

GT3

Organização da informação e do conhecimento no século XXI

A ELABORAÇÃO E USO DE VOCABULÁRIO CONTROLADO EM SISTEMA DE GERENCIAMENTO ELETRÔNICO DE INFORMAÇÕES ARQUIVÍSTICAS: a experiência da Embrapa Soja

FREITAS, L¹. M., SIMONATO, A.C²., ALBUQUERQUE, A.C³.

RESUMO: Este estudo tem como objetivo descrever a metodologia de elaboração de vocabulário controlado para representação de conteúdo em um sistema de gerenciamento eletrônico de informações arquivística da Embrapa Soja, na perspectiva da arquivística integrada. Relata os benefícios alcançados no uso de vocabulário controlado para organização, tratamento e recuperação de informações no sistema. Perpassa as questões conceituais de vocabulário controlado e Sistemas de GED – Gerenciamento Eletrônico de Informações. Foi constatada melhora significativa na pesquisa/busca de informações, por conseguinte da adoção de vocabulário controlado, que permite padronização no uso de termos que representam o conteúdo informacional dos documentos.

Palavras-chave: Controle de vocabulário. Vocabulário controlado. Sistema de gerenciamento eletrônico de informação arquivística. Arquivística integrada.

ABSTRACT: This study aims to describe the vocabulary development methodology controlled for content representation on an electronic management system of records to the Embrapa Soja, from the perspective of integrated archiving. Reports the benefits achieved in the use of controlled vocabulary for organizing, processing and retrieval of information in the system. It covers the conceptual issues of controlled vocabulary and GED Systems - Electronic Management of Information. Significant improvement was observed in the research / search for information, therefore the controlled vocabulary of adoption, enabling standardization in the use of terms representing the informational content of the documents.

Keywords: vocabulary control. Controlled vocabulary. Electronic management system of archival information. Integrated archival.

1 INTRODUÇÃO

Os sistemas informatizados trouxeram muitos benefícios às organizações no que concerne a agilidade na movimentação de informações, acesso remoto de dados, armazenamento, entre outros. No entanto, a recuperação de informações é ameaçada quando a busca se dá de forma dissociativa do registro, ou seja, a busca é feita de maneira diferenciada da entrada de dados no sistema.

¹ Mestranda do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina (UEL) e arquivista da Embrapa Soja. Endereço Residencial: Rua Hélio Antônio A. Bevenho, 45, Avelino Vieira, Londrina/PR Telefone: (43) 3348-2095 e 9994-6381, e-mail: lidiane.uel@gmail.com

² Professora do Departamento de Ciência da Informação da Universidade Estadual de Londrina (UEL)

³ Professora do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina (UEL)

Essa dissonância impacta na qualidade do resultado das informações recuperadas, que tende a ser imprecisas e generalizadas. Quando não se observa a linguagem usada, surgem ruídos de comunicação, que, geralmente, são atribuídos aos problemas de comunicação e processamento técnico, não se relaciona a questão da linguagem utilizada, dos termos que foram inseridos como representantes do conteúdo informacional (AGUIAR, 2008).

O uso da linguagem natural no discurso comum dificulta a rapidez da recuperação de informações e a precisão dos resultados dessa busca. As informações inseridas de acordo com o pragmatismo da linguagem natural dão margem à dispersão e inconsistência de informações. De acordo com Smit e Kobashi (2003) e Aguiar (2008), o controle da linguagem, dos termos utilizados no sistema é uma forma de produzir confiança para os usuários na organização e busca de informações.

Na linguagem natural as palavras podem assumir diversos significados e gerar ambiguidade de sentido. Uma mesma palavra pode ter significados completamente diferentes quando inseridas em contextos ou áreas distintas.

Conforme afirma Aguiar (2008) ao normalizar o uso da linguagem, tem-se a oportunidade de minimizar as ambiguidades e plurivocidade dos termos nomeados para representar as atividades organizacionais e tipos documentais, permitindo-se assim o compartilhamento eficaz e a satisfação dos usuários.

Nesse sentido, destaca-se a importância do vocabulário utilizado nos sistemas informatizados para tratamento, organização e recuperação da informação. Pois, do que vale estabelecer um complexo sistema de organização de dados, critérios de preservação, agilidade no processamento técnico, uso de equipamentos transmissor de última geração, se não pensar nos critérios e normalização da linguagem utilizada?

Assim, com a finalidade de contribuir com o pensar e as práticas arquivísticas e da observação de carência de pesquisas sobre o fazer arquivístico, este trabalho se propôs a apresentar o controle de vocabulários em um sistema de gerenciamento eletrônico de informação arquivística da Embrapa Soja, sob a perspectiva da arquivística integrada, corrente que estabelece o tratamento das informações arquivísticas desde a produção até seu destino final.

2 CONTROLE DE VOCABULÁRIO NA PERSPECTIVA DA ARQUIVÍSTICA INTEGRADA

Informação arquivística é toda informação produzida ou recebida no contexto do exercício das funções e atividades das organizações ou pessoas e está materializada ou registrada em algum tipo de suporte. De acordo com Moreno (2006) as informações arquivísticas se constituem em prova documental de extrema necessidade à vida das organizações, pois servem como fonte de pesquisa da memória institucional e também trazem autenticidade e transparência de suas ações.

Na perspectiva da arquivística integrada, corrente contemporânea da arquivística, desenvolvida por Rousseau e Couture (1982), o tratamento das informações arquivísticas se dá desde a produção até seu destino final.

A arquivística integrada, desenvolvida no Canadá por Rousseau e Couture (1982, 1998), propõe uma arquivística que se preocupa com o tratamento da informação desde seu nascimento até seu destino final. A arquivística é tratada como a disciplina que agrupa todos os princípios, normas e técnicas que regem as funções de gestão dos arquivos, tais como a criação, a avaliação, a aquisição, a classificação, a descrição, a comunicação e a conservação. (GARCIA e SCHUCH JUNIOR, 2002, p.46).

É nesse sentido que o controle de vocabulários desde a produção/registro de informações no sistema de gerenciamento eletrônico de informações arquivísticas da Embrapa Soja se justificou, pois o controle da linguagem documentária (vocabulário controlado, tesouros, índices) é muito utilizado como instrumento de pesquisa em arquivos históricos e centro de documentação, no tratamento de documentos de valor secundário. Pouco se observa o uso de vocabulário controlado em arquivos correntes, como tratamento e organização de documentos de valor primário.

A produção de documentos no sistema da Embrapa Soja é feita por usuário não especializado, profissionais das diversas áreas da empresa. Diante disso, a inexistência de um vocabulário controlado, para o registro e recuperação da informação, torna o sistema ineficiente.

O vocabulário controlado, artefato das práticas de organização e tratamento da informação arquivística, apresenta-se como ferramenta de apoio a recuperação de informações. Sua principal função é servir na representação de conteúdos com a finalidade de facilitar a identificação e recuperação das informações. Smit e Kobashi (2003, p. 20) esclarece que o controle de vocabulário é o processo de desenvolvimento

de um instrumento, “um objeto que se deseja atingir” e o vocabulário controlado, o resultado do processo de controle, o instrumento para nomear as atividades/funções.

Portanto, o vocabulário controlado é um instrumento importante para os sistemas informatizados, pois “têm por objetivo principal o controle da polissemia que ocorre na linguagem natural” (SOUZA, 2011, p. 425). Na perspectiva da recuperação da informação, o vocabulário controlado é um instrumento de controle terminológico (AGUIAR, 2008, p. 222). Pois, é um instrumento construído com uma linguagem artificial, constituído de termos organizados em estrutura relacional. (KOBASHI, 2008).

Sendo assim, controlar o vocabulário para representação da informação arquivística é fundamental para assegurar a recuperação rápida e precisa das informações, principalmente nos sistemas informatizados como sistemas de gerenciamento eletrônico de informações arquivísticas.

2.1 Metodologia de elaboração

Aguiar (2008) ressalta que a padronização não é sinônimo de normalização com base terminológica. Portanto, para elaboração do vocabulário controlado identificou-se que o plano de classificação das atividades meio do CONARQ, utilizado no controle dos arquivos, não seria satisfatório para representar, de forma controlada e de fácil usabilidade para os usuários do sistema de gerenciamento eletrônico de informações arquivísticas, o conteúdo informacional das funções, atividades e documentos da Embrapa Soja. A classificação de documentos contém um padrão de assuntos hierarquizados, mas não contempla a normalização dos termos. Como a manipulação do sistema é feita por usuário não especializado e o foco do vocabulário controlado é o usuário, evidenciou-se a necessidade de controle de vocabulário.

De acordo com Smit e Kobashi (2003) o controle de vocabulário deve compor o desenvolvimento de procedimentos macros e micros. Os procedimentos micros são a introdução de controle dos termos ou expressões, o estabelecimento da normalização gramatical, as opções de grafias, o controle e padronização dos nomes de pessoas, lugares e eventos, o controle de sinonímia, homonímia e quase-sinonímia e a padronização no uso de termos compostos. Após a execução dos procedimentos micros, são feitos os procedimentos macros para gerar sentido ao todo, por meio da ordenação dos termos, sendo estabelecidas categorização, estruturação e hierarquização que dê sentido e significação ao vocabulário controlado.

Assim, a elaboração do instrumento contou com profissionais de diversos setores da empresa, todos da área administrativa, pois o desenvolvimento de vocabulário controlado requer um trabalho colaborativo e integrado, “envolvendo tanto os gerenciadores do sistema de informação quanto os usuários da informação” (KABOSHI, 2008, p.2).

As etapas de construção seguiram a metodologia proposta por Smit e Kobashi (2008), assim descritas:

1. Criação de um grupo de trabalho envolvendo profissionais da informação, arquivistas, e de outras áreas, como: recursos humanos, finanças, patrimônio e suprimentos, secretariado-executivo, comunicação, entre outros;
2. Identificação e descrição das tipologias documentais produzidas na área administrativa da empresa;
3. Construção de uma lista de categorias, utilizando-se do plano de classificação das atividades meio do Conselho Nacional de Arquivos – CONARQ;
4. Elaboração de lista livre de termos utilizados tanto do plano de classificação, como das tipologias (nomes dados aos documentos, processos e dossiês da área administrativa);
5. Controle dos termos a ser utilizado (normalização e padronização);
6. Definição de categorias (termos gerais) e estabelecimento de hierarquia (termos específicos – vinculação as categorias gerais);
7. Submissão dos termos ao grupo de trabalho e usuários do sistema.

Vale lembrar que o processo de controle de vocabulário não é um processo pronto e acabado, pois depende de contínua atualização. Com essa preocupação foi delegada a alguns profissionais a responsabilidade de identificação de novas demandas, necessidade de inclusão de novos termos, cujo vocabulário não atende à representação e à renomeação ou exclusão de termos defasados.

Segundo Smit e Kobashi (2003, p. 42) um vocabulário controlado sob a óptica arquivística tem que necessariamente contemplar:

- Hierarquização das funções e respectivas atividades;
- Inclusão de remissivas;
- Relacionamento de termos;
- Descrição dos termos e condições de uso.

Como o vocabulário controlado foi elaborado para uso no sistema de gestão eletrônica de informações arquivísticas e é utilizado por usuário não especializado, nem todos os aspectos descritos pelos autores Smit e Kobashi (2003) foram contemplados no módulo de entrada de dados no sistema.

A hierarquização das funções e atividades encontra-se bem nítida no cadastro de documentos, de modo que o usuário pode utilizá-la facilmente. As remissivas também são passíveis de visualização pelo usuário não especializado, por meio de uma listagem alfabética de termos. O relacionamento dos termos, a descrição e as condições de uso foram incluídos no módulo da área de arquivo, por serem informações mais técnicas e específicas, considerou-se mais adequada para uso de profissionais especialistas, como arquivistas.

O objetivo de controlar o vocabulário é orientar a organização e tratamento de informações arquivísticas e contribuir para evitar a confusão na nomenclatura terminológica dos documentos, das funções e atividades e dos assuntos representativos. O uso de vocabulário controlado de acordo com Aguiar (2008, p.229) se justifica “para superar as dificuldades terminológicas provocadas pelo uso indiscriminado da linguagem natural”.

O vocabulário controlado, além de servir a sua função principal de facilitar na recuperação de informações, também orienta na organização física das séries documentais.

3 APLICAÇÃO E USO DO VOCABULÁRIO CONTROLADO NO SISTEMA DE GERENCIAMENTO ELETRÔNICO DE INFORMAÇÕES ARQUIVÍSTICAS DA EMBRAPA SOJA

O sistema de gerenciamento eletrônico de informações arquivísticas da Embrapa é responsável por gerar, controlar, armazenar e disponibilizar informações e documentos produzidos em meio eletrônico ou digitalizados.

O GED é um sistema que converte informações em voz, texto ou imagem para a forma digital. Funciona com softwares e hardwares específicos que permitem a captação, o armazenamento, a localização e o gerenciamento das versões digitais das informações. (SILVA; et al., 2003, p.2)

Segundo Villalobos e Oliver (2014, p. 28), o sistema de GED é uma “alternativa para o tratamento das informações não estruturadas”, ou seja, “aquelas informações que não estão cadastradas em bancos de dados formais ou corporativos”. Neles podem ser

inseridos documentos gerados em meio físico e transformados em documentos digitalizados, utilizando-se de hardware e software de captura, ou ainda, documentos nativamente digitais, aqueles gerados por e-mails ou programas como Word, Excel, Power Point, entre outros.

Acredita-se que os benefícios gerados por um GED colaboram com a busca das organizações pelo sucesso administrativo, pois, informação é recurso estratégico para conversão de uma ideia em ação (FIGUEIREDO, 1989).

O GED permite aos usuários acessarem documentos e informações de forma rápida e segura. No entanto, a rapidez é comprometida quando a forma de buscar pelas informações é dissonante da forma como entraram essas informações no sistema.

A Figura 1 ilustra o esquema de entrada e saída de dados no sistema de gerenciamento eletrônico de informações arquivísticas da Embrapa Soja com o controle de vocabulário e sem vocabulário controlado.

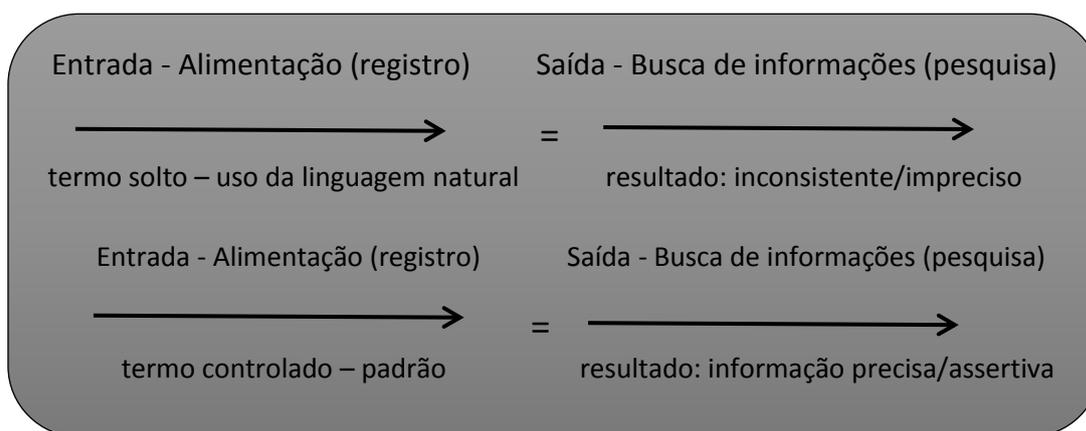


Figura01: Entrada e saída de dados de sistema informatizado com e sem uso de vocabulário controlado.

Fonte: Sistema eletrônico de informações arquivísticas da Embrapa Soja

Conforme aponta Smit e Kobashi (2003), o controle de vocabulário deve garantir que haja uma única forma de designação na entrada de dados/informações nos sistemas, o usuário deve saber como dar entrada nos dados para quando buscar informação, para que a recuperação seja assertiva, precisa e eficaz.

Sem o uso de vocabulário controlado no momento da entrada de dados, o usuário pode utilizar-se de termos aleatórios, termos soltos. Mas, essa medida dificulta no momento da busca, pois, se torna exaustiva, com a necessidade de dedução do termo utilizado na indexação. Portanto, anteriormente a aplicação do vocabulário controlado no sistema, a qualidade do resultado das buscas era inferior à qualidade obtida com o seu uso.

A título de exemplo, antes da adoção do vocabulário controlado, foi encontrado no sistema eletrônico de informações arquivísticas, o uso de diversos termos distintos para designar um mesmo documento ou processo, causando dispersão informacional (Tabela 1).

Tabela 1: Relação de termos soltos e padronizados utilizados para representação informacional dos documentos e processos no sistema eletrônico de informações arquivísticas da Embrapa Soja.

ÁREA	DOCUMENTO OU PROCESSO	TERMO SOLTO	TERMO CONTROLADO
Setor de Compras	Processo de compras	<ul style="list-style-type: none"> • Compra(s) • compra • aquisição • contratação. 	<ul style="list-style-type: none"> • compra
Setor de Recursos Humanos	Processo de capacitação no exterior	<ul style="list-style-type: none"> • capacitação • treinamento • curso no exterior • curso fora da Embrapa 	<ul style="list-style-type: none"> • capacitação no exterior
	Folha de frequência	<ul style="list-style-type: none"> • folha de frequência • folha ponto • cartão ponto 	<ul style="list-style-type: none"> • frequência
	Atestado médico	<ul style="list-style-type: none"> • atestado médico • atestado de saúde • justificativa médica • justificativa de falta • justificativa de ausência 	<ul style="list-style-type: none"> • justificativa de ausência
Setor Financeiro	Planilha orçamentária	<ul style="list-style-type: none"> • proposta orçamentária • programação orçamentária • previsão orçamentária • programação financeira 	<ul style="list-style-type: none"> • programação orçamentária
	Boleto de pagamento	<ul style="list-style-type: none"> • boleto • fatura • contas a pagar • despesa(s) • despesa 	<ul style="list-style-type: none"> • despesa

Fonte: Sistema eletrônico de informações arquivísticas da Embrapa Soja

Com o controle do vocabulário utilizado no momento do registro/produção de documentos ou processos, o usuário, mesmo não especializado, é induzido à seleção de

termos previamente estabelecidos, cuja significância representa o conteúdo informacional do assunto dos documentos, propiciando a padronização dos termos.

O sistema de gerenciamento eletrônico de documentos arquivísticos da Embrapa mantém uma interface amigável com uma lista de termos (descritores) que representam processos de trabalho da área administrativa da organização. A função do controle de vocabulários no sistema é representar atividades da área administrativa da empresa. Foram selecionados termos que representassem de forma padrão e normalizada o conteúdo informacional das funções e atividades da área administrativa, bem como os assuntos emanados dos documentos. O uso de vocabulário controlado no sistema facilitou na compreensão e linguagem igualitária sobre a nomenclatura das atividades e documentos.

4 CONSIDERAÇÕES FINAIS

A complexidade da estrutura da organização, as muitas mudanças no seu organograma, a multiplicidade de termos utilizados para denominar as áreas, as atividades e os documentos dificultam na recuperação de informações arquivísticas. Foi nesse sentido, que a elaboração do vocabulário controlado para o sistema eletrônico de documentos arquivísticos da Embrapa se justificou.

O vocabulário controlado apresentou-se como insumo estratégico para minimizar o problema de ambiguidades e plurivocidades dos termos na nomeação das funções e atividades e das tipologias documentais.

A adoção do vocabulário controlado gerou mais confiabilidade no uso do sistema eletrônico de documentos arquivísticos da Embrapa, porque propiciou maior satisfação nas respostas de buscas, que antes eram carregadas de incompreensão e confusão, devido às variações dos termos utilizados.

A metodologia utilizada para elaboração do vocabulário controlado abrangeu os procedimentos macro e micro sugeridos por Smit e Kobashi (2003), com enfoque da perspectiva da arquivística integrada, que propõe o tratamento da informação desde a sua produção até seu destino final.

Assim, o controle de vocabulário está sendo executado desde a produção/registro de informações no sistema para proporcionar mais eficiência na recuperação de informações e também para auxiliar na orientação da organização das séries documentais.

REFERÊNCIAS

AGUIAR, F. L. de. **O controle de vocabulário como dispositivo para organização e tratamento e recuperação da informação arquivística**. 2008. 267 f. Dissertação (Mestrado em Ciências da Informação) – Pontífica Universidade Católica de Campinas, Campinas, 2008.

FIGUEIREDO, N. Inovação, produtividade e sistema de informação. **Ciência da Informação**, Brasília, DF, v. 18, n. 1, p. 83-95, jan./jun. 1989. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/viewFile/1696/1300>>. Acesso em: 1 jul. 2014.

GARCIA, O. M. C.; SCHUCH JUNIOR, V. F. A aplicação da arquivística integrada, considerando os desdobramentos do processo a partir da classificação. **Informação & Informação**, Londrina, v. 7, n. 1, p. 41-56, jan./jun. 2002.

KOBASHI, N. Y. **Vocabulário controlado: estrutura e utilização**. [S.l.]: ENAP, 2008. Texto preparado para subsidiar debate na Reunião da Rede de Escolas do Governo. Disponível em: <http://www2.enap.gov.br/rede_escolas/arquivos/vocabulario_controlado.pdf>. Acesso em: 15 set. 2014.

MORENO, N. A. **A informação arquivística no processo de tomada de decisão em organizações universitárias**. 2006. 184 f. Tese (Doutorado em Ciências da Informação) – Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

ROUSSEAU, J. Y.; COUTURE, C. **Os fundamentos da disciplina arquivística**. Lisboa: Dom Quixote, 1998.

ROUSSEAU, J. Y.; COUTURE, C. **Les archives au XX siècle**. Montreal: Université de Montréal, 1982.

SILVA, D. P. et al. **Gerenciamento eletrônico de documentos: a tecnologia que está mudando o mundo**. Itera soluções. Disponível em: <http://www.iterasolucoes.com.br/Site/images/stories/Itera/SalaLeitura/ged_gerenciamento_eletronico_de_documentos.pdf>. Acesso em: 03 jul. 2014.

SMIT, J. W.; KOBASHI, N. Y. **Como elaborar vocabulário controlado para aplicação em arquivos**. São Paulo: Arquivo do Estado e Imprensa Oficial do Estado de São Paulo, 2003. (Projeto como fazer, 10).

SOUZA, T. A. R. F. **O vocabulário controlado como instrumento de organização e representação da informação na FINEP**. Comunicação oral apresentada no 12. Encontro Nacional de Pesquisa em Ciência da Informação, Brasília, DF, out. 2011. Disponível em: <<http://repositorio.ibict.br/bitstream/123456789/88/1/RosaliTatianaEnancib2011b.pdf>>. Acesso em: 15 set. 2014.

VILLALOBOS, A. P. de O.; OLIVER, P. **A gestão informatizada de documentos no Tribunal de Contas dos municípios da Bahia**. Ponto de Acesso, Salvador, v. 8, n. 1, p. 24-38, abr. 2014. Disponível em: <http://www.iterasolucoes.com.br/Site/images/stories/Itera/SalaLeitura/ged_gerenciamento_eletronico_de_documentos.pdf>. Acesso em: 16 set. 2014.