
A INDEXAÇÃO E CRIAÇÃO DE TAXONOMIAS PARA DOCUMENTOS DE ARQUIVO: PROPOSTA PARA A EXPANSÃO DO ACESSO E INTEGRAÇÃO DAS FONTES DE INFORMAÇÃO

Indexing and creating taxonomies for records: a proposal to expand the access and integration of information sources

Renato Tarciso Barbosa de Sousa (1), Rogério Henrique de Araújo Júnior (N)

(1) Universidade de Brasília, Campus Universitário Darcy Ribeiro, Brasília-DF, renasou@unb.br (N) araujojr@unb.br

Resumo

O cenário encontrado na Administração Pública brasileira é de uma grande acumulação de informações e o fenômeno do desperdício que acomete outros aspectos da sociedade também se faz presente no campo informacional. O desafio que se coloca é o de buscar mitigar esse problema do desperdício. E a palavra-chave para isso se chama integração. A pergunta problema a ser respondida é a seguinte: é possível aproximar a gestão de documentos da gestão da informação para otimizar o uso da informação contida nos documentos de arquivo das organizações públicas e privadas brasileiras? O objetivo do trabalho é a verificação da possibilidade de fazer a integração da gestão de documentos e da gestão da informação em um único conjunto de esforços, a partir da recuperação das informações contidas nos documentos de arquivo. A pesquisa é de natureza teórica e aplicada, pois pretende, a partir da literatura sobre o tema, trazer à lume os elementos necessários que permitam a efetiva integração entre a gestão de documentos e a gestão da informação. O método de pesquisa empregado é a revisão sistemática da literatura referente à gestão de documentos, gestão da informação, taxonomia e indexação. A realização da proposta passa pela integração de três instrumentos: plano de classificação, lista de termos controlados e linguagem documentária (taxonomia).

Palavras-chave: Classificação; Taxonomia; Indexação; Recuperação Da Informação.

Abstract

The scenario found in the Brazilian Public Administration is a great accumulation of information and the phenomenon of waste that affects others in their society is also not present in the informational field. The challenge is the ideal place to mitigate this problem of waste. And the key word for this is called integration. A question problem to know how is a problem: is it possible to approach a document management of the information administration to optimize or use the information contained in the archival documents of the Brazilian public and private organizations? The objective of the research project is an analysis of the possibility of integrating document management and information management into a set of efforts, from the retrieval of information contained in the archive documents. The research is of theoretical and applied nature, upon request, from the literature on the subject of the elements that allow the effective integration between document management and information management. The research method used is a systematic review of the literature regarding document management, information management, taxonomy and indexing. The realization of the proposal through three instruments: classification plan, list of controlled products and documentary language (taxonomy).

Keywords: Classification; Taxonomy; Indexing; Information Retrieval

1 Introdução e Proposta

A questão do volume documental produzido e/ou recebido pelas organizações é um fenômeno contemporâneo. Na década de 1980, o arquivista francês Bruno Delmas fez uma interessante projeção sobre o estoque documental produzido em todas as épocas da história do homem. Segundo o autor, mais da metade dos documentos existentes no mundo tinham sido acumulados nos últimos trinta e sete anos, isto é, de 1950 a 1987. Seguindo a projeção feita por Delmas, podemos afirmar que, atualmente, mais da metade da massa documental acumulada mundialmente foi produzida e/ou

recebida nos últimos vinte anos. A produção crescente de documentos de arquivo parece ser uma característica comum às organizações contemporâneas (SOUSA, 2014).

Schellenberg (2005), afirmou na década de 1950 que o volume dos documentos oficiais tem sofrido um tremendo aumento nos últimos 150 anos, mas via esse crescimento como resultado do caráter das administrações modernas, que alcançou um nível de formalidade nunca antes visto. É claro, inclusive, que os meios tec-

nológicos de produção e reprodução de documentos facilitaram o desenho desse cenário.

Dessa forma, a produção diária de grandes volumes de documentos das instituições públicas ou privadas, posiciona a organização dos documentos na agenda da Arquivística contemporânea. Essa afirmação ganha vulto com a preocupação, cada vez maior, de nossas sociedades pela transparência das ações do Estado.

O balanço feito pelo Ministério da Transparência e Controladoria-Geral da União (CGU) dos cinco primeiros anos da implementação da Lei de Acesso à Informação demonstrou o crescimento da transparência ativa. Entretanto, esse aumento tem um limite, há um teto que pode ser compreendido como o nível de tratamento, organização e de recuperação da informação.

Uma maior sofisticação do instrumental técnico-científico é vital para possibilitar uma busca rápida e eficiente da informação. Além disso, precisamos integrar esforços, pois não há mais espaço nas organizações para criação e manutenção de inúmeros sistemas de informação que não se integram e têm como consequência a duplicação de custo e energia.

O problema da pesquisa refere-se a este contexto de produção e gestão de grandes volumes de documentos e informações com os quais lidam as organizações contemporâneas, ocasionando, na maioria das vezes, em desperdício que pode ser considerado um fenômeno comum entre as empresas no Brasil. A imprensa brasileira não se cansa de produzir matérias sobre o desperdício de material de construção, de alimentos, no transporte da safra, de energia e, mais recentemente, de água. Entretanto, fala-se muito pouco do desperdício das informações nas organizações públicas brasileiras.

As questões que motivam este trabalho são: Quantas informações são produzidas nas instituições públicas que poderiam ser utilizadas em outros órgãos para melhorar a tomada de decisão, possibilitando uma melhor atuação na prestação de serviços mais eficientes à população? Ou mesmo entre unidades político-administrativas da mesma organização pública? Quanto efetivamente de informação é produzida ou recebida por uma organização no desenvolvimento de suas atividades? Essas informações são reaproveitadas ou reutilizadas no processo decisório das organizações?

Entende-se que a ausência de respostas a essas perguntas impõem dificuldades para garantir direitos e deveres, a perda na qualidade dos serviços, a demora na tomada de decisão, a insegurança quanto às rotinas a serem seguidas, os prejuízos na preservação da memória e, além disso, o alto custo social e financeiro para o país e sua população.

Apontar para a importância estratégica dos recursos informacionais nas organizações contemporâneas é, atualmente, um lugar comum em um ambiente com uma infinidade de tecnologias da informação, que são

diariamente oferecidas pelo mercado. O que não é comum é o entendimento de que uma parte significativa desses recursos informacionais é acumulada pelas próprias organizações.

O arquivo é parte considerável da base de conhecimento da organização, que entendemos no âmbito deste trabalho, como o conjunto de informações utilizado pela organização para tomar suas decisões, garantir seus direitos e deveres e planejar suas ações presentes e futuras, ou como propõe Duranti (1994), o arquivo é o canal de comunicação que permite aos registros transitarem entre gerações. Por isso, deve ser neutro evitando induzir, privilegiar ou distorcer fatos, pois só assim a análise de boa qualidade será possível.

Parte dessa base de conhecimento, talvez a maior, esteja dentro da própria organização. Uma parcela é de informação/conhecimento registrado. Isso pode ser encontrado na biblioteca, no arquivo e nas bases de dados não institucionais. Outra parcela é o chamado conhecimento tácito, que quando for de alguma forma registrado tornar-se-á arquivo. Assim, podemos inferir que o arquivo compreende o principal estoque informacional da base de conhecimento da organização.

A maior parte da base de conhecimento formal produzida para a administração de uma organização encontra-se no seu arquivo, que por sua vez, constitui o acervo de informações que dá suporte aos processos e atividades que registram e servem de subsídio para o processo de tomada de decisão.

Parte importante do acúmulo de documentos (produzidos e/ou recebidos) representa mera formalidade administrativa ou é resultado da irracionalidade dos procedimentos de gestão adotados pelas organizações.

Todavia, os problemas não são tecnológicos e nem de ausência de informação, pois as limitações tecnológicas e a falta pontual de um item informacional não constituem o problema principal. Por outro lado, a classificação, isoladamente, não garante precisão na recuperação dos documentos de arquivo e das informações neles contidas. As estratégias de acesso à informação oferecidas pelos motores de busca na *internet*, aquela livre e sem nenhum tipo de estruturação, que ainda se mostram dominantes, têm influenciado bastante o comportamento do usuário em todos os sistemas de informação, inclusive nos arquivísticos. A solução parece residir no aperfeiçoamento contínuo da gestão de documentos, sobretudo no se refere ao uso da classificação combinada com uso de linguagens documentárias para apoiar o processo de indexação dos documentos. De acordo com Sousa e Araújo Júnior (2017), caracterizar o processo de classificação no âmbito da prática arquivística associando-o às funcionalidades de uma linguagem documentária que contribua para a melhoria do processo de busca e recuperação da informação é hoje uma variável importante na melhoria dos processos de representação da informação arquivística.

O desafio que se coloca é o de buscar mitigar esse problema do desperdício. E a palavra-chave para isso se chama integração. Integração de esforços e integração de sistemas de informacionais. Para Guedes e Araújo Júnior (2014), a clareza de quais são as informações essenciais, as unidades básicas que formam a essência da atividade desenvolvida, deve estar acompanhada de práticas de armazenamento dos documentos que possam ser úteis para trabalhos futuros. Esta assertiva torna clara a necessidade de buscar os usuários da informação produzida, a fim de criar sistemas de informação mais precisos no atendimento das necessidades informacionais dos seus usuários.

A pergunta problema que nos propomos a responder é a seguinte: é possível aproximar a gestão de documentos da gestão da informação para otimizar o uso da informação contida nos documentos de arquivo das organizações públicas e privadas brasileiras?

O objetivo da pesquisa é a verificação da possibilidade de fazer a integração da gestão de documentos e da gestão da informação em um único conjunto de esforços, a partir da recuperação das informações contidas nos documentos de arquivo. Recuperação da informação que se inscreve no processo de localizar 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 (Araújo Júnior, 2007).

Para atingir o objetivo proposto precisamos apresentar o entendimento sobre o documento de arquivo e a necessidade de trabalhar com duas funções arquivísticas: a classificação e a descrição.

O documento arquivístico é um artefato humano com pressupostos e características específicas. O ambiente e o conteúdo são delimitados e definidos pelo sujeito acumulador, que pode ser uma pessoa física ou jurídica. Então quando falamos de arquivo, estamos nos referindo a um conjunto finito de documentos acumulados, que tem suas fronteiras demarcadas pela missão do criador, no caso das instituições, e pela área de atuação, no caso das pessoas físicas. Ao contrário daqueles encontrados em bibliotecas, por exemplo, os documentos arquivísticos não constituem um conjunto formado em vista de uma finalidade específica: eles representam, mais que tudo, o produto da atividade do sujeito criador (Sousa, 2014).

Entender o modo como as instituições se estruturam e como executam suas funções e atividades é compreender como os documentos são acumulados. Ele é resultado de um ato desenvolvido e, na maioria dos casos, cotidianamente repetido. A gênese se dá quando a organização tem algo a cumprir, a provar, a determinar. Surge naturalmente como resultado das ações desenvolvidas pelo sujeito criador. Após o registro das informações em suportes (papel, mídia magnética, microfilme, películas fotográficas, películas cinematográficas

etc.), é necessário mantê-los pelos valores administrativos, técnicos, legais, fiscais, probatórios, culturais e históricos que possam conter (Sousa, 2014).

Se o documento é o resultado da atividade de uma pessoa física ou jurídica, podemos falar do caráter orgânico desse registro. A organicidade é revelada pelo interrelacionamento e pelo contexto de existência e de criação. Porém, nem todos os documentos orgânicos são de caráter arquivístico, pois essa qualificação é limitada em termos de suportes (convencionais ou eletrônicos). Por exemplo, é comum encontrar, principalmente nas indústrias, informações orgânicas tridimensionais que não são arquivísticas. O suporte, nesse caso, não permite o reconhecimento desse documento como de caráter arquivístico, apesar de entendermos que as características físicas não sejam os atributos mais seguros para definição do caráter arquivístico de um documento orgânico.

A acumulação dos documentos orgânicos de caráter arquivístico dá origem aos arquivos da organização. Sob essa designação são agrupados todos os documentos, seja qual for o seu suporte, produzidos e recebidos no exercício das suas funções. A título de exemplo, temos os relatórios, os projetos de pesquisa, as atas, os contratos, os convênios, as correspondências, os balanços financeiros, entre outros. Estes documentos podem ser encontrados, em geral, nos setores de trabalho. O documento arquivístico nasce como resultado do cumprimento de uma atividade e é mantido como prova dela. E, também, com o objetivo de decidir, de agir e de controlar as decisões e as ações empreendidas e, ainda, para efetuar pesquisas retrospectivas que ponham em evidência decisões ou ações passadas. Reduzindo, assim, a incerteza e tornando as tomadas de decisões mais seguras, a partir do aprofundamento do conhecimento da cultura institucional e do próprio processo decisório.

Portanto, a metodologia para a classificação de documentos arquivísticos deve estar assentada nesses traços e características que distinguem o objeto, não no seu formato (suporte) e nem na multiplicidade do conteúdo informacional (matéria ou assunto) que eles carregam. A partir da apropriação dos pressupostos da Teoria da Classificação, entendemos que os princípios de divisão do conjunto (arquivo) devem ser procurados na própria gênese documental, no processo de criação, nas relações estabelecidas e na reprodução das conexões originais.

Na Arquivística, a classificação foi reconhecida pelos canadenses (Couture e Rousseau, 1999) como uma das sete funções arquivísticas. E, hoje, podemos entendê-la como uma função matricial do que-fazer arquivístico. O coração das práticas arquivísticas, que para Lopes (1996) é formado pelas funções classificação, avaliação e descrição, tem a classificação como ponto de partida para a realização das outras duas funções.

Parte-se do pressuposto, neste trabalho, que a classificação de documentos de arquivo não pode limitar sua finalidade somente ao acesso aos documentos. A classificação de documentos de arquivo, diferentemente do que ocorre, por exemplo, nas bibliotecas, precisa fundamentar outras operações do que-fazer arquivístico. Encontramos três objetivos a serem atingidos com a classificação: manter o contexto de produção dos documentos; fundamentar a avaliação e a descrição; e permitir o acesso por meio do processo de busca e a recuperação da informação.

O documento de arquivo, do ponto de vista do potencial informativo, pode ser percebido em dois momentos: a informação que está contida no próprio documento e a informação contextual, que é obtida no agrupamento dos documentos resultados da mesma ação, rotina, tarefa ou atividade. Por exemplo, como é possível saber como foi o processo licitatório para compra de um computador a não ser pelo conjunto de documentos que foram produzidos e/ou recebidos no âmbito do desenvolvimento desse processo.

Enquanto o documento tem valor primário, observa-se que a busca é feita pela informação que está contida no documento. No arquivo permanente prevalece a pesquisa contextual, fato que não aponta para uma exclusividade, mas demonstra uma tendência.

A ordem em que são apresentados os objetivos da classificação de documentos de arquivo não representa um julgamento de valor. Os três objetivos têm o mesmo tamanho em termos de importância.

É evidente a importância da classificação de documentos de arquivo para permitir a localização da informação necessária. Por outro lado, temos percebido que, por mais consistente que seja o plano de classificação, não é possível atender todas as possibilidades de busca utilizando, apenas, esse instrumento. A verticalização da estrutura hierárquica, isto é, o excessivo desdobramento das unidades de classificação não facilita a busca ao documento/informação, muito pelo contrário, o dificulta.

Assim, defendemos a utilização de outros instrumentais para garantir acesso rápido e eficiente à informação, seja aquela contida em determinado documento ou aquela obtida de um conjunto de documentos. Para isso, precisamos considerar a utilização das técnicas de indexação para proporcionar mais pontos de acesso, facilitando a busca e a recuperação da informação.

A estrutura hierárquica é fundamental para o agrupamento físico e intelectual dos documentos de arquivo, garantindo, assim, uma informação importante para a avaliação documental. Não é razoável realizar a avaliação de documento a documento, ao agrupá-los nas unidades de classificação obtemos novas informações que vão subsidiar as tomadas de decisão em relação aos prazos de guarda e a destinação final desses documen-

tos. O agrupamento funcional dos documentos permite que na etapa de avaliação seja analisado não documento a documento, mas a função que os originou.

Portanto, o abandono da estrutura hierárquica de classificação dos documentos de arquivo levando em conta tão somente as estratégias de acesso aos documentos/informação, empobreceria a avaliação e a descrição de documentos, além de dificultar a manutenção do contexto de produção dos documentos. Assim sendo, os três objetivos mencionados só serão atingidos, por enquanto, com a estrutura hierárquica obtida a partir da aplicação de níveis de classificação que são representações dos vários níveis dos processos de trabalho (função, atividade, rotina, tarefa, transação).

A outra função que precisamos considerar no âmbito do problema proposto é a descrição. Ela é objeto de inúmeros estudos e em 2004, o Conselho Nacional de Arquivos lançou uma publicação elencando 232 referências sobre o tema. São textos em espanhol, português, inglês e francês. Em 2013, a então Associação dos Arquivistas Brasileiros lançou uma publicação intitulada: A produção científica da Arquivologia em classificação, descrição e recuperação: o estado da arte (Sousa e Araújo Júnior, 2013), com 64 indicações de leitura em português, inglês e espanhol sobre esse tema.

Os canadenses, quando discorrem sobre as funções arquivísticas, referem-se à descrição juntamente com a indexação. E, também, destacam dois fenômenos que marcaram a evolução dessa função: a proximidade com a Ciência da Informação, que favoreceu a elaboração de normas e o desenvolvimento das tecnologias da informação (Couture e Lajeunesse, 2014, p. 152).

O Conselho Internacional de Arquivos (2001), define a descrição arquivística como a elaboração de uma acurada representação de uma unidade de descrição e de suas partes, caso existam, por meio da extração, análise, organização e registro da informação que sirva para identificar, gerir, localizar e explicar documentos de arquivo e o contexto e o sistema de arquivo que os produziu. A *Society of American Archivists*, em seu glossário *on line*(1) define como “o processo de análise, organização e detalhes do registro sobre os elementos formais de um documento ou conjunto de documentos, como criador, título, datas, extensão e conteúdo, a fim de facilitar o trabalho de identificação, gestão e compreensão”.

No Brasil, o Arquivo Nacional (2005), em seu dicionário de terminologia, entende a descrição arquivística como o “conjunto de procedimentos que leva em conta os elementos formais e de conteúdo dos documentos para elaboração de instrumentos de pesquisa”. E instrumento de pesquisa como o meio que permite a identificação, localização ou consulta a documentos ou a informações neles contidas.

Verifica-se que, em nenhum momento, as definições restringem o uso dessa função a uma das idades documentais, isto é, aos arquivos permanentes. É evidente que nas fases encontramos usuários diferentes, ou melhor, novos usuários vão surgindo em cada uma das idades. Oliveira (2012, p. 41) aborda essa questão e defende que

A descrição arquivística tem sido apresentada na literatura da área de forma redutora em relação ao que de fato significa e representa, seja porque sua pertinência é tratada como restrita aos arquivos permanentes, seja porque é identificada com a própria elaboração dos instrumentos de pesquisa, ou seja, ainda porque mais recentemente, nas últimas décadas, está vinculada ao cumprimento de padrões que visam à elaboração de instrumentos de pesquisa ou dispositivos tecnológicos para acesso às informações sobre os acervos arquivísticos.

A autora aprofunda essa discussão

[...] as duas principais finalidades da descrição tradicionalmente identificadas na literatura da arquivologia, encontramos: controle e acesso. Seu eu compreendo que dentre as funções de descrição arquivística estão inseridos a elaboração de mecanismos de acesso; a explicitação dos relacionamentos entre os documentos de arquivo e as funções e atividades que lhes deram origem; e ainda o controle do acervo, é possível perceber que, em todas as fases documentais (corrente, intermediária e permanente), a implementação de um processo descritivo é fundamental, uma vez que controle e acesso são necessários e relevantes. (Oliveira, 2012, p. 41)

Portanto, trabalhamos nesta pesquisa com a ideia de que a descrição é uma função que passa por todas as idades documentais, ou melhor, que pode ser utilizada enquanto os documentos tenham valor primário ou secundário.

A Norma brasileira de descrição arquivística (Nobrade), publicada em 2006, que foi baseada na Norma geral internacional de descrição arquivística (ISAD-G), do Conselho Internacional de Arquivos, definiu 28 elementos de descrição disponíveis. Esses elementos estão distribuídos em oito áreas, que são as seguintes: identificação, contextualização, conteúdo e estrutura, condições de acesso e uso, fontes relacionadas, notas, controle da descrição e a novidade da norma brasileira em relação à norma internacional, que foi a área de pontos de acesso e descrição de assuntos, onde se registram os termos selecionados para localização e recuperação da unidade de descrição. (Conselho Nacional De Arquivos, 2006).

Os elementos considerados, pela Nobrade, como obrigatórios, são os seguintes: código de referência, título, data (s), nível de descrição, dimensão e suporte, nome do produtor e condições de acesso (Conselho Nacional De Arquivos, 2006). Observa-se que o Conselho Nacional de Arquivos estabeleceu como objetivo da norma facilitar o acesso e o intercâmbio de informações em âmbito nacional e internacional, mas os elementos con-

siderados obrigatórios não atendem ao primeiro objetivo, isto é, facilitar o acesso, pois nenhum deles “registra os termos para localização e recuperação da unidade de descrição”.

Duas áreas da Nobrade fazem referência às preocupações quanto à busca pelo conteúdo dos documentos, são elas: “Conteúdo e Estrutura” e “Pontos de Acesso e Descrição de Assuntos”. A primeira tem como regra informar o âmbito (contexto histórico e geográfico) e o conteúdo (tipologia documental, assunto e estrutura da informação) da unidade de descrição. A segunda, Pontos de Acesso e Descrição de Assuntos, 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. Neste caso a regra é identificar os pontos de acesso que exigirão maior atenção na geração de índices e realizar a indexação de assuntos de maneira controlada sobre elementos de descrição para a pesquisa. (Conselho Nacional De Arquivos, 2006).

A indexação de assuntos, segundo a Nobrade, deve contemplar os nomes de entidades, eventos, áreas geográficas, períodos e *assuntos tópicos*. Cunha; Cavalcanti (2008), definem como assunto tópico o termo geral, nomes de eventos ou objetos, atribuído a um registro, com a finalidade de proporcionar acesso ao mesmo. Não há clareza de como trabalhar, na prática, com esses elementos como pontos de acesso e descrição de assuntos.

Em outra importante referência para o trabalho arquivístico, o e-ARQ Brasil que é o modelo de requisitos para sistemas informatizados de gestão arquivística de documentos - SIGAD (Conselho Nacional De Arquivos, 2011), encontramos em duas das sete operações técnicas do sistema, captura e pesquisa, localização e apresentação dos documentos, a preocupação com a descrição arquivística.

A captura consiste em declarar um documento como um documento arquivístico, incorporando-o ao sistema de gestão arquivística por meio das seguintes ações: registro; classificação; indexação; atribuição de restrição de acesso; e arquivamento.

Nas operações relacionadas à Pesquisa, Localização e Apresentação dos Documentos são definidas que um SIGAD tem que permitir a recuperação do documento por meio das formas de identificação implementadas, incluindo: identificador; título; assunto; datas; procedência/interessado; autor/escritor/originador; e classificação.

Para Cunha e Cavalcanti (2008), assunto é a matéria sobre a qual se fala ou se escreve, a matéria de que trata um documento, um elemento usado para descrever o conteúdo de um recurso informacional. Vocabulário controlado, palavras-chave ou frases podem ser utiliza-

dos para descrever o assunto ou conteúdo do documento.

Lembramos, ainda, que o Sistema Eletrônico de Informações (SEI), *software* utilizado na maior parte dos órgãos públicos federais, é muito frágil no que se refere à recuperação de informações. Trabalha, fundamentalmente, com a linguagem natural, que não garante eficácia, eficiência e segurança na busca à informação.

A partir dessas considerações, apresentamos a metodologia utilizada na pesquisa, os resultados e as considerações finais.

2 Metodologia

A pesquisa é de natureza teórica e aplicada, pois pretende, a partir da literatura sobre o tema os elementos necessários que permitam a efetiva integração entre a gestão de documentos e a gestão da informação, a partir de estratégias de recuperação das informações contidas nos documentos de arquivo. O método de pesquisa empregado é a revisão sistemática da literatura referente à gestão de documentos, gestão da informação, taxonomia e indexação. A revisão sistemática busca identificar estudos sobre um determinado tema em questão, aplicando métodos explícitos e sistematizados de busca e avaliar a qualidade e validade desses estudos, eliminando vieses (Costa Filho, 2016, p. 23-24).

A base empírica da pesquisa é formada por órgãos da Administração Pública Federal, instalados em Brasília, que têm grandes volumes documentais acumulados, uma acumulação cotidiana expressiva de informações e um papel importante na estrutura político-administrativa do Estado brasileiro.

Os principais pressupostos teóricos são os princípios da proveniência, isto é, o respeito à origem do acervo. Um conjunto documental acumulado (produzido e recebido) por um sujeito juridicamente reconhecido e que tenha missão, funções e atividades que lhe dão individualidade e características que o distinguem dos outros deve ser tratado como uma unidade e separado de outros conjuntos. Trata-se, conforme defende Bellotto (2004, p. 130).

Trabalha-se aqui com o conceito das três idades documentais, que é a operacionalização de outro conceito: o ciclo vital dos documentos. Jardim (2015, p. 29) entende que parece incontornável o reconhecimento da noção de ciclo vital de documentos quando nos detemos nas possibilidades teóricas e práticas da gestão de documentos.

As três idades documentais dispõem o arquivo em três fases: corrente, intermediária e permanente. Essas fases não são descontínuas ou estanques. Fazem parte de um mesmo processo, que se inicia no setor de trabalho da organização com a produção ou recebimento de informações registradas em suportes definidos (papel, meio

magnético, película fotográfica, película videográfica, entre outros) em decorrência do desenvolvimento de uma atividade e que são mantidos nesses locais, por um determinado período, pela possibilidade de uso administrativo ou técnico que possam ter e que são transferidas, com a diminuição ou extinção do uso citado, para um outro local (arquivo intermediário) onde devem ser mantidas por questões legais, jurídicas, fiscais e para a garantia de direitos e deveres. E que podem ser guardadas permanentemente pelo valor de testemunho da trajetória e da história da instituição.

Este trabalho fundamenta-se, ainda, no conceito de tipologia documental. Ela representa a união que se dá no momento do cruzamento entre a função administrativa e a espécie documental que induza ou que comprove o seu cumprimento. A tipologia documental é importante, também, para a classificação, pois adequa as unidades documentais às funções e atividades da instituição, e para a avaliação, porque as tabelas de temporalidade só atingirão seu objetivo se baseadas em uma identificação das espécies documentais e das respectivas funções que só a tipologia pode proporcionar.

A pesquisa sustenta-se, também, no conceito de gestão de documentos. O fenômeno da gestão de documentos (*records management* em inglês) buscava a eficácia e a economia de recursos. Outrossim, essa nova área tinha como escopo a intervenção nas fases iniciais, na produção e no uso dos documentos (Araújo, 2014).

A gestão de documentos propiciou à Arquivologia uma sinergia maior com as administrações. A visão anterior destinava grande atenção apenas aos arquivos permanentes. A partir de então, a maior parte da teoria arquivística passou englobar todo o ciclo documental, desde sua produção, para atender os interesses administrativos legais, até sua destinação final: guarda permanente, para atender à sociedade, à pesquisa; ou descarte, como forma de economia de recursos. (Alves, 2017, p. 39).

Trabalha-se, também, com o conceito de taxonomia, que surgiu, de acordo com Campos e Gomes (2007), "como Ciência das leis da classificação de formas vivas e, por extensão, ciência das leis da classificação. No ambiente dos sistemas de classificação, das ontologias, da inteligência artificial, é entendida como classificação de elementos de variada natureza". A taxonomia é, por definição, classificação, sistemática, completa as autoras.

Tradicionalmente, as taxonomias tiveram por função a classificação das espécies em botânica, e zoologia, adotando uma nomenclatura binária. O resgate da taxonomia nos sistemas de informação considera a unidade sistemática (*taxon*) não mais família, gênero, espécie, mas conceitos. Aqui, as classes se apresentam segundo uma ordem lógica, apoiada igualmente em princípios classificatórios.

O campo de teste para a verificação da viabilidade de associação de uma linguagem documentária (taxonomia) ao plano de classificação, está representado nos órgãos da administração pública brasileira e seus instrumentos de gestão arquivística, principalmente aqueles elaborados pelo Conselho Nacional de Arquivos (Conarq).

As etapas de verificação seguiram o seguinte roteiro:

- Realização de estudos para definir as bases teóricas para a elaboração de vocabulários controlados a serem aplicados em documentos de arquivo;
- Revisão de literatura de classificação de documentos de arquivo;
- Realização de estudos, a partir da teoria da classificação, para estabelecer os requisitos necessários à construção de instrumentos de classificação de documentos de arquivo;
- Identificação das propostas metodológicas para a elaboração de instrumentos de classificação de documentos de arquivo;
- Elaboração de metodologia para construção de vocabulário controlado em arquivo;
- Definição de um vocabulário controlado para as atividades-meio da Administração Pública Federal; e
- Vinculação do vocabulário controlado à estrutura do Código de Classificação da atividade-meio, elaborado pelo Conselho Nacional de Arquivos.

O uso da linguagem documentária como complemento da atividade de classificação de documentos, tem como objetivo padronizar os termos, palavras-chave ou descritores resultantes do processo de indexação, cumprindo a função de criar rótulos que, junto com o número de classificação dos documentos, representem pontos de acesso para a recuperação dos documentos demandados pelos usuários.

A opção pelo uso das taxonomias para a padronização dos termos no processo de indexação de documentos, está baseada na estrutura hierárquica de termos por camadas e pela similaridade da sua disposição lógica com as áreas funcionais previstas na estrutura organizacional das instituições. Essa constatação facilita a vinculação do código de classificação com a taxonomia, possibilitando classificar e padronizar a escolha dos descritores que melhor representem a essência conceitual dos documentos.

A taxonomia pode ser considerada como uma estrutura que possibilita classificar objetos, seres vivos, coleções de livros ou documentos em grupos ordenados hierarquicamente, a fim de possibilitar sua identificação, localização e acesso. Pode também ser definida como um sistema de classificação que apoia o acesso à in-

formação, permitindo alocar, recuperar e comunicar informações em um sistema, de maneira lógica.

Os objetivos das taxonomias, a seguir elencados, estão dispostos a partir de quatro variáveis: I. Como linguagem; II. Facilitadoras da organização das informações; III. Facilitadoras do acesso à informação e IV. Para gestão do conhecimento.

Finalmente, a taxonomia deverá ser usada como instrumento complementar à atividade de classificação de documentos, devendo ser pautada pelos seguintes passos:

1. Classificar o documento conforme o plano de classificação de documentos de arquivo;
2. Identificar no primeiro ou segundo nível de termos da taxonomia a área utilizada na classificação dos documentos;
3. Utilizar os termos dos diversos níveis da taxonomia como descritores (termos de indexação) em complemento à classificação dos documentos; e
4. Verificar a coerência da classificação combinada com a indexação nas dimensões funcionais e temáticas, respectivamente com os documentos.

Como regra geral para a classificação e a descrição dos documentos, a seleção dos descritores da taxonomia deve ser combinada com os diferentes níveis de termos, desde que seja preservada a unidade temática combinada com o código de classificação atribuído a cada um dos documentos processados.

3 Resultados

A importância do conceito de classificação para a Arquivística e o papel fundamental e matricial que ele tem para o fazer arquivístico, fica evidenciado e enseja esforços de perceber a classificação com um objetivo maior do que somente o de manutenção do vínculo arquivístico. O volume documental, as alterações ocorridas nas formas de acessar à informação, proporcionadas pelos motores de busca com o advento da *internet*, além das dificuldades de operacionalização da classificação dos documentos de arquivo nos ambientes de trabalho, limitam sobremaneira o papel da classificação no acesso aos documentos de arquivo. A defesa que se procura fazer aqui nesses apontamentos é que a classificação não pode responder mais sozinha pela importante e atualizada tarefa de busca e recuperação da informação contida nos documentos de arquivo.

O que se propõe então, é a da integração da gestão de documentos e da gestão da informação em um único conjunto de esforços, a partir da recuperação das informações contidas nos documentos de arquivo. A realização da proposta passa pela integração de três instrumentos: plano de classificação, lista de termos controlados e uso de linguagem documentária (taxonomia).

O plano de classificação sendo construído pelo método funcional, a lista de termos controlados a partir da função do documento e a taxonomia pelas áreas de atuação ou de negócio do sujeito acumulador.

A identificação do conteúdo do documento de arquivo deve considerar a elaboração de um modelo para representá-lo em sistemas informatizados de gestão de documentos de arquivo ou em qualquer outro instrumento que tenha como finalidade facilitar a busca e a recuperação do documento, por meio da tradução documentária, ou seja, da tradução da linguagem natural para a linguagem documentária.

Quanto à lista de termos controlados, ela seria construída por meio de um modelo de identificação e representação do conteúdo de documentos de arquivo, composto de três partes: identificação da atividade (razão de ser do documento); ações que incidem sobre a atividade (identificação dos termos que especificam a atividade); e complemento (dados adicionais e relevantes).

Na concepção do modelo, a identificação da atividade e as ações que incidem sobre a atividade, são obrigatórias e o complemento é facultativo. Estes itens de descrição devem apoiar a identificação dos conteúdos dos documentos e gerar, como consequência, os produtos documentários que permitirão compor metadados (classificação, termos de indexação e resumo), para possibilitar a recuperação das informações e dos documentos em um sistema ou base de dados.

A proposta do Modelo de identificação e representação do conteúdo de documentos de arquivo é um caminho possível para a efetivação de controle de vocabulário, ou seja, é um modelo que por meio da gênese documental, busca atender um dos aspectos importantes da atividade de descrição, que é a criação de pontos de acesso para facilitar a busca e a recuperação dos documentos a partir dos elementos assunto/conteúdo/sumarização.

O entendimento do formato e o subseqüente registro das características físicas dos conjuntos documentais (descrição), a apreensão dos conceitos e temas relativos aos assuntos dos documentos (classificação/indexação), podem ser considerados aspectos capitais para a organização de um sistema informatizado de gestão de documentos. Na Arquivística, a organização e a gestão da informação só serão efetivas para a recuperação da informação (acesso físico dos documentos), se estas duas vertentes forem integradas em um modelo de gestão estratégica da informação arquivística, onde haverá a concretização do vínculo arquivístico considerando o documento em suas dimensões capitais, ou seja, forma (diplomática) e conteúdo (informação).

A elaboração da taxonomia se dará por meio de um número previamente determinado de termos onde a relação hierárquica será definida do primeiro nível em relação aos demais. Para a seleção dos termos devem

ser usados alguns documentos importantes: nomenclatura das áreas funcionais da organização, bem como o seu organograma; regimentos internos de cada área funcional descritos no organograma da instituição; normas, portarias e regulamentação das atividades do órgão; e plano de classificação de documentos de arquivo da instituição.

A criação dos níveis hierárquicos de termos deve também refletir a preocupação em detalhar os grandes temas e temas específicos do âmbito de atuação da organização, as áreas funcionais com os respectivos macroprocessos, processos de apoio e processos finalísticos, assim como a natureza de classificação dos seus documentos de arquivo. Assim, a criação dos níveis da taxonomia deve ser ponderada por meio das seguintes diretrizes: flexibilidade de uso dos termos, por meio de uma oferta adequada de pontos de acesso, refletindo as áreas funcionais, temas relacionados às atividades de apoio e finalísticas, bem como a nomenclatura utilizada na classificação dos documentos da instituição; normalização e padronização dos descritores a serem empregados na indexação dos documentos; uso da taxonomia em conjunto com o código de classificação de documentos, objetivando criar mais pontos de acesso aos documentos, providência que facilitar a busca e recuperação dos documentos; e instrumento complementar às funções e aplicações do código de classificação de documentos da instituição.

A taxonomia deverá ser usada como instrumento complementar à atividade de classificação de documentos com a finalidade que Sousa e Araújo Júnior (2017), apresentam, ou seja, a taxonomia vinculada à classificação de documentos, visa estruturar e sistematizar o conjunto de assuntos correspondentes às atividades desempenhadas pelas áreas de trabalho de uma organização para garantir aos tomadores de decisão e usuários em geral, o acesso mais rápido aos documentos e às informações necessárias ao processo decisório da instituição informações necessárias ao processo decisório da instituição.

Como regra geral para a classificação e a descrição dos documentos, a seleção dos descritores da taxonomia deve ser combinada com os diferentes níveis de termos, desde que seja preservada a unidade temática combinada com o código de classificação atribuído a cada um dos documentos processados.

Em um modelo de aplicação da Taxonomia combinada à classificação dos documentos de arquivo, o instrumento deve trazer um aporte de termos de indexação para complementar a classificação dos documentos, a fim de criar pontos de acesso adicionais para a recuperação dos conjuntos documentais.

A inclusão, em um sistema informatizado de gestão de documentos, do plano de classificação, da lista de termos controlados e da linguagem documentária (taxonomia), concretizará a possibilidade de acesso rápido,

seguro e eficiente à informação, possibilitando o cumprimento integral da função social do arquivista quer seja na garantia do acesso aos acervos documentais, quer seja na manutenção e preservação de um estoque informacional que deve servir às pessoas e organizações. Com isso, é possível equacionar a preservação do vínculo arquivístico com uma maior eficiência na recuperação dos documentos e das informações arquivísticas.

O esforço a ser empreendido está em agregar novas formas de busca, respeitando o vínculo arquivístico permitindo que as várias formas de explorar o potencial informativo dos acervos (informação contida no próprio documento e informação contextual, obtida do conjunto de documentos acumulados em uma transação) sejam alcançadas.

O terceiro e último elemento é a taxonomia, que possibilitará uma transversalidade na busca e agrupamento temático das informações na organização, viabilizando uma pesquisa mais completa para o usuário e garantindo a gestão racional dos recursos informacionais

4 Considerações finais

Os resultados preliminares da pesquisa apontam que quando acrescentamos a taxonomia no momento do registro do documento de arquivo, ou melhor, na inclusão dos metadados em um sistema informatizado de gestão de documentos, ampliamos o conceito de proveniência. E, também, fugimos das limitações que, às vezes, a classificação funcional apresenta, apesar da pacificação dessa ideia existente na literatura arquivística, conforme destacam Couture e Lajeunesse (2014, p. 147).

As limitações da classificação funcional, principalmente no que se refere à recuperação das informações contidas nos documentos e que ficaram mais evidentes no ambiente digital, não a invalida. Autores verticalizaram essa fragilidade e apostaram em experiências como o plano de classificação misto, que se fundamenta nas funções e atividades no nível das grandes unidades de classificação (classes, divisões, séries ou subséries) e nos assuntos ou temas no nível do dossiê e da peça. Outras experiências apareceram, tais como: a classificação por facetas e a classificação no nível da peça. (COUTURE, LAJEUNESSE, 2014, p. 148-152).

Entretanto, esses autores esqueceram que a classificação na Arquivística não tem como objetivo somente a recuperação de documentos. Ela visa, também, a manutenção do vínculo arquivístico e a fundamentação de outras funções arquivísticas como a avaliação e a descrição.

A proposta esboçada neste trabalho partiu da ideia desses três objetivos da classificação e buscou, na taxonomia, mitigar as limitações da classificação funcional na recuperação de informações contidas nos documen-

tos de arquivo. Entende-se que ela, a proposta, permite uma busca vertical, horizontal e transversal dos documentos pelas várias funções desenvolvidas nas organizações públicas.

Essa taxonomia pode ser utilizada nos outros sistemas de informação e na própria estrutura das páginas de internet das organizações, permitindo, assim, um primeiro nível de integração dos vários sistemas de informação. Há um longo caminho a trilhar, a história está apenas começando.

Notas

- (1) SOCIETY OF AMERICAN ARCHIVISTS. *A glossary of archival and records terminology*. Disponível em: <<http://www2.archivists.org/glossary/terms/a/archival-description>>. Acesso em 01 nov. 2017

Referências

- Alves, Ítalo Henrique (2017). O processo administrativo no âmbito do programa de gestão de documentos. Brasília: Universidade de Brasília (2017) Dissertação de Mestrado.
- Araújo Júnior, Rogério Henrique de (2007). Precisão no processo de busca e recuperação da informação. Brasília: Thesaurus, 2007. 175 p.
- Araújo, Carlos Alberto Ávila (2014). Arquivologia, biblioteconomia, museologia e ciência da informação: o diálogo possível. Brasília: Briquet de Lemos Livros, 2014. 200 p.
- Bellotto, Heloisa Liberalli (2005). Arquivos permanentes: tratamento documental. Rio de Janeiro: FGV, 2005. 320 p.
- Brasil. Arquivo Nacional (1990). Cadastro nacional de arquivos federais. Rio de Janeiro: Arquivo Nacional, 1990.
- Brasil. Arquivo Nacional (2006). Norma brasileira de descrição arquivística. Rio de Janeiro: Arquivo Nacional, 2006.
- Campos, Maria Luiza de Almeida; Gomes, Hagar Espanha (2007). Taxonomia e classificação: a categorização como princípio. // Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (Enancib), 8 (2007) Anais. Salvador: Ancib, 2007.
- Conselho Nacional de Arquivos (2004). Descrição arquivística: referências bibliográficas. Rio de Janeiro: Arquivo Nacional, 2004.
- Costa Filho, Cássio Murilo (2016). Possíveis limitações do ciclo vital dos documentos ante ao pós-custodialismo: o modelo australiano records continuum como instrumento de elucidação. Brasília: Universidade de Brasília (2016) Dissertação de Mestrado.
- Couture, Carol; Rosseau, Jean-Yves (1999). Les fonctions de l'archivistique contemporaine. Quebec: Presses de l'Université du Québec, 1999.
- Couture, Carol; Lajeunesse, Marcel (2014). L'Archivistique à l'ère du numérique: les éléments fondamentaux de la discipline. Québec: Presses de l'Université du Québec, 2014.
- Cunha, Murilo Bastos da; Cavalcanti, Cordélia Robalinho de Oliveira (2008). Dicionário de biblioteconomia e arquivologia. Brasília: Briquet de Lemos, 2008.
- Duranti, Luciana (1994). The concept of appraisal and archival theory. // The American Archivist. Chicago 57: 2 (1994) 328-344.
- Guedes, William; Araújo Júnior, Rogério Henrique de (2014). O fluxo de informações na auditoria pública e a teoria matemática da comunicação. // Encontros Bibli: Revista Eletrônica de Bibli-

otecnologia e Ciência da Informação. Florianópolis 19: 40 (2014) 33-50.

Jardim, José Maria. Caminhos e perspectivas da gestão de documentos em cenários de transformações. // *Acervo*. Rio de Janeiro 28: 2 (2015) 19-50.

Lopes, Luís Carlos (1996). *A informação e os arquivos. teorias e práticas*. Niterói: Eduff, 1996.

Oliveira, Lúcia Maria Velloso de (2012). *Descrição e pesquisa: reflexões em torno dos arquivos pessoais*. Rio de Janeiro: Móbile, 2012.

Schellenberg, Theodore Roosevelt (2005). *Arquivos modernos. Princípios e técnicas*. Rio de Janeiro: Fundação Getúlio Vargas, 2005.

Sousa, Renato Tarciso Barbosa de; Araújo Júnior, Rogério Henrique de (Orgs.) (2013). *A produção científica da Arquivologia em classificação, descrição e recuperação: o estado da arte*. Rio de Janeiro: Associação dos Arquivistas Brasileiros, 2013.

Sousa, Renato Tarciso Barbosa de (2014). Alguns apontamentos sobre a classificação de documentos de arquivo. // *Brazilian Journal of Information Science: Research Trends*, Marília 8: 1-2 (2014) 1-24.

Sousa, Renato Tarciso Barbosa de; Araújo Júnior, Rogério Henrique de (2017). A classificação e o vocabulário controlado como instrumentos efetivos para a recuperação da informação arquivística. // *Vaquinhas*, Nelson; Caixas, Marisa; Vinagre, Helena (eds.). *Da produção à preservação informacional: desafios e oportunidades*. Évora: Publicações do Cidehus, 420-442.

Copyright: © 2017. Souza e Araujo. This is an open-access article distributed under the terms of the Creative Commons CC Attribution-ShareAlike (CC BY-SA), which permits use, distribution, and reproduction in any medium, under the identical terms, and provided the original author and source are credited.
