

## Estratégias de Leitura Documentária para Indexação: um estudo de caso com o Centro de Informações Nucleares (CIN)

Milena Polsinelli Rubi

SciELO Books / SciELO Livros / SciELO Libros

RUBI, M. P. Estratégias de Leitura Documentária para Indexação: um estudo de caso com o Centro de Informações Nucleares (CIN). In.: FUJITA, M. S. L., NEVES, D. A. B., and DAL'EVEDOVE, P. R., eds. *Leitura documentária: estudos avançados para a indexação* [online]. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2017, pp. 69-91. ISBN: 978-85-7983-917-7. Available from: <http://books.scielo.org/id/3pk5m>.  
<https://doi.org/10.36311/2017.978-85-7983-917-7.p69-92>.



All the contents of this work, except where otherwise noted, is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença [Creative Commons Atribuição 4.0](https://creativecommons.org/licenses/by/4.0/).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia [Creative Commons Reconocimiento 4.0](https://creativecommons.org/licenses/by/4.0/).

# ESTRATÉGIAS DE LEITURA DOCUMENTÁRIA PARA INDEXAÇÃO: UM ESTUDO DE CASO COM O CENTRO DE INFORMAÇÕES NUCLEARES (CIN)

*Milena Polsinelli Rubi*

## 1 INTRODUÇÃO

A leitura documentária, realizada pelo indexador na fase de análise, visa propiciar a identificação de conceitos para posterior representação nos produtos documentários – número de classificação, resumos e índices – que satisfaçam a demanda dos usuários.

Para isso, o leitor documentalista, apesar de não ser o especialista do assunto, interage com o texto mediante o domínio de uma linguagem documentária, da estrutura textual e da intenção do sistema de informação para leitura.

Desenvolvido dentro do Projeto Integrado “Leitura em Análise Documentária<sup>1</sup>”, esse estudo foi realizado com o Centro de Informações Nucleares da Comissão Nacional de Energia nuclear (CIN/CNEN) que é, desde 1970, responsável pela coleta e disseminação de informações na área de energia nuclear no Brasil.

---

<sup>1</sup> Desenvolvido pela Dra. Mariângela Spotti Lopes Fujita, o referido projeto visou, principalmente, à obtenção de subsídios sobre os serviços de análise documentária em Centros de Informações especializados no Brasil. A primeira parte do projeto realizou estudo de caso junto à Sub-Rede Nacional de Informações em Ciências da Saúde Oral do Sistema BIREME/ OPAS para a produção da base de dados LILACS.

Durante a realização da pesquisa, o CIN apresentava dez indexadores especialistas em áreas de assunto relacionadas à Energia Nuclear (Biologia, Energia, Energia Nuclear, Engenharia Elétrica, Engenharia Química, Física, Geologia, Medicina Nuclear, Meio Ambiente e Química e outros) e em procedimentos de indexação sistematizados para a inserção de registros na base de dados do *International Nuclear Information System* (INIS), do qual o Brasil é país-membro cooperante através do CIN. Desses indexadores apenas um era funcionário do CIN, os outros nove eram indexadores *freelancers*, tendo seus serviços de indexação terceirizados. Nenhum era bibliotecário.

Foram coletados e analisados os dados referentes ao CIN e seus indexadores. Dessa forma, e com o conhecimento teórico da atividade de análise documentária, foi possível investigar, no contexto de um centro referencial em Energia Nuclear, a prática da atividade de indexação e caracterizar o profissional que a executa, com os seguintes objetivos:

- Identificar os procedimentos de leitura documentária utilizados para exame de documento e identificação de conceitos;
- Examinar as estratégias de leitura usadas pelos indexadores para a identificação de conceitos que melhor representem o documento.

Como metodologia, foram utilizados o manual de indexação do INIS e a técnica de coleta de dados introspectivos denominada “Protocolo Verbal” ou “Pensar Alto” nos moldes de Ericsson e Simon (1987), seguida da entrevista retrospectiva. Esta técnica permite a observação das estratégias utilizadas pelos indexadores, pois privilegia o Pensar Alto durante a realização da tarefa.

## **2 ESTRATÉGIAS DE LEITURA DOCUMENTÁRIA PARA INDEXAÇÃO**

A leitura documentária constitui-se a atividade principal da Análise Documentária. Na indexação, o resultado será a identificação e seleção dos termos que irão representar o documento para o usuário.

O termo indexação (*indexing*) remonta à corrente inglesa e, de acordo com os “Princípios de indexação” do *World Information System for Science and Technology*<sup>2</sup> (UNISIST, 1981, p. 84) é “[...] a ação de descrever e identificar um documento de acordo com seu assunto.”

Essa publicação do UNISIST originou a primeira norma a esse respeito publicada em 1985 pela *International Standardization for Organization* (ISO), sob número 5963, com o título “*Documentation - methods for examining documents, determining their subjects, and selecting indexing terms*”.

Em 1992, a Associação Brasileira de Normas Técnicas (ABNT) publicou a tradução dessa mesma norma, sob número 12676, intitulada “Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação”. A indexação é definida pela Norma 12676 como “Ato de identificar e descrever o conteúdo de um documento com termos representativos dos seus assuntos e que constituem uma linguagem de indexação.” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992, p. 2).

Para Chaumier (1988, p. 63) “[...] a indexação é a parte mais importante da análise documentária. Consequentemente é ela que condiciona o valor de um sistema documentário.” Ainda segundo o autor, uma indexação inadequada ou uma indexação insuficiente representa 90% das causas essenciais para a aparição de “ruídos” (os documentos não pertinentes à questão que são recuperados em uma pesquisa bibliográfica) ou de “silêncios” (os documentos pertinentes à questão existentes no acervo que não são recuperados).

O termo indexação é definido por Van Slype (1991) como a operação que consiste em enumerar os conceitos sobre os quais trata um documento e representá-los por meio de uma linguagem combinatória – lista de descritores livres, lista de autoridades e o thesaurus de descritores – tendo como finalidade a busca documental que será realizada a partir dos índices ou dos catálogos. Nessa definição, referido autor destaca a representação dos conceitos, por meio de uma linguagem específica com vistas ao

---

<sup>2</sup> Sistema internacional vinculado à Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) e conhecido pela sigla UNISIST.

processo de recuperação da informação por meio de índices ou catálogos, colaborando com nossa intenção de pesquisa ao estudarmos a política de indexação para construção de catálogos.

Lancaster (2004, p. 1) explica que “[...] os processos de indexação identificam o assunto que trata o documento [...]” e eles implicam “[...] a preparação de uma *representação* do conteúdo temático dos documentos.” (LANCASTER, 2004, p. 6, grifo do autor).

Consideramos, portanto, que a indexação diz respeito à identificação do conteúdo do documento, por meio do processo de análise, e à representação desse conteúdo através de conceitos. Esses conceitos, por sua vez, serão representados ou traduzidos em termos advindos de uma linguagem documentária, com vistas à intermediação entre o documento e o usuário no momento da recuperação da informação, seja em índices, catálogos ou bases de dados.

O processo de indexação é composto por diferentes etapas, cujo número varia de acordo com os autores, como veremos a seguir.

O UNISIST (1981) afirma que o processo de indexação compreende dois estágios:

- o estabelecimento de conceitos tratados num documento: envolve a compreensão do conteúdo do documento como um todo; a identificação dos conceitos representativos desse conteúdo; a seleção dos conceitos válidos para recuperação.
- a tradução dos conceitos nos termos da linguagem de indexação;

Segundo Chaumier (1986) existem também duas etapas:

- o reconhecimento e extração de conceitos informativos composto pela apreensão do conteúdo do documento por meio da leitura; pela identificação dos conceitos tendo em vista os objetivos do sistema de informação as necessidades dos usuários e pela seleção de conceitos segundo a exaustividade e especificidade do sistema de informação;
- a tradução desses conceitos em linguagem natural.

Para Van Slype (1991) a indexação comporta quatro operações distintas, a saber:

- conhecimento do conteúdo do documento;
- escolha dos conceitos a serem representados, baseando-se na aplicação da regra da seletividade e exaustividade;
- tradução dos conceitos selecionados da forma em que aparecem impressos no documento, para os descritores do tesauro aplicando a regra da especificidade;
- incorporação dos elementos sintáticos.

Assim como para o UNISIST (1981) e para Chaumier (1986), para Lancaster (2004) a indexação também envolve duas etapas:

- a análise conceitual;
- a tradução.

A Norma ABNT 12676 descreve três estágios da indexação:

- exame do documento e estabelecimento do assunto de seu conteúdo;
- identificação dos conceitos presentes no assunto;
- tradução desses conceitos nos termos de uma linguagem de indexação.

Os autores divergem quanto ao número de etapas da indexação. Porém, elas tratam basicamente das mesmas operações:

- análise: leitura e segmentação do texto para identificação e seleção de conceitos;
- síntese: construção do texto documentário com os conceitos selecionados;
- representação: por meio de linguagens documentárias.

Observamos, portanto, que a etapa inicial da indexação é a análise. Essa análise é realizada por meio da leitura feita pelo indexador que procura compreender de maneira geral o documento para identificar e selecionar os termos que representarão o documento para a recuperação.

Concordando com os autores Fujita (1998; 2001; 2003) e Dias e Naves (2007), acreditamos que a leitura documentária é a etapa primordial da indexação, pois corresponde à primeira fase de abordagem do leitor indexador e desencadeará o processo de análise de assunto do texto para identificação e seleção de conceitos.

No entanto, na prática profissional do indexador, a leitura total de um documento é impraticável, já que a leitura documentária está sujeita a condições específicas: limite de tempo, propósito definido, geração de produtos e serviços e grande volume de documentos de uma mesma área de assunto.

Sendo assim, o indexador deve lançar mão de seu conhecimento prévio – linguagem documentária utilizada pelo sistema de informação, estrutura textual do documento e assunto – e acionar estratégias durante a leitura documentária a fim de que seu objetivo seja atingido: identificação e seleção de conceitos de um documento.

As estratégias de leitura, as ações que o leitor realiza no ato de ler, têm sido definidas por vários autores. Essas estratégias segundo Faerch e Kasper (1980, citados por NARDI, 1993) são planos potencialmente conscientes do leitor para resolver algo que se apresenta como um problema na compreensão.

Para Brown (1980) estratégia é qualquer controle deliberado e planejado das atividades que levam a compreensão. A autora ainda acredita que as ações são intencionalmente selecionadas.

Kato (1987) distingue dois tipos de estratégias que definem o comportamento do leitor: as *estratégias cognitivas* que são aquelas que são automáticas e subconscientes utilizadas durante a leitura fluida e as *estratégias metacognitivas* que são conscientes do leitor frente a um problema. As estratégias cognitivas de Kato (1987) são denominadas por Cavalcanti (1978) como estratégias automáticas e as metacognitivas são chamadas estratégias controladas.

Sob o ponto de vista de Nardi (1993), Brown (1980) considera estratégia cognitiva como “*skill*” (uma estratégia que teria sido adquirida em algum momento da aprendizagem e se tornando automática) e estratégia metacognitiva como estratégias simplesmente (o uso consciente de uma “*skill*” ou de uma nova tática, em momentos de solução de problemas).

Na nossa concepção, estratégias são ações empreendidas pelo leitor no ato de ler, potencialmente conscientes, geralmente direcionadas para a solução de um problema.

Brown (1980) lista atividades consideradas de natureza metacognitiva:

- explicitação dos objetivos da leitura;
- identificação de aspectos importantes da mensagem;
- alocamento de atenção a áreas importantes;
- monitoração do comportamento para ver se está ocorrendo compreensão;
- engajamento em revisão e auto-indagação para ver se o objetivo está sendo atingido;
- tomada de ações corretivas quando são detectadas falhas na compreensão;
- recobrimento de atenção quando a mente se distrai ou faz digressões.

Para Kato (1987) existem apenas duas estratégias metacognitivas básicas, sendo as demais apenas subtipos dessas estratégias. São elas:

- estabelecimento de um objetivo explícito para leitura;
- monitoração da compreensão tendo em vista esse objetivo.

Em se tratando de leitura documentária, consideramos importante abordar as noções de estratégias de leitura em documentação apresentadas por Cintra (1989), em que são apontados vários fatores que con-



correm para a qualidade de um texto: a manutenção do tema, a correção gramatical, a adequação lexical, a estrutura do texto.

Segundo a referida autora, na leitura para fins documentários é preciso que haja uma cooperação autor/leitor, já que o autor não prevê quem poderá ler o que ele publicou.

Sabe-se também que na leitura para fins documentários e tratamento da informação (leitura técnica), não é necessário, nem aconselhável uma leitura linear, letra por letra, palavra por palavra, o leitor avança no texto na medida que consegue predizer o que vem a seguir.

O leitor deve buscar detectar a estrutura do texto. O leitor, que tem facilidade de reconhecer as superestruturas textuais, capta melhor as ideias principais do texto. Diferentemente, um leitor que lê linearmente, faz esforços desnecessários para compreender trechos isolados, mas sem apoiar-se na estrutura textual e seus conhecimentos prévios para inferir significados, e levantar hipóteses que o ajudarão a apreender a temática global. O apoio na estrutura textual permite ao leitor ser seletivo e alocar atenção a trechos importantes.

As estratégias ascendentes (*bottom-up*) e descendentes (*top-down*) citadas por Cintra (1989) mostra que o leitor, durante o processo de leitura, pode realizar dois movimentos: movimento *bottom-up*, em que o leitor vai lendo na dependência do contexto escrito, ou seja, vai extraindo, linearmente, dos símbolos impressos o significado, caminhando das partes para o todo, e movimento *top-down*, no qual há maior dependência de conhecimento prévio do leitor, pois ele vai fazendo generalizações e previsões a partir de “esquemas” que tem armazenados em sua memória, formulando hipóteses que ajudarão na compreensão do texto.

Kato (1987) considera leitor experiente aquele que utiliza os dois tipos de estratégias, as ascendentes (dependentes do texto, da análise cuidadosa do *input* visual) e as descendentes (baseadas no conhecimento prévio do leitor e na sua capacidade de inferência, de previsão), relacionando ora um tipo, ora outro, de maneira consciente, no momento em que cada uma delas se fizer necessária. Há momentos na leitura em que um trecho difícil para o leitor exige que ele leia linear e cuidadosamente, e há outros em que apenas inferências pelo contexto permitem a compreensão sem problemas.

Além disso, acredita-se que o leitor estratégico proficiente seja aquele que além de utilizar apropriadamente estratégias ascendentes e as descendentes, mantém em mente o objetivo da leitura.

Para essa verificação, foi realizado diagnóstico e caracterização do indexador do Centro de Informações Nucleares (CIN) por meio de suas atividades de análise e leitura documentária para indexação.

### 3 METODOLOGIA

#### 3.1 Centro de Informações Nucleares (CIN)

O Centro de Informações Nucleares (CIN), departamento da Comissão de Nacional de Energia Nuclear (CNEN), faz do Brasil um país membro cooperante do *International Nuclear Information System* (INIS) que é *coordenado pela International Atomic Energy Agency* (IAEA). O principal produto do desse sistema de informação é a base de dados bibliográfica INIS que, atualmente possui mais de 3 milhões de registros sendo a fonte mundial mais completa sobre literatura técnico-científica na área nuclear. (BRASIL, 2015).

O CIN “[...] é o ponto de contato nacional para encontrar informações específicas sobre a área nuclear através do acesso a serviços automatizados e à base de dados bibliográficas.” (BRASIL, 2015).

A opção por um estudo de caso com esse sistema de informação partiu do pressuposto de que as atividades de análise documentária realizadas para geração da Base de Dados INIS contaria com recursos humanos treinados e com metodologia padrão.

Com a permissão e o interesse do Sr. Luiz Fernando Passos de Macedo, então gerente do Centro de Informações Nucleares, foi possível a coleta de informações necessárias à caracterização funcional desse sistema.

O funcionamento do INIS depende de seus membros cooperantes que se encarregam de:selecionar a literatura nuclear gerada em seu país;preparar os dados de entrada de acordo com as regras do INIS;enviar uma cópia do texto completo de artigos da literatura que não está disponí-

vel através dos canais comerciais convencionais; disponibilizar os serviços e produtos de informação do INIS aos usuários e/ou membros do Sistema.

Para isso, a atividade de análise documentária, compreendida pelas fases de análise, síntese e representação, é realizada no CIN para a alimentação da Base de Dados INIS. Esses registros representam toda a literatura (teses, artigos, livros, capítulos de livros, trabalhos de congressos etc.) gerada e publicada por pesquisadores brasileiros.

Para a coleta desses trabalhos são feitos contatos com as instituições da pesquisa, universidades, bibliotecas da área nuclear e até mesmo com os próprios autores para que sejam enviados ao CIN os trabalhos publicados na área de energia nuclear e afins.

Os documentos que chegam ao CIN passam por um processo de seleção de assunto para garantir que eles estejam dentro do escopo da Base de Dados INIS.

Para manter a uniformidade dos arquivos informacionais foram elaborados manuais de referência que contém regras, padrões, formatos e diretrizes para catalogadores e indexadores nos quais o Sistema está baseado. Conhecidos como INIS Reference Series, esses manuais estavam disponíveis em CD-ROM e, atualmente estão disponíveis no site [www.iaea.org](http://www.iaea.org).

Os indexadores do CIN realizam o processamento temático dos documentos utilizando o INIS: Manual for Subject Analysis (IAEA-INIS-12) produzido em conjunto pelo INIS e ETDE (Energy Technology Data Exchange) e publicado em 1995 para substituir o INIS: Manual for Indexing (IAEA-INIS-12, Rev.2 ), além do INIS Thesaurus.

O Manual descreve em sequência todos os passos da análise de assunto em 6 capítulos: 1 Introdução à Análise de Assunto; 2 Análise Preparatória; 3 Classificação de Assunto; 4 Resumo e Aumento de título; 5 Indexação de Assunto; 6 Diretrizes e Exemplos para Análise Temática em campos de Assuntos Particulares.

Nos procedimentos sugeridos para indexação existe apenas um item destinado à leitura do documento. Sendo assim houve a necessidade de leitura de todo o Manual para reconhecimento e síntese de aspectos relacionados à leitura.

Após esse reconhecimento, os aspectos sintetizados foram ordenados para garantir que a sequência “antes da leitura - durante a leitura - após a leitura” fosse mantida.

### **RECOMENDAÇÕES DO SISTEMA PARA IDENTIFICAÇÃO DE CONCEITOS**

1. O entendimento do conteúdo técnico-científico do documento é o pré-requisito básico para indexação, isto significa que o indexador deverá ser especialista na área do assunto da publicação a ser indexada, e ao mesmo tempo, familiarizado com o Thesaurus e com as regras de indexação.
2. O critério a ser sempre utilizado para interpretação do documento deverá ser inteiramente baseado no usuário.
3. O indexador deverá estar atento para identificar cada parte da literatura sob o ponto de vista do usuário
4. Identificar os conceitos principais que geralmente constituem a razão essencial da publicação do material.
5. Identificar uma, duas ou talvez três ideias principais que constituem a razão essencial da publicação do material. Cada uma dessas ideias deverá ser representada por descritores coordenados.

### **LEITURA E IDENTIFICAÇÃO DE CONCEITOS**

1. Leia cuidadosamente o título e resumo.
2. O mais comum é começar a indexação pelos conceitos contidos no resumo e no título e depois esquadrihar o texto todo para encontrar itens informacionais que estão faltando no resumo ou que requerem mais precisão.
3. Esquadrinhe o texto todo para encontrar conceitos mais detalhados, prestando cuidadosa atenção ao Índice de Assunto, Introdução, Resumo, Conclusões, Gráficos e Tabelas.

4. O indexador deve decidir se a publicação consiste de partes de procedimentos com tópicos diferentes suficientemente para que uma **subdivisão** seja feita no nível bibliográfico e cada parte seja tratada como um documento separado.
5. Para encontrar o descritor mais específico apropriado, o indexador deverá checar *wordblock* do descritor considerado para ver se algum dos termos menores associados a ele são apropriados.
6. Sugere-se que uma lista de trabalho dos itens informacionais seja feita primeiramente. Durante a leitura do título e resumo e esquadramento do texto todo, o indexador lista aqueles itens que representam a informação significativa. Feita esta lista, ela deverá ser “traduzida” em descritores.

#### REPRESENTAÇÃO DOS CONCEITOS POR DESCRITORES DO THESAURUS

1. Cada um dos conceitos identificados deve ser representado por descritores retirados do Thesaurus
2. Selecione somente aqueles descritores que representem conceitos **realmente discutidos em detalhe** no documento e conceitos sobre os quais alguma informação útil pode ser obtida pela leitura do documento.

#### 3.2 Protocolo verbal

Para a observação das estratégias de leitura de indexadores do CIN foi aplicada, como método de observação de processos, a técnica de coleta de dados introspectivos “Protocolo Verbal” ou “Pensar Alto” nos moldes de Ericsson e Simon (1987). Segundo os autores, as observações do processo são como observações que fornecem informações sobre passos de processamento individual, tais como verbalizações espontâneas, sequência de movimentos com os olhos, exteriorizando seus processos mentais, mantendo a sequência das informações processadas.

Cohen (1987) cita o “Pensar Alto” dentre três tipos básicos de dados provenientes da técnicas introspectivas:

- Auto-Relato: afirmações generalizadas sobre o comportamento durante a leitura;
- Auto-Observação: inspeções de comportamentos específicos de leitura, enquanto a informação ainda está sob o foco de atenção;
- Auto-Revelação: “Pensar alto”, o pensamento é direta e automaticamente externalizado.

Para Ericsson e Simon (1987), o leitor pode exteriorizar seus processos mentais enquanto a informação processada está sob o foco de sua atenção. O indivíduo lê e interpreta ao mesmo tempo, exteriorizando em voz alta tudo o que “passa pela sua cabeça” durante a leitura.

Nardi (1993) esclarece que, na procura de métodos para acesso direto aos processos mentais dos indivíduos, os pesquisadores encontraram apoio na Linguística, Sociologia e Psicologia. Essas áreas utilizam métodos introspectivos (Relatos, Protocolos, Declarações Verbais), ou seja, declarações dos próprios informantes sobre como organizam e processam a informação.

Segundo Cavalcanti e Zanotto (1994), protocolos são geralmente definidos como relatos verbais dos processos mentais conscientes do informante. Em outras palavras, eles se referem ao “pensar alto” do informante enquanto realiza uma tarefa de qualquer natureza.

Os protocolos verbais foram introduzidos na pesquisa qualitativa em Psicologia a partir de 1980 e desde então sua validade como reveladores de processos mentais tem sido questionada. Durante a era do behaviorismo, eles foram banidos da ciência. Segundo Ericsson e Simon (1987), quando o cognitivismo entrou em evidência como um novo paradigma, eles ressurgiram como principal fonte de dados para a pesquisa cognitiva. Este renascimento ocorreu dentro do arcabouço teórico do processamento da informação relacionado a estudos de solução de problemas.

Eles ultrapassaram os limites da Psicologia Cognitiva e passaram a ser usados na Linguística Aplicada, área em que marcaram presença forte na pesquisa sobre leitura em língua estrangeira.

Segundo Ericsson e Simon (1987) alguns autores criticam a técnica do “pensar alto”, por acreditarem que ela pode modificar os processos mentais, fazendo com que as informações dos sujeitos não sejam precisas, completas, não sejam confiáveis. Em defesa do Relato Verbal, Ericsson e Simon (1987) apoiam-se na teoria do processamento da informação e explicam que a informação é armazenada na memória a curto prazo (STM) - duração intermediária de retenção e acesso rápido à informação e na memória de longo prazo (LTM) - armazenamento e duração relativamente permanente, lenta a recuperação da informação.

O modelo proposto por Ericsson e Simon (1987) prevê que a informação recém-apreendida pelo processador central é mantida na memória de curto prazo por algum tempo e é diretamente acessível para processamento subsequente, enquanto que na memória de longo prazo a informação precisa ser recuperada antes de ser relatada. Dessa forma, as informações coletadas em Relatos Verbais são as recém-apreendidas, diretamente acessíveis para processamento subsequente.

Cavalcanti e Zanotto (1994) argumentam que a questão da confiabilidade dos dados provenientes de técnicas introspectivas é um problema de outras técnicas também. Apesar de ainda controversa, essa técnica é, segundo Nardi (1993), o único instrumento de coleta, no momento disponível, que possibilita observar processos do leitor durante a compreensão de um texto, o que justifica a nossa escolha.

A seguir, descrevemos a metodologia do Protocolo Verbal para observação da leitura documentária.

### *1. Seleção do texto*

A seleção do Texto-Base foi solicitada ao CIN com a recomendação de que fosse um texto ainda não indexado e que não fosse de uma

área muito específica, podendo ser indexado por qualquer profissional independentemente de sua área de atuação.

O texto selecionado pelo CIN é da área de Química publicado como artigo no periódico *Química Nova*:

FIGUEIRA, R.C.L.; CUNHA, I.I.L. A contaminação dos oceanos por radionuclídeos antropogênicos. *Química Nova*, São Paulo, v.21, n.1, p.73-77, 1998.

O texto apresenta uma figura e três gráficos além da seguinte estrutura textual:

- Título;
- Resumo (somente em inglês);
- Palavras-chaves (omitidas durante a entrevista)
- Introdução
- Dois subtítulos:
  - Comportamento dos radionuclídeos no meio marinho;
  - Níveis de radionuclídeos nos oceanos;
- Conclusão;
- Referências bibliográficas.

## 2. Seleção dos sujeitos

A seleção dos sujeitos indexadores para o propósito da pesquisa teve como critérios o tempo de permanência da pesquisadora no Rio de Janeiro e a disponibilidade dos indexadores externos.

De acordo com a metodologia de observação do Protocolo Verbal, a entrevista realizada com 4 indexadores constou da leitura do texto-base. Para esta leitura, a entrevistadora solicitou que a fizessem naturalmente, conforme sua preferência e rotina diária, tendo como objetivo a identificação e seleção de conceitos para a indexação, uma vez que esse é o nível de leitura frequentemente mais exigido para a realização de tarefas de um indexador.



### *3. Contatos para conhecer a disponibilidade do indexador*

O próprio CIN contactou os indexadores para verificação de interesse pela pesquisa e disponibilidade para a entrevista, fazendo o agendamento da mesma.

### *4. Conversa informal com cada sujeito*

No dia de cada entrevista foi realizada uma conversa informal com cada sujeito antes da realização do Protocolo Verbal. Foram mencionados os objetivos da pesquisa mostrando sua importância para o desenvolvimento da área, evidenciando-se a preocupação de manter a identidade de cada um dos sujeitos oculta com o propósito de não comprometer os dados e deixá-los a vontade durante a realização da tarefa.

### *5. Familiarização com a tarefa do “ThinkAloud” (Pensar Alto) através de “Instruções aos Sujeitos”*

Antes da aplicação do Protocolo Verbal ou Pensar Alto, foi feita uma familiarização com a tarefa utilizando um texto contendo “Instruções aos Sujeitos”, elaborado com o propósito de descontrair e ao mesmo tempo apresentar procedimentos que poderão auxiliar o sujeito no desempenho da tarefa.

### *6. Gravação do “pensar alto” durante a leitura do texto-base*

Antes de começar a gravação, foi entregue ao sujeito o texto-base lembrando que é preciso pensar alto durante toda a leitura e exteriorizar seus processos mentais, procurando esquecer a presença da pesquisadora que estará presente apenas com o intuito de lembrar que é preciso pensar alto e também controlar o gravador.

### *7. Entrevista retrospectiva*

Logo após a aplicação do protocolo verbal, foi feita uma entrevista retrospectiva com cada sujeito com o objetivo de esclarecer alguns pontos considerados obscuros pela pesquisadora.

### 8. *Transcrição dos Protocolos Verbais dos sujeitos*

As transcrições foram feitas de maneira a destacar a compreensão dos sujeitos, suas dúvidas, equívocos, identificação e seleção de termos.

Com a gravação do “Pensar Alto” durante a leitura e entrevista retrospectiva dos quatro sujeitos indexadores entrevistados, foi possível a transcrição literal completa em texto escrito e, em seguida, a análise dos dados coletados.

## 4 RESULTADOS E DISCUSSÃO DA OBSERVAÇÃO DAS ESTRATÉGIAS DO INDEXADOR NA LEITURA DOCUMENTÁRIA

Considerando-se que as estratégias metacognitivas listadas por Brown (1987) e associadas à Norma são as mais representativas para a leitura com fins documentários, a discussão dos resultados obtidos realizou-se a partir dos seguintes aspectos:

- I. Sequência de operações (aspectos) da leitura de sujeitos;
- II. Manter em mente o seu objetivo de representar o texto para futura recuperação;
- III. Exploração de seu conhecimento de estruturas textuais;
- IV. Identificação de aspectos importantes da mensagem para seleção de conceitos;
- V. Domínio da linguagem do sistema como conhecimento prévio.

### Sequência de operações (aspectos) da leitura dos sujeitos

A sequência de operações, sintetizada a partir dos dados coletados na entrevista, dos quatro sujeitos, revelou pela análise documentária, que dos 10 tipos de operações consideradas indicativas de estratégias metacognitivas, o sujeito 1 utilizou 8 tipos, o sujeito 2 utilizou 6, o sujeito 3 utilizou 7 tipos e o sujeito 4 utilizou 4 tipos. Com relação ao padrão de sequências de operações, observou-se que, com exceção do sujeito 4 que repetiu somente 2 operações mais de uma vez, os outros sujeitos possuem um padrão de sequência, mesmo não sendo semelhantes ou cons-

tantes. Observa-se um ponto divergente entre os padrões: a operação 5 “Associação com linguagem” foi utilizada somente pelo sujeito 1. Essas constatações revelam que: a leitura feita pelos sujeitos é metacognitiva; os indexadores se utilizam muito mais do conhecimento prévio do assunto do que da linguagem documentária do sistema para a identificação de conceitos; há o domínio da estrutura textual de um artigo para ir direto às partes do texto em que foram identificaram aspectos importantes da mensagem e eles seguem a metodologia descrita no Manual INIS.

#### Manter em mente o seu objetivo de representar o texto para futura recuperação

De acordo com a análise das sequências das operações dos 4 sujeitos a operação “Explicitação dos objetivos da leitura” foi utilizada em 4 momentos pelo sujeito 1, em 3 momentos pelo sujeito 2, em 1 momento pelo sujeito 3 e em nenhum momento pelo sujeito 4. Como objetivo da leitura do indexador é representar o texto para futura recuperação pelo usuário do sistema, é extremamente necessário a manutenção deste objetivo em mente para identificação e seleção de conceitos que possam, realmente, expressar o conteúdo do documento, o que não ficou explícito durante a leitura do sujeito 4.

#### Exploração de seu conhecimento de estruturas textuais

Os sujeitos 1, 2, 3 foram os que melhor exploraram a estrutura textual, seguindo as recomendações do Manual INIS. O sujeito 1 explorou 6 partes da estrutura textual e dos 9 termos selecionados, 6 são iguais ao INIS Thesaurus. O sujeito 2 explorou 4 partes da estrutura textual, selecionando 6 termos dos quais 4 são descritores. O sujeito 3 também explorou 4 partes da estrutura textual e selecionou 8 termos dos quais 6 são descritores. Já o sujeito 4 explorou 3 partes da estrutura textual e selecionou 4 termos dos quais 3 são descritores. Os resultados demonstram que os sujeitos são muito proficientes, explorando de forma eficaz a estrutura textual resultando na identificação de conceitos. Isto mostra que o domínio da estrutura textual facilita a exploração de modo a garantir a estratégia de identificação de conceitos.

### Identificação de aspectos importantes da mensagem para seleção de conceitos

Na análise das partes da estrutura textual em que foram identificados conceitos, os resultados demonstram que o sujeito 1, além de ter sido o que mais explorou toda a estrutura textual, também utilizou o maior número de partes para identificação. Com relação ao sujeito 2, que também explorou toda a estrutura textual, observou-se que, para identificação de termos, limitou-se ao Título, Resumo, Introdução e Conclusão. O sujeito 3 explorou a estrutura textual, utilizando para identificação de conceitos, Introdução, Figura 1, Resumo e um subtítulo. O sujeito 4 limitou-se somente ao Resumo. A observação da sistemática de identificação de conceitos na leitura dos sujeitos demonstrou o uso de diferentes estratégias mas sem uma sistemática constante, ou seja, o sujeito 1 limitou-se a grifar conceitos que reconhece por associações com linguagem do sistema durante a leitura. Já o sujeito 2 relaciona partes importantes da estrutura textual com os termos identificados. O sujeito 3 somente grifa as palavras que reconhece como chave. O sujeito 4 não explicitou a utilização de nenhuma sistemática para identificação de conceitos. Os resultados dos sujeitos 1 e 2 combinam as estratégias de identificação de conceitos com o domínio da linguagem do sistema e da estrutura textual e conhecimento prévio do assunto, garantindo que os termos extraídos sejam mais representativos e ao mesmo tempo, compatíveis com a linguagem de recuperação.

### Domínio da linguagem do sistema como conhecimento prévio

A verificação da compatibilidade dos termos selecionados pelos sujeitos com a linguagem do sistema resultou em que, do total de 23 termos selecionados pelos sujeitos e pelo autor, 13 são descritores, 7 seriam adaptados, e 3 não foram encontrados. Correspondendo a 86,95% de compatibilidade com o sistema, conclui-se que os termos selecionados são **extremamente compatíveis**. Considerando-se que o indexador é **especialista**, o domínio do assunto se torna mais fácil e é utilizado como conhecimento prévio.

## 5 CONSIDERAÇÕES FINAIS

Conforme os objetivos de identificar os procedimentos de leitura documentária utilizados para exame de documento e identificação de conceitos e examinar as estratégias de leitura usadas pelos indexadores para a identificação de conceitos que melhor representem o documento, foi realizado referencial teórico inicial sobre estratégias de leitura em Análise Documentária, bem como estudo de caso com o Centro de Informações Nucleares (CIN) da Comissão Nacional de Energia Nuclear (CNEN) para caracterização do indexador e levantamento das atividades de Análise Documentária.

A partir da síntese dos resultados obtidos, pode-se concluir que o indexador do Centro de Informações Nucleares é bastante estratégico e proficiente na realização da leitura para indexação.

Embora não sejam bibliotecários, os indexadores recebem treinamento específico do Sistema e contam com uma linguagem documentária bem estruturada e flexível e com Manuais de Referências para realização da indexação.

Sendo especialistas na sua área de atuação, os indexadores:

- Utilizam o domínio do assunto como conhecimento prévio e não a linguagem documentária do sistema, selecionando assim termos relacionados com sua área de atuação e extremamente compatíveis;
- Pouco utilizam a sequência de operações, pois quanto maior o conhecimento prévio do assunto menor a sequência de operações utilizada;
- Utilizam muito a exploração da estrutura textual;
- Não apresentam dificuldades durante a leitura;
- Seguem as recomendações do INIS para indexação.

A leitura do indexador do CIN é metacognitiva ao explorar toda a estrutura textual do documento, ao utilizar estratégias de leitura com os aspectos previstos por Brown (1980) e, principalmente, ao manter em mente

o objetivo da leitura – representar para recuperar. É vital para a indexação bem sucedida que o indexador realize estratégias de compreensão da leitura interagindo com o texto e tendo em vista o seu objetivo que é representar para que o usuário recupere.

Quanto à observação de estratégias de leitura utilizadas, ficou explícito que existem duas operações distintas utilizadas pelos indexadores durante a leitura: *Identificação de conceitos* e *Seleção de conceitos*. Isso esclarece que o indexador realiza as duas operações *durante* a leitura e não *após* a leitura. A Norma indica a realização das operações, mas não explicita o *momento* em que acontece, somente indicando que o estágio de Identificação de Conceitos deve ser feita após o exame do documento. De acordo com essa indicação, entende-se que a identificação de conceitos presentes no conteúdo do documento será realizada *após* a leitura e os conceitos serão, depois de identificados, representados por descritores de uma linguagem adotada pelo sistema de informação.

Tendo em vista os resultados obtidos na análise das estratégias através da técnica do Protocolo Verbal conclui-se que a presença de estratégias metacognitivas é resultado da compreensão da leitura e o uso maior ou menor de estratégias significa uma proficiência maior ou menor na compreensão do conteúdo do documento.

## REFERÊNCIAS

- 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.
- BRASIL. Ministério da Ciência, Tecnologia e Inovação. Comissão Nacional de Energia Nuclear. Centro de Informações Nucleares. *Sobre o CIN*. 2015. Disponível em: <<http://www.cnen.gov.br/centro-de-informacoes-nucleares/sobre-o-cin>>. Acesso em: 11 jan. 2017.
- BROWN, N. Metacognitive development and reading. In: SPIRO, R. J.; BRUCE, B. C.; BREWER, W. F. (Org). *Theoretical issues in reading comprehension*. New Jersey: L. Erlbaum, 1980. p. 453-481.
- CAVALCANTI, C. R. *Indexação & tesouros: metodologia & técnicas*. Ed. Preliminar. Brasília: Associação dos Bibliotecários do Distrito Federal, 1978.

- CAVALCANTI, M. C.; ZANOTTO, M. S. Introspection in Applied Linguistics: meta-research on verbal protocols. In: BARBARA, L.; SCOTT, M. (Ed.). *Reflections on language learning*. Clevedon: Multilingual Matters, 1994. p. 148-156.
- CHAUMIER, J. *Análisis y lenguajes documentales*. Barcelona: Mitre, 1986.
- CHAUMIER, J. Indexação: conceito, etapas e instrumentos. Trad. José Augusto Chaves Guimarães. *Revista Brasileira de Biblioteconomia e Documentação*, São Paulo, v. 21, n.1/2, p. 63-79, jan./jun. 1988.
- CINTRA, A. M. M. Estratégias de leitura em documentação. In: SMIT, J. W. (Coord.). *Análise documentária: a análise da síntese*. 2.ed. Brasília: IBICT, 1989. p. 29-37.
- COHEN, A. D. Using verbal reports in research on language learning. In: FAERCH, C.; KASPER, G. (Ed.). *Introspection in second language research*. Clevedon: Multilingual Matters, 1987.p.82-95.
- DIAS, E. W.; NAVES, M. M. L. *Análise de assunto: teoria e prática*. Brasília: Thesaurus, 2007.
- ERICSON, K. A.; SIMON, H. A. Verbal reports on thinking. In: FAERCH, C.; KASPER, G. (Ed.). *Introspection in second language research*. Clevedon: Multilingual Matters, 1987.p.24-53.
- FUJITA, M. S. L. *A leitura documentária do indexador: aspectos cognitivos e lingüísticos influentes na formação do leitor profissional*. 2003. 321f. Tese (Livre-Docência em Análise Documentária e Linguagens Documentárias Alfabéticas) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2003.
- FUJITA, M. S. L. *A leitura em análise documentária*. 1998. 184 f. (Relatório de Pesquisa) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 1998.
- FUJITA, M. S. L. *Leitura em análise documentária*. 2001. 185 f. (Relatório de Pesquisa) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2001.
- KATO, M. *O aprendizado de leitura*. São Paulo: Martins Fontes, 1987.
- LANCASTER, F.W. *Indexação e resumos: teoria e prática*. 2.ed. Brasília: Briquet de Lemos, 2004.
- NARDI, M. I. A. *As expressões metafóricas na compreensão de texto escrito em língua estrangeira*. 1993. 260 f. Dissertação (Mestrado em Lingüística Aplicada ao Ensino de Línguas) – Pontifícia Universidade Católica, São Paulo, 1993.

UNISIST. Princípios de indexação. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v.10, n.1, p. 83-94, mar. 1981.

VAN SLYPE, G. *Lenguages de indización: concepción, contrucción y utilización en los sistemas documentales*. Trad. Pedro Hípola e Félix de Moya. Madrid: Fundación Germán Sánchez Ruipérez, 1991.