquisas Espaciais – INPE
Avenida dos Astronautas, 1758 – Jardim da Granja
São José dos Campos – SP

simone@sid.inpe.br