

## ÁREA: – ACESSO À INFORMAÇÃO

### PROJETO DE CONVERSÃO DE DADOS DAS BIBLIOTECAS DO INPE NO SISTEMA “SophiA”

*Maria do Carmo de Castro Nogueira*

[mcarmo@sid.inpe.br](mailto:mcarmo@sid.inpe.br)

Instituto Nacional de Pesquisas Espaciais

Av. dos Astronautas, 1758

12227-010 - São José dos Campos - SP – Brasil

#### **Resumo**

O objetivo é descrever a metodologia de como utilizar os “campos” do software “SophiA”, segundo às Regras de Catalogação Anglo Americano - AACR2 pelo grupo de catalogadores do INPE tendo como propósito manter a integridade dos dados das bibliotecas das unidades no sistema “SophiA”. Preliminarmente, discute-se a normalização dos “campos” mais simples, porém não menos importantes como: informação de publicação (tipo de material), notação de autor, descrição física e notas. Num contexto mais complexo, pretende-se trabalhar as tabelas dos campos da: entrada principal, imprensa, série, assunto e entradas secundárias contribuindo significativamente no tocante às Referências Bibliográficas quando se pesquisa na página Web da base de dados do INPE. A importância da migração dos metadados, seja de uma Instituição pública ou privada, está na discussão do trabalho em equipe para a padronização e a garantia na uniformidade das informações, minimizando esforços futuros. O objetivo final visa consolidar e validar o projeto de conversão da base de dados, embora, seja um trabalho dinâmico e contínuo com previsão para outros projetos de correções e manutenção do acervo bibliográfico.

Palavras-chave: Conversão de dados. Metadados. Normalização. Indexação (Ciência da Informação).

#### **Introdução**

O processo de migração dos metadados do INPE do sistema “Cali” para o sistema “SophiA” ocorreu de forma natural como os dados se apresentavam,

no final de 2003, pois tinha-se em mente a premissa de que haveria necessidade de correções no decorrer do tempo qualquer que fosse o sistema adotado, conforme descrito na publicação do INPE-14175-PRE/9298; Power Point: Estudo de Caso: Software SophiA Biblioteca no Instituto Nacional de Pesquisas Espaciais (INPE) <http://mtc-m17.sid.inpe.br/rep/sid.inpe.br/mtc-m17@80/2007/04.09.16.58?languagebutton=pt-BR&mirror=lcp.inpe.br/ignes/2004/02.12.18.39.49&searchsite=bibdigital.sid.inpe.br:80&searchmirror=sid.inpe.br/bibdigital@80/2006/04.07.15.50.13>(na

URLIB – software livre, desenvolvido por um pesquisador do INPE para as publicações internas).

Este trabalho visa um projeto de conversão dos dados, delineando as dificuldades enfrentadas, bem como a preocupação com relação à integridade dos dados e a apresentação coerente da base do acervo bibliográfico de três unidades do INPE. Um dos pontos relevantes desse projeto é que as Referências Bibliográficas consultadas pelos alunos de pós-graduação e pesquisadores no sistema SophiA pela “Biblioteca Física” da página da Web da Biblioteca do INPE saiam corretamente e sejam aproveitadas em seus trabalhos, relatórios, teses e dissertações de acordo com os padrões biblioteconômicos da Norma Brasileira ABNT NBR 6023, de 2002. O intuito é pesquisar em uma base de dados confiável, com informações padronizadas pelos padrões de Catalogação do AACR2, constatando e corrigindo as distorções dos dados de acordo com as Regras do Código Anglo Americano, no software SophiA. Incentivou-se o envolvimento dos indexadores Bibliotecários de Catalogação para um estudo minucioso do sistema desse Banco de Dados e consenso nas informações. O objetivo deste trabalho é descrever a metodologia e os procedimentos sobre como utilizar os “campos” do SOPHIA, segundo as Regras de Catalogação do AACR2 pelo grupo de catalogadores para uniformizar e validar o projeto final da conversão da base de dados do INPE, gerando um manual de utilização. Realizou-se uma reunião preliminar na Primasoft Informática, empresa que desenvolveu o sistema Sophia em que discutiu-se alguns itens da validação da conversão; como surgiram vários questionamentos, o resultado dos testes não foi o esperado. A conversão desse projeto de correções envolve também a migração para a versão SophiA Elo 2009 e a migração para uma versão mais recente do Oracle

elaborada entre os analistas de desenvolvimento da Primasoft Informática e o Serviço de Tecnologia de Informação - STI do INPE. Em fase de implementação também está o protocolo Z39.50 que disponibiliza, no sistema Sophia, a importação de registros com mais de 500 bases.

### **Histórico**

A Unidade do Serviço de Informação e Documentação (SID) da Biblioteca do Instituto Nacional de Pesquisas Espaciais (INPE) foi criada em 1965 para atender a um grupo de pesquisadores da comunidade técnico-científica em suas linhas de pesquisa, basicamente Engenharia e Tecnologia Espacial, Computação, Sensoriamento Remoto, Meteorologia, Ciência Espacial e a alunos de pós-graduação de mestrado e doutorado. No ano de 2003, a Biblioteca foi totalmente integrada e informatizada mediante a implantação de um novo sistema de informatização, o software "SophiA", que hoje está disponível na WEB. O patrimônio técnico-científico conquistado nesses 42 anos exige uma base de dados confiável e criteriosa para melhor atender aos pesquisadores e à comunidade nacional e internacional.

### **Objetivo**

Na consulta e revisão da literatura científica, constatou-se que não há um trabalho específico sobre conversão de dados de uma Biblioteca Especializada de um Instituto de Pesquisas. Os exemplos na literatura são estudos voltados às Bibliotecas Universitárias, Bibliotecas Escolares e à Catalogação propriamente dita, como menciona Ribeiro (2008, 2004).

O ponto relevante deste estudo é a estruturação do serviço de indexação e registro de dados, respeitando uma uniformidade comum a todas as três unidades, ou seja, ter o mesmo procedimento e a mesma linguagem. A continuidade da qualidade do serviço dos catalogadores não pode interferir na mudança ou substituição dos profissionais, evitando distorções no desempenho das atividades. O resultado final será a elaboração de um manual padronizado de procedimentos dentro de normas pré-estabelecidas.

### **Objetivo Geral**

O sistema Cadastramento de Livros (CALI), utilizado pelo SID até o início do ano de 2003, não permitia o uso de letras minúsculas nos campos. Quando foi

feita a migração do sistema CALI para o Sistema de Gerenciamento SophiA, não houve correção dos dados que estavam em caixa alta.

O objetivo geral é seguir um procedimento normalizado pelo grupo de profissionais de catalogação. É promover o desenvolvimento qualitativo de uma base de dados confiável; controlar as informações com precisão; mapear as informações utilizando as ferramentas do sistema e, através de pesquisas com os usuários, detectar se atende suas expectativas.

Para que a facilidade Referência Bibliográfica, que o SophiA disponibiliza para os usuários na versão Web, seja recuperada de acordo com a norma NBR-6023 da ABNT (Referências Bibliográficas), faz-se necessário a conversão e revisão de todos os campos.

A correção desses problemas possibilitará maior confiabilidade dos dados e permitirá aos usuários o uso efetivo desse recurso, um benefício a mais na elaboração dos trabalhos científicos. Ter um profissional constantemente atualizando e revisando a manutenção de tabelas do sistema garantindo a uniformidade dos dados é de extrema valia.

## **Metodologia**

As alternativas de escolha da metodologia a ser adotada para se pesquisar as obras do acervo geral e obter resultados significativos para a gestão dessa informação foram preponderantes.

Estabeleceu-se um cronograma rigoroso com reuniões periódicas com o grupo de indexadores sobre os critérios adotados para o preenchimento e o detalhamento dos “campos” e propostas de alterações. Analisou-se, previamente, a base de dados de Periódicos, com a alteração dos “Títulos de Periódicos”, processo de certa forma simplificado; converteu-se de caixa alta para as iniciais das palavras dos títulos de periódicos em maiúsculas conforme publicação de normas de transcrição de dados de coleções CCN, anexo III, IBICT (2000), <http://www.ct.ibict.br:82/ccn/admin/>. Na base de dados de “Obras”, o processo é bem mais complexo; alterou-se as tabelas desde o primeiro campo, embora de certa forma simples, porém não menos importante:

- Informação Publicação - tipo de “Material”: Atlas, Catálogo, Dados, Folhetos, Livro, Mapa, Norma Técnica, Periódico, Publicação do INPE,

Referência, Transitório INPE. Dessa tabela foram eliminados alguns itens similares e outros transferidos para “suporte”; são eles: CD-ROM, Disquete, DVD, Fita de Vídeo, Microficha, Pen Drive, Slide, Foto.

- Idioma - com apenas algumas exclusões, manteve-se a tabela de línguas: Alemão, Espanhol, Francês, Inglês, Italiano, Japonês, Mais de 2 línguas, Outras línguas, Português, Russo.
- ISBN - dar preferência primeiramente ao número que consta no verso da capa externa do livro, ou no verso da página de rosto nos Dados Internacionais de Catalogação na Publicação (CIP). No caso de uma obra apresentar os dois códigos - ISBN e ISSN - opta-se pelo código da indexação que for escolhida. Quando se tratar de um periódico, porém receber tratamento de livro, colocar no campo do ISBN o número e acrescentar entre parênteses (ISSN).
- Número de Chamada - a Classificação adota a UDC - Classificação Decimal Universal, Padrão Internacional em Língua Portuguesa, a tradução do inglês elaborada pelo IBICT(2007). Na Classificação de obras, além da Classificação normal de uma obra, existem certas peculiaridades dependendo do tipo de material a ser tratado, como: Folhetos, Publicações Internas, Referência, Catálogos, Dados, Atlas, Mapas (topográficos, carta imagem radar, mosaico e outros), e Normas Técnicas, bem como no caso de DVD, CD-ROM, Disquete, Fita de Vídeo, Microficha, Pen Drive, Slide e, futuramente, Fotografias; vale lembrar que este tipo de material especial usa o número do código de barras e não o número de tombo.
- Notação de Autor - fornecida pela Tabela de Cutter & Sanborn para autor pessoa, evento e entrada pelo título na falta de autor principal; e acrescida ao número a inicial do título em minúscula na maioria dos casos. Por apresentarem muitos detalhes e especificidades, optou-se por tratar posteriormente o campo de Autor Principal e Título.
- Edição - transcrever desde a primeira edição e ano, e se revisada e atualizada.
- Imprenta - estabeleceu-se que Local de Publicação - as Tabelas de Local Geográfico obedeceriam às iniciais das palavras em maiúscula;

em Editora – padronizou-se que as iniciais das palavras ficariam em letra maiúscula; e forneceu-se ao Suporte da PrimaSoft Informática uma lista das Siglas para serem incluídas no programa de conversão.

- Descrição Física - de acordo com Ribeiro (2004), descrito no livro de Catalogação de recursos bibliográficos pelo AACR2, manteve-se o caso de volume e partes em letra minúscula, bem como os demais itens.
- Notas - gerais e bibliográficas etc., sempre a primeira letra maiúscula e o restante em minúscula, utiliza-se esse campo também para informar que o material indexado acompanha um outro tipo de material.

Para o estudo de campos fundamentais e com maior especificidade, faz-se necessário especial cuidado, no que tange à:

- Entrada Principal - manutenção de tabelas de Autoridade (autor pessoa, evento, instituição), Entradas Secundárias (co-autor, editor, tradutor, colaborador...).
- Título - convencionou-se que a primeira letra do título seria em maiúscula e o restante das palavras e preposições ficariam em minúsculas para não prejudicar as pesquisas nas Referências Bibliográficas, muito embora exista uma grande quantidade do tipo de material específico do acervo bibliográfico do INPE, que são os Proceedings, Congressos, Conferências, Simpósios. Retirar o ponto, no final do título, para não duplicar na referência bibliográfica. Retirar o “Acompanha CD...” e colocar no campo de notas. Para nomes próprios e siglas não fornecidas ao Suporte, a conversão não atendeu às expectativas, pois seria necessário um programa em SQL para suprir essas necessidades. Não existe uma tabela para se normalizar.
- Série – padronizou-se só as iniciais de cada palavra do título da série. Exemplo: Lecture Notes in Computer Science, 25. No que se refere às siglas de série, forneceu-se uma lista completa ao Suporte da PrimaSoft para se obter o resultado desejado.
- Assuntos – alterou-se somente a primeira letra em maiúscula e as demais letras e palavras em minúsculas. Utiliza-se uma lista do Vocabulário Controlado de Termos Existentes, ou o THESAURUS da

NASA para a indexação dos descritores e palavras-chaves dos materiais.

De uma maneira geral, a PrimaSoft atendeu aos itens solicitados, embora os testes e o tempo fossem insuficientes para assegurar e consolidar a validação da base. A construção correta dos dados devem estar indexados sistematicamente dentro de regras e padrões biblioteconômicos, pois constitui um bem valioso. Nessa análise, o importante instrumento formal para definir critérios e prioridades assemelha-se como apontado por Ribeiro (2008, 2004). O conceito aqui representado por um processo cíclico e dinâmico, não difere dos critérios aqui descritos.

### **Conclusão**

O panorama contemporâneo exige respostas mais efetivas. Neste cenário, a comunidade técnico-científica necessita agregar valor ao desenvolvimento de suas pesquisas de forma confiável e, de excelência, são os desafios desta gestão da informação neste projeto de conversão de dados.

Neste contexto, é de suma importância a manutenção e revisão constante dos dados oferecidos para sua credibilidade e melhor utilização, levando-se em conta que:

1. Os pesquisadores e alunos das instituições públicas ou privadas precisam de informações uniformizadas para utilizá-las em seus trabalhos de pesquisa.
2. A correção e atualização das tabelas devem refletir a real necessidade dos indexadores em manter o processo técnico a quem se destina como um organismo vivo e dinâmico.
3. Um sistema de dados normalizado seguindo regras, traz benefícios para a Instituição que o promove.
4. Após a conversão uma vez por semana, acordou-se em se fazer a reindexação da base de dados da biblioteca levando um dia para atualização.

Disponibilizar à comunidade científica uma base de dados normalizada na medida certa é questão de sobrevivência.

O ambiente informacional uniforme permite inovar e crescer de maneira harmônica.

Como menciona Couto (2005), o aprimoramento e a capacitação dos bibliotecários responsáveis pela Catalogação dos Processos Técnicos nas

bibliotecas é um trabalho que deve ser efetivo, e acompanhar as mudanças e atualizações das regras de descrição estabelecidas pelo AACR2R.

É preciso investir na manutenção da base de dados do acervo para que haja informação padronizada. Assim, é fundamental obter um documento com os procedimentos e mantê-lo atualizado dentro das normas, pois são múltiplos os aspectos dos recursos da Biblioteconomia que é, inegavelmente, capaz e instrumentada para contribuir para este trabalho “em progresso”. É de extrema valia a articulação do profissional da informação com a comunidade científica.

**Title:** Project of converting data from INPE' libraries in the "Sophia" system

**Abstract:**

The objective is to describe the methodology of how to use the "fields" of "Sophia" software, according to the rules of Anglo-American Cataloging - AACR2 by the INPE group with the purpose to maintain the integrity of data from libraries units in the "Sophia" system. First, we discuss the normalization of the simpler "Fields" but not less important like: information of publication (type of material), author representation, physical description and notes. In a more complex context is to work the tables of the fields: the main entrance, publishing data, serie, subject and secondary entries contributing significantly to References when searching the Web page of the INPE database. The importance of migration of metadata from a public or private Institution, it is in teamwork discussion to guarantee the uniformity and standardization of information, minimizing future efforts. The final objective aims to consolidate and validate the project of converting the database, although, being a dynamic and continuous with further planning for other projects of revising and maintenance of bibliographic collection.

**Key-words:**. Conversion of data. Metadata. Standardization. Indexing (Information Science).

## **REFERÊNCIAS BIBLIOGRÁFICAS**

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: informação e documentação: referências: elaboração. Rio de Janeiro, 2002.

COUTO, A.M. de H.C. de et al. **Oficina de representação descritiva de documentos. Módulo I. Nomes pessoais e entidades coletivas.** Niterói, UFF/PROAC/NDC, 2005. 71p.

IBICT. **Normas de transcrição de dados de coleções CCN.** Anexo III. Brasília, 2000. Disponível em: <http://www.ct.ibict.br:82/ccn/admin/>

INPE. **Estudo de Caso: Software SophiA Biblioteca no Instituto Nacional de Pesquisas Espaciais (INPE).** São José dos campos: INPE, 2006. (INPE-14175-PRE/9298) Power point <<http://mtc-m17.sid.inpe.br/rep/sid.inpe.br/mtc-m17@80/2007/04.09.16.58?languagebutton=pt-BR&mirror=lcp.inpe.br/ignes/2004/02.12.18.39.49&searchsite=bibdigital.sid.inpe.br:80&searchmirror=sid.inpe.br/bibdigital@80/2006/04.07.15.50.13>>

RIBEIRO, A.M. de C.M. **Catálogo de recursos bibliográficos pelo AACR2R em MARC 21.** 3.ed., 1. reimpr. Com correções. –Brasília: Edição do Autor, 2008. 1 v.

RIBEIRO, A.M. de C.M. **Catálogo de recursos bibliográficos pelo AACR2R 2002.** 2.ed. rev. e acrescida de índice. –Brasília: Edição do Autor, 2004. 1 v.

PRIMASOFT Informática. **Sophia Biblioteca – Catálogo de tipos de materiais.** São José dos Campos: PrimaSoft Informática, 33p. (Apostila de exemplos).

UDC Consortium. **Classificação Decimal Universal.** 2.ed. Brasília: IBICT, 2007. 2 v.