



UNIVERSIDADE ESTADUAL DA PARAÍBA
Centro de Ciências Biológicas e Sociais Aplicadas (CCBSA)
Campus V – Ministro Alcides Carneiro
Curso de Bacharelado em Arquivologia

JANECELY SILVEIRA DE LIMA

**A (RE) CONSTRUÇÃO DA ACESSIBILIDADE: indexação
automatizada em arquivos jurídicos**

JOÃO PESSOA – PB
-2010-

JANECELY SILVEIRA DE LIMA

**A (RE) CONSTRUÇÃO DA ACESSIBILIDADE: indexação
automatizada em arquivos jurídicos**

Trabalho de Conclusão de Curso (TCC)
apresentado ao Curso de Bacharelado em
Arquivologia, do Centro de Ciências
Biológicas e Sociais Aplicadas da
Universidade Estadual da Paraíba, em
cumprimento às exigências para obtenção
do grau de **BACHARELA**.

Orientador: Prof. Dr. Josemar Henrique de Melo - UEPB

**JOÃO PESSOA – PB
-2010-**

L752r

Lima, Janecely Silveira de.

A (re) construção da acessibilidade: indexação automatizada em arquivos jurídicos / Janecely Silveira de Lima. – 2010.

61f. : il. color.

Digitado.

Trabalho Acadêmico Orientado (Graduação em Arquivologia) – Universidade Estadual da Paraíba, Centro de Ciências Biológicas e Sociais Aplicadas, 2010.

“Orientação: Prof. Dr. Josemar Henrique de Melo”.

1. Indexação Automatizada. 2. Recuperação da Informação. 3. Arquivo Jurídico. I. Título.

21. ed. CDD 025.04

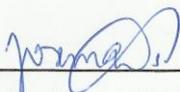
JANECELY SILVEIRA DE LIMA

**A (RE) CONSTRUÇÃO DA ACESSIBILIDADE: indexação
automatizada em arquivos jurídicos.**

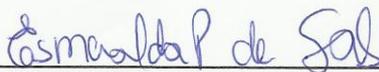
Trabalho de Conclusão de Curso (TCC)
apresentado ao Curso de Bacharelado em
Arquivologia, do Centro de Ciências
Biológicas e Sociais Aplicadas da
Universidade Estadual da Paraíba, em
cumprimento às exigências para obtenção
do grau de **BACHARELA**.

Aprovado em: 16/12/10

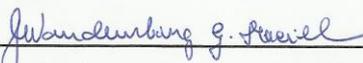
BANCA EXAMINADORA



Prof. Dr. Josemar Henrique de Melo – Orientador – UEPB



Prof. Ms. Esmeralda Porfírio de Sales - UEPB



Prof. Dr. João Wandemberg Gonçalves Maciel - UFPB

À minha mãe, por ela estar sempre ao meu lado nas horas que mais precisei, pela minha criação, pelos incentivos, e pelo seu exemplo de vida.

A Neto, que me incentivou sempre que precisei, e por mostrar que sou capaz de chegar onde eu quiser, basta querer!

DEDICO!

AGRADECIMENTOS

À Deus, por tudo que sou, e porque sem ele tenho a certeza que não chegaria ao alcance de meus objetivos. Foi ele quem fez eu me levantar nas horas mais difíceis, transformando-as em matérias-primas de aprendizado. A você, Deus, agradeço toda minha vida.

À minha querida mãe, que foi ao mesmo tempo pai e mãe, que não me deixou faltar nada, sempre fez tudo que estava ao seu alcance para que eu tivesse bons estudos, e se hoje estou me formando, devo a ela grande parte dessa conquista. Muito obrigada minha mãe, que eu amo tanto.

Agradeço imensamente a Neto, meu namorado amado, pelos inúmeros “puxões de orelha”, que sempre soaram como incentivo à minha vida acadêmica, inclusive na preparação deste trabalho. Desde que eu entrei na Universidade me mostrou um mundo novo, cheio de possibilidades, um mundo que eu não enxergava e me ensinou que o meu sucesso depende apenas de mim. Muito obrigada por toda a felicidade que você me proporcionou, companheirismo, principalmente nas horas difíceis, e pelo o amor que me foi concebido. Sim meu amor, eu fico com você para sempre!!!

Externo também agradecimento ao meu orientador, prof. Dr. Josemar Henrique de Melo, pela confiança, pelo olhar atento, pela orientação a mim concebida e por ter contribuído com sua experiência para a qualidade deste trabalho.

A banca, Prof. Dr. João Wandemberg e Prof. Ms. Esmeralda, por terem aceitado o convite de participar desta banca. Pois tenho muito apreço por ambos, admiro bastante seus trabalhos e por isso fiz questão que esses participassem desse momento tão importante para mim.

Agradeço todas as dificuldades que enfrentei; não fosse por elas, eu não teria saído do lugar. As facilidades nos impedem de caminhar. Mesmo as críticas nos auxiliam muito.

(Chico Xavier)

RESUMO

As Tecnologias Digitais de Informação e Comunicação (TDICs) vêm contribuindo para o desenvolvimento de processos e instrumentos aplicados as propriedades e o comportamento da informação em documentos. Com relação ao tratamento temático da informação, o uso dessas tecnologias permite que softwares auxiliem na construção, manutenção e gerenciamento das linguagens computacionais, e também na automação de atividades de indexação, dirigindo à otimização dos procedimentos e dos recursos aplicados na representação e recuperação da informação contida no documento. A partir desse contexto, o presente estudo visa avaliar o processo de indexação automatizada em um arquivo de escritório jurídico na cidade de João Pessoa (PB). Tais pretensões objetivam ainda apresentar um modelo de indexação automatizada feito nesse arquivo, além de verificar os procedimentos de busca e acesso a informação por parte dos usuários internos, mediante o uso da automatização da indexação. Em um sistema de recuperação da informação, os processos de indexação tornam-se elemento-chave para o seu desenvolvimento. O tipo de pesquisa utilizado tem o caráter exploratório, insere-se em um campo experimental e tem como abordagem o método qualitativo. Nas considerações finais, alvitram-se recomendações para futuras pesquisas no campo da indexação automatizada, uma vez que, a busca por uma ferramenta tecnológica deve ser vista como forma de se revitalizarem antigas formas de acesso à informação com o intento de melhorar ou estimular as metas de recuperação. Além disso, deverá agregar valores, inovando os programas e os processos já existentes.

Palavras-chave: Indexação Automatizada. Recuperação da Informação. Arquivo Jurídico.

ABSTRACT

The Digital Technologies for Information and Communication (DTIC) have contributed to the development of processes and tools used the properties and behavior of information in documents. With regard to the processing of thematic information, using these technologies allows software assist in the construction, maintenance and management of computer languages, and also in the automation of indexing activities, leading to optimization of procedures and resources used in the representation and retrieval of information contained in the document. From this context, this study aims to evaluate the indexing process in an automated file legal office in the city of Joao Pessoa (PB). Such claims still aim to present a model of automated indexing done in this file, and check the search and access to information by internal users through the use of automated indexing. In a system of information retrieval, the indexing processes become key to their development. The type of research has used the exploratory nature it fits in an experimental approach and its qualitative method. At the end, recommends setting up recommendations for future research in the field of automated indexing, since the search for a technological tool should be seen as a way to revitalize old forms of access to information with the intent to improve or enhance the goals recovery. Furthermore, it should add value through innovative programs and existing processes.

Keywords: Automated Indexing. Information Retrieval. Archive Legal.

LISTA DE ABREVIATURAS E SIGLAS

GED	Gestão Eletrônica de Documentos
KWIC	Keyword in Context
KWOC	Keyword out of context
LC	Linguagem Controlada
LE	Linguagem de Especialidade
LN	Linguagem Natural
PLN	Processamento de Linguagem Natural
SIC	Sociedade da Informação e do Conhecimento
SISA	Sistema de Indización Semi-Automático
SRI	Sistema de Recuperação da Informação
TDICs	Tecnologias Digitais da Informação e da Comunicação

LISTA DE FIGURAS

Figura 1 - Atalho para busca da informação no Excel	50
Figura 2 - Modelo adotado na realização da Indexação Automatizada de termos pelo uso do Excel	52

SUMÁRIO

1 INTRODUÇÃO	13
2 PROCEDIMENTOS METODOLÓGICOS	15
2.1 Caracterização da pesquisa	15
2.2 Objetivos	16
2.3 Problematização	16
2.4 Campo empírico	18
2.5 Instrumento de coleta de dados	18
3 OS SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO SUBSIDIANDO A GESTÃO INFORMACIONAL: UM ENFOQUE NOS ARQUIVOS DE ESCRITÓRIOS JURÍDICOS	20
3.1 A informação para o âmbito jurídico	20
3.2 As atribuições do arquivista em arquivos de escritórios de advocacia: da organização à recuperação da informação	22
4 INDEXAÇÃO: HISTÓRIA, CONCEITOS E TAXONOMIA	27
4.1 Breve histórico da indexação	27
4.2 A evolução dos métodos de indexação.....	31
4.3 A interdisciplinaridade da indexação com os processos de linguagens naturais e linguagens de especialidade: o uso de vocabulários controlados	35
5 A INDEXAÇÃO AUTOMATIZADA COMO FERRAMENTA DE RECUPERAÇÃO DA INFORMAÇÃO	39
5.1 Indexação automatizada: embasamento teórico e conjuntura atual	39
5.2 Vantagens e desvantagens da aplicabilidade da indexação automatizada em sistemas de recuperação da informação	42
6 A CONSTRUÇÃO E A APRESENTAÇÃO DO MODELO DE INDEXAÇÃO AUTOMATIZADA	45
6.1 A construção do modelo: o uso do Excel.....	45
6.2 A apresentação do modelo de indexação automatizada feito em um arquivo jurídico: suas interfaces	48

7 CONSIDERAÇÕES FINAIS	54
REFERÊNCIAS	56
APENDICE A – Entrevista	60

1 INTRODUÇÃO

Com a chegada e a consolidação da Sociedade da informação e do conhecimento (SIC), a humanidade sofreu mudanças significativas nas esferas sociais, políticas, econômicas e culturais, em um processo irreversível e cada vez mais veloz. Com essas mudanças o mundo passou a estar cada vez mais competitivo, e a informação, passou a ser insumo fundamental para as forças produtivas e determinantes das transformações nos modos de vida dos cidadãos e das organizações. Uma das causas dessas alterações está, sobretudo, marcada pelo advento e a consolidação das tecnologias digitais da informação e da comunicação (TDICs).

No tocante ao funcionamento das organizações, com o crescimento e o aperfeiçoamento das inovações tecnológicas, a economia passou a focar mais na informação e no conhecimento do que na força industrial para o desenvolvimento de suas ações. É neste período que surgem discussões relacionadas à automatização nos processos de tratamento da informação, não apenas para facilitar a gestão, mas para garantir resultados mais eficazes no tocante à organização da informação e conseqüentemente a tomada de decisões.

O arquivista, por sua vez, diante da virtualização da informação, proporcionada pelas TDICs, não poderá ficar alienado a esta questão, sendo assim, é preciso que o mesmo compreenda as possibilidades de aplicar as mais modernas técnicas aprimoradas no cotidiano das unidades informacionais/arquivos.

A partir desse contexto, o presente estudo tem por objetivo geral avaliar o processo de indexação automatizada em um arquivo de escritório jurídico na cidade de João Pessoa (PB). A indexação automatizada consiste num resultado de uma atividade intelectual para averiguar os valores informacionais atribuídos a um documento, ou seja, é feita por um profissional indexador que insere os termos extraídos do documento numa base de dados automatizada.

Especificamente, deseja-se verificar os procedimentos de busca e acesso a informação por parte dos usuários internos¹ do arquivo a partir da automatização da indexação e apresentar um modelo da indexação automatizada feito a partir do Sistema de Recuperação da Informação.

¹ Utilizamos o termo usuários internos para designar os indivíduos que trabalham no escritório e que buscam a informação no arquivo

Tais pretensões visaram discutir a utilização da automatização no processo de tratamento dos documentos, potencializando os procedimentos otimizados de busca e acesso da informação por parte dos usuários internos do arquivo, tornando ainda mais ágeis.

A hipótese central formulada sobre o objeto de estudo é a de que ***a indexação automatizada potencializa a recuperação da informação em arquivos.***

Dessa forma, como futura arquivista e preocupada com as questões atinentes à área, tais como, acesso, uso, representação, indexação e recuperação da informação, em espaços jurídicos – arquivo de escritório de advogados - far-se-á necessário contribuir com pesquisas voltadas a esses atributos. Esta pesquisa ainda irá contribuir significadamente para o meu desenvolvimento intelectual como profissional, tendo em vista que minhas experiências profissionais foram nestes ambientes.

Para a Arquivologia, esta pesquisa irá fortalecer mais ainda este campo do conhecimento que tem por objetivo organizar as informações seja qual for o suporte, necessitando de pesquisas que venham a garantir que grandes massas documentais sejam representadas e recuperadas de forma eficaz para o usuário, em particular nesta pesquisa os de escritórios de advocacia.

2 PROCEDIMENTOS METODOLÓGICOS

2.1 CARACTERIZAÇÃO DA PESQUISA

Quanto à classificação desta pesquisa, esclarecemos que a mesma se insere no campo do experimento. Segundo Dalen e Meyer (1974, p. 267),

a experimentação consiste em modificar deliberadamente a maneira controlada das condições que determinam um fato ou fenômeno e, em observar e interpretar as mudanças que ocorrem neste último. O Estado experimental estabelece as causas dos fenômenos, determinando qual ou quais são as variedades que atuam produzindo modificações sobre outras variáveis.

Em linhas gerais, de acordo com Triviños (2006, p. 112), as etapas de um processo de experimentação “parte de uma exata formulação do problema e das hipóteses que permitem uma delimitação precisa das variáveis que atuam sobre o fenômeno, fixando com exatidão a maneira de controlá-las”.

A pesquisa tem como abordagem o método qualitativo. A abordagem qualitativa compreende atividades de investigação que podem ser denominadas específicas (TRIVIÑOS, 2006). Ou seja, é denominada abordagem qualitativa aquela que a pesquisa faz parte à obtenção de dados descritivos mediante a relação direta e interativa do pesquisador com a situação objeto de estudo, valendo da razão discursiva.

Qualitativa é a abordagem que visa ponderar, analisar e interpretar dados à natureza dos fenômenos de estudos. Admite-se que esse tipo de pesquisa é a que tem por objetivo a interpretação dos fatos. Para tanto, Rodrigues (2007, p. 40) afirma que “forçoso é – ou seria – admitir, por via oblíqua, que tal pesquisa se limita ao esforço de compreensão. Isto porque compreensão é o esforço que visa a elucidar os significados dos fatos e ações para os sujeitos pesquisados”.

Com base nisso, o tipo de pesquisa desenvolvida pelo presente projeto tem o caráter descritivo, cujo desígnio é o estudo que apresenta informações, dados, inventários de elementos constitutivos ou contíguos ao objeto, dizendo o que ele é, do que se compõe em que lugar está localizado no tempo e no espaço, relevando periodicidades, indicando possíveis regularidades ou irregularidades, mensurando, classificando segundo semelhanças ou diferenças, situando-o conforme as circunstâncias (RODRIGUES, 2007, p. 29).

2.2 OBJETIVOS

O objetivo geral desta pesquisa pode ser proclamado da seguinte maneira:

- **Avaliar o processo de indexação automatizada em um arquivo de escritório jurídico na cidade de João Pessoa – PB.**

Esse objetivo pode ser desdobrado nos seguintes objetivos específicos, a saber:

- a) Verificar os procedimentos de busca e acesso a informação por parte dos usuários internos do arquivo a partir da automatização da indexação.**
- b) Identificar vantagens e desvantagens do uso da indexação automatizada.**
- c) Apresentar um modelo da indexação automatizada feito em um arquivo jurídico a partir do Sistema de Recuperação da Informação.**

2.3 PROBLEMATIZAÇÃO

Em todos os aspectos sociais, a humanidade teve os avanços tecnológicos como agentes determinantes nas mudanças em diversos campos de atividades. Hoje, com a consolidação da sociedade da informação e do conhecimento (SIC), passamos a estar inseridos em um mundo repleto de informações digitais, proporcionados pelos ambientes virtuais, a exemplo de bibliotecas, arquivos e centros de informação *online*.

Devido ao grande acúmulo de massas informacionais, proporcionadas pela pelo o momento atual, os métodos de indexar, armazenar e recuperar informações têm sido uma tarefa árdua para os sistemas manuais, havendo uma necessidade da

criação de novas tecnologias aplicadas a essas funções para facilitar o acesso e a busca pela informação desejada pelo usuário.

As Tecnologias Digitais de Informação e Comunicação (TDICs) vêm contribuindo para o desenvolvimento de processos e instrumentos aplicados ao tratamento da informação em documentos. Com relação ao tratamento temático da informação, o uso dessas tecnologias permite que *softwares* auxiliem na construção, na manutenção e no gerenciamento das linguagens computacionais, e também na automação de atividades de indexação, dirigindo à otimização dos procedimentos e dos recursos aplicados na representação e na recuperação da informação contida no documento.

O profissional da informação, nesse caso, o arquivista com conhecimentos de documentação jurídica, rotina e aspectos cognitivos envolvidos no trabalho de advogados deverá estar atento a essas questões, dirigindo seu olhar para essas mudanças quanto às migrações dos suportes da informação, uma vez contida em espaços físicos, hoje se configurando no âmbito virtual, proporcionados pelo o avanço das TDICs.

As automações dos arquivos, em especial os jurídicos, estão cada vez mais se tornando realidade, buscando melhorias no que diz respeito à organização de sua documentação, passando a ser uma exigência dessas organizações.

Em um Sistema de Recuperação de Informação (SRI), o processo de indexar, objeto de nosso estudo, torna-se como o elemento mais importante para sua eficácia, uma vez que o ato de indexar consiste em depositar dados que auxiliem a busca pela informação, ou seja, o ato de classificar e descrever os termos de um documento através de vocabulários controlados.

Os arquivos de escritório de advocacia contemplam uma infinidade de informações, por isso exige do profissional da informação uma descrição que forneça representações significativas dos documentos numa forma que preste a satisfação no momento da inclusão na base de dados.

Vale destacar que o propósito de uma busca numa base de dados é encontrar documentos que sejam úteis para satisfazer a uma necessidade informacional, além de evitar a recuperação de itens inúteis.

Com isso, entendemos que, com a automatização da indexação, essa poderá ser uma ferramenta que facilite o acesso e a busca de documentos em um arquivo

jurídico. Este aspecto leva à seguinte questão de pesquisa: ***Como entender e realizar o processo de indexação automatizada em um arquivo de jurídico?***

2.4 CAMPO EMPIRÍCO

A área jurídica, ao longo da última década, sofreu mudanças em suas características principalmente pela hegemonia da globalização que aumentou significativamente as áreas de atuação dos escritórios, em especial no mercado brasileiro.

Ultimamente, com o acréscimo de informação em instituições jurídicas, o advogado, passar a possuir também o papel de gestor, preocupando-se com atividades da produção da informação com valor agregado, recorrendo muitas das vezes ao arquivista na tarefa de buscar a informação correta para a tomada de decisões, baseado em experiências já realizadas e documentadas.

Para tanto, tomaremos como campo empírico da pesquisa o segmento jurídico, em especial escritórios de advocacia da cidade de João Pessoa – PB. Esclarecemos que será escolhido apenas um escritório, que está no mercado há mais de cinco anos, contando com o acervo todo organizado, em fase corrente e intermediária. O arquivo, embora pequeno, possui em média 1.000 processos de diversas causas. A estrutura do arquivo é toda acondicionada, preservando a integridade do documento. Atualmente o escritório possui quatro funcionários sendo eles: o gestor (advogado), dois estagiários de Direito e uma secretária.

2.5 INSTRUMENTO DE COLETA DE DADOS

Para este estudo foram utilizados procedimentos qualitativos, que segundo Minayo (2008, p. 22) “a metodologia inclui as concepções teóricas de abordagem, o conjunto de técnicas que possibilitam a apreensão da realidade e também o potencial criativo do pesquisador”. As técnicas que foram utilizadas na pesquisa são: entrevista e a observação participante.

Para Minayo (1993, p. 107), a entrevista é uma “conversa a dois, feita por iniciativa do entrevistador, destinada a fornecer informações pertinentes a um objeto de pesquisa”. A entrevista realizada junto ao gestor foi à estruturada, aquela em que o entrevistador segue um roteiro previamente estabelecido; as perguntas feitas ao indivíduo são predeterminadas (MARCONI; LAKATOS, 2008, p. 199).

A observação participante ou observação ativa, conforme Gil (1996) consiste no tipo de observação na qual existe a real participação do observador na vida da comunidade, do grupo ou de uma situação determinada. O observador assume o papel de um membro do grupo (GIL, 1996). Corroborando com este entendimento, Richardson (1989, p. 215) aponta que na observação participante “o observador não é apenas o espectador do fato que está sendo estudado, ele se coloca na posição e ao nível dos outros elementos humanos que compõem o fenômeno a ser observado”.

A observação foi feita durante a atuação no arquivo no período de estágio como arquivista, quando foi elaborado e aplicado o Sistema de Recuperação da Informação do arquivo, bem como a elaboração e a aplicabilidade da indexação automatizada.

3 OS SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO SUBSIDIANDO A GESTÃO INFORMACIONAL: UM ENFOQUE NOS ARQUIVOS DE ESCRITÓRIOS JURÍDICOS

Este capítulo tem por objetivo discutir sistemas de recuperação da informação em programas de gestão da informação, em especial, nos ambientes arquivísticos – escritórios de advocacia – focando nos arquivos. O estudo está dividido em duas partes. A primeira trata de discutir o conceito e a importância da informação para o âmbito jurídico. A segunda parte trata das atribuições do arquivista para organização da informação em arquivos jurídicos e a contribuição desse profissional para a elaboração e aplicação de um sistema de recuperação informacional no auxílio às buscas pela informação.

3.1 A INFORMAÇÃO PARA O ÂMBITO JURÍDICO

Com inúmeras mudanças ocorridas na sociedade nas últimas décadas, dentre elas, o avanço tecnológico, que a cada dia vem em crescimento desenfreado, à palavra informação é vital e tem habitado corriqueiramente a vida dos seres humanos, e também no funcionamento das instituições. Nesse contexto, Castells (2005, p. 39), concebe que:

uma revolução tecnológica concentrada nas tecnologias da informação começou a remodelar a base material da sociedade em ritmo acelerado. Economias por todo o mundo passaram a manter interdependência global, apresentando uma nova forma de relação entre a economia, o Estado e a sociedade em um sistema de geometria variável.

O sociólogo Manuel Castells considera que o processo complexo de emergência de uma nova sociedade pós-industrial, o eixo fundamental da economia passou a centrar-se na informação e não mais na força industrial para o desenvolvimento da sociedade. Ele ainda ressalta que, a revolução atual, refere-se às tecnologias da informação e que está ligada à expansão e reestruturação do capitalismo desde a década de 80 (CASTELLS, 2005).

Ainda seguindo o pensamento de Castells (2005, p. 53),

o processamento da informação é focalizado na melhoria da tecnologia do processamento da informação como fonte de produtividade, em um círculo

virtuoso de interação entre as fontes de conhecimentos tecnológicos e a aplicação da tecnologia para melhorar a geração de conhecimentos e o processamento da informação: é por isso que, voltando à moda popular, chamo esse novo modo de desenvolvimento de informacional, constituído pelo surgimento de um novo paradigma tecnológico baseado na tecnologia da informação.

A informação, segundo Souto (2006, p. 2), “passa a ser o principal fator de produção, capaz de interferir em qualquer contexto social”. Essa questão tem sido elemento de amplas discussões na área da Ciência da Informação e da Arquivologia, focalizando em múltiplos fatores e demonstrando várias propensões, como a promulgação rápida de informações na área social e na área técnica.

De modo similar, Barreto (1997, p. 2), assevera que a informação pode ser vista como “uma reconstrução das estruturas mentais do indivíduo através de sua competência cognitiva”. Essa ideia pode ser resumida, quando colocada a informação como um instrumento que modifica a consciência do homem e da comunidade.

Trazendo essa realidade para área jurídica, campo empírico de nossa pesquisa, a informação é algo que está sempre na ativa, pois ela não cai em desuso e sempre é reutilizada no desenvolvimento de atividades da jurisprudência, uma vez que, bem gerenciada, a informação dará suporte aos especialistas da área jurídica, servindo como base a tomada de decisões para situações antes já vivenciadas.

Baseado nas observações feitas, a grande massa documental existente nos arquivos jurídicos é algo que chama atenção, pois se não houver uma gestão eficiente que organize e recupere daquelas informações ali representadas, a quantidade de dados perdidos será incalculável.

Davenport (2002, p. 19) atesta que informação não é “facilmente arquivada em computadores e podem ter muitos significados em uma organização, além de estar baseada na maneira como as pessoas criam, distribuem e usam a informação”.

Na atualidade, a quantidade de informações produzidas é tão ampla, que nenhum indivíduo pode alimentar a esperança de estar ao corrente delas, mesmo que seja de uma infinita parcela. O problema que temos enfrentado é o de possibilitar que as pessoas que precisam de informações, tenham a sua obtenção com o mínimo de gastos (de tempo e de dinheiro) e sem que sejam assoberbadas por grandes quantidades de material irrelevante (FOSKETT, 1973).

Ao se tornar evidente que a informação é elemento imprescindível para tomadas de decisão no âmbito jurídico, as TDICs podem ser consideradas uma das principais ferramentas a serem utilizadas para aquisição nos resultados da qualidade e da produtividade, principalmente se considerarmos sua rapidez e sua confiabilidade.

Em qualquer entidade, os documentos de arquivo são a vida da própria instituição. A produção documental em todos os setores retrata todas as atividades. Os arquivos são a razão de ser da própria entidade, fonte permanente de informações (CASTRO, 1988). Para Rosseau e Couture (1994, p. 41) a “informação orgânica como estratégica é essencial para o desenvolvimento de qualquer organização”. A informação acrescenta mais valor quando permite a uma organização criar novas oportunidades, e o reconhecimento dos problemas que são acarretados através da demanda de informação. Para isso, é necessário que um profissional da informação, em especial o arquivista, mostre seu potencial e adentre nesta nova realidade.

Ainda, de acordo com Rousseau e Couture (1998, p. 63), qualquer organismo, “independentemente de tamanho, missão ou setor de atividade, para existir, funcionar e se desenvolver, necessita de informação. Nesse sentido, todos os membros do organismo têm necessidade de informação para desempenhar suas funções”.

De acordo com Lopes (2000, p.103), a natureza das informações arquivísticas é específica; trata-se de informações registradas em suporte definido, acumuladas por um indivíduo ou por um organismo que é, ao mesmo tempo, produtor e receptor.

Diante do exposto, um arquivo organizado em um escritório de advocacia é, sem dúvida, elemento impar, pois o profissional manterá as informações em fácil acesso, além de evitar contestações e duplicidade de informações proporcionadas, gerando segurança e confiabilidade das informações por parte dos clientes.

3.2 AS ATRIBUIÇÕES DO ARQUIVISTA EM ARQUIVOS DE ESCRITÓRIOS DE ADVOCACIA: da organização à recuperação da informação

Antes de darmos prosseguimento a esta sessão, a guisa de uma melhor visão da temática em tela, é preciso definir alguns conceitos importantes que utilizaremos

ao longo do texto como é o caso da palavra arquivo, designado comumente como um conjunto de documentos e/ou local de guarda.

Para este trabalho, utilizaremos o conceito de Alves (1993, p. 7), definido como um

conjunto orgânico de documentos, independentemente de sua data, forma e suporte material, produzidos e recebidos por uma pessoa jurídica singular ou coletiva, ou por um organismo público ou privado, no exercício de uma atividade e conservados a título de prova ou informação.

A grande massa documental acumuladas nas instituições e as novas tecnologias inseridas nas produções de arquivos e o crescente interesse na preservação da informação, levaram a criação da gestão de documentos, no intuito de buscar economia na guarda, no acesso, na recuperação e no uso da informação.

Não é preocupação nossa discutir as maneiras de como fazer uma gestão da informação adequada às exigências de um escritório de serviços jurídicos, mas, chamar a atenção que administrar, organizar, indexar e gerenciar a informação deverá ser uma frequente preocupação nesses escritórios, pois de acordo com Silva e Cunha (2002, p. 78),

nesta era do conhecimento, em plena sociedade da informação [e do conhecimento], perante a revolução informacional – o volume de informações geradas e consumidas aumenta a cada dia, num contexto em que a mudança tecnológica é a regra.

A importância de pesquisas realizadas na área jurídica por meio de arquivistas vem crescendo devido ao aumento da massa documental nesses arquivos. Esse aumento e a necessidade do uso da informação fizeram com que surgissem diversos problemas, a exemplo da perda da informação, que é causada pelo acúmulo documental sem que haja um profissional adequado, o arquivista, para controlar essa massa de documentos. É ingênuo afirmar que somente o arquivista proporcionará a gestão eficiente, sem o apoio do gestor e de ferramentas que favoreçam o controle da documentação.

Para controlar essas demandas, a fim de serem recuperadas com mais facilidade as informações, que derivam de ações doutrinárias, legislativas e jurisprudenciais, dos relatórios, dos dados dos clientes, carecem muito mais que um simples controle de informações, mas de pessoas capacitadas para que as informações possam se tornar mais acessíveis.

O arquivista passa a ter um papel fundamental neste ambiente, pois entendem que informação e conhecimento são pontos importantes para o desenvolvimento de qualquer organização na obtenção de benefícios como qualidade, produtividade e competitividade.

Dentre as principais competências de um arquivista estão: o alcance de uma economia de tempo, no que diz respeito à uma busca a determinado documento; à organização do acervo documental da instituição; à disponibilização de informações contidas no documento, seja qual for o suporte, aumentando a qualidade e agilidade na localização; à avaliação da autenticidade do material que fora coletado; o mantimento de atualizações de novos suportes da informação, dentre outras.

Dessa forma, nos escritórios de advocacia, a função do arquivista é fazer um trabalho intelectual e prático com a ajuda dos funcionários e dos gestores da organização nas tarefas de preservação da informação, contida nos processos, além de disponibilizar métodos de acesso e de busca da informação com mais rapidez e eficácia, pois são profissionais habilitados a desenvolverem métodos para a recuperação da informação.

Com a automatização, o uso dos Sistemas de Recuperação da Informação (SRI) tornou-se imprescindível no auxílio ao indexador humano, pois este é um sistema que norteia ações que visam desde a organização, a recuperação da informação, visando tornar disponível a informação a exemplo de catálogos, base de dados, etc. Os processos de produção de documentos ficaram mais rápidos, principalmente no que tange ao acesso e ao uso dos mesmos. A automação dos arquivos tornou a prática arquivística, mais ágil, porém, mais complexa.

Para a elaboração de um SRI digital, o arquivista deve está sempre se atualizando na questão das novas tecnologias, para que suas ferramentas em uso não se tornem obsoletas, e sempre que preciso realizar o *backup* para novos suportes, para não ocasionar a perda de informações. Sob o ponto de vista do sistema de recuperação da informação, a indexação é reconhecida como a parte, talvez, mais importante dentro dos procedimentos desenvolvidos e usados para o tratamento da informação, pois condiciona os resultados das estratégias de busca.

Os sistemas tradicionais utilizam união entre as palavras-chave da consulta do usuário e dos documentos do acervo a serem recuperados. Isso implica que o usuário tem de traduzir sua necessidade de informação em uma consulta composta

por palavras-chave, nesse caso, o nosso trabalho focaliza no gestor como usuário principal da recuperação da informação desejada.

Os termos da descrição do documento são armazenados em uma base de dados para busca e acesso à informação, por um profissional, neste caso o arquivista, que deve conhecer os fundamentos teóricos e técnicos do tratamento temático da informação.

Mas não é só isso o papel do arquivista perante aos SRI em um arquivo jurídico, eles também tem que conhecer o assunto do documento, não só para inseri-los nos sistemas, mas também, para receber o pedido de informação do usuário. Nesse processo, um usuário com uma determinada necessidade informacional, fará perguntas ao sistema por meio de estratégias de busca (LANCASTER, 2004).

O arquivista é um profissional que facilita a busca e o acesso à informação de forma eficaz e ágil. Há de convir que neste processo o arquivista, devido sua formação específica, é o profissional especializado que possui capacidades para lidar com a informação e o seu tratamento no documento. Esta mudança coloca em discussão o papel do arquivista, que, para sobreviver, devem se adequar e também participar ativamente de novos processos de gerenciamento da informação.

Nesse sentido, cabe, ainda, a esse profissional da informação o desafio da preservação da memória coletiva e individual, pois, com o grande aparato tecnológico existente, em meio a diversos sistemas informacionais, documentos eletrônicos e informações são perdidos com a mesma facilidade com que são gerados (INNARELLI, 2007).

Dessa maneira, diante das novas perspectivas das organizações, os escritórios de advocacia estão se adaptando às necessidades informacionais, precisando de mais rapidez no processo de recuperação da informação.

O termo automação/automatização pode ser pronunciado como a substituição do trabalho manual, por sistemas previamente programados que se auto-controlam, ou por sistemas que ainda necessitem do auxílio do homem para sua manutenção.

Para Negreiros (2007, p. 72) a automação em arquivos pode-se referir:

- À utilização de computadores para realizar as práticas arquivísticas (produções de índices, descrição, pesquisa, acesso, etc.);
- Ao processo de seleção de sistemas eletrônicos de gerenciamento de documentos arquivísticos, ou seja, a busca por um *software* que

gerencie a produção, uso e destinação dos documentos arquivísticos produzidos em uma organização;

- À inserção dos documentos eletrônicos no ambiente arquivísticos como um “sintoma” de automação das unidades arquivísticas; e
- Ao fornecimento de serviços relacionados aos documentos arquivísticos através da Internet.

Isso nos permite perceber que o arquivista, terá que realizar a automação dos arquivos junto aos profissionais da organização, já que partilham conhecimentos técnicos diferentes na definição dos termos a serem indexados no sistema, ou seja, no banco de dados – a plataforma que guardará os termos e informações extraídas dos documentos.

Diante do exposto, espera-se que as instituições responsáveis pela administração dos arquivos jurídicos, adotem nessa nova postura “ditada” pela SIC: a automação dos arquivos. Espera-se também que haja um trabalho multidisciplinar e um diálogo entre os profissionais de áreas distintas; e, sobretudo, a revisitação dos princípios arquivísticos dentro de uma nova estrutura tecnológica.

4 INDEXAÇÃO: HISTÓRIA, CONCEITOS E TAXONOMIA

Neste ponto procuraremos fazer um levantamento breve da história do surgimento da indexação, assim como a evolução de seus métodos - do manual até o automatizado. É feito ainda uma busca discursiva na interdisciplinaridade da indexação com os processos de linguagens naturais e linguagens de especialidade no uso de vocabulários controlados. Merece destacar que a partir da segunda sessão deste capítulo, será iniciado o confronto das respostas colhidas na aplicação da entrevista, diante ao nosso quadro conceitual de autores da temática.

4.1 BREVE HISTÓRICO DA INDEXAÇÃO

A indexação é uma prática muito antiga que surgiu a partir da elaboração de índices, remetendo-se às origens da escrita e da biblioteca. O índice é um instrumento de armazenamento e recuperação da informação, que se deu origem no momento em que o homem passou a se preocupar a torna acessível à informação contida em um documento. O avanço nos métodos e nas técnicas de construção de índices transformou este no mais importante instrumento de recuperação da informação, a mais poderosa ferramenta de busca (ROBREDO, 1994).

Antigamente, nas bibliotecas, já existiam listas dos documentos armazenados naquele local. Mas, com a necessidade de ordenar estas listas, estabeleceu-se uma profunda organização por assunto, causando mudanças no ato mecânico de elaborar índices, ou seja, inseriu-se um processo de análise do conteúdo dos documentos.

A indexação tem seu início com a história da bibliografia e com o processo de análise documentária, que é realizado mais profundamente em sequência ao aumento de publicações periódicas, que impulsionaram a necessidade de criação de mecanismos de controle bibliográfico em centros de documentação especializados.

No século XIV, a elaboração de inventários e/ou catálogos dos livros que os mosteiros possuíam, tornaram-se comuns e já obtinham um objetivo definido. É a partir deste momento que surgem guias para os livros.

Nesta época, os índices tinham o significado de uma lista de conteúdo, lista de resumos, dentre outros, não alcançando o significado que o índice tem nos dias atuais. A necessidade da elaboração de índices segundo Collison (1971)

apresentou-se logo que surgiu “a Bíblia inglesa e a indexação surgiu, então, em grande escala em 1737 com a compilação da primeira concordância completa da Bíblia por Alexandre Cruden”.

No século XVII, Cruden indexou a Bíblia, elencando quais os termos que deveriam entrar neste caminho, e formando verdadeiros padrões de clareza e coerência para a indexação. Na Alemanha, com a divulgação dos seus procedimentos, a indexação passou a contribuir na ideia de palavra-chave na representação de um item de sistema, que concebeu uma grande melhoria na busca de informações.

Foi no século XIX que a indexação deu uma engajada, aprimorando sua execução e atendendo às necessidades do público, que a buscavam para controlar grandes massas de documentos. Neste mesmo século e durante o século XX, a indexação ganha uma maior importância, sendo conceituada e elaborada cientificamente como um processo capaz de tratar a explosão documental que estava em percurso. No início do século XX, começam a aparecer as primeiras ferramentas de descrição de documentos com fins de pesquisa documentária.

Nos dias atuais, a indexação é caracterizada como o ato de descrever ou identificar um documento em termos de seu conteúdo. Com base nesse assunto, o desenvolvimento de técnicas de tratamento da informação está ligado à análise documentária como expansão do tratamento temático que permite a geração de resumos e a indexação.

Exibindo essa extensão e concretização, Lancaster (1993) afirma que

a indexação deve identificar, descrever ou caracterizar o conteúdo informacional de um documento para posterior exposição em linguagem natural ou em índices. Haja vista que, isso pode ser feito na estrutura de uma indexação sintética (de acordo com seu assunto) ou de uma indexação analítica (de acordo com os conceitos), em que também deve ser sempre analisada sua finalidade.

O principal objetivo de um serviço de indexação é assegurar a recuperação de qualquer documento ou informação no momento em que o usuário busca um assunto em um sistema de informações (STREHL, 1998). A indexação tem como objetivo essencial a seleção e/ou definição de termos (palavras ou expressões) que irão descrever o conteúdo de um determinado documento, sempre levando em consideração as necessidades informacionais do usuário.

O conceito de indexação surgiu a partir da elaboração de índices, mas na atualidade, ela está mais relacionada ao conceito de análise de assuntos. Para Chaumier (1980, p. 42) “a indexação é a parte mais importante da análise documentária”. De concórdia ao exposto, alguns teóricos ingleses e norte-americanos afirmam que a indexação é a própria análise documentária, ambas possuindo o mesmo objetivo, que é a representação do conteúdo informacional contido no documento para uma futura elaboração de índices e a análise de assunto como etapa inicial.

Por análise documentária compreende-se, no sentido apresentado por Silva e Fujita (2004, p. 138) a “área teórica e metodológica com o objetivo de tratamento temático de documentos, que abrange as atividades de indexação, classificação e elaboração de resumos, considerando as diferentes finalidades de recuperação da informação”.

A indexação como procedimento de análise documentária é desempenhada com mais intensidade a partir do aumento de publicações periódicas e da literatura técnico científica de modo geral, que incentivaram a necessidade de criação de mecanismos de controle bibliográfico em centros de documentação especializados.

Seguindo a corrente Inglesa, a indexação segundo Borko e Bernier (1978, p.8), é um “processo de analisar o conteúdo informacional dos registros do conhecimento e sua expressão na linguagem do sistema de indexação”.

A Norma 12.676, da Associação Brasileira de Normas Técnicas (1992, p. 2), indicou para a indexação três fases ou estágios:

- a) Exame do documento e estabelecimento do assunto de seu conteúdo;
- b) Identificação dos conceitos presentes no assunto;
- c) Tradução desses conceitos nos termos de uma linguagem de indexação.

Vale destacar que uma mesma publicação poderá apresentar conjuntos diferentes de termos de indexação dependendo do grupo de usuários ao qual se destina e aos interesses particulares desse grupo, ou seja, não há uma única forma correta de indexar.

As etapas são constituídas da análise conceitual e da tradução. Na saída do sistema, há a pergunta do usuário, formulada por meio de estratégias de busca, feita com a finalidade de recuperar documentos. Ambos os processos são mediados por

um vocabulário controlado, que será elaborado pelo indexador, necessitando também, na saída do sistema, que seja feita a análise conceitual e a tradução da pergunta do usuário para a linguagem do sistema (LANCASTER, 2004). Para tanto, o indexador tem que entender a linguagem do usuário para a elaboração deste vocabulário.

Desta feita, nota-se a importância da utilização de uma política de indexação e de um vocabulário controlado para orientar as atividades do profissional no momento da representação temática dos documentos. É o vocabulário controlado que estabelece o método de importação de termos e assuntos que compõem uma área limitada do conhecimento, tornando coerente a indexação.

O indexador deve estar atento às necessidades de busca do usuário, sempre acatando a cultura organizacional da instituição. Este ao indexar tem que utilizar termos ou palavras-chaves que sejam pertinentes a recuperação evitando colocações inúteis, que poderão causar dificuldades no processo da recuperação da informação desejada pelo usuário.

Segundo o UNISIST (1981), há de se considerar a indexação sobre dois pontos de vista distintos: “enquanto processo que consiste em descrever e identificar um documento com a ajuda de representações dos conceitos neles contidos e quanto à sua finalidade, permitindo busca e acesso à informação armazenada”.

Há um elemento de suma importância na indexação: a linguagem. A indexação, enquanto atividade, que elabora aspectos de conteúdos explícitos e implícitos dos documentos, emprega-se de uma linguagem composta por termos que podem ser armazenados ou utilizados para busca em linguagem natural ou convertidos para o vocabulário do sistema, ou seja, uma linguagem documentária (LD). A partir da linguagem utilizada na indexação que será construída a comunicação entre o usuário e o sistema de informação, no ato da busca.

De um modo geral, as linguagens documentárias, enquanto linguagens de tratamento e de recuperação da informação podem ser classificadas de acordo com três pontos de vista (GIL URDICIAIN, 1996, p. 22 *apud* por GIL LEIVA, 1999, p. 49):

- **Tipo de controle de vocabulário** – nesse sentido podem ser livres (listas de descritores livres) ou controladas (classificações listas de cabeçalhos de assuntos e tesouros).
- **Pela Coordenação** – pré-coordenadas (classificações coordenadas, lista de descritores livres e listas de palavras chave e tesouros).

- **Pela sua estrutura** – podem apresentar-se como hierárquicas (classificações), alfabéticas (listas de cabeçalhos de assunto), ou ambas (tesauros).

A indexação é também avaliada por meio de recuperação da informação, pelos índices de precisão e revocação. Lancaster (2004, p. 4), afirma que, com relação ao uso de revocação e de precisão, apesar de existirem outras abordagens, “elas são medidas a serem utilizadas para expressar os resultados de qualquer busca que simplesmente divida uma base de dados em recuperados e não recuperados”. Esta ação é designada como a capacidade de recuperar informações úteis; e precisão é a capacidade de impedir a recuperação de informações inúteis.

Procurou-se, nesse ponto, elencar, de forma breve, uma síntese do surgimento da indexação e sua característica multifacetada.

4.2 A EVOLUÇÃO DOS MÉTODOS DE INDEXAÇÃO

A indexação tem o intuito de substituir o texto de um documento por palavras-chave ou termos apropriados de seu conteúdo. Essa representação é feita a partir da análise do conteúdo do texto, no qual, o procedimento deve ser conduzido por um especialista, sob um olhar cauteloso nas metodologias.

Cabe frisar que existem pelo menos cinco formas para indexar essa análise feita a partir da documentação pesquisada, são elas: indexação manual, indexação semi-automática, indexação automática, indexação assistida por computador durante o armazenamento e a indexação automatizada, objeto de nosso estudo.

A indexação manual, chamada também de intelectual ou humana, como o próprio nome o diz, é desempenhada pelos seres humanos. Esse método de indexar é baseado, no julgamento comumente intuitivo do especialista, em função do texto e do interesse para a sua comunidade de usuários.

Mas, segundo Lancaster (1993, p. 20) “ao indexador raramente é dado o luxo de poder ler um documento do começo ao fim”. Por sua vez, o conhecimento do indexador sobre o domínio de identificar as características específicas do documento é de suma importância para a qualidade dessa análise e para atender ao usuário.

Hjorland (2001, p.776), concorda que a indexação deve ser moldada para atender as necessidades de determinados usuários,

uma vez que qualquer documento, em princípio, proporcionará respostas a uma infinidade de questões, as análises de assuntos devem estabelecer prioridades baseadas nos grupos de usuários específicos atendidos (ou serviços específicos proporcionados na ecologia da informação). [...] A melhor análise temática é a que faz o melhor prognóstico quanto ao uso futuro do documento.

Este o aspecto também é analisado por Battes (1998, p. 1187), quando,

o desafio para o indexador é tentar antecipar quais os termos que as pessoas possuem lacunas de informação de vários tipos procurariam nos casos em que o registros de que dispõem, de fato, fica a meio caminho de satisfazer a necessidade de informação do usuário.

A indexação manual vem se tornado um pouco inadequada, pois seu processo é muito caro e demorado. Além disso, há o fato de que o indexador tem que sempre estar atualizando o processo de indexação para que este permaneça possuindo termos úteis na busca dos usuários. Para isso, o indexador tem que possuir um grau de conhecimento sobre o assunto indexado, tornando-se outro fator que dificulta o processo da indexação manual.

Com referência à este aspecto, perguntamos ao gestor e a auxiliar administrativa (que na ocasião é a atual responsável pelo arquivo do escritório de advocacia), sobre como eram indexados os termos para a localização dos documentos e/ou processos no escritório, sem a presença do profissional arquivista e também sem o processo de indexação dos termos? Ambos entrevistados responderam que, “os termos eram indexados manualmente apenas pelo nome do cliente. Não sabíamos que existiam termos específicos que poderiam nos ajudar na busca. Não reconhecíamos a importância do profissional arquivista”.

Dando continuidade, a entrevista, indagamos-lhes, entretanto se a busca manual desenvolvida no escritório era considerada rápida diante da grande massa documental? Os entrevistados foram bem precisos, afirmando que, “em hipótese alguma era rápido essa busca, demorava muito quando se buscava um documento.”

Ao chegarmos ao arquivo, deparamo-nos com documentos misturados nas pastas, por não haver uma organização correta. Assim, ao procurar a informação desejada era bem difícil encontrá-la. Constata-se de acordo com o a voz dos entrevistados, para quem: “buscar manualmente e até mesmo indexar manualmente

é bem mais difícil, pois requer mais tempo e não agilidade ao se realizar essa tarefa”.

Nesse primeiro momento fica nítido que a indexação manual não atende as necessidades do arquivo sujeito de nossa discussão. Até o desenvolvimento da indexação automatizada, realizada neste acervo todo o processo de busca era feita manualmente.

A outra forma de fazer a análise do conteúdo é a indexação semi-automática, esta é feita por sistemas que indexam automaticamente o documento e, se necessário, dão a possibilidade de edição e de validação dos termos pelo especialista.

Segundo Pinto (2001, p. 227) “a indexação semi-automática seria a combinação da indexação manual com a indexação automática”. No primeiro momento é realizada a indexação automática dos documentos, levando em conta que os termos a serem indexado, serão os que mais se repetirem no texto. No segundo momento, o indexador humano irá ajustar a lista dos descritores definida pelo sistema.

Outro método é a indexação automática, que tem se tornado um tema bastante abordado pelos pesquisadores da Ciência da Informação e da Arquivologia.

É definida por Anderson e Perez-Carballo (2001, p. 256) como a

análise do texto por meio de algoritmos de computador. Ou seja, a indexação automática é feita por sistemas sem nenhuma validação por parte do profissional; os termos indexados são armazenados diretamente como descritores do documento.

A indexação automática é aquela realizada diretamente por sistema de computador que analisa o texto, reconhece e constrói índices para a recuperação do mesmo em pesquisa (VIEIRA, 1984).

Diversos *softwares* são bastante conhecidos para a indexação automática e semi-automática. É o caso do Sistema de Indización Semi-Automático (SISA) desenvolvido pelo Prof. Dr. Isidoro Gil Leiva da Universidade de Murcia na Espanha como resultado do seu estudo sobre automatização da indexação.

O SISA tem como principal função, selecionar automaticamente descritores que serão atribuídos a um texto, a partir de sua comparação com o vocabulário controlado de um sistema de informação e com uma lista de palavras consideradas

sem significados. Além disso, ela relaciona em lista à parte os termos considerados possíveis descritores, ou seja, palavras que não foram incluídas no vocabulário controlado, mas que, são repetidas diversas vezes no texto.

Para Lancaster (2004, p. 52), “vários programas de computador foram desenvolvidos para gerar, automaticamente, um conjunto de entradas de índices a partir de uma sequência de termos”. São exemplos de modelos o PRECIS, o KWIC e o KWOC, expostos a seguir.

O método PRECIS produz índice impresso fundamentado na ordem alfabética e na ‘alteração’ sistemática de termos para que ocupem a posição de entrada. É um sistema de indexação que apresenta sua análise conceitual como metodologia de identificação de conceitos baseados em um questionamento dirigido unicamente ao texto (FUJITA, 1989).

Já o *Keyword in Context* (KWIC) (palavra-chave no contexto) foi desenvolvido por Luhn em 1959 e foi o primeiro método de indexação automática. Seu intuito é fazer a indexação por palavra-chave no contexto, ou seja, era considerado a frequência das palavras dos títulos dos documentos, compondo um índice permutado. O KWIC é um programa que reconhece as palavras que não são palavras-chave, baseado em uma lista de palavras impedidas de entrar no sistema, que decididas na entrada.

O KWIC é um programa simples, barato e que obtém, em certo nível, acesso temático ao conteúdo de uma coleção. É nítido, porém, sua qualidade está diretamente relacionada à qualidade dos títulos, no sentido em que estes devem ser bons indicadores do conteúdo dos textos (LANCASTER, 2004).

O método *Keyword out of context* (KWOC) (palavra-chave fora do contexto) é semelhante ao KWIC, entretanto, as palavras-chave se tornam pontos de acesso e são repetidas fora do contexto. O KWOC é a iniciativa mais simples da indexação automática, pois é baseada na extração de palavras, geralmente extraídas do título.

Outro modelo é a indexação assistida por computador durante o armazenamento, para quem são sistemas que auxiliam o processo de armazenamento dos termos de indexação extraídos pelo indexador na etapa de análise conceitual. São facilitadores do processo de indexação uma vez que proporcionam, por meio de janelas de ajuda, notas explicativas sobre os termos e seus relacionados e, às vezes, acesso a documentos já indexados, para solução de dúvidas (SANTOS, 2009).

Por fim, a indexação automatizada, entendida como uma operação que requer uma base de dados e um indexador humano, que realizam a atividade intelectual de indexação.

Na literatura específica da área de indexação ainda não há uma consolidação no tocante a homogeneidade quanto aos termos automática e automatizada, tendo muitas vezes uma relação sinonímia. Para nosso trabalho utilizaremos o termo automatizado.

Destarte, espera-se que neste tópico haja uma contribuição em relação à abordagem da indexação, e também um favorecimento na literatura tendo em vista a falta de teoria relacionada, em especial a diferenciação de automático e automatizado. Nesse sentido, a indexação automática é feita totalmente pelo computador, e após a escolha de um programa adequado, esta será realizada sem que o ser humano interfira no processo. Já o segundo – indexação automatizada – há a ligação do ser humano com essas novas ferramentas de busca utilizadas no computador. O profissional a partir de uma leitura no documento elencará termos propícios para melhor busca, e assim os indexará em uma base de dados, tendo a interferência sempre que preciso do indexador.

4.3 A INTERDISCIPLINARIDADE DA INDEXAÇÃO COM OS PROCESSOS DE LINGUAGENS NATURAIS E LINGUAGENS DE ESPECIALIDADE: o uso de vocabulários controlados.

A indexação é uma ferramenta interdisciplinar e possui uma forte relação com a linguagem, especificamente as linguagens naturais e de especialidade, relacionada ao uso de vocabulários controlados, também conhecidos como linguagem controlada.

Para dá início a este ponto, conceituaremos a princípio, o que é linguagem. A linguagem deve ser estudada em toda a variedade de suas funções. Assim sendo, para um melhor entendimento destas funções, é necessário observar os fatores que compõem todo o processo linguístico, ou seja, todo ato de comunicação verbal (JAKOBSON, 1974).

Ainda sobre o campo das definições, Sager (1980) concebe que a linguagem

é utilizada por diferentes grupos de pessoas. No entanto, a linguagem se fará diante ao conhecimento do mundo, ou seja, indivíduos utilizarão a linguagem de acordo com a influência dos assuntos abordados no cotidiano, devido a seu lugar na sociedade e a localização geográfica.

Como podemos observar, a linguagem possui conceituações complexas. Há de se considerar, que há outras associações de linguagens que tentaremos a seguir estabelecer algumas disjunções entre elas, a saber: linguagem natural e de especialidade.

A Linguagem natural (LN) é concebida como sinônimo de discurso comum, ou seja, é a linguagem usada normalmente na fala e na escrita do senso comum. A linguagem natural possui duas técnicas essenciais na realização de buscas, são elas: a de truncagem de termos e a busca com operadores de proximidade. A truncagem de termos permite ao intermediário operacionalizar e usar a raiz do termo sem especificar todas as possíveis variações desse termo (prefixos e/ou sufixos). Já a técnica de busca com operadores de proximidade permite especificar, na estratégia, a posição relativa de dois ou mais termos entre eles próprios.

Rothman (1983) afirma que, “com a busca na LN, a base de dados está efetivamente auto-indexada, pois cada palavra no documento indexada é candidata a termo de busca e identifica, ainda, a unidade do texto no qual se encontra”. Sendo assim, a LN dos documentos contém termos de indexação ou pontos de acesso imediato, e os usuários podem interagir diretamente com os itens incluídos na base de dados.

Um das principais vantagens no uso da linguagem natural, é que essa permite o imediato registro da informação em uma base de dados, sem necessidade de consultar a uma linguagem de controle e os termos de entrada de dados são extraídos diretamente dos documentos que vão constituir a base de dados.

Por outro lado, a linguagem de especialidade (LE) são sistemas de significados, ou seja, signos, utilizados por indivíduos que comportam na comunidade linguística para comunicação, compreensão e compartilhamento de uma realidade. Para uma maior abrangência da LE, alguns autores consideram que as palavras, unidades de linguagem verbal, são signos neutros, ou seja, podem assumir diferentes significados de acordo com o contexto do uso, pois esses podem sofrer distorções de entendimentos passados por mudanças sofridas pelas sociedades.

A partir desse contexto, a linguagem de especialidade pode sofrer alterações de acordo com o contexto. Sager (1980 *apud* Galvão, 2003, p. 40) considera que a linguagem de especialidade:

- Indica possibilidades no emprego de suas unidades constituintes (termos) e significados que carregam (conceitos);
- Tem por objetivo a construção de signos monossêmicos, mediante a relação de um significante a um significado e vice-versa;
- Resultam de consensos conceituais existentes dentro do campo científico ou tecnológico. Se os consensos sofrem alterações, seja pela criação de novas teorias, seja pela identificação de novos fenômenos, as línguas da especialidade são alteradas;
- Tem por propósito a educação especializada e a comunicação entre especialistas do mesmo campo ou de campos relacionados;
- Não são linguagens são linguagens artificiais. A linguagem artificial é uma linguagem de especialidade sem nenhum elemento de linguagem geral;
- Não são apenas conjuntos de termos. Elas possuem uma dimensão pragmática, uma dimensão semântica e uma dimensão sintática;
- Para compreendê-las, é preciso ter uma educação especializada ou um treinamento específico.

Entende-se então, que a linguagem de especialidade, genericamente, é o conjunto de marcas lexicais, sintáticas e discursivas que indicam o uso de um código linguístico em qualquer ambiente de influência social.

Tanto a linguagem natural quanto à linguagem de especialidade estão relacionadas à linguagem controlada (LC), ou seja, um vocabulário feito por especialistas e/ou membros da organização que vise a padronização dos termos.

No tocante ao escritório, esse vocabulário foi elaborado pela estagiária de Arquivologia e pelo advogado (gestor), que favoreceu a padronização dos termos a serem extraídos pelo modelo de indexação automatizada aplicado neste ambiente.

Com relação ao vocabulário, perguntamos na entrevista ao gestor e à auxiliar se com a elaboração de um banco de dados² para subsidiar na busca mais ágil aos processos e conseqüentemente a elaboração de um vocabulário controlado,

² Esclarecemos que os termos utilizados foram elencados após uma reunião entre nós e o gestor para a escolha das palavras-chave que recuperassem, da melhor forma possível a informação desejada por eles, tornando o vocabulário controlado mais preciso, tendo em vista as especificidades do tipo de trabalho realizado.

construído pelo o gestor e pelo arquivista, foi fundamental nas buscas documentais?
E ainda mais, se a linguagem utilizada foi precisa?

As respostas colhidas foram similares, ambas afirmaram que:

a linguagem foi precisa, e de suma importância para este trabalho, tanto é que grandes escritórios, frente à experiência realizada aqui no arquivo, e diante das necessidades de organização dos documentos, iram levar a ideia para discussão em seus locais de trabalhos. Vale destacar, que a estagiária auxiliou bastante junto com o profissional do escritório na elaboração do vocabulário que em conjunto conversamos sobre ele e sobre os documentos em si.

Desta forma, é lícito afirmar que, com base no relato acima, um vocabulário controlado interagiu na elaboração do banco de dados e conseqüentemente da facilidade no momento da busca e da recuperação da informação

A LC mostra-se mais precisa no seu método de busca, através de termos apropriados e produz uma melhor recuperação. Esta pode ser definida como um conjunto de termos organizados de forma hierarquizada e/ou alfabética, com o objetivo de permitir a recuperação de informações temáticas, reduzindo substancialmente a diversidade de terminologia. Uma base de dados que utilize um vocabulário controlado possibilita, ao mediador no planejamento da estratégia de busca, a recuperação, no campo específico de descritor, apenas palavras-chave dos termos a partir de um vocabulário controlado, formando índices.

Diversos estudos sobre o uso da linguagem natural, linguagem de especialidade e linguagem controlada, uma vez subsidiando a recuperação da informação, têm se concentrado no uso do conjunto dessas linguagens na estratégia de busca, demonstrando que a utilização simultânea dessas linguagens proporciona melhor desempenho nos resultados.

5 A INDEXAÇÃO AUTOMATIZADA COMO FERRAMENTA DE RECUPERAÇÃO DA INFORMAÇÃO

Neste capítulo, primeiramente é apresentado alguns conceitos sobre indexação automatizada e sua relação com os métodos de recuperação da informação. Já no segundo momento, elencaremos e discutiremos algumas vantagens e desvantagens do uso da indexação automatizada.

5.1 INDEXAÇÃO AUTOMATIZADA: embasamento teórico e conjuntura atual

As Tecnologias Digitais de Informação e Comunicação (TDIC's) colaboram no desenvolvimento dos processos e instrumentos apostos ao tratamento da informação documentária. Em linhas gerais, essas novas tecnologias subsidiam a automatização da indexação, portanto, caberá ao arquivista o objetivo de verificar e compreender a análise do conteúdo efetuada por *softwares* e/ou ferramentas que colaboram no processo da indexação automatizada.

Pesquisas sobre a automatização deram início em meados do século XX. Para melhor compreensão de como se desenvolveu a automatização, verificamos algumas literaturas sobre o assunto, e diante a isto, se apresentou três momentos históricos que estão diretamente ligados à sua evolução.

Para Gil Leiva (1999, 2008) “o desenvolvimento da automatização da indexação ocorreu em três momentos históricos: o dos métodos estatísticos, o dos métodos linguísticos e o dos métodos mistos ou híbridos”.

Resumidamente, os métodos estatísticos segundo Vieira (1988) é o “método de frequência que consiste na contagem automática do aparecimento da palavra, que pode estar no título, resumo, título das referências citadas, texto e em diversas combinações dessas fontes”.

Já os métodos linguísticos, diante aos limites dos métodos estatísticos, procurou enfatizar nos critérios de discernimento da linguística nas palavras. Nesse sentido, Gil Leiva (1999, p. 82) explica que:

a partir do início dos anos sessenta associam-se as técnicas de Processamento da Linguagem Natural (PLN) que consiste no estudo e análise dos aspectos linguísticos de um texto mediante a utilização de programas informáticos – e a automatização da indexação.

Sendo assim, nota-se que os estudos linguísticos desenvolveram-se em direção à compreensão da estrutura textual, suas semelhanças e seu sentido. Ainda, de acordo com Gil Leiva (2008, p. 339) “os primeiros analisadores linguísticos surgiram na década de 1960 para o processamento automático de informação”.

Os aspectos linguísticos estão presentes na leitura para indexação, por meio da linguística textual que, segundo Fávero e Koch (1988, p.14) é a área responsável por “determinar o que faz com que um texto seja um texto e diferenciar as várias espécies de texto”.

A indexação automatizada, assim como a manual, tem sua preocupação e seus alvos focados na recuperação da informação que estão contidos em um determinado documento. Porém, diferentemente da manual, esta será representada por um computador que através do ser humano indexará termos contidos no documento, representando seu conteúdo.

Há de considerar o que já foi discutido anteriormente, a diferença entre a indexação automática e a indexação automatizada, muitas vezes utilizadas como sinônimos. Em concórdia com Naves (2004, p.22),

há uma controvérsia quanto ao uso das terminologias indexação automática e automatizada, podendo-se sintetizar que, na primeira, o trabalho desenvolvido para indexar assuntos é totalmente feito pelo computador, como indexação por palavras (KWIC), pela frequência com que as palavras aparecem no texto, entre outros. Já a indexação automatizada, a primeira etapa é extrair o conteúdo do documento (análise do assunto) é feita por um indexador humano, que após seu trabalho intelectual, insere os termos numa base de dados automatizada. No entanto, não há um consenso quanto a essa diferença apresentada, podendo-se verificar os dois termos usados, também, como sinônimos.

Atualmente, os computadores não geram de modo independente, mensagens ou necessidades de informação, a menos que sejam especificamente programados para esse fim por seres humanos, mas talvez chegue o dia em que executarão isso (LANCASTER, 1993).

No sentido denotativo, de acordo com o Dicionário Houaiss da Língua Portuguesa, automatizar é “prover de máquinas ou de dispositivos mecânicos ou eletrônicos, para agilização e otimização da produção, dos serviços etc.” Assim, destaca-se que a automatização da indexação é o emprego de dispositivos que agiliza e otimiza o processo de recuperação da informação por meio da indexação.

A recuperação da informação em uma base de dados elaborada num computador é mais fácil, flexível e menos trabalhosa do que a busca em métodos manuais, pois os sistemas informatizados disponibilizam diversos recursos, diferente dos manuais, fazendo com que o usuário que for buscar uma informação compreenda tais recursos, mesmo não estando familiarizados com os mesmos, otimizando a utilização desta base.

Ao aplicarmos a entrevista, indagamos aos entrevistados (o gestor e a sua auxiliar), quais foram as suas maiores dificuldades no processo manual de recuperação de informação, e o porquê das dificuldades.

A auxiliar asseverou que:

as dificuldades eram diversas, mas o que eu posso destacar é que, como os documentos não tinham identificação, sendo guardados tudo misturado na pasta, mesmo sendo do mesmo cliente a tal pasta, ao querer buscar um 'tipo de documento' [sic], tínhamos que pegar todo processo, pois não havia a discriminação dos documentos na pasta.

Por outro lado, o gestor afirmou que, “a localização, em virtude da falta de sistematização, e diante do grande volume de documentos, era muito complicada, pois alguns documentos, apenas eu conseguia localizar”

Com a experiência realizada no escritório de advocacia e a criação e a aplicação do sistema de recuperação da informação e consequentemente o modelo de indexação automatizada, perguntamos ao gestor e à sua auxiliar se a indexação automatizada atendeu às necessidades do arquivo, de forma eficaz, tais como, rapidez no processo de busca pela documentação desejada. Ambos responderam que, “com certeza, rapidez é a palavra-chave desse momento que passou a caracterizar o arquivo, deixando para traz o nosso tormento de buscar um documento”.

A indexação é bastante favorável para os usuários, pois sua consulta se realizará por sistemas e não acarretará alterações, a não ser que o profissional indexador junto com o gestor note que existam outros termos mais pertinentes para a indexação facilitando a recuperação da informação daqueles que a buscam.

A indexação automatizada pode ser definida como o processo realizado pelo homem, que identificará palavras relevantes (descritores) nos documentos e armazená-las-á em uma estrutura de índice em um banco de dados.

Pode-se concluir que, a indexação automatizada seria, portanto, aquela resultante do trabalho intelectual de um profissional para checagem do valor dos termos atribuídos a um documento por um programa de computador (SILVA, FUJITA, 2004, p.145).

5.2 VANTAGENS E DESVANTEGENS DA APLICABILIDADE DA INDEXAÇÃO AUTOMATIZADA EM SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO

Neste ponto, procuramos identificar na literatura da Ciência da Informação e da Arquivologia, algumas discussões relacionadas às vantagens e às desvantagens na automatização da indexação.

Nesse sentido, Gil Leiva (2008), aponta algumas vantagens e desvantagens da indexação automatizada, são elas:

Vantagens	Desvantagens
<ul style="list-style-type: none"> • Caracteriza a indexação humana como lenta, subjetiva e de alto custo; • Afirmam que, com a automatização da indexação tem-se diminuição de erros que repercutirá positivamente na recuperação das informações em bases de dados; e • A indexação automatizada parece ser mais precisa, permitindo uma recuperação dos documentos mais rica. 	<ul style="list-style-type: none"> • A indexação automática reconhece palavras e não conceitos defendem que se deve perseguir a captação de terminologias dos textos, porque esta cumpre a função representativa, cognitiva e comunicativa que apresentam os conceitos e, por tanto, o conhecimento; • Na maioria das ocasiões a automatização da indexação restringe-se às áreas específicas do conhecimento; • Uma das principais razões abordadas contra a automatização do processo é a impossibilidade, no estado atual da investigação, conseguir indexação totalmente automática.

Fonte: Gil Leiva (2008)

Já os autores Méndez Rodríguez e Moreira González (1999, p. 4-8) identificam quatro fatores que levaram às experiências de indexação automatizada:

- O alto custo da indexação humana, em termos de tempo, suscitou a ideia de explorar de maneira eficaz, a um custo e tempo reduzidos, o volume constantemente crescente de informação. Essa questão indexação automática.
- Aumento exponencial da informação eletrônica e a proliferação de textos completos.

- A Gestão Eletrônica de Documentos (GED) e a informatização dos processos documentais.
- A automatização de processos cognitivos e a pesquisa crescente e os avanços em Processamento de Linguagem Natural (PLN). A automatização de processos cognitivos permite o surgimento de sistemas inteligentes, que somados ao PLN, podem lidar com a atividade de indexação. Porém, os autores alertam para complexidade da linguagem e afirmam que um sistema não pode lidar globalmente com ela, sendo capaz apenas de reconhecer cadeias de caracteres.

Alguns autores, que estudam a automatização da indexação, chegam a citar que a GED é um dos fatores que fortalecem o emprego da indexação automatizada, pois este processo cria maiores condições de trabalho dos profissionais da informação.

A indexação manual é outro grande fator levantado em benefício à automatização da indexação. A indexação pode diferenciar de indexador para indexador. Outro forte argumento em favor dos sistemas automáticos é que eles são mais objetivos, auxiliando o profissional quando este for indexar em meio eletrônico.

É de suma importância destacar que a indexação automatizada pode ser vista como um instrumento facilitador da atividade de indexação. Concorde-se aqui com esta abordagem para serviços de informação já formados, como os arquivos, posto que a indexação “automática” já é um fato e pode ser considerada uma solução com relação a grandes massas informacionais.

No entanto, os sistemas automáticos ainda não lidam satisfatoriamente com a linguagem humana, necessitando do auxílio do profissional indexador, para que tal atividade obtenha uma melhor qualidade. O indexador também pode contribuir nas atividades de construção e avaliação dos sistemas automáticos (GIL LEIVA, 2008).

Lancaster (2004) e Moreiro González (2004) afirmam e concordam com a proposta dos *softwares* de indexação semi-automáticos na medida em que defendem o uso de *software* para indexar documentos, com a ressalva de que é necessário ter em mente as limitações em seu processo, as quais só a análise apurada de um ser humano pode solucionar, sendo clara a necessidade de avaliação final dos termos de indexação por um profissional.

Ainda sobre a indexação automatizada, Farmer (2006, p. 99), quando se refere às ferramentas de categorização automática de documentos digitais, confirma a “necessidade da parceria homem-máquina para somar à capacidade de

processamento de textos dessas ferramentas, a inteligência, os julgamentos e a experiência humana”. Essa parceria, sem sombra de dúvidas traria uma grande melhoria na efetividade das taxonomias, na atuação dos sistemas e na precisão da indexação.

6 A CONSTRUÇÃO E A APRESENTAÇÃO DO MODELO DE INDEXAÇÃO AUTOMATIZADA

Neste capítulo, são apresentadas o desenvolvimento do modelo de indexação automatizada, sua implementação e um banco de dados e os processos subjacentes a esse trabalho. O modelo consiste em representar os termos extraídos dos documentos para a recuperação da informação. O modelo foi aplicado em um arquivo jurídico.

6.1 A CONSTRUÇÃO DO MODELO: O USO DO EXCEL

A busca pela informação na atual sociedade configurou-se em uma busca mais frequente, uma vez que a informação passou a ser fundamental nas tomadas de decisões da maioria da população, em especial nas organizações. Por isso, far-se-á necessário a busca de sistemas de recuperação da informação que invistam na prática de oferecer a informação para quem dela necessita.

Com o passar dos anos e com o avanço do desenvolvimento tecnológico, sistemas informacionais estão se tornando cada vez mais reais e já passam a fazer parte do cotidiano das organizações. Diante as necessidades da adesão das novas tecnologias em arquivos, os arquivistas deverão estar atentos na elaboração e na sustentação dos sistemas de informação, junto a profissionais da informática.

De acordo com o Dicionário sobre informação e informática, do Instituto Europeu de Softwares (ESI)³, encontramos a seguinte definição para sistemas de informação:

Organizações que fornecem, usam e distribuem informação. Considera-se que inclui os recursos organizacionais relacionados, tais como os recursos humanos, tecnológicos e financeiros. É de fato um sistema humano, que inclui provavelmente recursos computacionais para automatizar determinados elementos do sistema.

Os sistemas de informação são compostos por partes interligadas (classificação, catalogação, indexação etc.) com um único objetivo, o de disponibilizar a informação da melhor maneira possível.

O que se observa, nos sistemas de informações tradicionais, é uma eterna e penosa procura pelo que se deseja em meio a uma grande quantidade de

³ Ver: http://www.esi.es/Help/Dictionary/Definitions/Information_System.html.

informações emaranhadas. Sistemas de filtragem de dados esforçam-se para tornar estas tarefas mais amenas na tentativa de busca pelas informações de forma a subsidiar o usuário com as informações requeridas, a tempo e a hora, para a tomada de decisão (LAND, 1982).

Todos os profissionais que sejam interessados em conseguir informações com mais velocidade, devem se preocupar em não apenas conhecer, mas principalmente saber explorar essas novas tecnologias para que aproveitem ao máximo, sem que sintam impedidos de realizar seus trabalhos com tamanha rapidez.

Para tanto, Rezende (1998) assevera que,

Não basta, portanto, ter disponíveis as informações, em alguns casos isto pode até dificultar ainda mais os resultados. É preciso que as informações disponibilizadas estejam organizadas e acessíveis a tempo de vencer os prazos e a concorrência.

Rezende salienta ainda sobre o envolvimento de vários fatores nos escritórios de advocacia:

Nos escritórios de advocacia, vistos como indústrias que processam e transferem conhecimento, na busca de soluções para os problemas do cliente, a essência do trabalho é a manipulação, reelaboração e transferência de informações. O conhecimento acumulado e transferido no ambiente de um escritório de advocacia envolve tecnologia, conteúdo informacional e, principalmente, interação entre pessoas (REZENDE, 1998).

É diante a tais dificuldades que apresentamos o Excel, como uma ferramenta que subsidiará a busca por informações através da indexação automatizada. A recuperação da informação em uma base de dados elaborada num computador é mais fácil⁴, flexível e menos trabalhosa do que a busca em métodos manuais, pois os sistemas informatizados disponibilizam diversos recursos, diferentes dos manuais, fazendo com que o usuário que for em busca da informação, compreenda tais recursos, mesmo não estando familiarizados com os mesmos, otimizando a utilização desta base.

Para a elaboração de uma base de dados, destinada para a representação dos arquivos, utilizamos o Excel, uma das ferramentas produzidas pela Microsoft. As

⁴ Nosso intuito é que as pessoas que trabalhem diretamente com a base de dados sejam competentes em informação, que para Perrenoud (2000), umas dessas competências é utilizar as novas tecnologias.

planilhas são ferramentas versáteis usadas para manipular e analisar dados tanto numéricos quanto alfabéticos. A habilidade no seu uso é útil em uma ampla variedade de áreas específicas.

Uma das vantagens mais importantes das planilhas é sua facilidade de uso, se comparado a outros sistemas com interfaces variadas e repletas de gráficos. Outra vantagem é que em apenas um clique ela poderá colocar em ordem alfabética tudo que o indexador havia colocado aleatoriamente. Ela pode ser ainda, alterada, sem que haja perda de informação, além de trabalhar diretamente com o ser humano.

A indexação automatizada tem sua preocupação e seu alvo focados na recuperação da informação que estão contidos em um determinado documento. Porém, diferentemente da manual, a automatizada será representada por um computador que através do ser humano indexará termos contidos no documento, representando seu conteúdo informacional.

Entretanto, a indexação automatizada ou manual só poderá ser realizada se todo acervo documental estiver organizado para que o processo de recuperação se dê também de maneira ordenada sobre todas as séries documentais.

Diante ao exposto, perguntamos na entrevista realizada, se depois de organizado o arquivo, somente a organização por si só, sem a presença de uma indexação automatizada, era bastante suficiente para a acessibilidade dos processos. Para tanto, os entrevistados, em concordância, responderam que, “com a chegada da estagiária no arquivo e conseqüentemente a organização do acervo, mesmo ficando ótima a organização, era preciso mais, algo que agilizasse a busca pelo documento”.

Neste trabalho entendemos que a aquisição automática pode e deve ter participação ativa dos arquivistas na indexação automatizada e também na sua atualização, ao utilizar de uma ferramenta de fácil uso, por isso utilizamos o Excel. Ainda que, em princípio, sejamos partidários do uso de *softwares* gratuitos, o Excel, por fazer parte do Office, costuma estar disponível na maior parte dos escritórios de advocacia. No Excel é possível executar tarefas com poucas linhas de programação, ganhando com isso tempo e facilidade de uso.

6.2 A APRESENTAÇÃO DO MODELO DE INDEXAÇÃO AUTOMATIZADA FEITO EM UM ARQUIVO JURÍDICO: SUAS INTERFACES

Nesta sessão apresentaremos o experimento aplicado no arquivo sujeito da pesquisa. Antes, falaremos como foi todo o processo de criação do banco de dados. Foi preciso primeiramente nos situar perante a empresa e aos documentos produzidos por ela.

Após conhecer um pouco da empresa, buscamos organizar a informação contida em cada pasta dos clientes. Por diversas vezes encontramos documentos fora de ordem que não correspondiam aos clientes daquelas pastas. O gestor, por não ter conhecimento dos métodos arquivísticos, ao chegar uma informação de um cliente, abria uma pasta aleatoriamente e colocava todas as novas informações recebidas de outros clientes nesta pasta, e isso acarretou perda de documentos. Ao organizarmos as pastas e as gavetas a partir do levantamento dos clientes do escritório, decidimos seguir o método de ordenação alfanumérico dentro da pasta e da gaveta.

O procedimento de organização acontecia da seguinte forma: ao chegar documentos do cliente, da pasta em questão, organizávamos de acordo com a data que constava no documento, ou seja, os documentos com a data mais recente era o primeiro da pasta. Assim facilitaria a busca pela informação, pois o gestor utilizava mais os documentos recentes para produzir informações a partir deles. Não permitíamos também que documentos de outros clientes fossem colocados indevidamente em outras pastas.

Cada cliente, assim identificado, passou a ter pasta de documentos e sub-pastas para cada um dos processos em que figura como parte. A entrada é feita por: nome-sobrenome do cliente, tipo (e número) de ação, e nome da outra parte (esses dados estão contidos numa folha de rosto, elaborada por nós). Com isso se tem à mão a informação do andamento do processo de cada cliente, possibilitando, inclusive, manter o cliente informado sobre a tramitação do seu interesse.

Após esse processo, procuramos identificar quais tipos documentais que o escritório produzia, para inserirmos no banco de dados a ser construído. A meta que se tem é de manter um sistema que permita recuperar a informação de forma mais rápida e eficiente: O processo ou documento a documento. No caso, o arquivo agora

passa a permitir ao advogado localizar pastas, processos ou documentos de clientes, com maior eficiência e eficácia.

Na figura 2, que mostraremos ao longo do texto, exploraremos o esquema que utilizamos no arquivo para o uso da indexação dos termos extraídos dos documentos. Convém ressaltar que, após um grande período de observação no modo pelo qual se organizava, classificava e indexava os documentos anteriormente ao uso da indexação automatizada, tais como dificuldade em localizar, perdendo-se tempo, e muitas vezes dinheiro; além de perdas irreparáveis de documentos, necessitou a adoção de ferramentas e ações que favorecessem uma maior precisão e velocidade na busca por informações, uma vez que, com o uso da planilha por meio do Excel, possibilitaria a criação de tópicos que descrevessem o local de arquivamento, local de trâmite do processo e as tipologias documentais, levando em consideração um vocabulário controlado.

Esclarecemos que foi realizado um trabalho bastante minucioso, em que a informação depositada nesse banco de dados elaborado, como em qualquer outro, procurou-se ter bastante habilidade para não inserir dados errôneos. Merece destacar que a documentação colocada no banco de dados fora toda organizada no ambiente físico e que nas pastas e nas gavetas havia uma notação⁵ respectivamente.

A planilha elaborada possui os seguintes campos:

- 1) **Número do processo:** Campo destinado a designar o número gerado pelas instancias responsáveis do julgamento das ações;
- 2) **Constituinte:** Campo informativo dos clientes do escritório;
- 3) **Partes:** Designa qual é a outra parte do processo;
- 4) **Ação:** Designa sobre qual é a natureza do processo;
- 5) **Tipologia Documental:** Campo que configura na descrição das tipologias documentais, uma vez que, o advogado muitas vezes se interessa apenas por um tipo documental, e estando esse descrito, facilita a rapidez na busca documental;

Notação é o elemento de identificação das unidades de arquivamento, constituída de números, letras, ou combinação de ambas, que permite sua localização (PAES 2004).

6) Local de Arquivamento: Lugar destinado a informar qual local está arquivado o documento descrito, bem como a gaveta e a pasta, representados pelas seguintes variáveis:

6.1) Documentos Ativos em fase Corrente: Esses documentos são representados pela letra A e número 1; que no momento da busca já informa qual é a fase que o documento se encontra, nesse caso corrente.

6.2) Documentos Ativos em fase Intermediária: Esses documentos são representados pela letra A e número 2; que no momento da busca, também já informa qual é a fase em que o documento se encontra, nesse caso intermediário.

6.3) Pastas: Representado pela letra P, informa qual é a pasta em que se encontra o documento.

6.4) Gaveta: Representado pela letra G, informa qual é a gaveta em que se encontra o documento.

7) Local de Trâmite: Informa qual é a vara, o PROCON, o Fórum, ou seja, os locais em que será julgado os processos.

Para realizar a busca por algum processo classificado no banco de dados, utilizamos o atalho do teclado *Ctrl+L*, assim ao digitarmos o que procuramos, esse localizará com eficácia a informação desejada (ver figura abaixo).

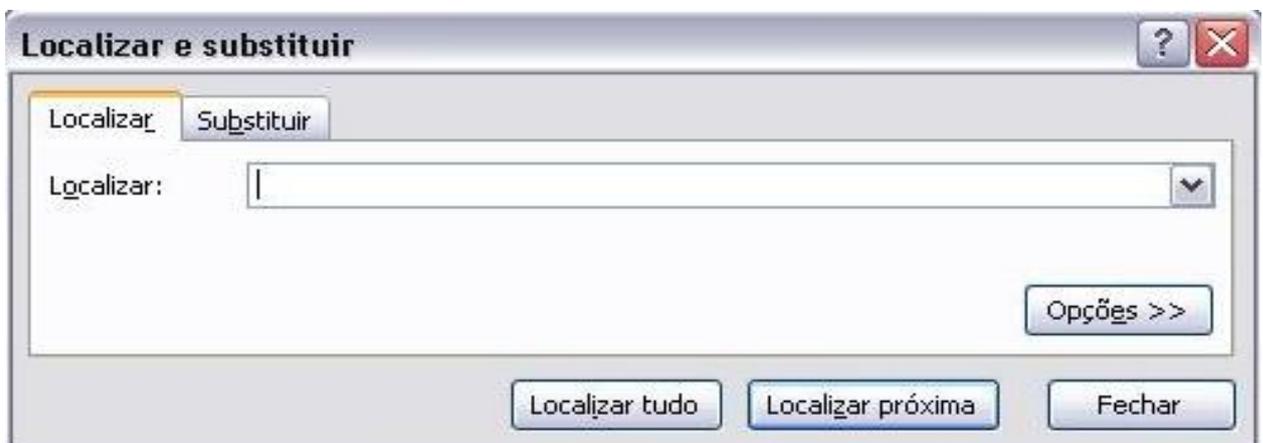


Figura 1 – Atalho para busca da informação no Excel.

O usuário ao realizar a busca, tanto ele pode localizar o que lhe for pertinente, quanto substituir um elemento, através deste atalho. Vale destacar que, quando altera a informação em um computador que tenha este arquivo, os demais também se modificaram, pois estão em rede. Por isso é imprescindível que haja comunicação entre aqueles que têm a permissão.

Quanto essas permissões, ou seja, quem poderá incluir ou retirar os documentos da representação, ficou apenas destinado ao Gestor e o Estagiário do Arquivo, uma vez que o banco encontra-se em Rede, e protegido por senha, já que se trata de um arquivo privado e de informações de caráter de Justiça.

Na página seguinte é explicitado através da figura 2, o modelo que foi utilizado, bem como seus campos descritos acima e alguns exemplos dos termos indexados⁶.

⁶ Lembramos que esses termos são dados hipotéticos.

Microsoft Excel

Área de Trabalho: A1, Nome da Função: Nº Processo

BANCO DE DADOS [Modo de Compatibilidade]

	A	B	C	D	E	F	G
1	Nº Processo	Constituinte	Partes	Ação	Tipologia Documental	Local de Arquivamento	Local do Trâmite
2	200.2007.019.372-3	Abílio Gonçalves	Maria Aparecida Feitosa	Reparação de Dano Material	Documentos Pessoais; Procução; Petição Inicial; Termo de Audiência de Conciliação; Audiência de Instrução e Julgamento; Recurso Inominado; Contra-Razões aos Embargos de Declaração;	A1; G4; P1	6ª VARA CÍVEL DE JOÃO PESSOA
3	200.2006.015.898-3	Adiel Fonseca	Caixa Econômica Federal	Ordinária	Petição; Sentença; Extrato	A1; G3; P1	7ª VARA CÍVEL DE JOÃO PESSOA
4	200.2009.000.760-4	Adiel Fonseca	Caixa Econômica Federal	Ordinária	Petição; Recibos; Contestação; Termo de Audiência.	A1; G3; P1	8ª VARA CÍVEL DE JOÃO PESSOA
5	200.2002.782.673-1	Adiel Fonseca	João Jacinto de Lira	Cobrança	Petição Inicial; Contestação; Sentença.	A1; G3; P1	4ª JEC JOÃO PESSOA
6	200.2008.929.988-7	Adiel Fonseca	VISA Cartões de Crédito	Anulatória	Consulta Processual	A1; G3; P1	5ª JEC JOÃO PESSOA
7	200.2009.000.002-5	Adiel Fonseca	Banco do Nordeste do Brasil	Cobrança	Petição Inicial; Requerimento.	A1; G3; P1	10ª VARA CÍVEL DE JOÃO PESSOA
8	200.2009.000.780-2	Albertina Madureira	Banco do Brasil	Cobrança	Contrato de Honorário Advocatício; Petição Inicial; Impugnação a Contestação; Requerimento	A1; G3; P1	16ª VARA CÍVEL DE JOÃO PESSOA
9	200.2006.000.009-2	Alberto Henrique Tavares de Melo	Michel le Tenorio	Despejo por Falta de Pagamento	Petição;Procução;Mandado de Citação; Mandado de Imissão de Posse; Contestação; Petição Inicial.	A1; G1; P1	11ª VARA CÍVEL JOÃO PESSOA
10	200.2006.000.860-2	Alberto Henrique Tavares de Melo	Michel de Lima Soares	Reparação Danos Morais	Impugnação a Contestação; Petições;	A1; G1; P1	11ª VARA CÍVEL JOAO PESSOA
11	200.2006.010.879-2	Alberto Henrique Tavares de Melo	Michel de Lima Soares	Reintegração de Posse	Petição Inicial; Procução...	A1; G1; P1	6ª VARA CÍVEL JOAO PESSOA
12	200.2007.008.475-4	Alberto Henrique Tavares de Melo	Rita Tereza Vilar	Indenização	Petição Inicial; Termo de Audiência de Conciliação; Sentença; Mandado de Intimação.	A1; G1; P1	1ª JEC JOÃO PESSOA
13	200.2009.002.201-1	Alisson Diogenes	Education do Brasil	Cobrança	Procução; Petição Inicial; Contestação	A1; G1; P1	7ª VARA CÍVEL JOAO PESSOA
14	200.2008.023.058-7	Alisson Diogenes	Zuleide Cabral Souza	Indenização	Mandado de Citação e Intimação; Parecer; Nota de Foro; Relatório; Auto de Prisão em Flagrante; Estatuto Social da Seguradora; Petição Inicial; Termo de Audiência Preliminar; Recibos; Procução;	A1; G1; P2	18ª VARA CÍVEL JOÃO PESSOA
15	200.2003.044.000-8	Alisson Diogenes / Wlademir Calixto	Zuleide Cabral Souza	Lei 9503/ Trânsito	Termo de Audiência Preliminar; Laudo de Exame de Corpo de Delito; Mandado de Intimação; Requerimentos; Parecer; Auto de Prisão em Flagrante; Relatório; Documentos Pessoais.	A1; G1; P1	JECRIM JOÃO PESSOA
16	200.2007.740.600-2	Ana Lucia	Banco Santander	Cobrança	Petição Inicial; Contestação; Impugnação a Contestação; Petições; Minuta de agravo de Instrumento; Agravo de Instrumento; Contra-Razões ao recurso.	A1; G1; P1 A2; ARQUIVO	11ª VARA CÍVEL DE JOÃO PESSOA

Pronto

Figura 2 – Modelo adotado na realização da Indexação Automatizada de Termos pelo uso do Excel.

O processo de recuperação é bem simples de ser executado, foi nesta perspectiva que foi pensado e articulado o modelo, para que qualquer pessoa pudesse acessar a informação desejada.

Para mostrar a importância do banco de dados, perguntamos aos entrevistados se hoje, após a aplicação experimental deste modelo de banco de dados, e conseqüentemente da indexação automatizada, como se encontra as buscas aos documentos? Há atualização dessas informações nessa base de dados?

Eles afirmaram que:

Ao ser realizado o Banco de Dados no computador e ali foram inseridas relações de processos, que podem ser utilizados para a localização dos documentos. A identificação das partes e organização dos documentos em si dentro delas, pela ordem, dentro de uma sistematização para mim foi muito importante.

Como coloca a própria teoria sobre indexação, procuramos, na criação do banco de dados, elaborá-lo com bastante clareza, simplicidade e coerência: fácil manuseio, sem ambigüidades ou sobreposições; flexibilidade e agilidade; expansibilidade: permitindo que novos documentos sejam incorporados; e eficácia na entrada e na busca de informações além do baixo custo ao utilizarmos um *software* disponível nos próprios computadores da instituição.

Assim sendo, para Wolthers (1998) *apud* Rezende (2000, p. 52), “o desafio é criar sistemas de consultas que levem ao usuário a documento procurado”. A questão, entretanto, não é apenas organizar os arquivos, e sim mantê-los em perfeito funcionamento. É nesta parte que se torna importante e indispensável a presença do profissional arquivista.

7 CONSIDERAÇÕES FINAIS

Com relação à Indexação Automatizada, o maior dilema enfrentado pelos arquivos hoje em dia para a implementação dessa ferramenta, pode ser o de não saber exatamente o que é uma indexação por si só, nem muito menos o que seria uma indexação automatizada e em quais situações ela pode ser útil.

Neste trabalho, nosso objetivo primordial foi o de chamar a atenção para os profissionais da área jurídica quanto ao uso da indexação automatizada e alguns dos suportes dados a partir de seu uso, em ambientes arquivísticos para a recuperação e a representação da informação. Mostramos o quanto é fácil aplicar e buscar a informação no banco de dados elaborado.

Esperamos ainda, que haja uma maior valorização quanto ao profissional arquivista nas práticas de automação de arquivos, e a participação do mesmo nos planos de gestão da informação do arquivo, bem como nos processos de gerenciamento de documentos eletrônicos.

Esperamos ainda, que a indexação automatizada possa ser vista como um instrumento facilitador da atividade de indexação. O arquivista também pode contribuir nas atividades de construção e de avaliação dos sistemas automáticos. Para tanto, espera-se que a ferramenta Excel, venha a contribuir na vida dos profissionais da informação – os arquivistas - uma vez que, os recursos para a indexação automática são bastantes altos e diante a isto, a indexação automatizada se torna como uma opção a mais na recuperação da informação obtendo precisão e agilidade em suas buscas, com um baixo custo.

Finalizamos a entrevista perguntando aos entrevistados se: Depois de realizada a aplicabilidade da indexação automatizada, você sugere a adoção dessa ferramenta em outros escritórios de advocacia? Por quê?

Sugiro o trabalho do profissional arquivista em escritório, ou de estagiário em sendo o caso, pois o trabalho deste profissional ou futuro profissional é de grande importância para o arquivo do escritório, e controle, portanto, das informações, dos documentos, dos processos. E a indexação dos dados, fazendo com que estes dados possam ser localizados em uma relação existente e prévia, seja manual, seja eletrônica, através do computador, é muito interessante para o trabalho, e agiliza a busca pelos documentos, processos, agilizando o trabalho, sobretudo quando o escritório for crescendo e for aumentando a demanda do arquivo.

Asseveramos ainda que, a indexação automatizada, a partir de sua adesão, irá facilitar a busca das informações desejadas pelos usuários.

Portanto, deve haver um vocabulário controlado, e os gestores junto a um profissional arquivista, devem elencar as palavras-chave do processo, esclarecendo na parte de tipologias documentais, quais serão os tipos documentais que se encontrarão ali, assim ao buscar a informação no banco de dado, o interessado já saberá onde está a sua informação, ou seja, qual gaveta e pasta ela se encontra.

Na verdade, é um trabalho que exige integração de ambas as partes, arquivista e advogados, que mutuamente se orientam para a melhor indexação dos documentos e da preservação do sistema.

Intentamos à guisa de conclusão, que esse trabalho contribua para futuras pesquisas sobre a importância de indexação automatizada no processo de recuperação da informação, uma vez que a busca por uma ferramenta tecnológica deve ser vista como forma de se revitalizarem antigas formas de acesso à informação com o intento de melhorar ou estimular as metas de recuperação. Além disso, deverá agregar valores, inovando os programas e os processos já existentes do uso do Excel quanto elemento de um Sistema de Recuperação a Informação.

REFERÊNCIAS

ALVES, Ivone *et al.* **Dicionário de terminologia arquivística**. Lisboa: Instituto da Biblioteca Nacional, 1993.

ANDERSON, James D.; PEREZ-CARBALLO, José. The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part 2: Machine indexing, and the allocation of human versus machine effort. **Information Processing and Management**, v. 37, n. 2, 2001.

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

BARRETO, Aldo Albuquerque. **O rumor do conhecimento**. 1997. Disponível em: < <http://aldoibct.bighost.com.br/rumor/orumor.pdf> >. Acesso em: 20 abr. 2010.

BATTES, Marcia J. Indexing and access for digital libraries and the internet. **Journal of the American Society for information Science**, 49, 1998.

BORKO, Haroldo; BERNIER, Charles. **Indexing concepts and methods**. New York: Academic Press, 1978.

CASTELLS, Manuel. **A sociedade em rede**. São Paulo: Paz e Terra, 2005. v. 1

CASTRO, Astréa de Moraes e *et al.* **Arquivística Arquivologia: técnica e ciência**. Brasília: Ao Livro Técnico, 1988.

CHAUMIER, Jacques. **Travail et methodes du/de la documentaliste: connaissance du problème**. Paris : ESF/Libraries Techniques, 1980. Exposé 3, Chap.3: L'indexation.

COLLISON, Robert Lewis. **Índices e indexação**: guia para indexação de livros, e coleções de livros, periódicos, e coleções de livros, periódicos, partituras musicais, com uma seção de referência e sugestões para leitura adicional. Tradução de Antônio Agenor Brinquet de Lemos. São Paulo: Polígono, [1971].

DALEN, Deonol B.; MEYER, William J. **Manual de técnica de la investigación educacional**. Buenos Aires: Paidós, 1974.

DAVENPORT, Thomas H. **Ecologia da informação**: porque só a tecnologia não basta para o sucesso na era da informação; tradução Bernadete Siqueira Abrão. São Paulo: Futura, 2002.

FARMER, Linda. Automatic categorization: what's it all about? **Serials Librarian**, v. 51, n. 2, 2006.

FÁVERO, Leonor Lopes; KOCH, Ingedore G. Villaça. **Linguística textual: introdução**. São Paulo: Cortez, 1988.

FOSKETT, Antony Charles. **A abordagem temática da informação**. Tradução de Antonio Agenor Briquet de Lemos. São Paulo: Polígono, 1973.

FUJITA, Mariângela Spotti Lopes. **PRECIS na língua portuguesa: teoria e prática**. Brasília: Universidade de Brasília, 1989.

GALVÃO, Maria Cristina Barbosa. **Manual de Softwares: facilitando a comunicação empresa-sociedade**. Brasília. Universidade de Brasília. 2003.

GIL, Antonio Carlos. **Como elaborar Projetos de Pesquisa**. 3. ed. São Paulo: Atlas, 1996.

GIL LEIVA, Isidoro. **Manual de indización**. Teoría y práctica. Gijón: Trea, 2008, 429 p. Sumário e prólogo da obra disponível em: webs.um.es/isgil.

_____. **La automatización de la indización de documentos**. Gijón (Astúrias): Eiciones Trea, 1999.

HJORLAND, Birger Toward. A theory of aboutness, subject, topicaliy, theme, domain, Field, content...and relevance. **Journal of the American Society for information Science and Technology**, 51, 2001.

Houaiss Antonio, Villar M de S, Franco FM de. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva; 2001.

INNARELLI, Humberto Celeste. Preservação digital e seus dez mandamentos. *In*: SANTOS, Vanderlei Batista; INNARELLI, Humberto Celeste; SOUSA, Renato Tarciso Barbosa (Orgs.). **Arquivística: temas contemporâneos: classificação, preservação digital, gestão do conhecimento**. Distrito Federal: SENAC, 2007.

JAKOBSON, Roman. **Lingüística e comunicação**. São Paulo: Cultrix, 1974.

LANCASTER, Frederic Wilfrid. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos, 2004.

_____. **Indexação e resumos: teoria e prática**. Trad. de Antonio Agenor Briquet de Lemos. Brasília: Briquet de Lemos, 1993.

LAND, F. Adapting to Changing User Requirements, **Information & Management**, 5(1), 1982.

LOPES, Luís Carlos. **A nova arquivística na modernização administrativa**. Rio de Janeiro : [s. n.], 2000.

MARCONI, Marina de Andrade. LAKATOS, Eva Maria. **Fundamentos da Metodologia Científica**. 6. Ed. São Paulo: Atlas. 2008.

MÉNDEZ RODRÍGUEZ, Eva Maria; MOREIRO GONZÁLEZ, José Antonio. Lenguaje natural e indización automatizada. **Ciencias de la Información**, v. 30 , n. 3, 1999.

MINAYO, Maria Cecília de Souza. **O desafio do conhecimento**: pesquisa qualitativa em saúde. 11. ed. São Paulo: HUCITEC; Rio de Janeiro: ABRASCO, 2008.

_____. **O desafio do conhecimento**: pesquisa qualitativa em saúde. 2. ed. São Paulo: Hucitec, 1993.

MOREIRO GONZÁLEZ, José Antônio. **El contenido de los documentos textuales**: su análisis y representación mediante el lenguaje natural. Gijón (Astúrias): Trea, 2004.

NAVES, Madalena Martins Lopes. **Curso de Indexação**: princípios e técnicas de indexação, com vistas à recuperação da informação. 2004. (Desenvolvimento de material didático ou instrucional - Apostila).

NEGREIROS, Leandro Ribeiro. **Sistemas eletrônicos de gerenciamento de documentos arquivísticos**. Dissertação (Curso de Mestrado em Ciência da Informação). Escola de Ciência da Informação da Universidade Federal de Minas Gerais de Minas Gerais: Belo Horizonte, 2007.

PAES, Marilena Leite. **Arquivo**: teoria e prática. Rio de Janeiro: FGV, 2004.

PERRENOUD, Philippe. **Dez novas competências para ensinar**. Porto Alegre: Artes Médicas Sul, 2000.

PINTO, Virgínia Bentes. Indexação documentária: uma forma de representação do conhecimento registrado. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 6, n. 2, 2001.

REZENDE, Ana Paula; MACHADO, Valéria Maria. Centro de informação jurídica eletrônico e virtual. **Ciência da informação**, Brasília, v. 29, n. 1, 2000.

REZENDE, Martha Izabel de Souza Duarte. Gerenciamento da informação e do conhecimento no escritório de advocacia. *In...* / CIBERÉTICA. **Simpósio** Internacional de Propriedade Intelectual, Informação e Ética. Florianópolis, 19 a 21 de novembro de 1998.

RICHARDSON, Roberto Jarry. **Pesquisa Social**: métodos e técnicas. 2. ed. São Paulo: Atlas, 1989.

ROBREDO, Jaime. **Documentação de hoje e amanhã**: uma abordagem informatizada de biblioteconomia e dos sistemas de informação. 2. ed. ver. ampl. Brasília: Edição do autor, 1994.

RODRIGUES, Rui Martinho. **Pesquisa acadêmica**: como facilitar o processo de preparação de suas etapas. São Paulo: Atlas, 2007.

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

_____. **Les fondements de la discipline archivistique**. Québec: Université du Québec, 1994.

ROTHMAN, J. Is indexing obsolete? Keyword indexing and free-text searching. In: **FEINBERG**, H. Indexing specialized formats and subjects. London : Scarecrow, 1983.

SAGER, Juan C. **English special languages: principles and practice in science and technology**. Wiesbaden: Brandstetter, 1980.

SANTOS, Viviane Neves dos. **Indexação automática de documentos textuais: iniciativas dos grupos de pesquisa de universidades públicas brasileiras**. 2009. 72 p. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) - Departamento de Biblioteconomia e Documentação, Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2009.

SILVA, Edna Lúcia da; CUNHA, Miriam Vieira da. A formação profissional no século XXI: desafios e dilemas. **Ciência da Informação**, Brasília, v. 31, n. 3, 2002.

SILVA, Maria do Remédios da; FUJITA, Mariângela Spotti. Lopes. A prática de indexação: análise da evolução e tendências teóricas e metodológicas. **Transinformação**, Campinas, v. 16, n. 2, 2004. Disponível em: < <http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=65> >. Acesso em: 15 abr. 2009.

SOUTO, Sônia Miranda de Oliveira. **O profissional da informação frente às tecnologias do novo milênio e as exigências do mundo do trabalho**. 2006. Disponível em: <<http://www.cinform.ufba.br/ivanais/artigos/texto16.htm>> Acesso em: 23 out. 2008.

STREHL, Letícia. Avaliação da consistência da indexação realizada em uma biblioteca universitária de artes. **Ciência da Informação**, v. 27, n. 3, Brasília, 1998.

TRIVIÑOS, Augusto N. S. **Introdução à pesquisa em ciências sociais: a pesquisa qualitativa em educação**. São Paulo: Atlas, 2006.

UNISIST. Princípios de indexação. **Escola de Biblioteconomia, Universidade Federal de Minas Gerais**. Belo Horizonte, v. 10, n. 1, 1981.

VIEIRA, Simone Bastos. Indexação automática e manual: revisão de literatura. **Ciência da Informação**, Brasília, v. 17, n. 1, 1988.

_____. **Análise comparativa entre indexação automática e manual da literatura Brasileira de ciência da informação**. Dissertação (Mestrado em Ciência da Informação) Departamento de Ciência da Informação e Documentação da Universidade de Brasília, Brasília, 1984.

APÊNDICE

A entrevista realizada

1. Antes do processo de automatização da indexação, e até mesmo sem a presença de um profissional arquivista, como eram indexados os termos para a localização dos documentos e/ou processos no escritório?
2. Essa busca (manual) era considerada rápida diante da grande massa documental?
3. Quais foram as suas maiores dificuldades no processo manual de recuperação de informação, e por quê?
4. Depois de organizado o arquivo, você acreditaria que, somente a organização por si só, sem a presença de uma indexação automatizada, era bastante suficiente para a acessibilidade dos processos?
5. A experiência da indexação automatizada atendeu as necessidades do arquivo, de forma eficaz, tais como, rapidez no processo de busca pela documentação desejada?
6. Com a elaboração de um banco de dados para subsidiar na busca mais ágil aos processos, a elaboração de um vocabulário controlado, escolhidos entre o gestor e o arquivista, foi fundamental nas buscas documentais? Você acha que a linguagem utilizada foi precisa?
7. Hoje, pós-aplicação experimental deste banco de dados, e conseqüentemente da indexação automatizada, como se encontra a busca aos documentos? Há atualização dessas informações nessa base de dados?
8. Depois de realizada a aplicabilidade da indexação automatizada, você sugere a adoção dessa ferramenta em outros escritórios de advocacia? Por quê?