

PROPOSTA PARA A CRIAÇÃO DE UM VOCABULÁRIO CONTROLADO A PARTIR DO SISTEMA DE INFORMAÇÕES DO ARQUIVO NACIONAL DO BRASIL (SIAN)

Mariane Costa Pinto¹

RESUMO: O artigo propõe a criação de um vocabulário controlado a partir de termos já existentes no Sistema de Informações do Arquivo Nacional (SIAN). A indexação, atividade normalmente associada à Biblioteconomia, pode ser considerada, junto à descrição, uma das partes mais importantes da recuperação da informação e deve ser feito por especialistas. Para a elaboração de um vocabulário controlado é necessária a criação de uma política de indexação. Esta política compreende a elaboração de procedimentos para a transcrição da linguagem natural para a documentária. A partir de decisões, como entradas pré ou pós-coordenadas e indexação exaustiva ou específica e da comunicação entre integrantes de diferentes equipes é possível desenvolver uma recuperação de informação eficiente e eficaz.

Palavras-chave: Vocabulário controlado. Arquivo Nacional. Indexação.

PROPOSAL FOR THE ESTABLISHMENT OF A CONTROLLED VOCABULARY FROM NATIONAL ARCHIVE INFORMATION SYSTEM (SIAN) IN BRAZIL

ABSTRACT: The article proposes the creation of a controlled vocabulary from the terms that exists in the National Archives Information System (SIAN). Indexing, usually associated with librarianship activity, can be considered, with description, one of the most important pieces of information retrieval and should be done by experts. For the elaboration of a controlled vocabulary an indexing policy need to be created. This policy includes the development of procedures for the transcription of the natural language for the documentary one. From decisions as pre or post-coordinated entries, comprehensive or specific indexing and communication between members of different teams is possible to develop an efficient and effective information retrieval.

Keywords: Controlled Vocabulary. National Archives. Indexing.

1. INTRODUÇÃO

Os documentos arquivísticos são organizados em categorias: fundo/coleção, seção, série, subsérie, dossiê e item documental, de acordo com a sua proveniência.

Fundo é o conjunto de documentos de uma mesma proveniência, este termo equivale a Arquivo. Coleção é o conjunto de documentos com características comuns, reunidos intencionalmente. Seção é a subdivisão da estrutura hierarquizada de organização que corresponde a uma primeira fração lógica do fundo ou coleção, em geral reunindo documentos produzidos ou acumulados por unidades administrativas com competências específicas, também chamada grupo ou subfundo. Série é uma subdivisão que corresponde a

¹Bibliotecária da Coordenação de Documentos Audiovisuais e Cartográficos mariane@arquivonacional.gov.br

uma sequência de documentos relativos à mesma função, atividade, tipo documental ou assunto. Subsérie é a subdivisão da série. Dossiê é a unidade de arquivamento constituída de documentos relacionados entre si por assunto (ação, evento, pessoa, lugar, projeto). Item documental é o documento que compõe o dossiê ou processo. (NOBRADE, 2006).

A partir da descrição dos documentos, é possível criar pontos de acesso para sua recuperação. A indexação aparece como atividade responsável pela recuperação do conteúdo dos documentos, exigindo uma transcrição da linguagem natural para a documentária.

O Arquivo Nacional tem como sua maior finalidade “o pleno acesso à informação”. Para que se cumpra este objetivo, seu acervo deve estar organizado. Por organização, entende-se o processamento técnico, que engloba atividades de recolhimento, tratamento técnico, preservação, divulgação, controle de acesso e uso de seu patrimônio arquivístico (ARQUIVO NACIONAL, 2012).

No tratamento técnico do acervo permanente da instituição, concentram-se duas atividades essenciais à recuperação da informação: a descrição e a indexação. É por meio delas que o usuário pode fazer uso dos instrumentos de pesquisa para o acesso aos documentos.

No Arquivo Nacional aplica-se a descrição multinível, ou seja, a representação dos elementos e informações contidos nos documentos, levando-se em consideração a estrutura de organização de um acervo e integrando-o em diferentes níveis e em relações horizontais e verticais (ARQUIVO NACIONAL, 2005).

Indexação é um processo que visa à recuperação da informação. A partir dela é possível a elaboração de índices. Nos sistemas informatizados de recuperação da informação, “uma indexação de qualidade assegura uma recuperação mais precisa.” (SOUZA, 2009, p. 7).

De acordo com a Associação Brasileira de Normas Técnicas (1992, p. 2), “NBR 12676: métodos para análise de documentos: determinação de seus assuntos e seleção de termos de indexação”, a indexação compreende três estágios: “exame do documento e estabelecimento do assunto de seu conteúdo; identificação dos conceitos presentes no assunto; tradução desses conceitos nos termos de uma linguagem de indexação.”

Ao atribuir um descritor [isto é, um termo de indexação] a um documento, o indexador declara que tal descritor possui alto grau de relevância para o conteúdo do documento; quer dizer, ele declara que o significado do descritor está fortemente associado a um conceito incorporado ao documento, e que é adequado à área temática do documento. (LANCASTER, 2004, p. 11).

Lancaster (2004) afirma que neste processo, o responsável pela indexação deve estar

atento a três perguntas: Do que trata o documento? Por que foi incorporado ao acervo? Quais de seus aspectos serão de interesse para os usuários?

o indexador tem a função primordial de compreender a leitura ao realizar uma análise conceitual que represente, adequadamente, o conteúdo de um documento para que ocorra correspondência com o assunto pesquisado pelo usuário. No entanto, para que essa correspondência aconteça, a adoção de uma política de indexação torna-se imprescindível, pois ela será norteadora de princípios e critérios que servirão de guia na tomada de decisões para otimização do serviço e racionalização dos processos. (RUBI; FUJITA, 2003, p. 67).

Essa política deve levar em conta três fatores: as características e objetivos da organização, que determinam o serviço oferecido; a identificação dos usuários, para o atendimento de suas necessidades; e os recursos humanos, materiais e financeiros disponíveis na instituição, que norteiam o funcionamento do sistema de informações, suas especificidades e limitações (CARNEIRO, 1985, p. 221).

Cabe destacar que a indexação é um dos possíveis pontos de acesso, ou seja, um “elemento de informação, termo ou código que, presente em unidades de descrição, serve à pesquisa, identificação ou localização de documentos.” (NOBRADE, 2006, p.17). Outros devem ser utilizados para o preenchimento de informações que não se referem ao seu conteúdo, como suporte, forma de escrita, espécie e gênero.

2. A INDEXAÇÃO DO ARQUIVO NACIONAL

O Arquivo Nacional possui um sistema de informação integrado, conhecido como Sistema de Informações do Arquivo Nacional (SIAN). Cada equipe responsável pela organização documental realiza a descrição arquivística de acordo com a *Norma brasileira de descrição arquivística* (NOBRADE), porém não há uniformidade nos procedimentos de tratamento técnico. Isso compromete a qualidade da indexação.

Para aqueles que fazem as descrições dos documentos no sistema de informação, existe uma equipe encarregada de avaliar os termos de indexação inseridos. Como não é sua atribuição analisar a descrição como um todo, os termos chegam isolados às suas mãos, o que dificulta avaliar se o termo é pertinente ou não.

Por não existir no Arquivo Nacional um vocabulário controlado próprio, a recomendação é que o indexador pesquise o termo a inserir em quatro fontes principais: a apostila *Regras para indexação dos documentos da CDA*, de Nilda Sampaio Barbosa, (ex-bibliotecária da instituição); *Terminologia de assuntos* do sítio da Biblioteca Nacional;

Vocabulário controlado básico do Senado Federal e o Catálogo online da *Library of Congress*. Isso acontece para que, mesmo com descrições diferentes, a instituição possa apresentar termos de busca iguais nos diferentes acervos, em benefício do pesquisador usuário.

Ainda assim, esta recomendação nem sempre é o suficiente, o que torna cada vez mais difícil a construção de uma política de indexação e um vocabulário controlado próprios.

Para que se construa uma política de indexação, é necessário que se crie estratégias para a recuperação da informação. A começar pela escolha de indexadores que possam compreender o documento, fazer uma análise conceitual e representar adequadamente o seu conteúdo. Não basta construir uma lista de procedimentos a ser seguidos, deve-se, acima de tudo, entender a filosofia dos objetivos da recuperação da informação.

A indexação, sob este ponto de vista, é uma das partes mais importantes do tratamento técnico, junto à descrição, e deve fazer parte do planejamento dos sistemas de informação e, principalmente, ser realizado por indexadores capacitados. (FUJITA; RUBI, 2003).

Com este objetivo, aqui apresenta-se uma proposta de criação de vocabulário controlado tomando como exemplo uma parte da listagem de termos de indexação autorizados do SIAN. Neste caso, selecionou-se os termos iniciados com a letra N, exceto nomes próprios.

3 OBJETIVOS

3.1 OBJETIVOS GERAIS

Apresentar os termos de indexação do Sistema do Arquivo Nacional, acrescentando termos específicos, gerais e relacionados.

3.2 OBJETIVOS ESPECÍFICOS

A partir da análise dos termos, ser capaz de diagnosticar inconsistências na indexação das informações na base de dados;

Reunir em ordem alfabética os termos autorizados do SIAN para identificar a necessidade do estabelecimento de relações entre eles.

4 METODOLOGIA

Para melhor exemplificar a necessidade da avaliação do processo de indexação no Arquivo Nacional, e para que possamos dar um primeiro passo rumo a criação de uma nova política, elaborou-se um exemplo de vocabulário controlado com termos iniciados pela letra N.

Neste caso, utilizou-se somente termos inseridos no SIAN até o dia 21 de junho de 2016. De acordo com a listagem *Termos de Indexação* da base de dados, existem 1405 (mil quatrocentos e cinco) termos de indexação que começam com N. Dos quais, apenas 35 (trinta e cinco) são termos possíveis para um vocabulário controlado, o restante trata-se de nomes próprios.

Na formação dos termos relacionados, gerais e específicos, primou-se pela utilização dos termos já existentes na base de dados. As fontes de referência para auxiliar tal processo foram o Catálogo de Terminologia de Assuntos do sítio da Biblioteca Nacional e o *Vocabulário Controlado Básico* do Senado Federal. Também dessas fontes foram retirados termos que melhor suprissem a formação das cadeias e renques quando ausentes no SIAN. Alguns foram passados do singular para o plural.

Os termos vem acompanhado das seguintes siglas:

TG (Termo Genérico): o termo que se segue refere-se a um conceito com conotação mais ampla.

TE (Termo Específico): o termo que se segue refere-se a um conceito com conotação mais específica.

TR (Termo Relacionado): o termo que se segue está associado, mas não é nem um sinônimo, nem um termo genérico ou termo específico.

NE (Nota Explicativa): uma nota que se junta a um termo para indicar seu significado específico dentro de uma linguagem de indexação.

USE (indica o termo preferido): o termo que se segue é o termo preferido quando se deve escolher entre sinônimos ou quase-sinônimos.

5 SUGESTÃO DE VOCABULÁRIO CONTROLADO

Nação brasileira

T.G Nação

Nacionalidade

N.E. Usado para obras que tratam de um aspecto do Direito civil que vincula a pessoa ao Estado através do nascimento ou obtida por meio de naturalização.

T.G. Direitos humanos

T.R. Naturalização

T.E. Estrangeiros

Nariz

N.E. Pode ser usado com a subdivisão - Doenças

T.G. Face

T.R. Olfato

T.R. Otorrinolaringologia

T.E. Cavidades nasais

Nascimento

T.G. Ciclo de vida

T.R. Aborto

T.R. Gravidez

T.R. Natalidade

T.E. Parto (Obstetrícia)

Natação

T.G. Esportes aquáticos

T.R. Piscina

T.E. Nado borboleta

T.E. Nado *crawl*

T.E. Nado sincronizado

Natal

T.G. Datas comemorativas

T.G. Feriados e festas cívicas

T.E. Papai Noel

Natimorto

T.R. Morte

T.R. Nascimento

Naturalismo

T.G. Filosofia

T.G. Materialismo

T.R. Mecanismo (Filosofia)

T.R. Natureza

Naturalista

USE Naturalismo

Naturalização

T.G. Direito civil

T.G. Direito internacional público

T.G. Migração

T.R. Nacionalidade

T.E. Dupla nacionalidade

Naus

N.E. Segundo Houaiss (2004), Nau é “designação genérica que se aplicava a navios de grande porte, acastelados à proa e à popa, e ger. com um único mastro envergando vela redonda; navio de grande porte, com três mastros, velas redondas, fortemente armado, integrante geralmente de uma linha de batalha”.

T.G. Embarcações

T.R. Barcos

T.R. Lanchas

T.R. Navios

Naufrações

T.R. Acidentes marítimos

T.R. Segurança marítima

Nave espacial

T.G. Aeronave

T.R. Espaço cósmico

T.R. Exploração espacial

T.R. Navegação espacial

Navegação

T.R. Comércio marítimo

T.E. Navegação a vapor

T.E. Navegação aérea

T.E. Navegação fluvial

T.E. Navegação marítima

Navegação a vapor

T.G. Navegação

T.R. Navegação aérea

T.R. Navegação fluvial

T.R. Navegação marítima

Navegação aérea

T.G. Navegação

T.R. Navegação a vapor

T.R. Navegação fluvial

T.R. Navegação marítima

Navegação fluvial

T.G. Navegação

T.R. Navegação a vapor

T.R. Navegação aérea

T.R. Navegação marítima

Navegação marítima

T.G. Navegação

T.R. Navegação a vapor

T.R. Navegação aérea

T.R. Navegação fluvial

Navios

N.E. Indica noção genérica. Será seguido da denominação quando a indicação for específica.

T.G. Embarcação

T.R. Barcos

T.R. Lanchas

T.R. Naus

T.E. Navios-escola

T.E. Navios de guerra

Nazismo

T.G. Doutrina política

T.R. Fascismo

T.R. Neonazismo

T.E. Movimento antinazista

Negligência

T.G. Ato ilícito

T.R. Imperícia

T.R. Imprudência

Negócios

N.E. Usado no sentido de transações comerciais.

T.G. Administração

T.G. Economia

T.R. Administração de empresas

T.R. Comércio

T.E. Empresas

T.E. Negócios eclesiásticos

Negócios eclesiásticos

T.G. Negócios

Negros

N.E. Usar, também, ligado a determinados assuntos. Ex: Negros na literatura.

T.R. Cultura negra

T.R. Escravidão negra

T.R. Discriminação racial

T.E. Trabalhador negro

Neutralidade

N.E. Usado para obras que tratam de países que se recusam a tomar partido entre os países beligerantes.

T.G. Direito de guerra

T.R. Guerra

T.R. Relações internacionais

Neve

T.G. Inverno

T.R. Chuva

T.E. Esqui

T.E. Gelo

Nome suposto

T.R. Contravenção

Nomeação

N.E. Usar como subdivisão de cabeçalho. Ex: Juízes – Seleção e nomeação.

Nota diplomática

USE Diplomacia – Notas

Notas

N.E. Usar como subdivisão de cabeçalho de assunto. Ex: Igrejas – Notas. Especificar o tipo

de nota. Ex: Nota fiscal; Nota promissória, Nota de rodapé.

Notificações

T.R. Carta precatória

T.R. Edital

Núcleos coloniais

T.R. Imigração

T.R. Divisões territoriais e administrativas

Nunciatura apostólica

T.G. Apostolado

Nutrição

T.R. Alimentação

T.R. Desnutrição

T.E. Nutrição infantil

T.E. Nutrição materna

Nutricionista

USE Nutrição

6 CONSIDERAÇÕES FINAIS

Em uma instituição como o Arquivo Nacional, a construção de uma política de indexação esbarra em alguns fatores como a dificuldade na identificação do usuário da instituição, já que é um órgão destinado a qualquer cidadão brasileiro. Além disso, falta na estrutura da organização, espaço no qual o tratamento técnico possa ser realizado de forma independente de tipos e suportes de acervos, ou seja, no qual técnicos possam padronizar procedimentos, analisá-los e discuti-los.

Na situação apresentada não foi possível determinar se o processo de indexação é exaustivo ou específico. Percebe-se, entretanto, uma tendência à exaustividade na seleção dos assuntos.

Já a linguagem adotada enfrenta sérios problemas de restrição por conta do sistema integrado não apresentar busca *booleana*, o que dificulta a utilização da indexação pós-coordenada. Ainda assim, muitos termos são inseridos ansiando-se a pós-coordenação, fato que pode resultar em dificuldade na recuperação da informação. Há uma enorme capacidade de revocação com precisão mínima. Muitos termos são registrados em linguagem livre. Em uma instituição na qual o volume documental é imenso e diversos fundos são compartilhados por diferentes equipes, este fato resulta na criação de diferentes termos para assuntos

semelhantes. O ideal para a instituição no momento, com o sistema na forma em que está, é a utilização da indexação pré-coordenada e da linguagem documentária.

Para que o Arquivo Nacional possa construir um vocabulário controlado próprio, uma série de novos procedimentos podem ser adotados, como a formação de um grupo de indexadores para discutir e analisar documentos e termos em conjunto, além de elaborar procedimentos para a realização da atividade, tais como: escolha de linguagem; limites para entrada de assuntos compostos e para inserções de termos de indexação; a avaliação do sistema para se conhecer o nível de satisfação do usuário; o levantamento dos assuntos centrais e periféricos dos acervos, para saber se a listagem de termos presente supre a necessidade dos usuários; por fim, revisão de todo o vocabulário controlado do SIAN e dos catálogos e sistemas das equipes; com análise termo a termo para iniciar o processo de montagem de cadeias e renques.

Este artigo anseia ser um instrumento na elaboração do aperfeiçoamento do sistema para que o mesmo responda de maneira eficiente o desejo cada vez maior da população de explorar todas as possibilidades de acesso à informação.

REFERÊNCIAS

APPOINTMENT. In: LIBRARY of Congress Online Catalog. Washington, DC, [20--?]. Disponível em: <<http://catalog.loc.gov/>>. Acesso em: 21 dez. 2012.

ARQUIVO NACIONAL (Brasil). **Acesso à informação**: institucional. Rio de Janeiro, [2012?]. Disponível em: <http://www.arquivonacional.gov.br/cgi/cgilua.exe/sys/start.htm?sid=129>. Acesso em: 15 out. 2012.

_____. **Dicionário brasileiro de terminologia arquivística**. Rio de Janeiro, 2005.

_____. **SIAN**: Sistema de Informações do Arquivo Nacional. Rio de Janeiro: Arquivo Nacional, 2001. Disponível em: <<http://capela/sian/Seguranca/Principal.asp>>. Acesso em: 20 jun. 2016.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676**: métodos para análise de documentos: determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992.

BIBLIOTECA NACIONAL (Brasil). **Terminologia de assuntos**. Rio de Janeiro: Biblioteca Nacional, 2000. Disponível em: <www.bn.br>. Acesso em: 21 jun. 2016.

BRASIL. Conselho Nacional de Arquivos. **NOBRADE**: Norma Brasileira de Descrição Arquivística. Rio de Janeiro: Arquivo Nacional, 2006.

BRASIL. Senado Federal. **VCBS**: Vocabulário Controlado Básico. Brasília: Subsecretaria de

Biblioteca do Senado Federal, [200-]. Disponível em:

<http://biblioteca2.senado.gov.br:8991/F/7GX5KSFYUSNMBXKNA3MATXQASDJJDMRUAALAI2LDP675VC8NF-01026?func=find-b-0&local_base=sen10>. Acesso em: 21 jun. 2016.

CARNEIRO, Marília Vidigal. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 221-241, set. 1985.

CONSELHO NACIONAL DE ARQUIVOS (Brasil). **NOBRADE**: Norma Brasileira de Descrição Arquivística. Rio de Janeiro: Arquivo Nacional, 2006.

FUJITA, Mariângela Spotti Lopes; RUBI, Milena Polsinelli. Elementos de política de indexação em manuais de indexação de sistemas de informação especializados. In: **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 8, n. 1, p. 66-77, jan./jun. 2003. Disponível em:

<<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/viewFile/375/193>>. Acesso em: 21 jun. 2016.

HOUAISS: Dicionário da língua portuguesa. Rio de Janeiro: Objetiva, 2004. Disponível em: <<http://capela/houaiss/>>. Acesso em: 21 dez. 2012.

LANCASTER, F. W. **Indexação e Resumos**: teoria e prática. Brasília: Briquet de Lemos, 2004.

SOUZA, Vanessa Inácio. **Indexação**: teoria e práticas do corpo indexador da Universidade Federal do Rio Grande do Sul. Porto Alegre, 2009.