

A GERAÇÃO DA BASE DE DADOS DO INSTITUTO DE PESQUISAS ENERGÉTICAS E NUCLEARES: PROBLEMAS E PERSPECTIVAS

TEREZINE ARANTES FERRAZ
IRENE JOSEFA DE SOUZA

Divisão de Informação e Documentação Científicas
Instituto de Pesquisas Energéticas e Nucleares (IPEN)
05508 São Paulo, SP

MARINA ZAMPERETTI

Departamento de Processamento de Dados
Instituto de Pesquisas Energéticas e Nucleares (IPEN)
05508 São Paulo, SP

São descritos os antecedentes que levaram à criação da base de dados gerada pela Divisão de Informação e Documentação Científicas e pelo Departamento de Processamento de Dados do Instituto de Pesquisas Energéticas e Nucleares, bem como o tipo de documentos que compõem essa base de dados. São igualmente descritos, de forma não pormenorizada, os procedimentos utilizados para a introdução de dados na base. A base de dados é operacional desde 1982, sendo as informações acessíveis em linha. Os objetivos principais, com a geração dessa base, foram atingidos: gerar um instrumento que pudesse orientar o direcionamento, em termos de política científica, a ser seguido pela instituição, possibilitando o acesso em linha, bem como sob forma de bibliografia analítica impressa.

1. INTRODUÇÃO

O Instituto de Pesquisas Energéticas e Nucleares — IPEN foi criado com a finalidade de realizar pesquisas e desenvolvimento no campo das aplicações pacíficas da energia nuclear. Fundado em 1956, mediante convênio firmado entre o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Universidade de São Paulo, com o nome de Instituto de Energia Atômica (IEA), o IPEN está, atualmente, subordinado à Comissão Nacional de Energia Nuclear. escola cumpra seus elevados objetivos. No segundo, O Quadro, apresenta-se a Biblioteca Gen. Cordeiro de Farias - assim denominada em homenagem ao primeiro comandante da Escola Superior de Guerra - suas diretrizes, o

A geração da base de dados do Instituto de Pesquisas Energéticas e Nucleares...

A comunidade técnico-científica do IPEN é formada, atualmente, por cerca de 500 técnicos de nível superior, predominantemente engenheiros, físicos, químicos, biólogos, bioquímicos, médicos, matemáticos e geólogos, na sua grande maioria mestres e doutores.

O IPEN vem dando grande ênfase ao desenvolvimento de projetos e programas de pesquisa, visando ao estabelecimento de tecnologia própria para a produção de compostos de urânio, elementos combustíveis de reatores nucleares e outros materiais e produtos indispensáveis à utilização da energia nuclear, abrangendo todas as etapas do ciclo do combustível nuclear. Outras atividades prioritárias são as referentes a estudos e ensaios relacionados com reatores de potência e a pesquisas para desenvolver aplicações de métodos nucleares na solução de problemas da comunidade, nos campos de Medicina, Bioquímica, Engenharia, Indústria e na preservação e aproveitamento dos recursos naturais, além da produção de fontes radioativas, radioisótopos e substâncias marcadas. São também realizadas pesquisas nos campos das Ciências dos Materiais, Física de Reatores, Física e Química Nucleares, Física da Matéria Condensada e Radioquímica, com o objetivo de consolidar e expandir o nível atual de conhecimento e fornecer subsídios para o desenvolvimento de futuras tecnologias. Por outro lado, o IPEN tem contribuído significativamente na formação de recursos humanos para a área nuclear, em nível médio e em nível de pós-graduação, com a formação de mestres e doutores. Como elemento de conexão entre a universidade e a indústria, o IPEN desempenha essa função mediante a realização de cursos de especialização e de pós-graduação. O IPEN é responsável pelas áreas de concentração de tecnologia nuclear e pela de reatores nucleares de potência e tecnologia do combustível nuclear, a nível de mestrado e doutorado.

Grande parte da produção técnico-científica do IPEN é divulgada em publicações seriadas, denominadas, atualmente, *Publicação IPEN*, iniciada em 1958, e *Informação IPEN*, iniciada em 1962.

Dessas séries destaca-se a *Publicação IPEN*, onde são publicados resultados inéditos de pesquisas e desenvolvimento realizados no IPEN, nas áreas já mencionadas. Trata-se de publicação de circulação internacional, dirigida a instituições congêneres, sendo indexada nos periódicos de resumos *Nuclear Science Abstracts* (encerrado em junho de 1976) e no *INIS Atomindex* (corrente). Embora sejam publicações seriadas, são consideradas como relatórios pelas fontes secundárias e conhecidas pelas siglas *IPEN-Pub* e *IPEN-Inf*.

A biblioteca do IPEN, denominada Divisão de Informação e Documentação Científicas — DIDC, foi instituída no mesmo ano da criação do IPEN, com o objetivo de fornecer à comunidade técnico-científica informações na área nuclear.

O prédio onde está instalada a DIDC situa-se no epicentro da área onde estão localizadas as demais unidades técnico-científicas do IPEN, portanto de fácil acesso aos usuários.

TEREZINE ARANTES FERRAZ

A área atualmente ocupada pela DIDC é superior a 2.000m², dos quais 1.333 são destinados ao salão de leitura e saletas individuais de estudo; a administração, processos técnicos, documentação e armazenagem ocupam a área restante.

A divisão possui acervo representativo no campo nuclear e áreas afins, somando um total de 17.000 livros (entre monografias, teses, manuais, normas), 1.643 títulos de periódicos, 42.876 relatórios *full size*, 303.818 relatórios em microfichas e 1.302 folhetos e separatas.

As coleções da DIDC são de livre acesso, permitindo-se o empréstimo a domicílio de livros e relatórios⁷.

A DIDC, além dos serviços gerais de referência, atende a pedidos de levantamentos bibliográficos retrospectivos, automatizados e manuais, faz empréstimos a domicílio e interbibliotecas, promove cursos de metodologia da pesquisa bibliográfica e orientação individual sobre os serviços e publicações da biblioteca; mantém um serviço de reprografia e edita publicações do tipo serviço de alerta, com a finalidade de manter o usuário informado sobre o material recebido na quinzena.

Em 1975, após dois anos de estudos e planejamento, a DIDC deu início às atividades de automação e, em outubro de 1978, a coleção de livros estava toda controlada pelo computador. Seguiu-se a coleção de publicações periódicas e, finalmente, em 1979, o empréstimo. O sistema de automação da DIDC intitula-se *Integrated Library System (ILS)* e sua descrição consta de trabalho já publicado⁵.

Na época da sua implantação, o ILS, com as suas características de programa concebido para integrar as atividades de aquisição, processamento e empréstimo, constituiu uma iniciativa pioneira no Brasil. Os produtos emitidos pelo ILS consistem de relatórios impressos, que oferecem ao leitor grande variedade de abordagens: as tradicionalmente oferecidas pelas bibliotecas e outras que atendem a características específicas da literatura nuclear e peculiaridades de comportamento do usuário de bibliotecas dessa área⁴.

Vencida a etapa da automação das atividades da biblioteca, a DIDC se propôs, em 1980, a converter a publicação *Scientific and Technical Papers (STP)* — bibliografia analítica da produção científica do IPEN e até então feita por processos manuais — num sistema automatizado, visando o arquivamento, recuperação e divulgação da produção técnico-científica do IPEN. Esse sistema gera dois produtos: 1) uma base de dados criada com a utilização do programa STAIRS e 2) relatórios, cujo produto final é uma bibliografia analítica, a partir da qual se obtém a publicação *Scientific and Technical Publications — STP*, nos moldes do INIS Atomindex.

Assim, em 1982 foi publicada uma edição cumulativa do STP, comemorativa dos 25 anos de criação do Instituto, contendo os 1.345 trabalhos até então publicados pelo corpo técnico-científico do IPEN.

A geração da base de dados do Instituto de Pesquisas Energéticas e Nucleares...

A geração de uma base de dados contendo toda a produção científica do Instituto correspondia a um anseio da comunidade do IPEN, que necessitava ter acesso rápido e preciso às publicações geradas pelo próprio Instituto, as quais constituíam, conseqüentemente, a sua memória técnico-científica. O conhecimento das pesquisas até então efetuadas era essencial para o delineamento das futuras linhas de pesquisa a serem seguidas pelo Instituto. Em cumprimento a essa necessidade foi gerada a base de dados do *Scientific and Technical Publications*, o *STDB*.

2.A SITUAÇÃO MUNDIAL

Um número sempre crescente de bibliotecas em todo o mundo está localizando informação vital mais rapidamente em serviços de recuperação da informação em linha. Terminais de computador, linhas telefônicas e redes de informação promovem a ligação entre as bibliotecas e os serviços de recuperação da informação. Os serviços em linha provêem acesso a centenas de bases de dados contendo informação bibliográfica, pictórica, numérica, textual, etc.

Os EEUU, a partir da década sessenta, foram e têm sido o país gerador e detentor do maior número de bases de dados. A experiência mais marcante foi, sem dúvida, a da *Lockheed Missiles and Space Company*, de Paio Alto (Califórnia), que, em 1966, foi beneficiada com um contrato para projetar, programar e implantar um serviço de computação para dar suporte às publicações bibliográficas da *National Aeronautics and Space Administration (NASA)*³.

Dessa ocasião para o presente tem-se observado um número cada vez maior de bases de dados bibliográficas e não-bibliográficas. Atualmente há no mundo mais de 150 serviços de recuperação da informação, cobrindo, virtualmente, todos os ramos do conhecimento, com milhões de referências a artigos de revistas, relatórios técnicos, teses, dissertações, conferências, normas, especificações, patentes, livros, etc.

Os serviços de recuperação da informação desempenham, a cada dia, papel mais importante em grandes e modernas bibliotecas. Nos EEUU o número de bibliotecas utilizando esses serviços é estimado entre 4 a 5 mil, e esse número seria significativamente maior se o Canadá, Europa e Japão fossem incluídos nessa contagem². Presentemente esse número é superior a 10 mil, computados aí 30 países³.

Bibliotecas e centros de informação na área comercial constituem, provavelmente, o maior bloco de bibliotecas utilizando serviços comerciais em linha. A necessidade de acesso imediato à informação corrente é certamente o elemento responsável por essa alta representação, vez que a informação corrente é essencial para as áreas de P & D no mundo competitivo dos negócios. Seguem-se as entidades de pesquisa, agências governamentais e bibliotecas acadêmicas e públicas.

Além dos centros coletores e processadores da informação, existem os inter

TEREZINE ARANTES FERRAZ

mediários da informação (*information brokers*), que fazem o acesso aos centros/sistemas para os seus clientes e, em geral, combinam as atividades de busca com as de localização e obtenção de documentos⁶. O IBICT e a PROMON, entre outros, são usuários dos serviços de SDC e *Loockheed*.

Os sistemas estrangeiros de recuperação da informação explorados no Brasil, na área de C & T, não chegam a duas dezenas: IFIS, MEDLINE, IMIS, AGRÍCOLA, CAIN, CAB, BIOSIS, AGRIS, COMPENDEX, NUCLEAR SCIENCE ABSTRACTS (NSDB) e NTIS são alguns deles. Garcia⁶ relaciona exaustivamente essas e outras bases de dados, fornecendo, além do nome da entidade mantenedora, outros dados de interesse.

No Brasil são relativamente poucos os trabalhos sobre geração de bases de dados nacionais especializadas¹.

Tem-se notícia de que o IBICT, como órgão coordenador da informação em C & T no Brasil, está desenvolvendo esforços para a constituição de bases de dados nacionais especializadas, dentro de um sistema que assegura a compatibilidade e o intercâmbio de dados. Com esse programa de criação e utilização de bases de dados nacionais, o IBICT inicia a formação das seguintes bases de dados especializadas: Amazônia/Carajás; Química; Geociência/Tecnologia Mineral e Energia.

Charbel² cita, entre as bases de dados geradas no Brasil, as do SERPRO, ARUANDA, PROBAN, PRODASEN, SEADE, SIM, SBC, das Bolsas de Valores de São Paulo e Rio de Janeiro, TELCOMM, QUEM É QUEM (Grupo Visão) e algumas outras em âmbito restrito, como é o caso da PRODAM, IMLA da BIREME, DATAPREV, AGROBASE e outras.

No campo da energia nuclear, o Centro de Informações Nucleares da Comissão Nacional de Energia Nuclear (CIN/CNEN) gerou uma base de dados específica — FONTE, que inclui artigos de periódicos, relatórios, teses, anais de congressos e palestras sobre vários aspectos das fontes alternativas de energia. Essa base de dados destinava-se, inicialmente, ao uso interno da CNEN. A demanda externa de documentos, especialmente brasileiros, relativos a outras fontes de energia além da nuclear, fez com que a CNEN estendesse seu uso à comunidade em geral.

Em 1981, conforme relatado no início deste trabalho, a Divisão de Informação e Documentação Científicas (DIDC) e o Centro de Processamento de Dados do IPEN iniciaram estudos com o objetivo de colocar em máquina toda a produção técnico-científica gerada pelos pesquisadores do IPEN.

2. ENTRADA DE DADOS

Conforme mencionado, a produção técnico-científica do IPEN se compõe, em grande parte, de trabalhos divulgados nas séries do Instituto (*Publicação— IPEN e Informação—IPEN*) ou publicados em periódicos nacionais e estrangeiros ou, ainda, apresentados em simpósios, congressos, conferências, etc.

A geração da base de dados do Instituto de Pesquisas Energéticas e Nucleares...

A partir das características dessas publicações e visando a abranger as informações pertinentes a esses documentos, foram determinadas as categorias que compuseram o formulário de entrada. Abaixo, encontram-se relacionadas essas categorias, com as respectivas funções.

010	Assunto principal
020	Assunto secundário
030	Título primário
040	Autor(es) + afiliação
045	Departamento ou divisão
050	Idioma
060	Título original
070	Relatório primário
080	Relatório secundário
090	Publicação periódica
100	Fonte
110	Notas
120	Descritor
130	Resumo

3.1 Categoria 010 — Assunto principal

É informação obrigatória, que visa a determinar o assunto do documento e localizá-lo hierarquicamente na forma impressa.

Três dígitos numéricos foram previstos para essa informação, condificada a partir da *Tabela de Categoria de Assunto do INIS Atomindex*.

010	código da categoria de assunto
-----	--------------------------------------

(3)	(3)
-----	-----

3.2 Categoria 020 — Assunto secundário

Até 05 assuntos de três dígitos numéricos foram previstos para esta categoria, que visa a dar informação dos demais assuntos tratados no documento.

02	código da categoria de assunto;,,,,
----	--

(3)	(3)
-----	-----

TEREZINE ARANTES FERRAZ

3.3 Categoria 030 — Título primário

Definiu-se como título primário aquele vertido para o idioma inglês. Quando o documento original for escrito em inglês, a categoria 030 ficará vaga, pois essa informação será especificada na categoria 060.

3.4 Categoria 040 — Autor(es) e afiliação

Categoria obrigatória, onde são especificados os autores e a afiliação daqueles que publicaram trabalhos em colaboração e que não são vinculados ao Instituto.

3.5 Categoria 045 — Departamento ou Divisão

Essa informação identifica o centro e a área que deram origem à publicação. Está disponível somente em linha, tendo por finalidade determinar a quantidade de trabalhos publicados em cada departamento/divisão e o vínculo do pessoal técnico-científico quando da realização de suas pesquisas.

3.6 Categoria 050 — Idioma

Somente para publicações escritas em idioma diferente do inglês é que se especifica esta categoria.

3.7 Categoria 060 — Título original

O preenchimento da categoria 060 é obrigatório para que a unidade seja aceita pelo sistema. Deve ser transcrito no mesmo idioma da publicação.

3.8 Categoria 070 — Relatório primário

Definiu-se como relatório primário a sigla alfanumérica das publicações das séries IPEN.

3.9 Categoria 080 — Relatório secundário

Definiu-se como relatório secundário a sigla alfanumérica atribuída por outra instituição que não seja o IPEN.

3.10 Categoria 090 — Informações da publicação periódica

Esse parágrafo só será preenchido quando se tratar de um trabalho originalmente publicado como artigo de periódico. Os dados que compõem esta categoria são: título, volume, fascículo, paginação e data do periódico.

3.11 Categoria 100 - Fonte

Imprenta e colação são especificados nesta categoria.

3.12 Categoria 110 - Notas

Informação de conferência, tese, dissertação, reimpressão, etc. são especificados como notas.

A geração da base de dados do Instituto de Pesquisas Energéticas e Nucleares...

3.13 Categoria 120 — Descritores

Categoria obrigatória, que visa a identificação do documento na sua especificidade. Para fins de impressão do índice de assunto, eles são identificados como principal (M) ou qualificador (Q).

3.14 Categoria 130 — Resumo

O resumo é transcrito em inglês, uma vez que a publicação tem como finalidade principal divulgar internacionalmente a produção técnico-científica do IPEN.

4. FORMATO DE SAIDA

4.1 Para imprimir o STP nos moldes do INIS Atomindex, definiram-se as categorias que seriam reunidas e destacadas. Abaixo encontram-se relacionados alguns dos casos previstos, sendo que o conjunto *a* representa os elementos de identificação do trabalho, e o *b* e o *c* identificam o resumo e descritores, respectivamente:

1º Caso:

- a) Título traduzido (em negrito). Autor(es) + afiliação. (Idioma). Título original. Relatório primário (em negrito). Fonte. Relatório secundário. Notas.
- b) Resumo
- c) Descritores

2º caso:

- a) Título original (em negrito). Autor(es) + afiliação. Periódico (título e volume, em negrito). Notas.
- b) Resumo
- c) Descritores
- d)

4.2 Para a criação da base de dados e apresentação das informações na tela, determinou-se que as categorias apareceriam na mesma ordem definida no formulário de entrada. Nessa fase decidiu-se também pela formatação de todos os campos, visando não só a recuperação da informação em linha, como também alteração dos campos, via terminal.

5. DESCRIÇÃO GERAL DO SISTEMA

O sistema se compõe de três fases:

- criação dos arquivos magnéticos;
- atualização dos arquivos magnéticos e da base de dados;
- emissão dos relatórios finais.

5.1 Criação dos arquivos magnéticos

Para a criação dos arquivos magnéticos contendo as referências bibliográficas, as informações entradas em disquete são consistidas, gerando uma listagem de dados e um relatório de inconsistências. Após análise e acerto de possíveis er-

TEREZINE ARANTES FERRAZ

ros, esses dados são gravados em disco, num arquivo único, e em seguida passados para fita magnética, no formato exigido pelo programa TRC 260, para utilização do STAIRS e posterior criação da base de dados.

Para a criação dos arquivos magnéticos contendo as categorias e remissivas de assunto o procedimento é semelhante ao descrito acima, sendo que após a *consistência* dos dados as informações são gravadas em arquivos magnéticos distintos.

5.2 Atualização dos arquivos magnéticos

Para o arquivo magnético contendo as referências bibliográficas não se faz atualização no sentido exato da palavra, mas, sim, cria-se um outro arquivo, cujo processamento é o mesmo descrito na fase anterior.

5.3 Emissão dos relatórios finais

Quatro subfases constituem a emissão dos relatórios finais, a saber:

5.3.1 Criação dos arquivos temporários

A partir do arquivo permanente de referências bibliográficas é efetuado o processamento, para criação dos arquivos temporários para emissão dos relatórios finais.

5.3.2 Emissão dos resumos bibliográficos

A partir dos arquivos temporários criados na subfase anterior e do arquivo de categoria de assunto é efetuado o processamento, para emitir o relatório de referências bibliográficas. (Figura 1)

00552 Present state of technology for the production of metallic thorium.

Haydt, H.M.

(In Portuguese). Estado atual da tecnologia da produção de tório metálico. IEA/PUB-63, 1963. 11 refs. [Paper presented at the 17th Congresso Anual da Associação Brasileira de Metais. Rio de Janeiro, GB, Brasil, jul. 1962. Reprinted from Bol. Assoc. Bras. Met., 19(77): 571-8, 1963].

Because of the importance of the mineral reserves of Th in Brazil, the Comissão Nacional de Energia Nuclear has emphasized the development of thorium-plutonium reactors for energy production. The present state of technology for the production of metallic Th, beginning with thorium oxide, is summarized. The principal steps of the process-the fluorination with gaseous HF, the reduction by metallic Ca of the ThF₄ in the presence of zinc fluoride, and the subsequent steps of pre-fusion and final fusion with ingot production in electrical furnaces with consumable electrode-are described. (NSA).

THORIUM: fabrication

A geração da base de dados do Instituto de Pesquisas Energéticas e Nucleares...

5.3.3. Emissão do índice de autor

Nesta subfase é emitido um relatório contendo um índice de autor(es) com o(s) respectivo(s) título(s) do(s) trabalho(s) e o(s) número(s) do(s) resumo(s) correspondente àquele documento. (Figura 2)

MONTEIRO, P.R.B.
Single channel pulse height analyzers.
(IEA/PUB-510) – 01264

MONTEIRO, W.A.
Effect fo fast neutron irradiation on the recovery and
recrystalilization of niobium.
(IEA/PUB-488) - 00582

Irradiation effect with fast neutrons in the mechanical and structural
properties of niobium. 00659

Figura 2 – Índice de autor.

5.3.4 Emissão do índice de assunto

O arquivo contendo remissivas de assunto é processado juntamente com um dos arquivos temporários criados na subfase 4.3.1., no sentido de emitir um relatório contendo um índice de assunto ordenado alfabeticamente e estruturado hierarquicamente. (Figura 3)

SELENIUM:

activation analysis

Application of activation analysis to the determination of pollutants
in the atmosphere. 00220
(IEA/PUB-326).

Determination of some atmospheric pollutants by neutron activation
analysis. 00253

neutron activation analysis

Interfering elements in activation analysis.
Study of instrumental analysis. Possibilities for the groups: Sc-Zn,
Hg-Se and As-Sb-Br. 00299
(IEA/PUB-533).

photonuclear reactions

Systematic study on nuclear resonance scattering. 00159
(IEA/PUB-351).

Figura 3 – Índice de assunto.

6. FLUXOGRAMA DO SISTEMA

No fluxograma que se segue são apresentadas as fases do sistema, sem entrar em pormenores quanto a programas e outros aspectos específicos de computação. (Figura 4)

7. CONCLUSÃO

A criação do sistema STP atendeu plenamente aos objetivos previstos, pois a forma impressa da publicação corresponde a todas as exigências da biblioteca e a

TEREZINE ARANTES FERRAZ

base de dados é capaz de fornecer informações de interesse para a administração do Instituto, em termos de delineamento da sua política científica.

Abstract:

The generation of data bases in the Brazilian Institute of Energetic and Nuclear Research: problems and perspectives

The antecedents of the creation of the data bases generated by the Division of Scientific Information and Documentation and by the Department of Data Processing of the Institute of Energetic and Nuclear Research are described, as well as the documents which compose the data base.

Describes also, in detail, the used procedures for the introduction of data in the bases. The data bases have been operational since 1982, and their informations are accessed on line. The main objective of the bases was fulfilled: to generate an instrument capable of establishing guidelines of scientific policy, to be followed by the institution, allowing access on line, and on batch (analytical bibliography).

REFERÊNCIAS

1. BASES de dados nacionais. Guia de serviços de consulta. Ed. 1983. Rio de Janeiro, EMBRATEL/1983/ 42p.
2. CHARBEL, J.L. Os sistemas de informação e a EMBRATEL. **Rev. Bras. Biblioteconomia e Doc.** 16(3/4): 55-56, jul./dez. 1983.
3. CUNHA, M.B. da. Uso de bases de dados por países em desenvolvimento. Problemas e perspectivas. **Rev. Bibliotecon. Brasília** 12(1) :25-34, jan./jun. 1984.
4. FERRAZ, T. A. Sistema integrado de automação de processos técnicos. **Rev. Bras. Biblioteconomia e Doc.** 11(3/4) 211-20, jul./dez. 1978.
5. FERRAZ, T.A.; GUEDES, O. & SOUZA, I. J. O Sistema de processamento de dados utilizado pela Divisão de Informação e Documentação Científicas do Instituto de Pesquisas Energéticas e Nucleares. **IPEN-Inf-1**, nov. 1979.
6. GARCIA, M. L. de A. A informação científica e técnica no Brasil. **Ci. Inf.**, Brasília, 9(1/2) :41-81, 1980.
7. TERADA, E. **Uso da informação e adequação da coleção de periódicos:** estudo aplicado ao Instituto de Pesquisas Energéticas e Nucleares, São Paulo, 1984.

A geração da base de dados do Instituto de Pesquisas Energéticas e Nucleares...



Figura 4 – Fluxograma do Sistema