

UMA BASE DE PESQUISADORES CIENTÍFICOS

Alice Ferry de Moraes¹
Rejane Machado²

RESUMO

Este trabalho aborda a criação de base com nomes dos produtores científicos da Fiocruz, visando facilitar os *links* entre esses nomes e suas produções registradas em bases de conhecimento, além de preparar a Instituição para o futuro uso dos *Topics Maps*, exigidos pela ISO/IEC 13.250 de 2000. Os anais de eventos científicos da Fiocruz e catálogos de teses são utilizados como fontes de pesquisa. A coleta dos nomes foi realizada tal como eles se apresentam nas fontes. A entrada básica dos nomes foi estabelecida pela AACR2 e as informações foram registradas, na base, de acordo com o Formato Marc. A base servirá para a identificação e avaliação da produção científica dos pesquisadores, não apenas nos eventos da Fiocruz, mas também nas fontes de comunicação científica sob as Normas de Vancouver. Os *links* com a produção citada na Plataforma Lattes e na sua versão institucional (Fiolattes) serão ser beneficiados.

Palavras-chave: Produção científica. Comunicação científica. Pesquisadores.

1 INTRODUÇÃO

A ciência, tal como se apresenta, desenvolve uma infra-estrutura para tornar informações acessíveis sob as mais variadas formas. Crescem os intermediários na troca de conhecimento e o processo informacional torna-se uma tarefa social e fundamental para controle, organização e consolidação da informação em níveis nacional e internacional, por meio das bases de dados e das bases de conhecimento.

¹ Bibliotecária – Centro de Informação Científica e Tecnológica/Fundação Oswaldo Cruz
Doutora em Ciência da Informação
Universidade Federal do Rio de Janeiro/ Instituto Brasileira de Informação em Ciência e Tecnologia
ferry@cict.fiocruz.br

² Bibliotecária – Centro de Informação Científica e Tecnológica/Fundação Oswaldo Cruz
Mestre em Informação e Comunicação em Saúde
Escola Nacional de Saúde Pública Sérgio Arouca/Fundação Oswaldo Cruz
rejane@cict.fiocruz.br

O desenvolvimento da informatização traz novos papéis sociais, novos serviços e criação de organizações como redes públicas para transferência de dados, agências de informação, plataformas de dados científicos. A organização e a confiabilidade da informação são elementos determinantes nesse momento.

Os produtores científicos da FIOCRUZ têm seus nomes registrados sob variadas formas de entradas, atendendo às normas de autoria de publicações e/ou das bases de dados. Isso impede, algumas vezes, o reconhecimento de trabalhos diferentes de uma mesma autoria. Há casos em que o autor aparece citado pelo prenome, outros com sobrenome e seus prenomes completos, outros com sobrenome seguido de iniciais do prenome e demais sobrenomes e ainda os sobrenomes compostos determinados pelos próprios autores, ressaltando ainda o caso das produtoras científicas, que apresentam variações de nomes de acordo com o seu estado civil.

Este projeto restringe-se ao quadro de produtores científicos da Fundação Oswaldo Cruz (FIOCRUZ). Mas, no futuro, poderá servir como parâmetro técnico para normalização de outros nomes, pesquisadores ou não, cujas produções científicas integrem acervos e bases desta e de outras instituições.

No primeiro momento, é possível apresentar como diagnóstico sobre a necessidade de construção desta base, as dificuldades já encontradas na Plataforma Lattes e Fiolattes sobre registros de nomes de produtores científicos em diferentes formas, inviabilizando os *links* com suas produções.

O objetivo geral deste trabalho é disponibilizar informações sobre o quadro de produtores científicos, integrantes e egressos das áreas de Pesquisa e Ensino da Fundação Oswaldo Cruz (FIOCRUZ), visando a identificação e a localização da produção científica, onde estiver publicada, e sob quaisquer normas diferentes, nos vários suportes informacionais e seus registros armazenados em bancos de dados e de conhecimento.

2 METODOLOGIA

O trabalho inicia-se com a coleta de nomes dos produtores científicos da FIOCRUZ sob todas as formas possíveis, a partir dos índices dos anais de eventos científicos promovidos pela Instituição, assim como dos catálogos de teses e dissertações.

Em seguida, é necessário identificar os nomes coletados e inseri-los em uma base construída, através da qual registros possam facilitar os *links* com suas produções em diversas publicações e bases de dados e de conhecimento onde se encontram. Em outras palavras, é realizada a identificação das diversas formas dos nomes dos produtores científicos da FIOCRUZ (*basic name* – por prenome). É realizado, também, o registro dos nomes coletados sob as diferentes normas nacionais e internacionais (ABNT e Vancouver) (*sort name* – sobrenome padronizado). Os nomes de família originariamente compostos, os sobrenomes compostos com diferenciadores, no caso de sobrenomes muito comuns, os sobrenomes indicados pelas produtoras científicas de acordo com a opção feita com base em seu estado civil são registrados tal como eles aparecem nas diferentes fontes de pesquisa (*display name* – sobrenome tal como o produtor científico deseja ser identificado).

Paralelamente, é iniciada a construção da estrutura de uma base piloto para testar as diversas formas de registro de nomes e as fontes de onde são coletados. Para tanto, é utilizada, como parâmetro de infra-estrutura tecnológica, o Formato MARC, de padrão internacional (ISO 2709/1975), que permite o intercâmbio de dados com quaisquer bases.

Utilizando os campos do Formato Marc, são feitos os registros das fontes de cada nome pesquisado.

3 FONTES DE PESQUISA

Ao iniciarmos a construção da base *Ensino e Pesquisa da FIOCRUZ: Produtores Científicos*, nos deparamos com um determinado universo de fontes para coletas de dados. Pensamos, inicialmente, em trabalhar com o ensino no nível de pós-graduação e temos como fontes a nosso dispor coleções dos Anais das Jornadas Científicas de Pós-Graduação (incompleta) e das Bienais de Pesquisa do FIOCRUZ.

As buscas às fontes que faltavam nos levam aos Anais das Jornadas Científicas e aos Anais das Reuniões da Iniciação Científica (coleções completas), alterando de forma substancial o processo e o cronograma de trabalho.

Apesar da importância, o registro dos nomes relacionados nos anais das reuniões de iniciação científica ficará para um segundo momento de trabalho.

Os nomes registrados nos anais levam a modificações nas formas de pesquisa inicialmente adotadas. Os nomes dos pesquisadores nas referidas fontes ora estão todos registrados nos índices; ora os nomes são observados apenas no corpo da publicação, as ausências de alguns nomes dos participantes do evento; ora os nomes dos pesquisadores registrados nos índices aparecem de forma diferente dos nomes registrados no corpo da publicação, ora há, no mesmo índice, o mesmo nome de pesquisador registrado de forma diferente, etc.

Há casos em que um mesmo pesquisador (a) apresentou diversas e diferentes entradas de nomes (ver Anexo), o que legitima a criação dessa base.

O que é verificado, durante o manuseio das fontes ratifica uma possível inoperância no momento de recuperação de informação nas bases de dados, assim

como uma inconsistência no levantamento de dados para uma avaliação da produção científica da FIOCRUZ.

É feito, então, um confronto de nomes com temas, laboratórios e demais pesquisadores autores de um mesmo trabalho para retificar ou ratificar um determinado registro.

Alguns nomes estrangeiros (orientadores externos) são pesquisados no *Google*.

Apesar dos esforços efetuados pela equipe de trabalho, alguns registros permanecem sem esclarecimentos quanto aos prenomes ou até mesmo sobrenomes dos pesquisadores.

4 FORMATO MARC

O formato *Machine Readable Cataloging Record* - Marc 21 é gerenciado pela *Library of Congress* (LC) dos Estados Unidos e é considerado padrão internacional adotado para registro bibliográfico. O Marc oferece diversos tipos de campos que possibilitam a descrição bibliográfica de documentos, independente de sua forma. Esse formato possui, entre outras divisões, a de autoridades contendo especificações que possibilitam o controle das autorias das produções científicas, objeto deste trabalho.

A Biblioteca Nacional (BN), que utiliza esse formato para registro de autoridades (nomes de pessoas físicas, no caso), teve essa sua base pesquisada de maneira a auxiliar na montagem da estrutura da base de produtores científicos. A partir dessa observação, alguns campos e sub-campos foram considerados essenciais para a elaboração de nossa planilha de coleta de dados (nomes dos pesquisadores).

Adota-se o campo **040** para o nome da Instituição que criou o registro, ou seja, a Fiocruz (BR-RjFOC). Esse código é fornecido pela LC.

O campo líder (**LDR**), utilizado com vistas à recuperação e gerenciamento dos registros da base, é adotado como o primeiro campo, no qual é possível identificar o *status* do registro, ou seja, se é novo ou alterado, possibilitando, também, identificar se o registro da autoridade está completo ou não.

O campo **100** é utilizado para a entrada dos nomes de acordo com as normas da AACR2 (*Anglo American Cataloging Rules, 2. Edition*). O indicador adotado é sempre o **1**, uma vez que o nome aqui é registrado pelo sobrenome do pesquisador.

Para registro das diversas formas de nomes coletados nos anais é utilizado o campo **400**.

O registro das fontes pesquisadas dá-se no campo **670**, cabendo unicamente para o caso da ausência do nome pesquisado na Plataforma Lattes, o uso do campo **675**.

Os sub-campos **a** são utilizados rotineiramente. O sub-campo **c** serve para indicar o nome do pesquisador, atuando como orientador. Já o sub-campo **q** é utilizado para o caso dos nomes incompletos, no qual poderá ser registrado o(s) nome(s) encontrado a partir de outra fonte, completando assim a falha no registro indicada por iniciais do(s) nome(s).

É importante ressaltar que a Plataforma Lattes é parâmetro de registro dos nomes de maneira plena e correta. O indicador referente ao seu registro é sempre o **0**, uma vez que a entrada se faz pelo prenome do pesquisador

5 SOFTWARE ALEPH

A instalação da base para o Formato Marc é realizada no *software* Aleph, que é utilizado pelas bibliotecas do Centro de Informação Científica e tecnológica (CICT) da Fiocruz. A primeira medida tomada é a distribuição de senhas de acesso à nova base. Em seguida, é instalada a planilha estabelecida pela equipe de trabalho, mas algumas questões ocorrem na customização da base, chamada FIO 13. As questões são repassadas à Ex Libris, empresa responsável pela assessoria técnica do Aleph. Novo teste de catalogação é realizado, dessa vez com sucesso. Posteriormente, constata-se a necessidade de customização do módulo OPAC nos dois segmentos, ou seja, na recuperação e na pesquisa pelo *browser*, e é providenciada a retirada da visualização dos campos que não pertencem ao registro de autoridades.

No estágio atual, a base, que já possui mais de seis mil registros está sendo customizada para sua inserção na WEB, de maneira a poder ser utilizada pelos usuários (público e pesquisadores) e catalogadores das bibliotecas da Fiocruz.

6 CONCLUSÃO

Avaliação do desenvolvimento do projeto nos leva a concluir que a disponibilização da base de nomes pesquisados e tratados será de grande valia para as bibliotecas da área, além das bibliotecas da Fiocruz. Também será útil para o acesso à produção científica registrada na Plataforma Lattes institucional, ou seja, a base Fiolattes.

Tornar disponível a base de produtores científicos, utilizando o Formato Marc, possibilitará o intercâmbio de registros informacionais de autoridades, aqui criteriosamente descritas, colaborando na alimentação de outros sistemas de informação.

Também é importante a disseminação da base de nomes, incluindo os registros sob as formas *de basic, sort e display names*, visando o futuro uso dos *Topic Maps*, que processarão as informações através de aplicações SGML (*Standard Generalized Markup Language*) baseadas em XML, conforme exigência da ISO/IEC 13.250 de 2000, para *links* com as bases de conhecimento na Internet.

A identificação dos produtores científicos contribuirá para os processos de análise, avaliação e divulgação da produção institucional, agregada às linhas e grupos de pesquisa da Fiocruz.

Com base na análise das fontes de pesquisa, será proposta uma uniformização dos registros nessas fontes, de maneira a evitar dificuldades na identificação dos produtores científicos na Instituição.

REFERÊNCIAS

BIENAL DE PESQUISA, 1., 1998, Rio de Janeiro. *Anais ...* Rio de Janeiro: Fiocruz, 1998.

BIENAL DE PESQUISA, 2., 2000, Rio de Janeiro. *Anais ...* Rio de Janeiro: Fiocruz, 2000.

BIENAL DE PESQUISA, 3., 2002, Rio de Janeiro. *Anais ...* Rio de Janeiro: Fiocruz, 2002.

CATÁLOGO de teses e dissertações, 1980-1994. Rio de Janeiro: Fiocruz, 1995.

CATÁLOGO de teses e dissertações, 1996-2001. Rio de Janeiro: Fiocruz, 2002.

CÓDIGO de catalogação anglo-americano – AACR2. 2. Ed. Brasília: FEBAB, 2002.

FUNDAÇÃO BIBLIOTECA NACIONAL. Disponível em <http://www.bn.br> . Acesso em: 00/08/2003.

JORNADA CIENTÍFICA DA FIOCRUZ, 1., 1980. *Anais ...* Rio de Janeiro: Fiocruz, 1980.

JORNADA CIENTÍFICA DA FIOCRUZ, 2., 1983. *Anais ...* Rio de Janeiro: Fiocruz, 1983.

JORNADA CIENTÍFICA DA FIOCRUZ, 3., 1985. *Anais ...* Rio de Janeiro: Fiocruz, 1985.

JORNADA CIENTÍFICA DA FIOCRUZ, 4., 1991. *Anais ...* Rio de Janeiro: Fiocruz, 1991.

JORNADA CIENTÍFICA DA FIOCRUZ, 5., 1995. *Anais ...* Rio de Janeiro: Fiocruz, 1995.

JORNADA CIENTÍFICA DE PÓS-GRADUAÇÃO DA FIOCRUZ, 1., 1994, *Anais ...* Rio de Janeiro: Fiocruz, 1994.

JORNADA CIENTÍFICA DE PÓS-GRADUAÇÃO DA FIOCRUZ, 2., 1995, *Anais ...* Rio de Janeiro: Fiocruz, 1995.

JORNADA CIENTÍFICA DE PÓS-GRADUAÇÃO DA FIOCRUZ, 3., 1996, *Anais ...* Rio de Janeiro: Fiocruz, 1996.

JORNADA CIENTÍFICA DE PÓS-GRADUAÇÃO DA FIOCRUZ, 4., 1998, *Anais ...* Rio de Janeiro: Fiocruz, 1998.

JORNADA CIENTÍFICA DE PÓS-GRADUAÇÃO DA FIOCRUZ, 5., 1999, *Anais ...* Rio de Janeiro: Fiocruz, 1999.

JORNADA CIENTÍFICA DE PÓS-GRADUAÇÃO DA FIOCRUZ, 6., 2000, *Anais ...* Rio de Janeiro: Fiocruz, 2000.

JORNADA CIENTÍFICA DE PÓS-GRADUAÇÃO DA FIOCRUZ, 7., 2001, *Anais ...* Rio de Janeiro: Fiocruz, 2001.

JORNADA CIENTÍFICA DE PÓS-GRADUAÇÃO DA FIOCRUZ, 8., 2003, *Anais ...* Rio de Janeiro: Fiocruz, 2003.

MARC Concise Format. Authority. Washington, D.C. : Library of Congress, 2004. Disponível em: <http://www.loc.gov/marc/>. Acesso em: 03 ago. 2004.

MOREIRO GONZÁLEZ, J.A. *et al.* Construcción automática de topic maps. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 20., 2002, Fortaleza. *Anais ...* Fortaleza: Febab, ACB, 2002. Ecard.

ABSTRACT

This paper approaches the basis creation with names of the scientific producers from Fiocruz, seeking to facilitate the links between those names and their productions registered in knowledge bases, besides preparing the Institution for Topics Maps future use, demanded by ISO/IEC 13.250/2000. The annals of scientific events from Fiocruz and catalogs of theories were used as research sources. The collection of the names was accomplished as they came in the sources. The basic entrance of the names was established by AACR2 and the information was registered, in the base, according to the Marc format. The base will be used for the identification and evaluation of researchers scientific production, not only at Fiocruz events, but also in scientific communication sources under the Norms of Vancouver. The links with the production mentioned in the Lattes and its institutional version (Fiolattes) will be benefited.

Key words: Scientific production. Scientific communication. Searching.

ANEXO

LDR 00212nz 2200121n 4500
008 000000nn acnnaabn fa aaa d
040 |a BR-RjFOC
|c BR-RjFOC
1001 |a SILVA, Patrícia Machado Rodrigues e
4000 |a Patrícia Machado Rodrigues e Silva
4000 |a Silvia, P. M. R.
|q (Patrícia Machado Rodrigues)
4001 |a Silvia, Patrícia M. R. E
|q (Machado Rodrigues e)
4000 |a Patrícia Machado Rodrigues e Silva
|c (Orientadora)
4000 |a Patricia Machado Rodrigues e Silva
4000 |a Patrícia M. R. e Silva
|q (Machado Rodrigues)
|c (Orientadora)
4000 |a Patricia Machado Rodrigues e Silva Martins
|c (Orientadora)
4001 |a SILVA, Patrícia Machado Rodrigues e
|c (Orientadora)
4001 |a SILVA, Patrícia Machado Rodrigues e
4000 |a Patrícia Silva
|c (Orientador)
4001 |a SILVA, Patrícia M. R.
|q (Machado Rodrigues)
4000 |a Patricia MR Silva
|q (Machado Rodrigues)
|c (Orientadora)
4000 |a P. M. R. Silva

|q (Patricia Machado Rodrigues)
670 |a AACR2
670 |a Plataforma Lattes, 2004
670 |a Anais da IV Jornada Científica, 1991
670 |a Anais da V Jornada Científica, 1995
670 |a Anais da V Jornada Científica, 1995
670 |a Anais da II Jornada Científica de Pós-Graduação, 1995
670 |a Anais da III Jornada Científica de Pós-Graduação, 1996
670 |a Anais da V Jornada Científica de Pós-Graduação, 1999
670 |a Anais da VII Jornada Científica de Pós-Graduação, 2001
670 |a Anais da VIII Jornada Científica de Pós-Graduação, 2003
670 |a Anais da I Bienal de Pesquisa, 1998
670 |a Anais da IV Jornada Científica de Pós-Graduação, 1998
670 |a Anais da V Jornada Científica de Pós-Graduação, 1999
670 |a Anais da IV Jornada Científica de Pós-Graduação, 1999
670 |a Anais da VI Jornada Científica de Pós-Graduação, 2000
670 |a Anais da II Bienal de Pesquisa, 2000
670 |a Anais da VII Jornada Científica de Pós-Graduação, 2001
670 |a Anais da III Bienal de Pesquisa, 2002
670 |a Anais da VIII Jornada Científica de Pós-Graduação, 2003
670 |a Anais da III Jornada Científica de Pós-Graduação, 1996
SYS 000004378