

A INDEXAÇÃO E O CONTROLE DE VOCABULÁRIO EM ARQUIVOS: UMA INVESTIGAÇÃO NO ÂMBITO DOS ARQUIVOS PÚBLICOS ESTADUAIS BRASILEIROS

Jéssica Ferreira de Sousa Azevedo

Arquivista e Mestranda em Ciência da Informação pela Universidade Federal Fluminense
e-mail: azevedo.jessicafs@gmail.com

Rodrigo de Sales

Professor do Departamento de Ciência da Informação da Universidade Federal de Santa Catarina
e-mail: rodrigo.sales@ufsc.br

Resumo: No ambiente arquivístico, uma maneira de aperfeiçoar o acesso aos documentos é o aprimoramento do processo de recuperação da informação. Dentre as perspectivas possíveis de se abordar a questão da recuperação da informação, a indexação ocupa lugar de destaque, uma vez que viabiliza a recuperação de documentos por meio da extração de termos que operam como pontos de acesso aos conteúdos dos mesmos. Tais termos, cuja função é atuar como descritores de conteúdos de documentos, podem ser provenientes de um vocabulário controlado que padroniza e facilita a busca por informações. O presente artigo apresenta os resultados de uma pesquisa que buscou analisar a utilização da indexação e do controle de vocabulário nos arquivos públicos estaduais brasileiros. Para tanto, adotou-se como procedimento metodológico um estudo de caso realizado nos arquivos públicos estaduais, viabilizado por meio de um questionário composto por questões abertas e fechadas, empregado para a etapa de coleta de informações. Tal procedimento permitiu compreender como os arquivos investigados abordam as questões teóricas e práticas atinentes à indexação e ao controle de vocabulário. Os resultados obtidos reiteram que, de fato, o processo de indexação e os instrumentos de controle terminológico ainda não é consenso nos arquivos públicos estaduais brasileiros estudados.

Palavras-chave: Indexação em Arquivos. Recuperação da Informação. Indexação. Vocabulário Controlado.



1 INTRODUÇÃO

Em um contexto de crescente produção documental, os arquivos têm um papel não somente de guarda de documentação, mas também de avaliação e de preservação da memória. Para que essa documentação seja eficientemente acessada, é necessário certo grau de controle nos processos de organização da informação e padronização das terminologias utilizadas para a representação dos fundos arquivísticos.

O presente artigo apresenta um estudo que teve como pressuposto o fato de que a indexação ainda é um assunto e uma prática bastante incipiente no fazer arquivístico, sendo feita, quando muito, de maneira intuitiva e não padronizada. Nesse sentido, formulou-se como problema de pesquisa a seguinte questão: como o processo de indexação vem sendo desenvolvido nos arquivos públicos estaduais brasileiros?

Desse modo, o objetivo geral da pesquisa foi investigar como a questão da indexação vem sendo abordada nos arquivos públicos estaduais brasileiros. Para alcançar tal finalidade, alguns objetivos específicos foram definidos: a) verificar se os arquivos públicos estaduais realizam o processo de indexação; b) constatar quais instrumentos são utilizados pelos arquivos investigados

para o processo de indexação; c) identificar como são selecionados os pontos de acesso aos documentos.

A relevância da pesquisa se sustenta na ideia de que o valor informativo do documento de arquivo tem função preponderante na medida em que as funções administrativas, jurídicas e históricas do documento só podem ser cumpridas mediante sua capacidade de informar. Não é descabido afirmar que a questão informativa do documento de arquivo, que aproxima a Arquivologia dos aspectos próprios da Ciência da Informação, é tratada ainda de maneira resistente na literatura arquivística mais tradicional. Entretanto, no Brasil, o acesso à informação pública já é uma exigência legal formalizada por meio da Lei n. 12.527 de 2011, e os arquivos públicos ocupam papel de destaque neste quesito. O processo de indexação e os instrumentos por ele criados e utilizados são fundamentais para a recuperação de conteúdos de documentos e os seus respectivos acessos.

Amparado por pesquisa bibliográfica para a delimitação de conceitos e definições a respeito da indexação em arquivos, este trabalho empregou um estudo de caso nos arquivos públicos estaduais brasileiros de modo a contextualizar a realidade nacional e constatar como a indexação em arquivos vem sendo abordada na esfera pública.

O artigo está estruturado por esta introdução e mais quatro seções. Na seção dois, explicitam-se conceitos e definições provenientes da literatura a respeito do documento de arquivo e o acesso à informação; na seção três, discute-se a questão da recuperação da informação em ambientes arquivísticos; na seção quatro, são apontados e discutidos os resultados obtidos pela pesquisa. Para finalizar, são expostas as considerações finais.

2 DOCUMENTO DE ARQUIVO E ACESSO À INFORMAÇÃO

A história da Arquivologia remonta à Antiguidade. Vestígios destes longínquos tempos “demonstram a existência de arquivos reais, religiosos, econômicos e privados no Oriente Próximo e na Ásia” (INDOLFO *et al.*, 1995, p. 9). Rousseau e Couture (1998, p. 29) atribuem “ao aparecimento da escrita [...] o nascimento dos arquivos e da arquivística, bem como as novas ocupações, entre as quais a de arquivista”. Com o passar do tempo, a Arquivologia foi se especializando enquanto disciplina e criando suas bases teóricas e metodológicas.

O conceito de arquivo, conforme conhecemos atualmente, vem desde a Segunda Guerra Mundial. Indolfo *et al.* (1995, p. 10) apresentam esse marco histórico como início de “uma expansão inédita da produção documental”. Schmidt (2012, p. 30) afirma que houve “um grande avanço científico e tecnológico pelo qual nem o fazer nos arquivos saiu ileso quanto menos o saber, que passa a ser cada vez mais requisitado”.

Cruz Mundet (2008, p. 231, tradução nossa) trata de um dos primeiros princípios fundamentais da Arquivologia: o respeito aos fundos. O princípio de respeito aos fundos significa “manter agrupados, sem mesclar com outros, os documentos de qualquer natureza procedentes de uma entidade, [...] respeitando a estrutura ou classificação própria da dita entidade”.

Tal princípio se expandiu e logo evoluiu para o princípio da ordem original dos documentos. Conforme Schellenberg (2006, p. 246), este princípio “estabelecia que os documentos de cada órgão deveriam ser mantidos, no arquivo de custódia, na ordem dada pelo serviço de registro do órgão, e não reorganizados por grupos de assuntos”.

Díaz e Ruipérez (2003) trazem o objeto de estudo da Arquivologia como sendo o arquivo, que, segundo os mesmos, seriam “conjuntos organizados de documentos”. Quando se pensa em Arquivologia, a primeira imagem que se tem é o documento. Mas há de se pensar que nem todo documento é arquivístico. Para definir um documento de arquivo, deve-se, primeiramente, entender suas especificidades.

Segundo Cruz (2013), os documentos de arquivo são assim considerados se possuem quatro características: a) precisam ter sido “recebidos, produzidos e/ou acumulados por pessoa física ou jurídica no desenvolvimento de uma atividade organizada, para a consecução de um determinado fim” (CRUZ, 2003, p. 12); b) caráter probatório: um documento é prova de uma atividade ou função e, posteriormente, pode adquirir “caráter legal ou informativo, para a pesquisa científica” (CRUZ, 2003, p. 13); c) organicidade (ou relação orgânica): os documentos devem estar sempre inter-relacionados de acordo com a tramitação necessária para cumprir sua função, devendo ser naturalmente acumulados em função desta e; d) originalidade: cada documento é único e insubstituível.

Sousa e Araújo Júnior (2017, p.49) são mais concisos e afirmam apenas que “o documento arquivístico nasce como resultado do cumprimento de uma atividade e é mantido como prova dela”.

O Conceito das Três Idades se baseia na ideia de que o documento possui um ciclo de vida dentro de uma instituição (entidade), constituído pelas: Fase corrente (primeira idade), Fase intermediária (segunda idade) e Fase permanente (terceira idade). Fase corrente pode ser definida, segundo a Lei n. 8.159 (1991, art. 8º, § 1º), como os documentos “em curso ou que, mesmo sem movimentação, constituam objeto de consultas frequentes”. Cruz (2013, p. 19) complementa afirmando que a fase corrente é composta por documentos “de uso exclusivo da fonte geradora, cumprindo ainda as finalidades que motivaram a sua criação”. Já na Fase intermediária, encontram-se os documentos “que, não sendo de uso corrente nos órgãos produtores, por razões de interesse administrativo, aguardam a sua eliminação ou recolhimento para guarda permanente” (LEI n. 8.159, art. 8º, § 2º). Enquanto aguardam, devem ser mantidos “em depósitos de armazenagem temporária” (CRUZ, 2013, p. 19). Por fim, fazem parte da Fase permanente “os documentos que já cumpriram

as finalidades de sua criação, porém, devem ser preservados em virtude do seu valor probatório e informativo para o Estado e para o cidadão” (CRUZ, 2013, p. 19). A Lei n. 8.159 (1991, art. 8º, § 3º) complementa que eles possuem valor histórico, devendo ser preservados definitivamente. É importante salientar que “essas fases não são descontínuas ou estanques” (SOUSA; ARAÚJO JÚNIOR, 2017, p. 52).

Quanto à pesquisa que usualmente é feita pelos usuários, Sousa e Araújo Júnior (2017, p. 50) afirmam que “enquanto o documento tem valor primário, observa-se que a busca é feita pela informação que está contida no documento [e já] no arquivo permanente prevalece a pesquisa contextual, fato que não aponta para uma exclusividade, mas demonstra uma tendência”.

A Constituição Federal de 1988 apresenta em seu art. 5º, inciso XIV, que é “assegurado a todos o acesso à informação e resguardado o sigilo da fonte”. A Lei n. 8.159, de 8 de janeiro de 1991, traz no seu art. 4º a afirmação que “todos têm direito a receber dos órgãos públicos informações de seu interesse particular ou de interesse coletivo ou geral, contidas em documentos de arquivos”. Desse modo, o direito de acessar, irrestritamente, os documentos de arquivo são amparados legalmente tanto pela Constituição Federal, quanto pela referida Lei.

Schellenberg (2006, p. 162) afirma que “a atividade de tornar os documentos acessíveis é, sem dúvida, a mais importante de todas as atividades executadas por um arquivista”. O autor define acesso como o fornecimento dos documentos ou reproduções, bem como prestação de “informações relativas aos documentos ou neles contidas, ao governo e ao público” (SCHELLENBERG, 2006, p. 162).

Segundo o Dicionário Brasileiro de Terminologia Arquivística (2005, p. 19), acesso pode ser definido como “possibilidade de consulta a documentos e informações”, assim como “função arquivística destinada a tornar acessíveis os documentos e a promover sua utilização”. Peterson (2011, p. 217) complementa que o “acesso é a disponibilidade de arquivos para consulta como resultado tanto de uma autorização legal quanto da existência de instrumentos de pesquisa”.

No Brasil, temos uma cultura de sigilo, na qual por diversas razões históricas e culturais, “a tradição brasileira é [...], preferencialmente, de restrição” (RODRIGUES, 2011, p. 264). Em 2011, com a criação da Lei de Acesso à Informação (LAI), houve uma regulamentação do acesso, visto que a mesma prevê os procedimentos que “destinam-se a assegurar o direito fundamental de acesso à informação e devem ser executados em conformidade com os princípios básicos da administração pública”.

Martins (2011, p. 234) discorre sobre as obrigações do governo, que são “publicar e disseminar informações essenciais”, bem como “receber do público pedidos de informação e respondê-los”. Segundo o Art. 5º da LAI, a informação deverá ser “franqueada, mediante procedimentos objetivos e ágeis, de forma transparente, clara e em linguagem de fácil

compreensão”, ou seja, não deve ser utilizada uma linguagem técnica, e sim uma linguagem que possa ser acessível para o usuário.

É evidente a importância de um aparato legal para a garantia e regulamentação do acesso aos documentos. O acesso é fundamental, pois a partir dele é exercida a plena cidadania. Apesar disso, para que as informações contidas em um arquivo sejam recuperadas e o acesso seja amplamente dado, é necessária a existência de um instrumento de pesquisa (RODRIGUES, 2003). Schellenberg (2006, p. 161) complementa que tais instrumentos (chamados também de instrumentos de busca) “contêm informações que são obtidas através da análise de documentos”. Segundo o autor, “alguns instrumentos [...] mostram o caráter e o valor dos arquivos em função de sua proveniência; outros [...] mostram o caráter e o significado dos documentos em função do assunto de que tratam” (SCHELLENBERG, 2006, p. 161). O conceito de assunto ainda não é um consenso para o universo arquivístico. Conforme Gonçalves (1998, p. 23, grifo da autora), “o emprego do termo ‘assunto’ gera inúmeras confusões, sendo ora entendido como ‘função’, ora como ‘tema’. Seria conveniente que o uso do termo “assunto” fosse evitado, pois se refere, mais propriamente, ao conteúdo estrito de um documento”.

Ribeiro (2005, p. 85) define instrumento de pesquisa como “representações (imagens) da informação, que servem de intermediárias entre os pesquisadores e o produto informacional que é procurado”. Já o Dicionário Brasileiro de Terminologia Arquivística (2005, p. 108) traz como sinônimo o termo **instrumento de recuperação da informação** que é o “meio que permite a identificação, localização ou consulta a documentos ou a informações neles contidas”. A publicação também cita como tipos de instrumentos: guia, catálogo, inventário, índice, entre outros.

Cada um desses instrumentos descreve partes da instituição e/ou do acervo. Ribeiro (1996, p. 15) afirma que

o primeiro instrumento de pesquisa a ser elaborado será o guia, com a finalidade de orientar o utilizador e fornecer uma visão geral sobre os fundos existentes num arquivo. Só depois o guia [...] se passará a descrições mais completas – o inventário – e apenas em situações especiais se elaborarão catálogos, os quais pressupõem uma descrição documento a documento, necessariamente morosa.

Heredia Herrera (1995 *apud* Barros, 2017, p. 38) aponta os índices, tesouros e indexações como instrumentos de pesquisa auxiliares de descrição, ou seja, “instrumentos facilitadores e especializados dentro do universo dos arquivos”.

Além disso, para que possam ser elaborados tais instrumentos, a documentação precisa estar organizada. De acordo com Araújo Júnior, Sousa e Albuquerque (2015, p.44), “a organização da informação na esfera pública é uma condição para o pleno acesso à informação”. Smit e Kobashi (2003, p.11) afirmam que “para que um arquivo, qualquer um, preencha sua missão de ser o ‘arsenal da administração e o celeiro da história’, pressupõe-se sua organização”.

Segundo Lima e Álvares (2012, p. 35), “o objetivo principal da organização da informação é fornecer acesso à informação, por meio da estruturação dos elementos de organização do conhecimento”. Brascher e Café (2008, p. 05) complementam que “o objetivo do processo de organização da informação é possibilitar o acesso ao conhecimento contido na informação”.

É importante ressaltar que “a organização física dos documentos [...] não é, de modo geral, suficiente para garantir a recuperação da informação em diferentes situações” (SMIT; KOBASHI, 2003, p. 14), isto é, embora a disposição dos documentos em suas séries e fundos (de acordo com os princípios arquivísticos) seja a forma mais correta de se arquivar documentos, é necessário mais do que isso na hora de recuperar com precisão a documentação, conforme veremos mais adiante.

3 A REPRESENTAÇÃO DA INFORMAÇÃO EM ARQUIVOS

Araújo Júnior (2007, p. 65) define recuperação da informação como a “comparação do que se solicitou com o que está armazenado, bem como o conjunto de procedimentos que este processo envolve”. Com isso, localizam-se “documentos e itens de informação que tenham sido objeto de armazenamento, com a finalidade de permitir o acesso dos usuários aos itens de informação, objetos de uma solicitação” (SOUSA; ARAÚJO JÚNIOR, 2017, p. 49).

A Norma Brasileira de Descrição Arquivística, elaborada pelo Conselho Nacional de Arquivos (2006), possui em sua área de número 8 as diretrizes para descrição de pontos de acesso e indexação de assuntos. Este elemento, não existente na *General International Standard of Archival Description*, a norma ISAD (G), tem como objetivo

registrar os procedimentos para recuperação do conteúdo de determinados elementos de descrição, por meio da geração e elaboração de índices baseados em entradas autorizadas e no controle do vocabulário adotado. (CONSELHO NACIONAL DE ARQUIVOS, 2006, p. 59)

Smit e Kobashi (2003, p. 12) definem pontos de acesso como pontos preestabelecidos que nomeiam determinados agrupamentos de documentos. Estes pontos permitem o acesso direto à tais documentos dentro de uma massa documental. Os usuários precisam acessar a documentação pelos assuntos (ou pontos de acesso) apenas no começo de sua busca, para definir os fundos e as séries que serão relevantes. Sem esse acesso, “muita informação não será disponibilizada aos utilizadores e a pesquisa sobre os temas gerais, que são objecto dos seus estudos, sairá muito prejudicada” (RIBEIRO, 1998, p. 22).

A partir disso, podemos perceber o quanto a etapa de representação da informação é necessária em um arquivo (bem como em qualquer instituição que demande acesso à documentação). O produto obtido através da representação da informação é uma linguagem documentária (LD) que é, segundo Novellino (1996, p.38), “um instrumento de comunicação ao permitir que indexadores e

usuários partilhem um mesmo vocabulário”. Este vocabulário compartilhado é alcançado na medida em que as LDs cumprem seu papel mediador entre a linguagem natural e a linguagem especializada (LARA, 1993, p. 74).

Vale ressaltar que as linguagens documentárias derivam de uma linguagem natural, que pode ser entendida como uma linguagem livre e sem controle, utilizada pelo agente das instituições. Sales (2008, p. 34) faz a distinção entre linguagem natural e linguagem documentária ao afirmar que

a primeira é composta por palavras e seus diversos significados, ao passo que a segunda é composta por palavras que assumem o papel de termos, ou seja, restringem seus significados de acordo com um contexto ou de acordo com uma área específica de conhecimento.

Também de acordo com Sales (2008, p. 33), LDs são “linguagens formais e artificialmente criadas para auxiliar o processo de tratamento da informação responsável pela descrição de conteúdos documentais”. Elas são “construídas para indexação, armazenamento e recuperação da informação e correspondem a sistemas de símbolos destinados a ‘traduzir’ os conteúdos dos documentos” (CINTRA *et al.*, 2002, p. 33). Cabe salientar que o conceito de linguagem documentária é tradicionalmente aceito e consolidado na área da Biblioteconomia, ao passo que na Arquivologia ainda não se utiliza com frequência.

Quando, em um arquivo, se selecionam termos de um documento através da utilização de uma linguagem documentária, têm-se o processo de indexação. Este processo, que é parte da função arquivística descrição, conforme Sousa e Araújo Júnior (2017, p. 51) afirmam “em nenhum momento, as definições restringem o uso dessa função [descrição] a uma das idades documentais, isto é, aos arquivos permanentes.” Por isso, o processo de descrição pode ocorrer tanto no arquivo corrente, logo que o documento é criado, quanto no arquivo permanente, caso seja tratada uma massa documental acumulada. Através da indexação é possível atribuir pontos de acesso de acordo com a NOBRADE. Barros (2017, p. 37) afirma “que descrever é um processo que permite ao arquivista representar – no sentido mais denotativo do termo – todo o conhecimento que é mantido em seu arquivo”.

A indexação é fundamental nos arquivos, pois apesar de eles serem organizados de acordo com os princípios arquivísticos,

há utilizadores [...] que procuram informação sobre os temas mais diversos, seja para efeitos de investigação, seja por mera curiosidade, seja porque não conhecem os fundos e comecem a fazer uma pesquisa genérica antes de se centrarem numa documentação particular (RIBEIRO, 1998, p. 20).

Pinto (2016, p. 141) define indexação como a construção de “representações do conteúdo do documento de forma que sua inclusão possa ser recuperada em algum tipo de base de dados”. Rubi e Fujita (2003, p. 67) trazem a função primordial do indexador, “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”, ou seja, a importância do

indexador analisar corretamente um documento de forma a representar seu conteúdo de maneira mais acessível para o usuário.

Conforme ressaltado por Pinto (2016, p. 146), “os documentos de arquivo possuem especificidades que devem ser levadas em consideração na hora de indexar e, principalmente, distribuir as informações em seus campos específicos”. A autora ainda enfatiza a tradição existente de negação da aplicação das práticas biblioteconômicas em arquivos, visto as peculiaridades dos documentos de arquivo.

No momento da indexação de documentos arquivísticos, “o indexador deve analisar o documento como parte do fundo e como organismo independente” (PINTO, 2016, p. 147), ou seja, deve ser considerado não só o documento, como também o seu contexto de produção, história administrativa, entre outros.

Segundo a NOBRADE (2006), os elementos que devem ser contemplados na indexação arquivística são: “nomes de entidades, eventos, áreas geográficas, períodos e assuntos tópicos” (p. 59). Entretanto, conforme ressaltado por Sousa e Araújo Júnior (2017, p. 51), “não há clareza de como trabalhar, na prática, com esses elementos como pontos de acesso e descrição de assuntos”.

Kobashi (2008, p. 02) traça a relação entre a indexação e o vocabulário controlado da seguinte forma: “um vocabulário controlado é utilizado para indexar documentos”, enquanto “indexar é caracterizar conteúdos de documentos por meio de descritores de um vocabulário controlado”.

De acordo com o procedimento 8.1.8 da NOBRADE (2006, p. 60, grifo nosso), “Para indexar os assuntos presentes nas unidades de descrição, extraia os termos em linguagem natural, transformando-os em linguagem controlada, de modo a construir um vocabulário controlado. Para isso, para cada termo de indexação, registre as fontes validadoras institucionais, como a *Library of Congress* e a Biblioteca Nacional, ou impressas, como tesouros especializados. Na medida do possível, registre também uma nota de escopo e uma definição do termo”.

Para normalizar o processo de indexação, a Associação Brasileira de Normas Técnicas (ABNT) publicou a NBR 12.676 (1992, p. 02), que dispõe “Métodos para análise de documentos – Determinação e seus assuntos e seleção de termos de indexação”. A Norma define indexar como: “ato de identificar e descrever o conteúdo de um documento com termos representativos dos seus assuntos e que constituem uma linguagem de indexação”.

Diante ao exposto, pode-se perceber que o processo de indexação é, ainda, pouco aplicado em ambientes arquivísticos. Apesar de não ser imprescindível para o funcionamento de um arquivo, a indexação é peça chave para a recuperação da informação, pois auxilia tanto o usuário quando o profissional arquivista. Com a elaboração de uma indexação e o emprego de um vocabulário adequadamente padronizado, é possível alcançar eficiência e eficácia para o acesso aos arquivos.

4 PROCEDIMENTOS METODOLÓGICOS

Após a compreensão conceitual dos aspectos atinentes à indexação e à recuperação da informação, optou-se pela abordagem de estudo de caso direcionado às instituições arquivísticas estaduais do Brasil, adotando como instrumento de coleta de informações um questionário com questões abertas e fechadas, que foi enviado via correio eletrônico aos arquivos estaduais de todas as unidades federativas do país. Com os dados levantados, foi realizada uma análise dos mesmos, com o intuito de chegar a um panorama nacional. Nesse sentido, o universo da pesquisa foi constituído pelos arquivos públicos estaduais brasileiros.

No quesito procedimental, realizou-se contato via e-mail solicitando aos responsáveis pela equipe técnica de cada arquivo o preenchimento do questionário (instrumento de coleta de informações). Tal questionário tinha como objetivo entender como é o funcionamento do processo de recuperação da informação dos arquivos analisados. Foram solicitadas informações sobre formas de recuperação, instrumentos utilizados, método de seleção de termos, entre outros. O questionário foi elaborado de modo a possibilitar a identificação referente aos seguintes aspectos: a) a forma de recuperação de documentos no arquivo; b) se o arquivo realiza o processo de indexação; c) que instrumentos são utilizados para a recuperação de documentos/informações.

Para contextualizar melhor os arquivos investigados, foram solicitadas uma identificação do arquivo e algumas informações básicas sobre eles, a saber: Qual a entidade tutelar do arquivo? Em que ano foi aberto ao público? Possui pessoal técnico qualificado? Caso tenha respondido SIM à questão anterior, especifique a quantidade de arquivistas, técnicos de arquivo e estagiários. Após esta etapa de identificação, partiu-se para a etapa mais voltada ao tema da pesquisa, cujas perguntas foram:

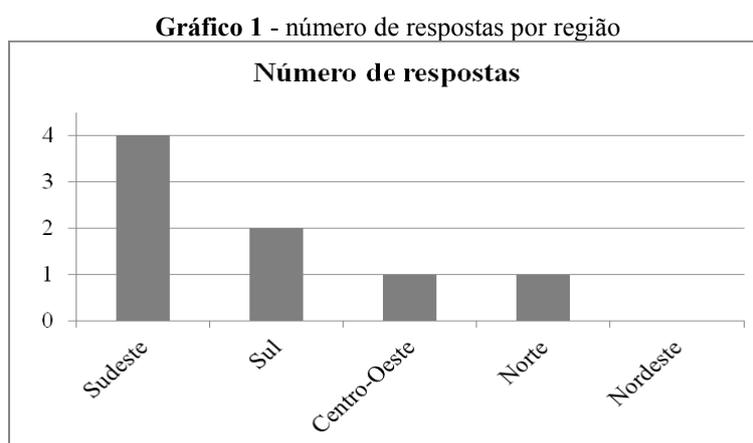
- 1) A recuperação da informação é feita, primordialmente, de qual forma? (Guias, inventários, catálogos, índices etc.)
- 2) Os instrumentos de pesquisa são manuais ou automáticos?
- 3) O arquivo desenvolve o processo de indexação dos documentos?
- 4) O arquivo recupera informações por quais pontos de acesso?
- 5) O arquivo trabalha com recuperação de informações por palavras-chave?
- 6) Caso tenha respondido à pergunta anterior com SIM, defina como as palavras-chave são selecionadas.
- 7) No acesso por palavras-chave, que tipo de linguagem (ou instrumento) é utilizado? Natural ou controlada?
- 8) Como os termos são selecionados para a recuperação da informação? Com base numa lista de termos elaborada durante a indexação. / Com base numa lista de termos já existente na instituição. / Com base em um tesauro. / Outras formas.

5 RESULTADOS ALCANÇADOS

Dos vinte e sete (27) arquivos estaduais que receberam o questionário enviado, apenas sete enviaram suas respostas via e-mail. Dada a localização geográfica, o questionário foi aplicado presencialmente no Arquivo Público do Estado do Rio de Janeiro (APERJ), totalizando, assim, oito questionários devidamente respondidos que compuseram o corpus de análise da pesquisa.

Dos oito arquivos respondentes, quatro Estados são da região sudeste (Rio de Janeiro, São Paulo, Minas Gerais e Espírito Santo), dois Estados são da região sul (Paraná e Santa Catarina), um Estado é da região norte (Acre) e um Estado é da região centro-oeste (Mato Grosso do Sul). Não foi recebida resposta alguma da região nordeste.

Os e-mails contendo o questionário foram enviados em 17 de julho de 2016. Foi estipulado um prazo para recolhimento dos dados de 30 dias, porém mesmo após o prazo não houve mais nenhum contato das instituições. A visita ao APERJ ocorreu em setembro de 2016.



Fonte: Elaborado pelos autores, dados levantados pela pesquisa (2018).

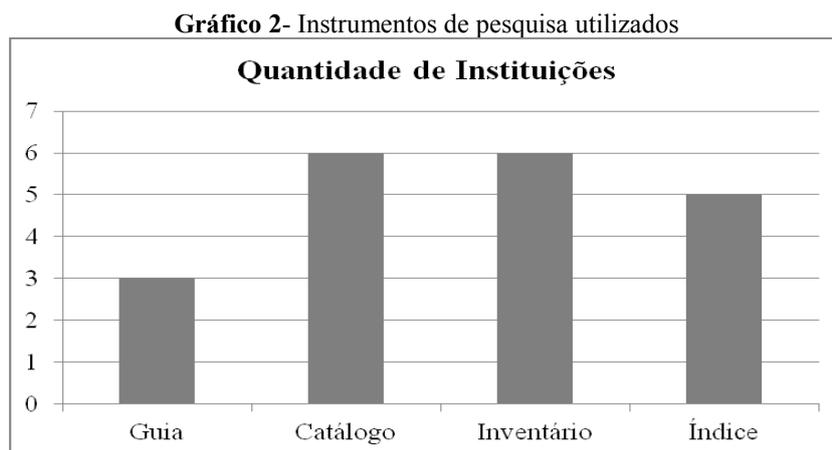
Para que a pesquisa ficasse mais completa, foi feito contato com um representante da Executiva Nacional de Estudantes de Arquivologia (ENEA). O responsável pela ENEA, que é estudante de Arquivologia no estado da Paraíba, informou que o referido estado não possui arquivo ativo por ora, e não poderia ajudar com relação à participação dos arquivos correspondentes aos outros estados nordestinos.

Vale destacar que, quanto à presença de arquivistas no quadro de funcionários, apenas três instituições confirmaram (Espírito Santo, Paraná e São Paulo), sendo que o restante dos arquivos destacou possuir técnicos, historiadores e bibliotecários como servidores.

Apresenta-se, a seguir, os resultados obtidos a partir da análise realizada sobre os questionários respondidos, respeitando cada item do questionário atinente ao processo de recuperação da informação.

1) *A recuperação da informação é feita, primordialmente, de qual forma?*

Três instituições declararam dispor de todos os instrumentos citados na pesquisa (guia, catálogo, inventário e índice). Dentre as demais instituições, uma afirmou que apenas o guia não era oferecido, enquanto outra instituição alegou não adotar nem guia nem índice em seus fazeres arquivísticos. Os outros três arquivos restantes afirmaram possuir apenas um instrumento de pesquisa, sendo um somente catálogo, outro somente índice e outro somente inventário. Dois dos arquivos investigados citaram ainda como instrumentos de pesquisa, bases de dados online.



Fonte: Elaborado pelos autores, dados levantados pela pesquisa (2018).

2) *Os instrumentos de pesquisa são manuais ou automáticos?*

Três arquivos afirmaram possuir apenas instrumentos manuais. As demais instituições apontaram a existência de ambos, manuais e automáticos.

3) *O arquivo desenvolve o processo de indexação dos documentos?*

Das oito instituições analisadas, seis desenvolvem o processo, ou seja, a maioria.

Gráfico 3 - Porcentagem de Instituições que desenvolvem o processo de indexação



Fonte: Elaborado pelos autores, dados levantados pela pesquisa (2018)..

4) *O arquivo recupera informações por quais pontos de acesso?*

Três instituições não responderam e uma afirmou que a definição dos pontos de acesso depende do acervo ou do fundo em questão. Uma das instituições afirmou estar desenvolvendo um vocabulário controlado para normalizar os pontos de acesso segundo a ISAD (G), ISAAR e ISDF, além da NOBRADE. As demais forneceram as seguintes respostas:

- a) Arquivo Público Mineiro (APM): Pelo nome dos órgãos ou instituições produtoras dos documentos, pela tipologia documental e por meio de índices onomásticos, toponímicos e de assuntos.
- b) Arquivo Público do Estado do Rio de Janeiro (APERJ): Assunto, local, tipo de documento, registro de autoridade, entre outros.
- c) Arquivo Público do Estado de Santa Catarina: Fundo, coleção, série documental, data-limite, notação (codificação).

5) *O arquivo trabalha com recuperação de informações por palavras-chave?*

Neste caso, utilizou-se palavras-chave entendidas como assunto dos documentos, ou seja, termos que representem tal assunto. Dois arquivos afirmaram não trabalhar com recuperação da informação por meio de palavras-chave.

Dos arquivos que responderam sim, três não explicitaram como essas palavras são selecionadas. Dos demais foram obtidas as seguintes respostas:

a) linguagem livre;

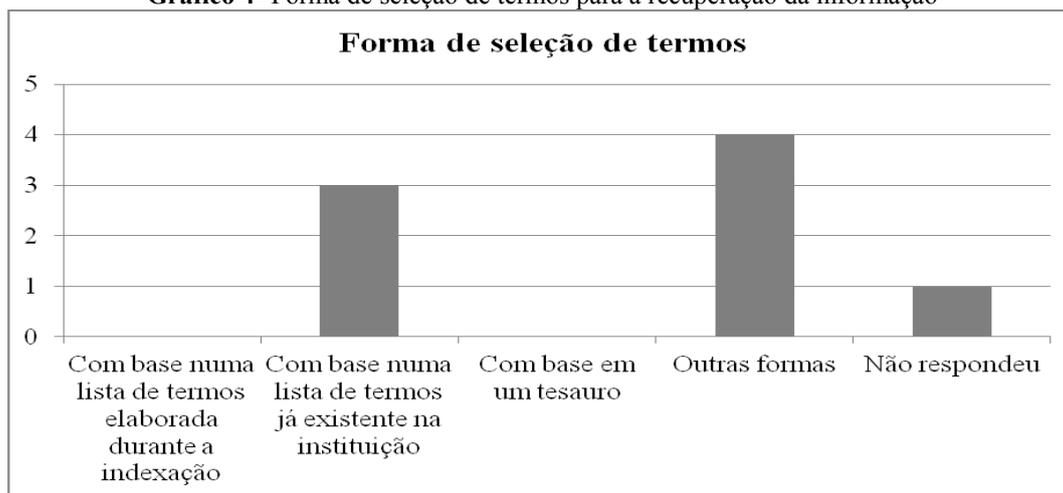
b) são selecionadas a partir de uma descrição sumária dos documentos e submetidas à confrontação com um vocabulário controlado para padronizar o uso dos termos;

c) são selecionadas de acordo com as características de cada acervo, segundo a NOBRADE. Além destas, uma instituição que afirmou não recuperar por palavras-chave acrescentou que utiliza uma tabela produzida pela instituição para codificação dos fundos, sub-fundos, coleções, séries, etc. Neste caso, não ficou explícito se essa tabela contém palavras-chave ou outro tipo de termo.

6) *Como os termos são selecionados para a recuperação da informação?*

Um arquivo não respondeu, três responderam que eram baseados numa lista de termos já existente na instituição e quatro arquivos disseram que era de outras formas. Nenhuma instituição investigada apontou a existência de um tesouro ou da elaboração de uma lista de termos no momento da indexação.

Gráfico 4- Forma de seleção de termos para a recuperação da informação



Fonte: Elaborado pelos autores. Fonte: dados levantados pela pesquisa (2018)

Houve casos de informações desencontradas nas respostas. Por exemplo, conforme apontado na questão e, em um mesmo questionário que afirmava não haver recuperação por palavras-chave, posteriormente, afirmava que as mesmas eram selecionadas por meio de linguagem natural. Casos como este não foram descartados. Foram aqui apresentadas as informações assim como foram recebidas, sem renúncia de nenhum dado.

Ficou evidente, a partir da análise das informações extraídas dos questionários, que não existe um controle efetivo de vocabulário nas instituições (arquivos pesquisados) e, as que desenvolvem o processo de indexação, nem sempre controlam os termos que serão utilizados no momento da recuperação dos documentos, dificultando, assim, a busca de informações nos arquivos.

A partir dos resultados obtidos na pesquisa, foi possível confirmar parcialmente o pressuposto inicial da mesma: não existe, no Brasil, uma cultura de controle de vocabulário e de indexação nos arquivos públicos estaduais. Apesar da importância e do papel fundamental da organização da informação (aqui abordada sob a ótica da indexação) para o aprimoramento da recuperação da informação e, conseqüentemente, da eficiência no acesso à informação, direito garantido por Lei Federal, a indexação e os instrumentos a ela ligados ainda são tratados de forma incipiente ou superficial, ou até mesmo negligenciados nas instituições pesquisadas. Tal fato, logicamente, acarreta na falta de padronização terminológica e de procedimentos de tratamento temático da informação, o que, conseqüentemente, dificulta o acesso à informação pública nos arquivos.

6 CONSIDERAÇÕES FINAIS

Conforme exposto, a indexação, processo que se caracteriza em definir palavras-chave que por sua vez atuam como pontos de acesso para recuperar documentos/informações, é um processo fundamental para que as instituições arquivísticas possam cumprir o papel de viabilizar o acesso à

informação, por meio do processo de recuperação de documentos. O presente estudo procurou compreender como a indexação vem sendo abordada no âmbito dos arquivos públicos estaduais brasileiros, uma vez que os mesmos se configuram em espaços centrais para o acesso à informação pública.

Retomando a questão que norteou este estudo, a saber: como o processo de indexação vem sendo desenvolvido nos arquivos públicos estaduais?, constatou-se que de fato a prática da indexação ainda está significativamente à margem do fazer arquivístico, ao menos na dimensão dos arquivos públicos estaduais. Um dos indicadores que sustentam esta afirmação está no pouco número de respondentes que participaram da pesquisa. Nesse sentido, o pressuposto levantado na introdução deste artigo – “a indexação ainda é assunto e prática bastante incipientes no fazer arquivístico, sendo feita, quando muito, de maneira intuitiva e não padronizada”, foi corroborado com o desenvolvimento deste estudo. Em outras palavras, não há, no universo arquivístico brasileiro (esfera pública), significativo investimento e aplicação do processo de indexação.

Pode-se observar que é necessário um maior investimento por parte dos arquivos na recuperação da informação, para que o usuário tenha seus objetivos alcançados com mais eficiência e rapidez. A indexação é fundamental para a recuperação de documentos, e poderia ter mais relevância no fazer arquivístico.

Conforme apresentado, este processo é amplamente discutido na Biblioteconomia, e deve ser adaptado para que possa ser utilizado também e em maior medida em arquivos. Barros (2016, p. 34) afirma que

A indexação enquanto parte de um processo de representação do conteúdo documental ainda é pouco utilizada no contexto teórico-metodológico dos arquivos e da Arquivística e seu uso muitas vezes ocorre de forma improvisada e incongruente. Contudo, os processos de indexação podem auxiliar no processo de análise e síntese que ocorre ao longo da descrição e representação de arquivos ou mesmo na definição de séries e identificação de conjuntos documentais [...].

Existe pouca produção literária sobre o tema na perspectiva arquivística. É necessário maior estudo e discussão para que sejam formuladas algumas diretrizes sobre a indexação em arquivos. O controle terminológico, discutido também neste trabalho, é uma prática raramente utilizada, e é fundamental para evitar perda de informações no momento da recuperação de documentos.

Enfim, é possível concluir que, no que se refere à prática arquivística em espaços públicos estaduais brasileiros, a ótica está ainda direcionada mais ao contexto de produção e organização dos documentos dos fundos arquivísticos, função fundamental na Arquivologia, do que no contexto de uso da informação, função predominantemente discutida no universo dos profissionais da informação. Assim, é possível perceber ainda certo distanciamento do fazer arquivístico no que diz respeito às discussões sobre o acesso à informação e o fazer do profissional da informação, dentre os quais, o arquivista deve ser parte integrante.

REFERÊNCIAS

ARAÚJO JÚNIOR, Rogério Henrique de. **Precisão no processo de busca e recuperação da informação**. Brasília: Thesaurus, 2007.

ARAÚJO JÚNIOR, Rogério Henrique de; SOUSA, Renato Tarciso Barbosa de; ALBUQUERQUE, Sérgio Farias de. Métodos, técnicas e instrumentos de organização e gestão da informação nas organizações. In: BATISTA, Dulce Maria Baptista; ARAÚJO JÚNIOR, Rogério Henrique de (org.). **Organização da informação: abordagens e práticas**. Brasília: Thesaurus, 2015.

ARQUIVO NACIONAL. **Dicionário brasileiro de terminologia arquivística**. Rio de Janeiro: Arquivo Nacional, 2005.

Associação Brasileira de Normas Técnicas (ABNT). **NBR 12676: métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação**. Rio de Janeiro, 1992.

BARROS, Thiago Henrique Bragato. A Indexação e a Arquivística: aproximações iniciais no universo teórico da organização e representação do conhecimento. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, v. 21, n. 46, p. 33-44, maio/ago., 2016.

BRASIL. Constituição (1988). Constituição da República Federativa do Brasil. **Diário Oficial da União**, Brasília, DF. 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm. Acesso em: 19 jul. 2016.

BRASIL. LEI 8.159, de 08 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. **Diário oficial da República Federativa do Brasil**. Brasília, DF. 1991. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/L8159.htm. Acesso em: 10 jun. 2016.

BRASIL. LEI n. 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei no 8.112, de 11 de dezembro de 1990; revoga a Lei no 11.111, de 5 de maio de 2005, e dispositivos da Lei no 8.159, de 8 de janeiro de 1991; e dá outras providências. **Diário oficial da República Federativa do Brasil**. Brasília, DF. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm. Acesso em: 18 jul. 2016.

BRASCHER, Marisa; CAFÉ, Lígia. **Organização da informação ou organização do conhecimento?**. Comunicação oral apresentada no IX ENANCIB ao GT-02 - Organização e representação do conhecimento, São Paulo, 2008.

CINTRA, Anna Maria Marques, *et al.* **Para entender as linguagens documentárias**. 2. ed. São Paulo: Polis, 2002.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). **NOBRADE**: norma brasileira de descrição arquivística. Rio de Janeiro: Arquivo Nacional, 2006.

CRUZ, Emília Barroso. **Manual de gestão de documentos**. 2. ed. Belo Horizonte: Secretaria de Estado de Cultura de Minas Gerais, 2013. Cadernos técnicos do Arquivo Público Mineiro. n. 3.

CRUZ MUNDET, José Ramón. **Manual de archivística**. Madrid: Fundación Germán Sanchez Rupieréz, 2008.

DÍAZ, Juan Carlos Galende; RUIPÉREZ, Mariano García. El concepto de documento desde una perspectiva interdisciplinar: de la diplomática a la archivística. **Revista General de Información y Documentación**. Madrid: 2003. n. 2.

GONÇALVES, Janice. **Como classificar e ordenar documentos de arquivo**. São Paulo: Arquivo do Estado, 1998. 37 p.

INDOLFO, Ana Celeste *et al.* **Gestão de documentos**: conceitos e procedimentos básicos. Rio de Janeiro: Arquivo Nacional, 1995. 49 p.

KOBASHI, Nair Yumiko. **Vocabulário controlado**: estrutura e utilização. Escola Nacional de Administração, 2008. (Texto preparado para subsidiar o debate sobre vocabulário controlado da Pesquisa Mapeamento para a reunião da Rede de Escolas do Governo).

LARA, Marilda Lopes Ginez de. Linguagens documentárias, instrumentos e mediação e comunicação. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 26, n. 1/2, p. 72-80, jan./jun. 1993.

LIMA, José Leonardo Oliveira; ALVARES, Lillian. Organização e representação da informação e do conhecimento. In: ALVARES, Lillian (org.). **Organização da informação e do conhecimento**: conceitos, subsídios interdisciplinares e aplicações. São Paulo: B4 Editores, 2012. 248 p. 21-48.

MARTINS, Paula Ligia. Acesso à informação: um direito fundamental e instrumental. **Acervo**, Rio de Janeiro, v. 24, n. 1, p. 233-244, jan./jun. 2011.

NOVELLINO, Maria Salete Ferreira. Instrumentos e metodologias de representação da informação. **Informação & Informação**, Londrina, v. 1, n. 2, p. 37-45, jul./dez. 1996.

PETERSON, Trudy Huskamp. Um Projeto Sobre Acesso. **Acervo**, Rio de Janeiro, v. 24, n. 1, p. 217-232, jan./jun. 2011.

PINTO, Mariane Costa. A indexação na recuperação da informação em arquivos: uma abordagem inicial. **Ágora**, Florianópolis, v. 26, n. 52, p. 141-150, jan./jun. 2016.

RIBEIRO, Cândida Fernanda Antunes. **Indexação e controlo de autoridade em arquivos**. Porto: Câmara Municipal, 1996.

RIBEIRO, Cândida Fernanda Antunes. **O acesso à informação nos arquivos**. Porto. 1998. Dissertação (Doutorado em Arquivística) – Faculdade de Letras, Universidade do Porto, Porto, 1998.

RIBEIRO, Cândida Fernanda Antunes. Organizar e representar informação: apenas um meio para viabilizar o acesso?. **Revista da faculdade de letras, ciências e técnica do património**, Porto, v. 4, p. 83-100, 2005.

RODRIGUES, Georgete Medleg. A representação da informação em arquivística: uma abordagem a partir da perspectiva da Norma Internacional de Descrição Arquivística. In: RODRIGUES, Georgete Medleg.; LOPES; Ilza Leite (orgs.). **Organização e representação do conhecimento na perspectiva da ciência da informação**. Brasília: Thesaurus, 2003. p. 210-229.

RODRIGUES, Georgete Medleg. Legislação de acesso aos arquivos no Brasil: um terreno de disputas políticas pela memória e pela história. **Acervo**, Rio de Janeiro, v. 24, n. 1, p. 257-286, jan./jun. 2011.

ROUSSEAU, Jean-Yves; COUTURE, Carol. **Os fundamentos da disciplina arquivística**. Lisboa: Publicações Dom Quixote, 1998.

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

SALES, Rodrigo de. **Tesauros e ontologias sob a luz da teoria comunicativa da terminologia**. Florianópolis. 2008. 164 f. Dissertação (Mestrado em Ciência da Informação), Universidade Federal de Santa Catarina, Florianópolis, 2008.

SCHELLENBERG, Theodore Roosevelt. **Arquivos modernos: teoria e prática**. 6. ed. Rio de Janeiro: FGV, 2006.

SCHMIDT, Clarissa Moreira dos Santos. **Arquivologia e a construção do seu objeto científico: concepções, trajetórias e contextualizações**. São Paulo. 2012. 320 f. Tese (Doutorado em Ciência da Informação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2012.

SMIT, Johanna Wilhelmina; KOBASHI, Nair Yumiko. **Como elaborar vocabulário controlado para aplicação em arquivos**. São Paulo: Arquivo do Estado, 2003. 56 p.

SOUSA, Renato Tarciso Barbosa de; ARAÚJO JÚNIOR, Rogério Henrique de. A indexação e criação de taxonomias para documentos de arquivo: proposta para a expansão do acesso e integração das fontes de informação. **Brazilian Journal of Information Science: Research Trends**. v. 11, n. 4, p.47-56, 2017.

THE USE OF INDEXING AND VOCABULARY CONTROL IN ARCHIVES: AN INVESTIGATION WITHIN THE BRAZILIAN STATE PUBLIC ARCHIVES

Abstract: *In the archival environment, one way to improve access to documents is to improve the information retrieval process. Among the possible perspectives to approach the issue of information retrieval, indexing occupies an important place, because it facilitates the retrieval of documents through the extraction of terms that operate as access points to their contents. Such terms, which function is to act as descriptors of document contents, may come from a controlled vocabulary that standardizes and facilitates the search for information. This article presents the results of a research that sought to analyze the use of indexation and vocabulary control in Brazilian state public archives. For that, was used as methodological procedure a case study, made possible through a questionnaire composed of opened and closed questions, sent to the Brazilian state public archives, used for the information gathering stage. This procedure allowed us to understand how the archives investigated address the theoretical and practical issues related to indexing and vocabulary control. The results point out that, in fact, the indexing process and the instruments of terminological control are still very incipient subjects in Brazilian archival practice, especially in public archives.*

Keyword: *Indexing in Archive. Information Retrieval. Indexing. Vocabulary Controlled.*

Originals recebidos em: 21/05/2018

Aceito para publicação em: 06/12/2018

Publicado em: 02/01/2019