

Comunicação Oral

PARÂMETROS SOCIOCOGNITIVOS DE CONSTRUÇÃO DE INSTRUMENTO DE REPRESENTAÇÃO TEMÁTICA DA INFORMAÇÃO DE ÁREAS TÉCNICO-CIENTÍFICAS

Greissi Gomes Oliveira – UFSCAR
Vera Regina Casari Boccato – UFSCAR

Resumo

Identificou-se parâmetros de construção de um instrumento de representação temática da informação, diante das literaturas técnica e científica na área de Organização e Representação do Conhecimento, no contexto sociocognitivo de bibliotecários e usuários e pela perspectiva das Unidades de Informação Técnico-Científicas federais (UITCs). O universo de pesquisa foi as UITCs do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP), *campi* de São Paulo, São Roque e Sertãozinho no desenvolvimento da metodologia qualitativa com abordagem sociocognitiva. Os procedimentos metodológicos constaram das aplicações de questionários com os bibliotecários responsáveis para a caracterização sócio-organizacional das UITCs e do Protocolo Verbal em Grupo com três bibliotecários, três docentes e seis discentes, totalizando doze sujeitos de pesquisa, com o intuito de observar as opiniões e registrar as declarações dos sujeitos participantes sobre as características e as necessidades de uso de uma linguagem documentária representativa do contexto das UITCs e de sua comunidade usuária. Os resultados obtidos foram: os bibliotecários e os usuários visualizam a importância de um instrumento de representação temática para indexação e recuperação da informação; a utilização da linguagem natural não representa satisfatoriamente os assuntos dos documentos; os bibliotecários identificam-se com o uso de uma linguagem pós-coordenada no SRI; os termos da linguagem documentária devem contemplar as áreas de especialidade e a de busca dos usuários, entre outras ocorrências. Diante de tais resultados e, subsidiado pela literatura científica em ORC e pelas declarações dos sujeitos participantes da pesquisa, foram estabelecidos onze parâmetros de construção de um instrumento de representação temática no contexto sociocognitivo de bibliotecários e usuários de Unidades de Informação Técnico-Científicas. Conclui-se que uso de uma linguagem documentária construída a partir de parâmetros sociocognitivos propicia a representação para a recuperação da informação técnico-científica com qualidade, confiabilidade e rapidez.

Palavras-chave: Linguagens documentárias alfabéticas. Sistemas de recuperação da informação. Metodologia sociocognitiva. Bibliotecas técnico-científicas. Instituto Federal de Educação, Ciência e Tecnologia de São Paulo.

Abstract

Identified construction parameters of an instrument of thematic representation of information on the technical and scientific literature in the area of Knowledge Representation and Organization in the context of social cognitive librarians and patrons and the prospect of Units Information Technical-Scientific federal (UITCs). The universe of the research was UITCs Federal Institute of Education, Science and Technology of São Paulo (IFSP), campuses of São Paulo, São Roque and Sertãozinho in the development of socio-cognitive approach to qualitative methodology . The methodological procedures consisted of the application of questionnaires with librarians responsible for the socio - organizational UITCs and Verbal

Protocol Group with three librarians, three teachers and six students, totaling twelve subjects research, in order to observe the opinions and record the participants' statements about the features and the need for use of an indexing language representative of the context of UITCs and its user community. The results were: librarians and users see the importance of an instrument of thematic representation for indexing and retrieval of information, the use of natural language is not satisfactorily subjects of documents; librarians identify themselves with the use of a language post-coordinated in SRI; terms of indexing language should cover specialty areas and user search, among other events. Given these results, and subsidized by the scientific literature on ORC and statements of subjects participating in the research were established Eleven parameters of building an instrument thematic representation in the context of social cognitive librarians and users of units scientific-technical information. We conclude that use of an indexing language built from social cognitive parameters provides representation for the recovery of scientific and technical information with quality, reliability and speed.

Keywords: Alphabetical indexing languages. Information retrieval systems. Sociocognitive methodology. Technical-scientific libraries. Federal Institute of Education, Science and Technology of São Paulo.

1 INTRODUÇÃO

O uso de sistemas de recuperação da informação (SRI) por unidades de informação, como bibliotecas, tem por objetivo permitir aos seus usuários o acesso aos conteúdos dos acervos destas unidades. A recuperação da informação nesses sistemas pode dar-se através da busca por nome do autor, por título da obra, por palavras do texto e através do tema ou assunto de uma obra. Para que ocorra a precisão na busca da informação por assunto, torna-se fundamental o uso de linguagens estruturadas, denominadas linguagens documentárias (LD), que são instrumentos com vistas a permitir representar os conteúdos do acervo. Nessa perspectiva, a temática de nossa pesquisa pautou-se nos parâmetros de construção de um instrumento de representação temática para recuperação de informação por assunto em unidades de informação técnico-científicas.

Nesta pesquisa, as Unidades de Informação Técnico-Científicas (UITCs) correspondem às bibliotecas do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo¹ (IFSP), autarquia de ensino pertencente ao governo federal do Brasil, que organizado em estrutura *multicampi* oferece cursos de nível técnico (equivalente ao ensino médio), tecnológico, licenciatura e bacharelado (equivalentes ao ensino superior).

A regulamentação dos Institutos Federais do país (BRASIL, 2008), à qual o IFSP é submetido, prevê que os mesmos destinem 50% de suas vagas a cursos técnicos e 20% a cursos superiores (de tecnologia, licenciatura, bacharelado, pós-graduação *stricto sensu e lato sensu*). Os Institutos Federais surgem como parte de políticas públicas para as regiões

¹ <http://www.ifsp.edu.br/>

geográficas em que estão inseridos já que os cursos oferecidos pelos *campi* procuram adaptar-se às necessidades profissionais desses locais de abrangência, orientando sua oferta formativa

em benefício da consolidação e fortalecimento dos arranjos produtivos, sociais e culturais locais, identificados com base no mapeamento das potencialidades de desenvolvimento socioeconômico e cultural no âmbito de atuação do Instituto Federal (BRASIL, 2008).

Dessa maneira, as UITCs dos Institutos possuem em seus acervos materiais que buscam contemplar todas as áreas técnico-científicas tratadas em cursos dos Institutos. Fazer uso da linguagem livre, ou seja, sem controle, para representação da informação de tais acervos, empregando as palavras tal como aparecem nos documentos pode, de acordo com Foskett (1973), apresentar falhas na recuperação tendo em vista a polissemia e a sinonímia. Ao passo que o tratamento temático desses acervos com o uso de uma linguagem documentária, oferece possibilidades de buscas de informação com maior qualidade em sistemas de recuperação da informação.

No que compete a construção de linguagem documentária, a participação do bibliotecário resulta na elaboração de um sistema de organização do conhecimento consistente, representativo dos diversos assuntos das áreas do conhecimento integrantes do acervo de uma unidade de informação. A contribuição do usuário conduz para a elaboração de uma linguagem com termos compatíveis com sua necessidade de busca e que permitirá padronizar o acesso ao SRI aumentando a precisão no processo de busca da informação.

Para isso, é importante a construção de uma linguagem documentária consistente, formada por vocabulário que represente o repertório das áreas científicas especializadas, a linguagem de busca do usuário e do contexto sócio-organizacional em que ela está inserida, seguindo, portanto, os princípios das garantias literária, de uso, organizacional e cultural (BOCCATO, 2009a, p. 119).

Nesse sentido, o sujeito bibliotecário e o sujeito usuário em suas tarefas de indexação e busca da informação, respectivamente, trazem consigo visões de mundo e conhecimentos específicos, exigindo, dessa forma, que a elaboração de ferramentas informacionais leve em conta o contexto sociocognitivo da informação, para terem reconhecimento por parte desses sujeitos.

A diversidade de áreas de conhecimento e de categorias de usuários às quais as unidades de informação do IFSP devem atender nos levou a definir o nosso problema de pesquisa como a ausência de parâmetros sociocognitivos para a construção de instrumento de representação temática da informação de áreas técnico-científicas.

Diante do exposto, nosso estudo tem por objetivo identificar parâmetros de construção de um instrumento de representação temática da informação, diante das literaturas técnica e científica na área de Organização e Representação do Conhecimento, no contexto sociocognitivo do bibliotecário e do usuário e pela perspectiva das unidades de informação técnico-científicas federais.

Nossa pesquisa justifica-se pela necessidade de uma linguagem documentária para representação e recuperação da informação em unidades de informação técnico-científicas que permitam a correta representação da informação por bibliotecários na atividade de indexação e o acesso a informação por usuários que buscam especificidade e qualidade na recuperação da informação.

2 AS LINGUAGENS DOCUMENTÁRIAS NA REPRESENTAÇÃO TEMÁTICA DA INFORMAÇÃO TÉCNICO-CIENTÍFICA

Os sistemas de recuperação da informação são formados por um grupo de atividades e componentes dedicados a promover o acesso ao conteúdo de documentos. (LANCASTER, 1987). Esses sistemas atuam como canal de comunicação entre documentos e usuários, buscando atender as necessidades de informação destes.

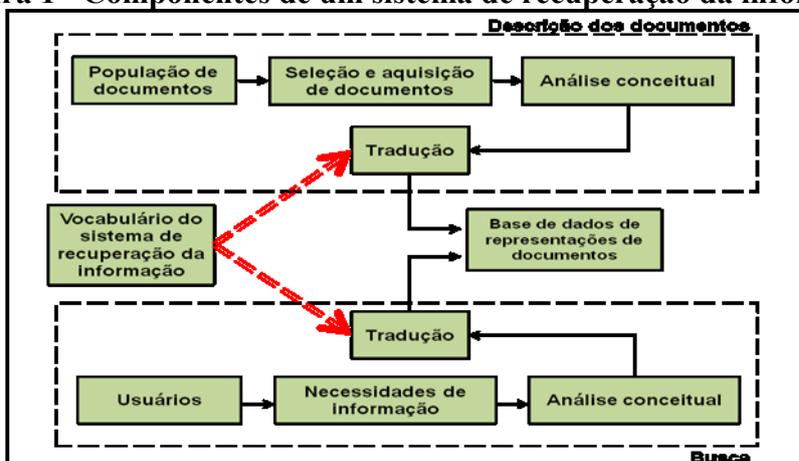
Cesarino (1985, p. 159) indica que os SRI podem ser estudados sob dois aspectos: enquanto um conjunto de operações realizadas para encontrar a informação desejada dentro de uma totalidade de informações ou “como parte de um modelo de comunicação dentro de um contexto social-cultural-histórico”, frisando que esses dois aspectos estão relacionados de maneira que o primeiro deve ser visto dentro do prisma do segundo.

Com base em Lancaster (2004), podemos dizer que um sistema de recuperação da informação é composto principalmente de: a) seleção e aquisição de documentos; b) descrição física e indexação dos documentos, enquanto processo de análise conceitual e tradução de documentos de acordo com o vocabulário do sistema; c) alimentação de base de dados e armazenamento do acervo; d) definição da estratégia de busca do usuário, onde também está presente o processo de análise conceitual das questões de busca e representação de acordo com a linguagem do sistema; e) interação usuário – sistema.

Temos que para a existência de um SRI é necessária (a) a seleção de documentos por uma unidade de informação onde tal seleção deve ocorrer de acordo com critérios que levem em conta a necessidade de seus usuários. Após a seleção, (b) são realizadas as representações descritiva e temática dos documentos compreendendo, a primeira, a descrição física, onde são indicadas informações como autoria, título, edição, editora entre outras e a segunda realizada

através da indexação desses documentos, enquanto análise conceitual e tradução. A partir dessas representações (c) são gerados registros que alimentam uma base de dados e os documentos são armazenados no acervo. Na busca o usuário, ao optar pelo uso da busca por assunto, interage utilizando a linguagem do SRI. Nessa fase faz-se presente novamente a análise conceitual para eleger os termos do vocabulário adotado pelo SRI que representem as necessidades do usuário e possibilitem elaborar a expressão de busca (d). Após a busca, são apresentados ao usuário os registros recuperados (e) para que seja possível analisar suas relevâncias e “pode ser que o usuário decida modificar sua estratégia de busca [...], reiniciando-se o processo” (CENDÓN, 2008, p. 64) (e). A Figura 1 demonstra os elementos de um SRI.

Figura 1 - Componentes de um sistema de recuperação da informação.



Fonte: Adaptado de Lancaster (2004, p. 18). Elaboração nossa.

Observando-se a figura 1, notamos que a linguagem documentária é um componente dos SRI, visando a Representação/Tradução dos conceitos identificados e selecionados previamente na Análise conceitual nos processos de indexação e de recuperação da informação.

Os catálogos de unidades de informação são um tipo de SRI e sofreram mudanças com relação ao suporte/forma de apresentação da informação: de manuais/fichas catalográficas para eletrônico (*online* ou não)/registros bibliográficos. Sobre o tratamento da informação, ele requer dois tipos de representação: 1) descritiva; 2) temática. No que compete à representação descritiva, temos nos códigos de catalogação diretrizes para elaborar suas representações, a exemplo do *Anglo-American Cataloguing Rules, 2nd edition (AACR2)*. Com relação à representação temática, ela pode ser realizada por meio da utilização de instrumentos de representação temática como as linguagens documentárias que permitem o controle de vocabulário de assunto.

Sobre as linguagens documentárias, Lara (2004, p. 232) expõe que o termo “[...] designa, de modo mais amplo e completo, a linguagem especialmente construída para organizar e facilitar o acesso e a transferência da informação [...]”. Diante disso, Boccato (2009a) expõe que elas possuem duas funções: 1) função pelo conteúdo: que consiste em representar o conteúdo dos documentos contidos em um SRI e; 2) função pelo uso: que procura mediar a recuperação da informação por meio da representação das perguntas formuladas pelos usuários na busca.

No cumprimento dessas funções, Cintra et al. (2002) e Boccato (2009a) revelam que as linguagens documentárias alfabéticas mais consistentes possuem um vocabulário formado por elementos da linguagem de especialidade, das terminologias e da linguagem natural expressa pelo usuário. Devido ao fato de serem linguagens construídas, o vocabulário (léxico) utilizado para a sua concepção é originário de discursos diferentes, representado por “palavras” preferenciais, numa relação combinatória de significados e significantes, denominadas de unidades lexicais de descritores.

Segundo Gardin et al. (1968 apud CINTRA et al., 2002), a construção de linguagens documentárias requer, a integração de três elementos básicos: 1) um léxico, que seria uma lista de elementos devidamente filtrados e depurados; 2) uma rede paradigmática que permita identificar as relações essenciais entre os descritores; 3) uma rede sintagmática que expresse as relações de sentido entre os descritores da LD. De acordo com Cervantes (2009), o léxico deve ser entendido como o conjunto de palavras que compõem um idioma ou área de especialidade, a rede paradigmática são as relações associativas que permitem uma padronização da linguagem, já a rede sintagmática refere-se às relações que podem ser empreendidas entre os termos.

As relações hierárquicas são caracterizadas pelas noções de superordenação e subordinação existentes entre os termos. As relações associativas identificam dependência entre os termos, porém sem estabelecer uma hierarquia (CERVANTES, 2009, p. 48). Já a relações de equivalência são aquelas que permitem o controle da sinonímia e da polissemia das linguagens, permitindo o estabelecimento de remissivas que orientam os usuários à utilização de termos preferidos (Ibid., loc. cit.).

Além disso, e de acordo com Guimarães (1990), elas podem ser pré-coordenadas, quando os conceitos são coordenados no momento da indexação, e pós-coordenadas, quando a coordenação de conceitos é efetuada no momento da busca (recuperação) de informação. Quanto a forma de apresentação de conceitos, são caracterizadas como hierárquicas e alfabéticas. Particularizando as linguagens documentárias alfabéticas, exemplificadas pelas

listas de cabeçalhos de assuntos, tesouros, entre outras, elas apresentam seus conceitos, representados pelos termos, organizados alfabeticamente, relacionados sintático-semanticamente e que indicam uma forma de interação entre eles.

Sobre as listas de cabeçalhos de assunto, vimos que Cutter (1876) foi o responsável pela elaboração das primeiras regras de construção de cabeçalhos de assunto, tendo por princípios básicos: a) especificidade: utilizar o termo mais específico para descrever um assunto e não utilizar o termo mais geral; b) entrada direta: para cabeçalhos compostos utilizar, preferencialmente, a forma direta e não a invertida. Exemplo: Organização territorial e não Território – Organização; c) referências cruzadas: para o relacionamento de assuntos próximos (CESARINO; PINTO, 1978, GIL URDICIAIN, 2004).

Para Boccato (2011a, p. 13) “as listas de cabeçalhos de assunto são linguagens pré-coordenadas, de estrutura associativa ou combinatória de palavras ou expressões, que têm a finalidade de representar os conteúdos documentários”. Elas “possuem regras específicas para as formas de entrada dos cabeçalhos, do uso de abreviaturas e, geralmente, arrolam palavras correspondentes a diversas áreas do conhecimento”. São exemplos a *Library of Congress Subject Headings*, construída pela Biblioteca do Congresso dos Estados Unidos e, no âmbito nacional, a Terminologia de Assuntos, elaborada pela Fundação Biblioteca Nacional e a Lista de Cabeçalhos de Assunto da Rede Bibliodata, desenvolvida pela Fundação Getúlio Vargas.

No que se referem as linguagens documentárias alfabéticas pós-coordenadas, Currás (1995) nos apresenta o tesouro, linguagem controlada composta por termos descritores e não-descritores (remissivos), representativos, normalmente, de uma área específica do conhecimento e aplicável mais a sistemas automatizados de recuperação da informação. Para Boccato, Ramalho e Fujita (2008, p. 201) os tesouros são

[...] linguagens de estruturas combinatórias e pós-coordenadas, constituídas de termos - unidades lingüísticas provenientes da linguagem de especialidade e da linguagem natural, denominados de descritores, providos de relações sintático-semânticas, referentes a domínios científicos especializados, possibilitando a representação temática do conteúdo de um documento, bem como a recuperação da informação.

São exemplos, *UNESCO Thesaurus*, o *AGROVOC Thesaurus*, o Thesaurus Brasileiro de Educação, o Tesouros em Ciência da Informação, entre outros tesouros internacionais e nacionais multilingues e monolingues.

Com isso, vimos que as diretrizes estabelecidas por normas internacionais - ANSI/NISO Z39.19, 2005 e ISO 25964, 2011 - e os subsídios teóricos e metodológicos oriundos de disciplinas e campos científicos interdisciplinares, com destaque para a

Terminologia, a Psicologia Cognitiva, a Linguística, a Ciência da Computação são elementos e fundamentos importantes na compilação de termos e no estabelecimento das relações conceituais entre eles (BOCCATO, 2011b) para a construção de uma linguagem documentária consistente e com o rigor que ela requer. Em associação, a cognição de bibliotecários e usuários, inseridos em um contexto organizacional técnico-científico, traz contribuições significativas à efetivação desse processo.

3 O CONTEXTO SOCIOCOGNITIVO DO BIBLIOTECÁRIO E DO USUÁRIO DE UNIDADES DE INFORMAÇÃO TÉCNICO-CIENTÍFICAS

As unidades de informação são instituições que têm como objetivo a aquisição, processamento, armazenamento e disseminação de informações, identificadas pelas bibliotecas, centros e sistemas de informação e documentação (TARAPANOFF, ARAÚJO JÚNIOR; COMIER, 2000). Sobre as bibliotecas, Oliveira (2008, p. 38) diz que elas são coleções de documentos bibliográficos e não bibliográficos que se organizam para a formação, consulta e recreação de todo o público ou de categorias específicas de usuários (OLIVEIRA, 2008), classificadas em bibliotecas escolares, universitárias, especializadas, públicas e comunitárias.

Tendo em vista que a instituição “biblioteca” tem por objetivo a disseminar o conhecimento no atendimento das necessidades informacionais específicas de cada categoria de usuário, isto é, oferecendo infraestrutura documental e informacional aos alunos e professores de ensino fundamental e médio (biblioteca escolar) e no ambiente universitário contemplando também as necessidades de busca e recuperação da informação de discentes e docentes de cursos de graduação e pós-graduação articuladas aos pilares da universidade, isto é, ensino – pesquisa – extensão.

Os Institutos Federais oferecem cursos técnicos de nível médio e cursos superiores de graduação em Tecnologias e Licenciaturas. Seu público é formado por discentes dos ensinos médio e superior e dos docentes que atuam nos respectivos cursos técnicos e superiores. Assim, vimos uma aproximação das bibliotecas dos Institutos Federais com as bibliotecas escolares e universitárias. Com relação à biblioteca escolar, a interlocução acontece pelo fato do IFSP terem alunos de ensino médio matriculados nos cursos técnicos. No contexto das bibliotecas universitárias, observamos que o Instituto Federal de São Paulo atua nas vertentes de ensino, pesquisa e extensão, ofertando cursos superiores e realizando atividades de pesquisa e de extensão, características originárias dessas bibliotecas.

Dessa forma, e no âmbito de nossa pesquisa, caracterizamos e nomeamos as bibliotecas do IFSP como Unidades de Informação Técnico-Científica (UITCs), já que o ensino técnico e superior em conjunto com as atividades de pesquisa e extensão fazem parte do foco também dessas Unidades e, tipologicamente, correspondem aos objetivos propostos pelas bibliotecas escolares e universitárias.

Acerca da abordagem sociocognitiva, também denominada de análise de domínio, Hjørland (2002) relata que os instrumentos, os conceitos, os significados e as estruturas, necessidades e os critérios de informação são modelados de acordo com as comunidades discursivas nas quais se inserem, caracterizadas por serem por serem científicas, acadêmicas ou profissionais. Hjørland e Albrechtsen (1995, p. 400) definem a análise de domínio como uma abordagem funcionalista, buscando compreender as funções implícitas e explícitas da informação e da comunicação em que, a partir dessa compreensão, traçam os mecanismos subjacentes ao comportamento informacional. Nesse sentido, as dimensões sociais, culturais e históricas influenciam os comportamentos informacionais dos sujeitos.

Partindo-se do pressuposto de que comunidades discursivas configuram-se num processo de comunicação ordenado e delimitado (HJØRLAND, 2002, p. 258), vimos, nesse contexto, as UICTs com a responsabilidade de disseminar a informação a seus usuários que, por sua vez, necessitam busca-la e recuperá-la com qualidade, precisão e rapidez. Os bibliotecários, responsáveis pela indexação dos conteúdos realizam a leitura do documento, que difere da leitura comum, tendo em vista a identificação e seleção de conceitos para a representação por meio de termos provenientes da linguagem documentária adotada pelo SRI. A determinação do assunto do documento é permeada, segundo Hjørland (1992), por cinco concepções: ingênua, idealismo subjetivo, idealismo objetivo, concepção pragmática de assunto e concepção realista/materialista de assunto.

O contexto sociocognitivo do bibliotecário durante a atividade de indexação, de acordo com Fujita (2006, p. 7), caracteriza-se por considerar o SRI e seus serviços de organização e recuperação da informação, bem como o indexador [no âmbito de nossa pesquisa, denominado bibliotecário], “seu conhecimento prévio profissional e objetivos em situação de interação durante o processamento textual para objetivos de indexação”.

Gonçalves (2008) relata que, na atividade de indexação, o contexto sociocognitivo do bibliotecário é formado pelo o conhecimento prévio que ele possui sobre sua ferramenta de trabalho – a política de indexação, as regras e procedimento do manual de indexação, a linguagem documentária e os interesses de busca dos usuários. Também, outros fatores contribuem na formação do contexto sociocognitivo do bibliotecário: a concepção de análise

emprega na determinação do assunto (simplista, orientada pelo conteúdo e orientada pela demanda), os objetivos da instituição, além da própria missão do bibliotecário, ou seja, “atuar como elo entre as informações constantemente difundidas nos mais diversos suportes e o usuário” (DAL’ EVEDOVE, 2010, p. 120). Diante desses fatores, a compreensão que se tem sobre eles, permite-nos agregá-los, também, como parâmetros colaborativos na construção de um instrumento de representação temática para as UITCs do IFSP.

Por outro lado e do ponto de vista do pragmatismo, observamos que suas necessidades de informação são a razão da existência de uma UITC e que influenciam a atividade de indexação para a busca e recuperação da informação. Tais necessidades são manifestadas por elementos de ordens social e cultural, o que caracteriza delinea o entendimento do contexto sociocognitivo do usuário como um processo de fundamental importância para nossa pesquisa, dada a diversidade de categorias de usuários que a UITCs atendem e suas particularidades de públicos.

Sob a abordagem da análise do domínio, as necessidades de informação dos usuários são “despertadas” por suas conexões sócio-culturais (HJØRLAND, 2002) que conduzem-no à uma interação com o SRI mediante a formulação das questões de busca, sejam elas assistidas pelo bibliotecário ou não. Como resultado e recorrendo-se à Boccato (2009a, p.123) observamos que

O julgamento realizado pelo usuário sobre a relevância ou não do documento recuperado está intimamente relacionado ao desempenho da linguagem documentária utilizada pelo sistema de recuperação da informação. A linguagem de busca do usuário deve ser compatível com a linguagem documentária do sistema, e esta deve representar as necessidades de informação do usuário, construídas por seus modelos mentais influenciados por seu meio.

Outro fato que reflete sobre tal julgamento é que um mesmo documento recuperado por diferentes usuários pode ser relevante para uns e irrelevante para outros e, por outro lado, um mesmo documento recuperado pode ter relevância, para um mesmo usuário, em um momento e não ter relevância em outro.

Gonçalves (2008) e Boccato (2009a) relatam que, no âmbito das bibliotecas universitárias, o contexto sociocognitivo do usuário é composto pela universidade, a grade curricular dos cursos oferecidos, os grupos de pesquisas, as iniciações científicas desenvolvidas e o catálogo *online*. Numa aproximação com as UITCs, vimos que os IFSP contemplam tais aspectos, já que eles são Instituições que relevam em seus objetivos o ensino, a pesquisa e a extensão, assim como as Universidades.

Dessa forma, entendemos que contexto sociocognitivo do bibliotecário e do usuário das UITCs pode auxiliar na determinação de parâmetros para a elaboração de instrumento de

representação temática. Portanto, as competências e as habilidades que o bibliotecário possui são variáveis que formam o contexto sociocognitivo do bibliotecário das UITC do IFSP juntamente com: concepções da análise de assunto que utiliza no processo de indexação, os objetivos da instituição e do próprio SRI; o conhecimento prévio que tem sobre suas ferramentas de trabalho e a mediação que realiza entre a informação e o usuário.

No que se refere aos usuários, o contexto sociocognitivo do usuário das UITCs é constituído por suas necessidades de informação, pelos conteúdos abordados nas disciplinas dos cursos de ensino médio/técnico e dos cursos de ensino superior e das pesquisas científicas que desenvolvem.

4 METODOLOGIA

A pesquisa tem caráter teórico-aplicado, de natureza descritiva e exploratória no contexto da metodologia qualitativa com abordagem sociocognitiva. A presente pesquisa foi aprovada pelo Comitê de Ética em Pesquisa em Seres Humanos da Universidade Federal de São Carlos (UFSCar) sob o número 120/2012, em reunião realizada em 27/03/2012.

A metodologia qualitativa com abordagem sociocognitiva possibilitou-nos “[...] o envolvimento direto com pessoas, lugares e processos interativos, originando uma coleta descritiva pautada no rigor que deve existir no momento da análise interpretativa dos dados” (BOCCATO; FUJITA, 2010, p. 30).

O universo de pesquisa foi formado pelas unidades de informação técnico-científicas do IFSP, representadas pelas bibliotecas dos *campi* São Paulo, São Roque e Sertãozinho. Esses *campi* foram escolhidos por terem ao menos um curso superior representante de uma das três áreas do conhecimento - Ciências Humanas, Ciências Exatas e Ciências Biológicas.

Como procedimentos metodológicos, efetuamos duas etapas de coletas de dados:

- 1) aplicação de questionário com os bibliotecários responsáveis pelas unidades de informação de cada *campus*, buscando identificar o contexto histórico-sócio-técnico-organizacional dessas unidades;
- 2) aplicação da técnica introspectiva Protocolo Verbal na modalidade em Grupo (PVG), desenvolvida no ambiente das bibliotecas do IFSP, *campus* de São Paulo, São Roque e Sertãozinho com três bibliotecários, três docentes que ministram disciplinas tanto nos cursos técnicos, como também nos cursos de nível superior e seis discentes, sendo três dos referidos cursos técnicos e três dos cursos de nível superior, totalizando doze sujeitos de pesquisa, com o intuito de observar as opiniões e registrar as declarações dos sujeitos participantes sobre as características e o uso de uma linguagem

documentária representativa do contexto sociocognitivo de bibliotecários e usuários das Unidades de Informação Técnico-Científicas.

Portanto, participaram dos PVGs um bibliotecário, um docente que ministra aula em curso de técnico e superior, um discente de graduação e um discente de ensino técnico vinculados em cada um dos três campi selecionados, isto é, São Paulo, São Roque e Sertãozinho.

Sobre a primeira etapa referente a aplicação do questionário, ele foi composto por dezessete questões, sendo seis questões abertas, cinco questões fechadas e seis semi-abertas e enviado por correio eletrônico aos bibliotecários responsáveis pelas UITC de cada *campus*. Após o retorno dos questionários, os dados foram sistematizados para a posterior análise dos dados. Até a conclusão desta pesquisa, o IFSP contava com vinte e uma unidades de informação e trinta bibliotecários². Dessa forma, esta pesquisa representa uma amostra de 14% do universo de unidades de informação do IFSP.

Acerca da segunda etapa, isto é, a aplicação do PVG ou como também nomeado de “Leitura como Evento Social”, é a reunião de um grupo de pessoas, normalmente entre três e seis, além do pesquisador, na execução de uma atividade de leitura de um texto e a discussão de temas suscitados por ele. A leitura torna-se um evento social específico, possibilitando a interação entre os sujeitos de pesquisa, o texto e o pesquisador, permitindo conhecer os pensamentos e as ações sobre o tema apresentado no texto, relacionado ao interesse de pesquisa em análise. Assim, por meio da leitura do “texto-base”, o grupo interagiu explanando seus pensamentos e os procedimentos que utilizam durante a execução da tarefa.

A aplicação do PVG seguiu os pressupostos metodológicos estabelecidos por Rubi (2004), constituído de três estágios: 1) Procedimentos anteriores à sessão de realização da coleta de dados. O texto-base selecionado para a leitura a “Recuperação de informações em bases de dados: usos de tesouro”, de autoria de Leonardo Fernandes Souto, publicado no periódico *TransInformação*, em 2003, trecho entre as páginas 74 a 79; 2) Procedimentos durante as sessões da coleta de dados; 3) Procedimentos após o término da sessão de coleta de dados: as declarações dos sujeitos foram gravadas, transcritas literalmente com suas identidades preservadas, para a posterior análise dos dados coletados.

Os dados dos questionários respondidos foram analisados buscando caracterizar o contexto histórico-sócio-técnico-organizacional das UITCs. Para a análise dos dados coletados mediante a aplicação dos PVGs, foram construídas oito categorias, considerando-se os objetivos de pesquisa, os fundamentos teóricos e as declarações dos sujeitos participantes, assim sistematizadas: 1) Procedimentos de indexação; 2) Importância da

² Dados apurados junto à lista de discussão dos Bibliotecários do IFSP em 10/02/2013.

linguagem documentária na indexação e recuperação da informação; 3) Escolha da linguagem documentária; 4) Características da linguagem documentária; 5) Importância da participação do usuário na construção de linguagem documentária; 6) Avaliação do sistema de recuperação da informação; 7) Capacidade de revocação e precisão do sistema; 8) Política de indexação.

A aplicação do questionário em conjunto com o PVG possibilitou realizar uma abordagem sociocognitiva de caracterização do cotidiano profissional e acadêmico dos bibliotecários e usuários a partir das opiniões que eles têm sobre a importância dos uso de uma linguagem documentária para a representação e recuperação por assunto em SRI nas unidades de informação técnico-científicas.

5 RESULTADOS E DISCUSSÃO

As informações colhidas permitiram tecer algumas considerações sobre as unidades de informações estudadas. Em todas as unidades há a presença de bibliotecário. O quadro de servidores é bastante diversificado. A unidade do *campus* São Paulo apresenta maior acervo. Nenhuma das unidades possui manual de indexação e vocabulário controlado. Apesar de todas terem um sistema automatizado de recuperação da informação, apenas a UITC do *campus* São Roque afirmou ter o acervo disponível na *Web* e apenas na do *campus* Sertãozinho o acesso ao sistema é restrito ao bibliotecário, dessa forma, neste último *campus*, a consulta do usuário ao SRI é intermediada pelo bibliotecário.

Com relação aos protocolos verbais em grupo, a análise dos resultados foi realizada a partir das oito categorias, apresentadas a seguir, elaboradas diante dos referenciais teóricos, objetivos da pesquisa e das falas dos sujeitos participantes.

Categoria 1 – Procedimentos de indexação. Resultados obtidos: utilização de linguagem hierárquica para arranjo físico do acervo; bibliotecários compreendem a importância ter em mente as necessidades dos usuários no momento da indexação; ausência de procedimentos metodológicos ou normativos orientadores do processo de indexação; ausência de instrumento de representação temática para indexação e recuperação da informação.

Categoria 2 - Importância da linguagem documentária na indexação e recuperação da informação. Resultados obtidos: bibliotecários e usuários visualizam a importância de um instrumento de representação temática para indexação e recuperação da informação.

Categoria 3 - Escolha da linguagem. Resultados obtidos: utilização de linguagem natural não representa satisfatoriamente os assuntos do acervo; preferência dos usuários pela consulta direta no acervo.

Categoria 4 – Características da linguagem documentária. Resultados obtidos: bibliotecários se identificam com o uso de linguagem pós-coordenada para SRI; termos devem contemplar as áreas de especialidade dos usuários; usuários e bibliotecários indicam que deve haver relações (hierárquicas e semânticas) entre os termos.

Categoria 5 - Importância da participação do usuário na construção de linguagem documentária. Resultados obtidos: bibliotecários consideram que a participação dos usuários, tendo em vista princípios de garantia de uso; tipologia de públicos aos quais a UICT deve abranger diz respeito a idade, níveis de conhecimento (ensino médio, ensino superior) e diversidade de áreas de assunto.

Categoria 6 - Avaliação do sistema de recuperação da informação. Resultados obtidos: sistema atual não atende as necessidades de informação; usuários preferem utilizar a busca direta no acervo físico; as informações visualizadas no catálogo podem ser ampliadas, com a apresentação de sumários, introdução, texto completo; o sistema deve ter recurso de sugestão automática de termos de busca; catálogo deve estar disponível para consulta na *Web*; catálogo deve permitir renovações e reservas *on line*.

Categoria 7 - Capacidade de revocação e precisão do sistema. Resultado obtido: sistema atual apresenta alta revocação e baixa precisão na busca de informação.

Categoria 8 – Política de indexação: Resultados obtidos: falta de uma política e de um manual de indexação.

A partir dos resultados obtidos e subsidiados pelas literaturas técnico-científica, estabelecemos onze parâmetros de construção de instrumento de representação temática em sistema de recuperação da informação no contexto sociocognitivo do bibliotecário e do usuário de UICT. São eles: 1) caracterização do perfil do usuário (público alvo) que fará uso da linguagem: discentes e docentes de cursos de nível superior e médio; 2) termos devem atender as necessidades de representação e recuperação da informação (garantias literária e de uso); 3) termos devem ter origem na linguagem natural e de especialidade (garantias de uso e literária); 4) termos devem representar o vocabulário de uso da organização (garantia organizacional); 5) a linguagem deve possuir tanto termos genéricos quanto específicos; 6) a linguagem deve promover o controle de sinônimos; 7) a linguagem deve identificar a homonímia com o uso de termos qualificadores; 8) estabelecimento de relações lógico-semânticas entre os termos de ordens hierárquica, equivalência e associativa; 9) inclusão de notas de escopo dos termos, quando necessário; 10) atribuição de termos deve ser contemplar o equilíbrio entre a exaustividade e a especificidade alcançada pelo sistema de recuperação da informação; 11) identificação/construção de sistema de recuperação da informação (catálogo)

que contemple, também, fatores, tais como: estar disponível *online*; oferecer os serviços de reservas e renovação *online*; permitir a visualização de informações como: capa, sumário, introdução e texto completo de materiais constantes do acervo das UITCs; possuir e ativar um recurso para sugestão de termos, na momento de realização da busca, tanto para correção da expressão de busca quanto para o armazenamento dos assuntos/termos procurados. Tal recurso é importante na coleta de termos, visando o processo de atualização da linguagem a partir também da perspectiva do usuário; disponibilizar e permitir a acessibilidade da linguagem para que o bibliotecário possa realizar a representação da informação com ela e a partir dela; disponibilizar e permitir a acessibilidade da linguagem para que o usuário possa realizar a busca por assunto, para a recuperação de informações úteis, com ela e a partir dela.

6 CONSIDERAÇÕES FINAIS

Vimos nos Institutos Federais, cuja existência tem o intuito de atender á políticas públicas, traços de uma aproximação proposta pelo contexto CTS. Neles, as unidades de informação “atuam como agentes fundamentais na concretização da missão dos Institutos, fomentando ensino, pesquisa e extensão, necessitando adequar-se a essa realidade, oferecendo produtos e serviços que contemplem como usuários alunos, docentes e servidores” (OLIVEIRA; AMARAL, 2012, p. 2).

Dessa forma, o uso de sistemas de recuperação da informação (SRIs) por unidades de informação tem por objetivo permitir aos seus usuários o acesso aos conteúdos dos acervos destas unidades. Ao empregar o uso da linguagem natural, sem controle, para representação da informação dos acervos nos SRIs, empregando as palavras tal como aparecem nos documentos, acarreta incoerências que resultam na alta revocação e baixa de precisão na recuperação de informações por parte dos usuários.

Diante da importância dessas Unidades de Informação e dos seus respectivos catálogos *online*, os onze parâmetros foram estabelecidos e norteiam a construção de um Vocabulário que permita o controle de sinonímias e homonímias e estabeleça relações lógico-semânticas entre os termos, criando, asism, uma linguagem documentária comum a ser utilizada pelo Sistema de Recuperação da Informação (catálogo) na indexação e na recuperação de documentos. Sobre a composição do vocabulário, vimos que ele deve ser proveniente das linguagens natural e de especialidade, como conta num tesouro, porém autores como Boccato (2009a; 2009b), Boccato, Ramalho e Fujita (2008) e Cervantes (2009), alegam que ele tem maior efetividade de aplicação para contextos especializados.

Já uma lista de cabeçalhos de assunto, de acordo com Gil Urdiciain (2004), pode ser definida essencialmente como linguagem pré-coordenada que utiliza termos da linguagem natural e pode apresentar algumas relações semânticas entre os termos indicadas por notações. Seu uso tem sido aplicado em unidades de informação com acervos generalistas como a Biblioteca Nacional e a *Library of Congress*.

Todavia, e conforme vimos, a diversidade de públicos apontada nos protocolos verbais diz respeito a fatores como idade, níveis de escolaridade (conhecimento) e áreas de especialidades distintas, o que nos faz refletir e recomendar às UITCs do IFSP a construção e uso de uma linguagem documentária com vocabulários advindos da linguagem natural e de especialidade (como ocorre em um tesouro), porém com a estrutura lógico-semântica entre os termos/cabeçalhos de uma lista de cabeçalhos de assunto, contemplando também as notas de escopo que se fizerem necessárias.

Em complementação, e ao lado do uso de uma linguagem documentária única por todos bibliotecários e usuários das UITCs, sugerimos, também a adoção de um catálogo coletivo, tendo em vista que ele proporcionará maior visibilidade dos acervos integrantes dessas bibliotecas e facilitar a busca integrada por parte de seus usuários reais e potenciais.

Com isso, acreditamos que essas ações fomentarão a implantação, futura, de um sistema integrado de bibliotecas técnico-especializadas. Colaborando com tal medida, julgamos pertinente a continuidade de pesquisas no universo do IFSP, voltadas para o estabelecimento de uma política para a construção de um manual de indexação, tendo como dois de seus elementos a definição de um catálogo coletivo e de uma linguagem documentária representativos do contexto sociocognitivo de bibliotecários e usuários de bibliotecas técnico-especializadas de institutos federais de ensino.

REFERÊNCIAS

ANSI/NISO. *Z39.19:2005: guidelines for the construction, format, and management of monolingual controlled vocabularies*. Bethesda: NISO, 2005.

BOCCATO, V. R. C. *Avaliação do uso de linguagem documentária em catálogos coletivos de bibliotecas universitárias: um estudo sociocognitivo com protocolo verbal*. 2009. 303 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2009a. Disponível em: <http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/boccatovrcdo_mar.pdf>. Acesso em: 12 jun. 2010.

BOCCATO, V. R. C. A linguagem documentária vista pelo conteúdo, forma e uso na perspectiva de catalogadores e usuários. In: FUJITA, M. S. L. (Org.). *A indexação de livros: a*

percepção de catalogadores e usuários de bibliotecas universitárias. Um estudo de observação do contexto sociocognitivo com protocolos verbais. São Paulo: Cultura Acadêmica, 2009b. Cap. 6, p. 119-135. Disponível em: http://www.culturaacademica.com.br/catalogo-detalle.asp?ctl_id=56. Acesso em: 12 jun. 2010.

BOCCATO, V. R. C. Linguagem documentária na representação e recuperação da informação pela perspectiva sociocognitiva em Ciência da Informação. In: BOCCATO, V. R. C.; GRACIOSO, L. S. (Org.). *Estudos de linguagem em Ciência da Informação*. São Paulo: Cultura Acadêmica, 2011a. Cap. 1, p. 9-34.

BOCCATO, V. R. C. Os sistemas de organização do conhecimento nas perspectivas atuais das normas internacionais de construção. *InCID: Revista de Ciência da Informação e Documentação*, Ribeirão Preto, v. 2, n. 1, p. 165-192, 2011b. Disponível em: <http://www.braperci.ufpr.br/download.php?dd0=18256>. Acesso em: 10 nov. 2011.

BOCCATO, V. R. C.; RAMALHO, R. A. S.; FUJITA, M. S. L. Título: A contribuição dos tesouros na construção de ontologias como instrumento de organização e recuperação da informação em ambientes digitais. In: GARÍA MARCO, F. J. *Avances y perspectivas em sistemas de información y documentación - IBERSID 2008*. Zaragoza: Universidad de Zaragoza, 2008. p. 199-209.

BOCCATO, V. R. C. FUJITA, M. V. L. O uso de linguagem documentária em catálogos coletivos de bibliotecas universitárias: um estudo de avaliação sociocognitiva com protocolo verbal. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 15, n. 3, p. 23-51, set./dez. 2010. Disponível em: <http://www.scielo.br/pdf/pci/v15n3/03.pdf>. Acesso em: 10 nov. 2011.

BRASIL. Ministério da Educação. *Instituto Federal de Educação, Ciência e Tecnologia: Concepção e Diretrizes*. Brasília, DF: Secretaria de Educação Profissional e Tecnológica, 2008. Disponível em: <http://rededefederal.mec.gov.br/>. Acesso em: 10 dez. 2011.

CENDÓN, B. V. Sistemas e redes de informação. In: _____. (Org.). *Ciência da Informação e Bibliotecnomia: novos conteúdos e espaços de atuação*. Belo Horizonte: Editora UFMG, 2008.

CESARINO, M. A. N. Sistemas de recuperação da informação. R. Esc. Bibliotecon., Belo Horizonte, v. 14, n. 2, p. 157-168, set. 1985. Disponível em: <http://www.braperci.ufpr.br/download.php?dd0=13794>. Acesso em: 13 jun. 2012.

CESARINO, M. A. N.; PINTO, M. C. M. F. Cabeçalho de assunto como linguagem de indexação. R. Esc. Bibliotecon., Belo Horizonte, v. 7, n. 2, p. 268-288, set. 1978. Disponível em: <http://www.braperci.ufpr.br/download.php?dd0=13794>. Acesso em: 13 jun. 2012.

CERVANTES, B. M. N. A construção de tesouros com a integração de procedimentos Terminográficos. 2009. 209 f. Tese. (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2009. Disponível em: http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/cervantes_bmn_do_mar.pdf. Acesso em: 12 out. 2010.

CINTRA, A. M. et al. *Para entender as linguagens documentárias*. 2 ed. São Paulo: Polis, 2002. 92 p. (Coleção Palavra-Chave, 4).

CURRÁS, E. *Tesauros: linguagens terminológicas*. [Thesaurus lenguajes terminologicos]. Antônio Felipe Corrêa da Costa (Trad.). Brasília: IBICT, 1995

CUTTER, C. A. *Rules for a dictionary catalog*. Washington: Government Printing Office, 1876.

DAL' EVEDOVE, P. R. *A perspectiva sóciocognitiva no tratamento temático da informação em bibliotecas universitárias: aspectos inerentes a percepção profissional*. 2010. 300 f. Dissertação (Mestrado em Ciência da Informação)–Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

FOSKETT, A. C. *A abordagem temática da informação*. São Paulo: Editora Universidade de Brasília; Editora Polígono, 1973.

FUJITA, M. S. L. *Abordagem cognitiva e sócio-cognitiva da leitura documentária na formação inicial do indexador: análise da perspectiva individual em contexto sócio-cultural*. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 7., 2006, Marília. *Anais...*, Marília: Unesp. Disponível em: <<http://www.ancib.org.br/media/dissertacao/gt2-22.pdf>>. Acesso em 10 jan. 2012.

GONÇALVES, M. C. *A indexação em catálogo on line em bibliotecas universitárias na percepção de usuários integrantes de grupos de pesquisa: uma contribuição ao desenvolvimento de política de indexação na rede de bibliotecas da UNESP*. 2008. 148 f. Dissertação (Mestrado em Ciência da Informação)–Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2008

GIL URDICIÁIN, B. *Manual de lenguajes documentales*. 2. ed. rev. y ampl. Madrid: Ed. NOESIS, 2004.

GUIMARÃES, J. A. C. *Recuperação temática da informação*. R. Bras. Bibliotecon. e Doc., São Paulo, v. 23, n. 3/4, p. 112-130, 1990.

HJØRLAND, B. The concept of 'subject' in Information Science. *Journal of Documentation*, London, v. 48, n. 2, p. 172-200, June 1992.

HJØRLAND, B. Epistemology and the socio-cognitive perspective in information science. *Journal of the American Society for Information Science and Technology*, Maryland, v. 53, n. 4, p. 257-270, 2002.

HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain analysis. *Journal of the American Society for Informacion Science*, Maryland, v. 46, n. 6, p. 400-425, 1995.

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE SÃO PAULO (Brasil). 2013. Disponível em: <www.ifsp.edu.br>. Acesso em: 01 jan. 2013.

LANCASTER, F. W. *Construção e uso de tesouros: curso condensado*. Brasília: IBICT, 1987. Disponível em: <<http://devds-01.ibit.br:8080/rlc/handle/1/781>>. Acesso em: 28 nov. 2012.

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. Tradução de Antonio Agenor Briquet de Lemos. 2. ed. Brasília: Briquet de Lemos, 2004. 452 p.

LARA, M. L. G. de. Linguagem documentária e terminologia. *Transinformação, Campinas*, v. 16, n. 3, p. 231-240, set./dez. 2004.

OLIVEIRA, M. de. Origens e evolução da Ciência da Informação. In: _____. (Org). *Ciência da Informação e Bibliotecnomia: novos conteúdos e espaços de atuação*. Belo Horizonte: Editora UFMG, 2008.

OLIVEIRA, G. G.; AMARAL, R. M. Mapeamento de processos em bibliotecas: estudo de caso em uma biblioteca do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 17., 2012, Gramado, *Anais...*, Gramado: CBBU, 2012.

ORTEGA, C. D. Relações históricas entre Biblioteconomia, Documentação e Ciência da Informação. *Datagramazero*. v. 5, n. 5, out. 2004. Disponível em: <www.dzg.org.br/out04/Art_03.htm>. Acesso em: 26 jun. 2012.

RUBI, M. P. *A política de indexação na perspectiva do conhecimento organizacional*. 2004. 135 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista; CAPES, Marília, 2004. Disponível em: <http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/rubi_mp_me_mar.pdf>. Acesso em: 12 out. 2010.

TARAPANOFF, K.; ARAÚJO JÚNIOR, R.; CORMIER, P. M. J. Sociedade da informação e inteligência em unidades de informação. *Ciência da Informação*, Brasília, v. 29, n. 3, p. 91-100, set./dez. 2000. Disponível em: <<http://www.scielo.br/pdf/ci/v29n3/a09v29n3.pdf>>. Acesso em: 13 jun. 2012.