

MAPEAMENTO DA PRODUÇÃO CIENTÍFICA DO INSTITUTO DE PESQUISAS ENERGÉTICAS E NUCLEARES DE SÃO PAULO¹

PRISCILA NOZAKI²

(prinozaki@yahoo.com.br)

RAIMUNDO N. M. DOS SANTOS³

(rnmacedo@uol.com.br)

MERY P. Z. IGAMI⁴

(mery@ipen.br)

JOSÉ C. BRESSIANI⁵

(jbressia@ipen.br)

NAIR Y. KOBASHI⁶

(nykobash@usp.br)

Diferentemente dos estudos tradicionais sobre produção científica, baseados em artigos científicos referenciados em bases de dados internacionais, voltados para estudos de impacto, este trabalho aborda a área de Energia Nuclear, mapeando a produção científica por meio do estudo das teses e dissertações produzidas no IPEN – Instituto de Pesquisas Energéticas e Nucleares, desde 1976 até 2005, depositadas na biblioteca da instituição, e dos projetos de mestrado e doutorado iniciados entre 2001 e 2006. Utilizando técnicas bibliométricas e cientométricas, mapeou-se essa produção, identificaram-se os temas mais recorrentes e as tendências de pesquisas. A representação gráfica dos resultados permitiu uma visualização global da informação. Os indicadores produzidos são informações que servirão como subsídios para a tomada de decisões quanto à política de investimento em pesquisa da instituição. Considera-se, assim, que a metodologia utilizada obteve informações relevantes para a gestão da informação e de recursos dentro do citado instituto.

Palavras-chave: bibliometria; cientometria; gestão da informação; produção científica.

¹ Pesquisa desenvolvida por meio do projeto de iniciação científica: Institucionalização da pesquisa científica no Brasil (em andamento, período de ago./2006 a jul./2007)

² Estudante de graduação do curso de Biblioteconomia e Documentação, Universidade de São Paulo

³ Professor adjunto, Universidade Federal de Santa Catarina

⁴ Bibliotecária chefe da divisão de informação, IPEN-CNEN/SP

⁵ Diretor da área de P&D, IPEN-CNEN/SP

⁶ Professora livre-docente, Universidade de São Paulo

1. INTRODUÇÃO

Estudos tradicionais de produção científica, baseados em artigos científicos referenciados em bases de dados internacionais, voltados para estudos de impacto, são questionáveis para quantificar e qualificar a ciência efetivamente produzida por um país ou uma instituição, isto porque, como no caso do Brasil, grande parte da produção científica não está publicada nos periódicos indexados por tais bases (MUELLER, 1999). Por isso, neste trabalho propõe-se o estudo de dissertações e teses, por meio de técnicas bibliométricas e cienciométricas para melhor representar essa produção científica.

O projeto global inclui estudos nas áreas de exatas, humanas e biológicas, cada uma delas representada por um *corpus* empírico específico, para que sejam caracterizadas as diferenças de tratamento necessárias para este tipo de estudo em cada área. Inserido neste projeto global, este trabalho aborda a área de exatas, mais especificamente a área de Energia Nuclear com o mapeamento da produção científica da área, por meio do estudo das teses e dissertações produzidas pelo programa de pós-graduação do IPEN – Instituto de Pesquisas Energéticas e Nucleares, desde 1976 até 2005, bem como dos projetos de pesquisa em mestrado e doutorado iniciados entre 2001 e 2006.

Assim, tal produção foi caracterizada por meio da identificação dos temas mais recorrentes, dos docentes com maior número de orientações, das áreas de concentração dos trabalhos e das tendências de pesquisa.

A partir dos dados colhidos no mapeamento, podem-se produzir indicadores e propor parâmetros metodológicos para analisar tematicamente a produção de ciência, tais informações poderão servir como subsídios para a tomada de decisões quanto à política de investimento em pesquisa da instituição.

Os mapas da produção científica, analisados no contexto em que está inserida essa produção, proporcionam um retrato da área num determinado período, permitindo uma avaliação aprofundada dessa produção e o questionamento quanto aos fatores, internos ou externos, que possivelmente a determinaram ou influenciaram. Proporcionam também a identificação das tendências de pesquisas que, analisadas em conjunto com o retrato do que já foi produzido, podem determinar quais áreas de pesquisa devem ser priorizadas e as necessidades de alterações na alocação dos recursos.

2. ESCOLHA DO *CORPUS* DA PESQUISA

Como a ciência se desenvolve num contexto de troca, a publicação dos resultados de pesquisas é muito importante para a manutenção dessa dinâmica. Tais resultados tornam-se informações acessíveis para a comunidade científica, o que permite seu uso por outros pesquisadores em outras pesquisas.

As dissertações de mestrado e as teses de doutorado são dois típicos documentos produzidos no contexto científico; tais trabalhos constituem o processo de treinamento e aprendizagem para o desenvolvimento de pesquisas e são formas conceituadas de publicação, dado o rigor metodológico e conceitual exigido para sua produção.

Assim, dissertações e teses podem ser consideradas formas representativas da produção científica de um país ou de uma instituição, não desconsiderando outras formas de publicação, mas consagrando o ambiente institucionalizado em que se produz esse tipo de trabalho. Dessa forma, a análise desses trabalhos pode caracterizar de forma confiável um tipo de produção científica, tanto quantitativa como qualitativamente.

Essa análise pode ser feita por meio de técnicas bibliométricas e cientométricas, produzindo-se os indicadores de produção científica, que poderão ser utilizados para subsidiar a tomada de decisões quanto à política científica de uma instituição.

A bibliometria é uma disciplina que analisa a comunicação impressa e compreende a aplicação de análises estatísticas e estudos quantitativos das características de uso e produção de documentos. Já a cientometria utiliza as técnicas bibliométricas para analisar a ciência como uma disciplina ou atividade econômica e preocupa-se com a identificação de domínios de interesse, de assuntos mais pesquisados, das redes de relacionamento entre os cientistas. É uma disciplina que analisa as características da investigação científica por meio de indicadores numéricos de publicação e patentes (SPINAK, 1998).

Os indicadores de produção científica de um país ou de uma instituição permitem uma visão detalhada dessa produção e através deles pode-se observar o estado da ciência e tecnologia, proporcionando algo como um raio-x da produção científica, numa linguagem que permite a comparação da produção de um país em relação a outro ou ao restante do mundo, de instituições e também de cientistas entre si.

Para a construção de mapas da ciência e de indicadores da produção científica, primeiramente é necessário identificar os repositórios de informações, analisar sua estrutura,

qualidade e pertinência para o objetivo do estudo. Atualmente, as bases de dados configuram-se como instrumentos de análise que viabilizam abordagens bibliométricas e cientométricas.

Por isso, para a realização deste estudo, foi selecionada a base de dados da *Biblioteca Terezine Arantes Ferraz*, biblioteca do IPEN, a qual contém os registros de toda a produção científica da instituição. A representatividade do IPEN, tanto no âmbito nacional quanto internacional, fica evidenciada na caracterização do instituto apresentada a seguir, o que justifica a sua escolha.

A base de dados da biblioteca do IPEN é estruturada em padrões internacionais, semelhante à base de dados do *INIS – International Nuclear Information System*, produzida pela *International Atomic Energy Agency – IAEA*, localizada em Viena, na Áustria, e sua indexação é efetuada com o tesouro da mesma agência (no idioma inglês). Por possuir tais características, esta base foi considerada adequada para a realização desta pesquisa.

3. CARACTERIZAÇÃO DA INSTITUIÇÃO

O Instituto de Pesquisas Energéticas e Nucleares – IPEN – é uma autarquia do Governo do Estado de São Paulo, vinculada à Secretaria de Ciência, Tecnologia e Desenvolvimento Econômico. É gerenciado técnica, administrativa e financeiramente pela Comissão Nacional de Energia Nuclear – CNEN – e associado, para fins de ensino de pós-graduação, à Universidade de São Paulo, onde está instalado, ocupando uma área de 500.000m².

A missão do instituto é contribuir com a melhoria da qualidade de vida da população brasileira, produzindo conhecimentos específicos, desenvolvendo tecnologias, gerando produtos e serviços e formando recursos humanos nas áreas nuclear e correlatas.

Sua atuação destaca-se em vários setores da atividade nuclear: aplicações das radiações e radioisótopos; reatores nucleares; materiais e ciclo do combustível; materiais e prestação de serviços de valor econômico e estratégico para o país, possibilitando estender os benefícios da energia nuclear a segmentos maiores da população brasileira.

A pós-graduação em *Tecnologia Nuclear* do instituto, conduzida em associação com a USP, comemorou 30 anos em junho de 2006 e tem recebido da CAPES os melhores conceitos, o que coloca os cursos da instituição entre os melhores do país. A cada ano são produzidas em média 120 novos itens.

4. METODOLOGIA

O *corpus* empírico da pesquisa foi constituído de informações bibliográficas consolidadas em duas bases de dados.

Os primeiros estudos, para a caracterização e mapeamento da área, foram realizados com as teses e dissertações produzidas no IPEN, desde a criação do programa de pós-graduação, em 1976, até 2005.

Após a avaliação dos primeiros resultados deste mapeamento, percebeu-se a necessidade de um estudo que permitisse a análise das tendências de pesquisa, por isso foi construída uma base de dados com informações dos projetos de mestrado e doutorado, iniciados entre 2001 e 2006, possibilitando, dessa forma, uma projeção da produção científica do instituto para anos seguintes.

Foram utilizados para os estudos dois *softwares*, em versão *DOS*, quais sejam: o *DATAVIEW*, *software* bibliométrico elaborado pelo *Centre de Recherche Rétrospective de Marseille (CRRM) da Universidade Aix-Marseille III*, na França, é utilizado para a produção de listas de frequência e matrizes de co-ocorrência sob a forma de tabelas e gráficos; e o *INFOTRANS*, *software* alemão da empresa IuK (Information und Kommunikation), utilizado para a transformação dos registros bibliográficos em registros bibliométricos que também permite a padronização automática de campos, complementação de campos com dados externos e reorganização geral da base de dados para a eliminação de registros duplicados (KOBASHI e SANTOS, 2006).

Os *softwares* mostraram-se úteis não somente para a preparação, correção dos dados e para as análises bibliométricas propriamente ditas, mas também para a avaliação da consistência dos dados inseridos nas bases.

4.1 A BASE DE DISSERTAÇÕES E TESES

A base de dados da biblioteca estava constituída em *WinIsis*, tal sistema foi considerado adequado devido à sua acessibilidade, facilidade no manuseio dos dados e na transposição dos mesmos para outros formatos, o que é necessário para a manipulação nos *softwares* de tratamento bibliométrico.

Dentre os registros que a compõem foram selecionados, a partir do campo “Tipo de documento”, somente as dissertações e teses, e com estes registros criou-se uma outra base de

dados, idêntica à base original, para servir unicamente aos fins deste estudo, com um total de 1190 dissertações e teses.

Para fins de análise temática, foram inseridos na base de dados, dois novos campos: categoria de assunto e área de concentração. As categorias de assunto foram identificadas por meio de pesquisa na base de dados do *INIS*, a busca foi efetuada pelo nome do autor e título do trabalho. As mesmas categorias atribuídas aos trabalhos na referida base foram adotadas na base de dados deste estudo; as áreas de concentração foram fornecidas pela comissão de pós-graduação, dessa forma, os dados coletados para os novos campos foram inseridos em todos os registros da base de dados.

Com todos os dados inseridos na base, emitiu-se um relatório em arquivo de texto e, utilizando esse arquivo, passou-se à análise de consistência destes dados com os *softwares*.

Por meio do *software DATAVIEW* foram emitidas listas, em ordem alfabética, dos conteúdos de cada campo selecionado para a análise: orientador, tipo de documento, local de defesa, ano de defesa, descritores de assunto, categorias de assunto e área de concentração. Os outros campos da base, como o número de localização no acervo ou dados complementares do documento, não foram analisados por não serem relevantes para os fins deste estudo.

Nestas listas, comparando-se os itens de cada campo, foram detectadas inconsistências nos seguintes campos: “orientador”, “local da defesa” e “tipo de documento”. A base continha diferentes formas de entrada para o nome de um mesmo orientador, instituição de ensino e para um mesmo tipo de documento, o que inviabilizava os estudos bibliométricos. Os campos “ano de publicação”, “categoria de assunto” e “área de concentração” não apresentaram problemas e o campo “descritores” continha um número irrelevante de inconsistências.

Para a correção dos erros foi utilizado o *INFOTRANS*, o qual possibilita a localização de uma seqüência de caracteres em um determinado campo e sua substituição. Por exemplo, no campo “tipo de documento” foram encontradas entradas como: Tese, Tese de doutoramento, Doutorado. Analisando-se tais entradas, optou-se pela adoção dos seguintes padrões para este campo: Tese (Doutoramento) e Dissertação (Mestrado). Assim, utilizando o *software INFOTRANS*, estes erros foram localizados e substituídos pelas entradas adotadas.

Os outros campos, onde também foram constatadas inconsistências, foram corrigidos da mesma forma: identificaram-se os erros por meio das listas emitidas pelo *DATAVIEW*, adotou-se um padrão de entrada e os dados incorretos foram substituídos pelos adotados, via *INFOTRANS*.

Este tipo de análise de consistência mostrou-se útil também para corrigir a base original da biblioteca, o que, por sua vez, possibilita atingir uma melhoria na função de busca e recuperação de dados na referida base.

Após esse tratamento, já com os dados padronizados, passou-se novamente à utilização do *DATAVIEW*, desta vez para a emissão de listas de frequência para as análises bibliométricas. Foram elaboradas as seguintes listas: *a)* assuntos pesquisados; *b)* distribuição dos principais assuntos por quinquênios; *c)* orientadores dos trabalhos; *d)* descritores de assunto; *e)* áreas de concentração; *f)* distribuição dos tipos de trabalho por quinquênios.

4.2 A BASE DE DADOS DE PROJETOS

Para a identificação dos temas que estão sendo estudados, idealizou-se a uma base de dados para a análise dos projetos de mestrado e doutorado iniciados no IPEN, no período de 2001 a 2006, totalizando 405 trabalhos.

Esta base, como a base de dissertações e teses, também foi construída em *WinIsis* e os campos selecionados para integrá-la foram: autor, título, categoria de assunto, data de início, área de concentração, tipo de trabalho, orientador, plano diretor e descritores de assunto.

Os dados foram fornecidos pela comissão de pós-graduação do instituto, no entanto, os projetos não possuem categorização e indexação, por isso cada trabalho passou pela análise das bibliotecárias para sua categorização e indexação.

A metodologia de análise dos dados foi similar àquela utilizada na base de dissertações e teses, com a diferença de que, por se tratar de uma quantidade muito menor de dados que foram registrados exclusivamente para este estudo, a etapa de análise de consistência e correção dos dados foi realizada num espaço de tempo bem mais curto.

Foram necessárias algumas alterações no campo “data de início”. Inicialmente neste campo foi inserida a data exata de início do projeto, entretanto, na análise de consistência observou-se uma dispersão muito grande destes dados. Portanto, para a análise bibliométrica, concluiu-se que seria mais interessante utilizar apenas o ano, para a obtenção de resultados mais representativos.

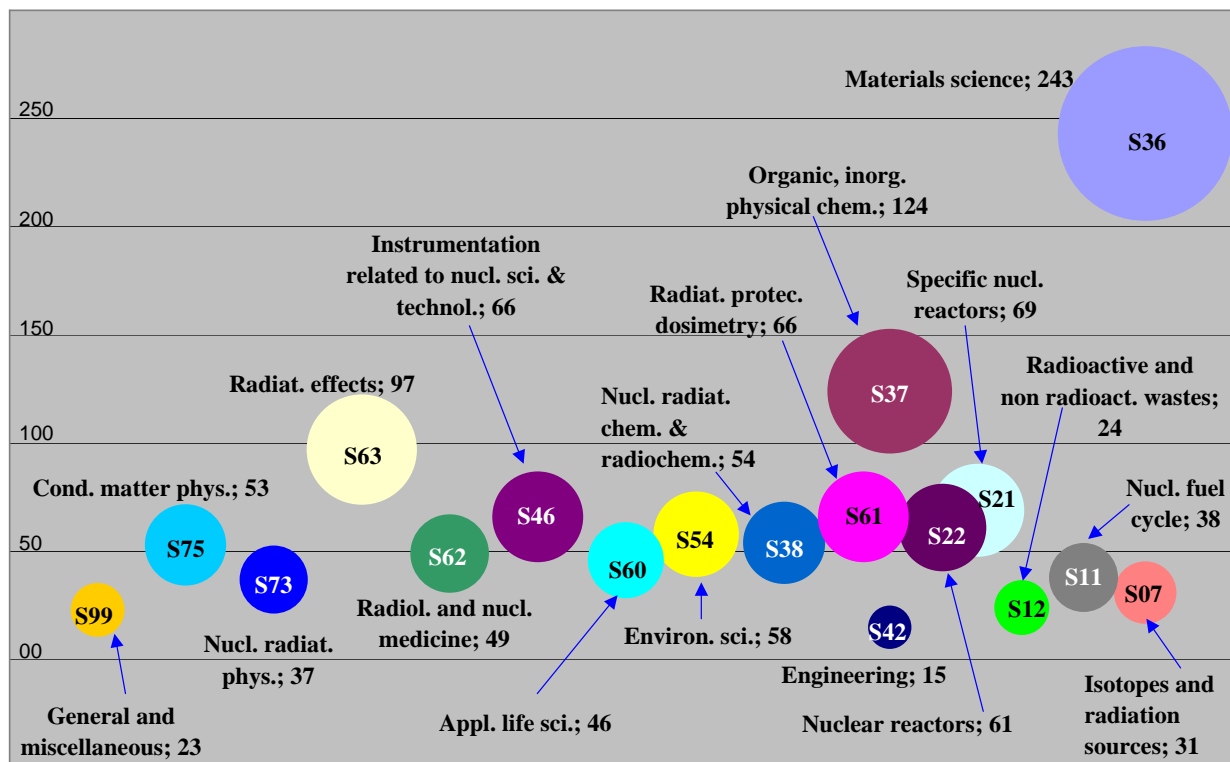
Procedeu-se, então, a substituição automática dos dados do campo por meio do *INFOTRANS*, assim a data, que estava no formato: ano, mês e dia (AAMMDD), foi substituída apenas pelo ano (AA).

Após estes procedimentos, com o uso do *DATAVIEW*, foram emitidas as seguintes listas de frequência: *a*) assuntos pesquisados (geral); *b*) assuntos pesquisados (por ano); *c*) áreas de concentração; *d*) orientadores dos trabalhos.

5. RESULTADOS

As listas emitidas das duas bases de dados foram, em seguida, exportadas para o *software EXCEL*, no qual foram construídos os gráficos necessários para a melhor visualização dos resultados.

Cada campo das bases, em particular, permite a análise de diferentes aspectos da produção científica do instituto, a seguir são apresentados alguns dos resultados obtidos neste estudo.

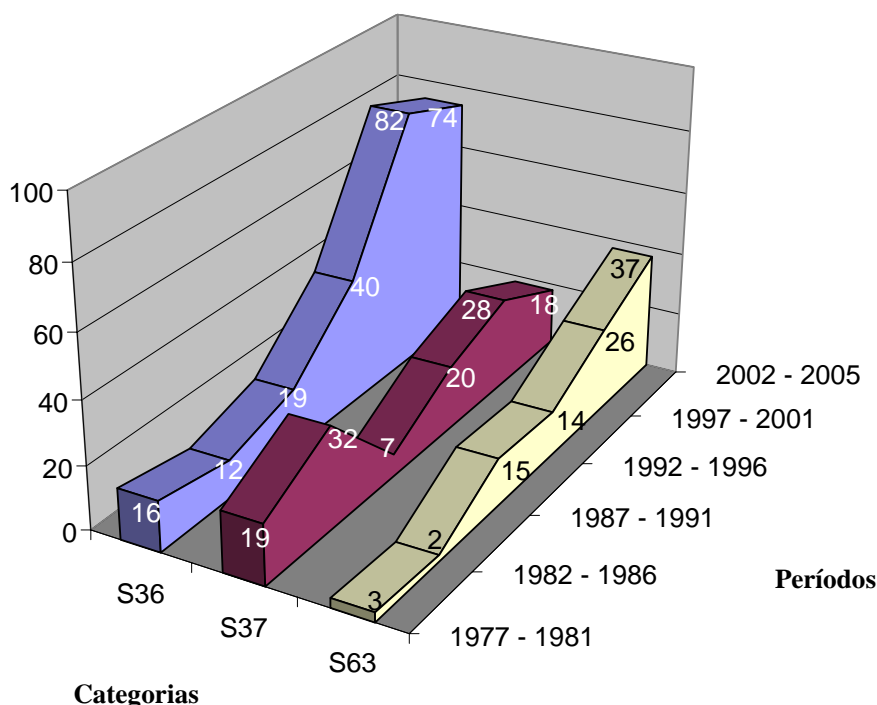


Gráf. 1 – Ranking dos assuntos mais pesquisados

A partir da análise do campo “Categoria de assunto”, da base de teses e dissertações, obteve-se uma lista com as 33 categorias de assunto contempladas no decorrer dos trinta anos da pós-graduação do IPEN. Com tais dados, foi construído o gráfico 1 com os dezoito assuntos mais pesquisados (SANTOS, et al), onde o tamanho de cada círculo e sua posição relativa no gráfico correspondem à quantidade de trabalhos realizados na referida categoria,

cujo número exato está indicado após a denominação da categoria. As denominações das categorias, bem como seus respectivos códigos (indicados no centro dos círculos), foram mantidos nos padrões da Lista de categoria de assuntos do INIS. (INIS, 2002).

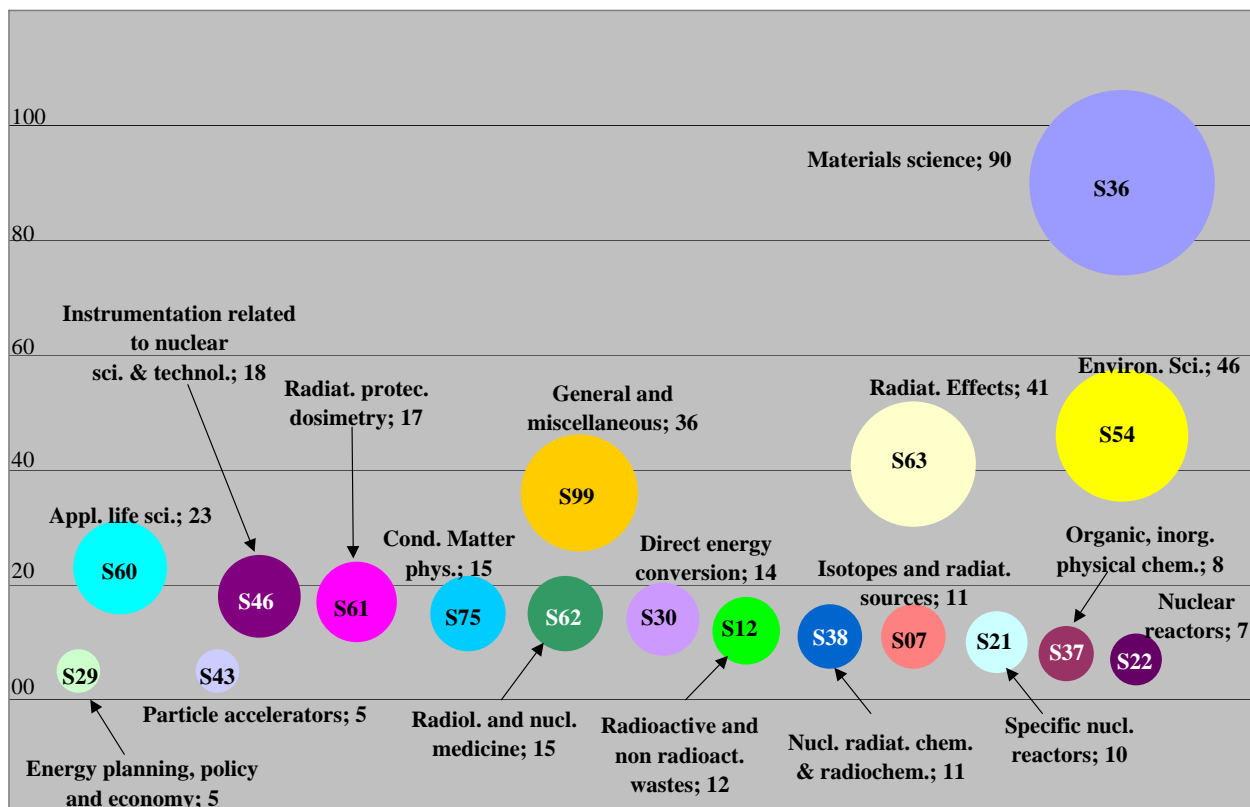
Estes resultados demonstram, em nível de assunto, as características da produção científica do instituto, o que auxilia na identificação de quais foram os temas que estiveram em foco, quais foram ou quais são as frentes de pesquisa, pode-se notar claramente a predominância de três categorias: Materials science (S36), Organic, inorganic physical chemistry (S37) e Radiation effects (S63), que somam 464 trabalhos correspondendo a 38% do total. O gráfico a seguir apresenta a evolução destas categorias, com os totais de trabalhos por quinquênio.



Gráf. 2 – Evolução das categorias líderes

Tais informações, correlacionadas aos cenários políticos, sociais e econômicos, transformam-se em informações estratégicas para a direção do instituto na tomada de decisões relacionadas às políticas de pesquisa.

A análise do mesmo campo na base de projetos, demonstrada no gráfico 3, possibilitou uma visão dos assuntos que estão sendo pesquisados e, dessa forma, uma previsão da produção científica para os próximos anos.



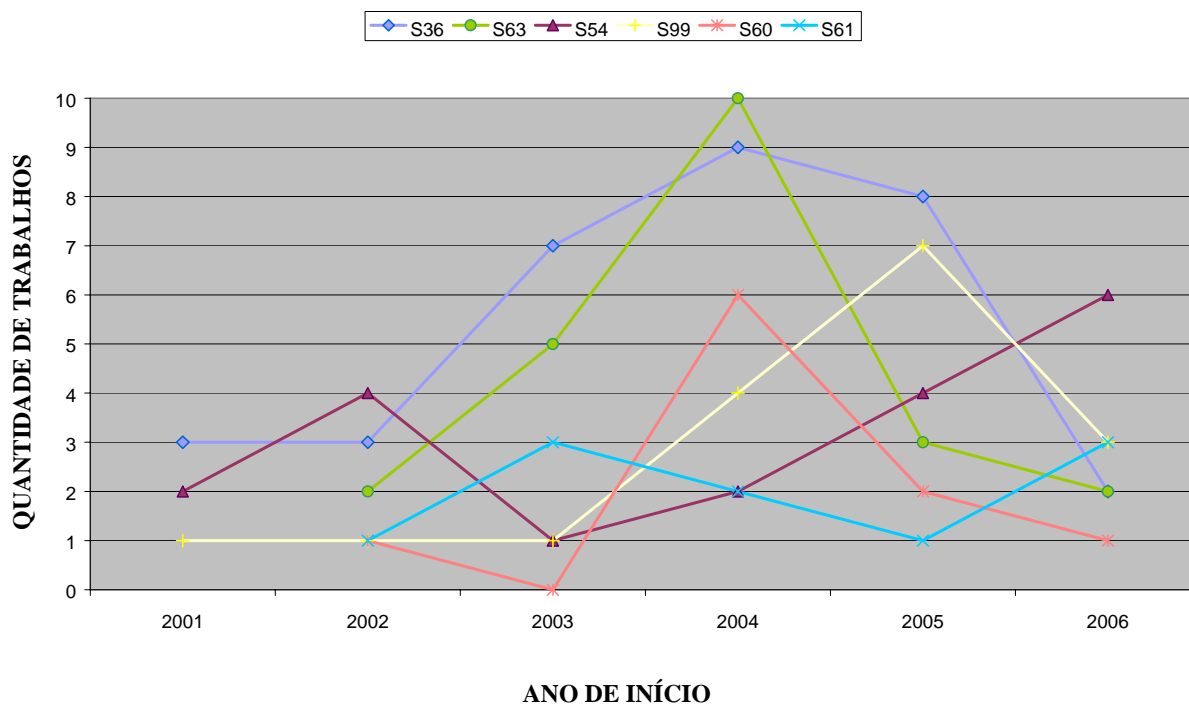
Gráf. 3 – Ranking dos assuntos mais pesquisados nos projetos

Este gráfico apresenta os dezoito assuntos mais pesquisados, suas informações podem ser comparadas com aquelas contidas no gráfico 1, permitindo uma análise das mudanças nos temas das pesquisas, dos novos assuntos que estão surgindo e daqueles que estão sendo menos abordados na atualidade.

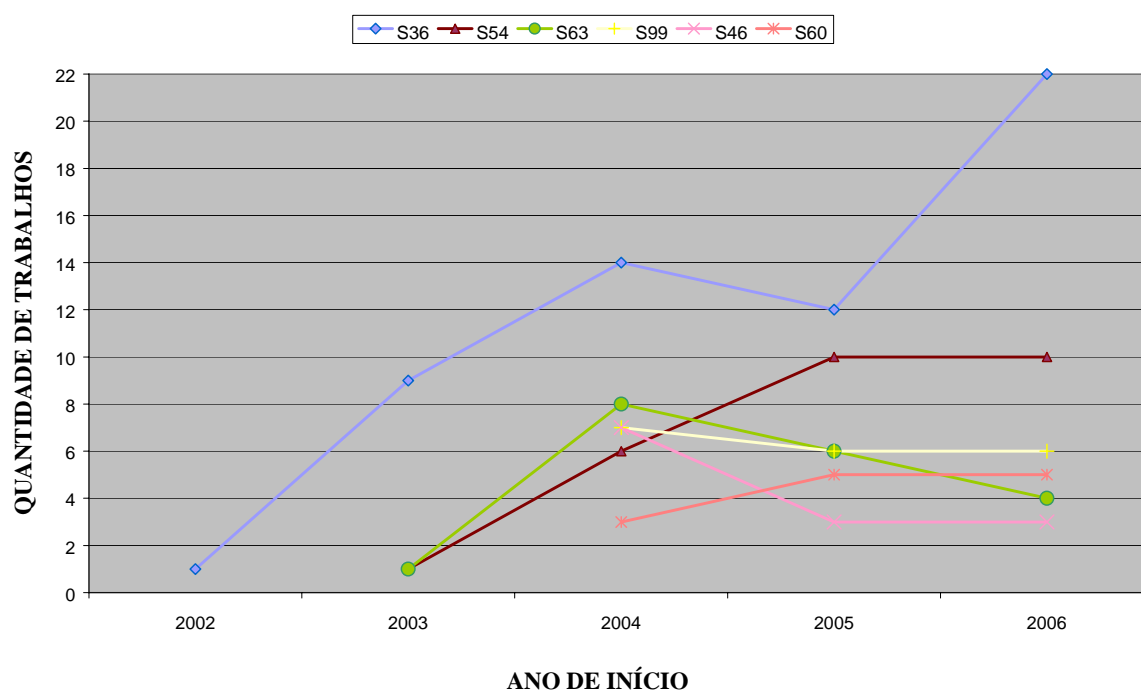
Esta forma de apresentação mostra a distribuição geral dos projetos por assunto, sem a diferenciação de projetos de mestrado e doutorado; na análise deste gráfico pelo especialista da área, detectou-se a necessidade de apresentar-se dados mais específicos: o tipo de trabalho, a categoria de assunto e o ano de início, dessa forma seria possível obter-se uma projeção mais detalhada desta produção científica.

Por este motivo, foram construídos os seguintes gráficos, contendo a quantidade de trabalhos por ano de início, das seis categorias de assunto mais pesquisadas, diferenciando-se os projetos de mestrado e de doutorado. Os resultados apresentaram características não evidenciadas no gráfico 3. Por exemplo, a categoria de assunto S36 (Materials science) está na primeira colocação na distribuição geral dos projetos, entretanto, quando as informações são analisadas separadamente, constata-se que, em quantidade de projetos, a categoria mantém-se em primeiro lugar, tanto em mestrado quanto em doutorado, porém, na evolução

dos projetos de doutorado, nota-se sua queda no período de 2004 a 2006, quando está em ascensão a categoria S54 (Environmental sciences).



Gráf. 4 – Evolução dos projetos de doutorado



Gráf. 5 – Evolução dos projetos de mestrado

Os gráficos 4 e 5 revelam que as tendências de pesquisa são diferentes para o mestrado e o doutorado, as informações neles contidas, interpretadas pelos especialistas do Instituto, constituem mais informações estratégicas para a definição de políticas de pesquisa e alocação de recursos.

Cabe esclarecer que as informações referentes aos projetos de doutorado iniciam-se em 2001, enquanto as dos projetos de mestrado iniciam-se em 2002, essa diferença ocorreu pois no levantamento dos dados, em 2006, foram inseridos na base para os estudos todos os projetos em andamento, quando ainda haviam projetos em aberto de doutorado, mas não de mestrado.

6. CONSIDERAÇÕES FINAIS

Os resultados do estudo permitiram uma visão global da produção científica do Instituto, com um detalhamento em nível de assunto que forneceu informações que podem ser utilizadas tanto na avaliação desta produção, quanto no planejamento e definição de políticas de pesquisa da instituição.

A análise e validação dos resultados pelo especialista da área foi imprescindível para o estudo, uma vez que nesse contato surgiram novas possibilidades de abordagem dos dados colhidos, com ênfase nas informações mais relevantes, segundo a visão de quem está inserido na área de Energia Nuclear e, por isso possui conhecimentos específicos importantes para a interpretação dos resultados de estudos desta natureza.

BIBLIOGRAFIA

CALLON, Michel; COURTIAL, Jean-Pierre; PENAN, Hervé. **Cienciometria**: la medición de la actividad científica: de la bibliometria a la vigilancia tecnológica. Gijón: Trea, 1995.

INSTITUTO DE PESQUISAS ENERGETICAS E NUCLEARES. **Plano diretor, 2004**. São Paulo: Divisão de Planejamento e Inovação tecnológica, IPEN/CNEN-SP, 2000-2005.

INTERNATIONAL NUCLEAR INFORMATION SYSTEM – INIS. **Subject categories and scope descriptions**. Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2002. (ETDE/INIS Joint reference series nº 2).

KOBASHI, Nair Y.; SANTOS, Raimundo N. M. dos. Institucionalização da pesquisa científica no Brasil: cartografia temática e de meios sociais por meio de técnicas bibliométricas. **Transinformação**, Campinas, v. 18, n. 1, p. 27-36, jan./abr., 2006.

MACIAS-CHAPULA, Cesar A. O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. **Ciência da informação**. Brasília, v.27, n. 2, p. 134-140, maio/ago. 1998.

MUELLER, Suzana P.M. O círculo vicioso que prende os periódicos nacionais. **DataGramZero - Revista de Ciência da Informação**, n. zero, dez. 1999, artigo 04. Disponível em: <http://www.dgz.org.br/dez99/F_I_art.htm>. Acesso em: 24 abr. 2007.

SANTOS, Raimundo N.M. dos; KOBASHI, Nair Y.; IGAMI, Mery P. Z.; BRESSIANI, José C. Brazil's insertion into nuclear energy research: bibliometric analysis of dissertations developed at the Instituto de Pesquisas Nucleares- São Paulo. In: GUERREIRO-BOTE, Vicente P. (Ed.). Current research in information sciences and technologies. Multidisciplinary approaches to global information systems. International conference on multidisciplinary information sciences and technologies, 1st, october 25-28th, 2006, Merida Spain. **Proceedings...** Badajoz, España: Open Institute of Knowledge, 2006. v.1, p.553-556.

TRZESNIAK, Piotr. Indicadores quantitativos: reflexões que antecedem seu estabelecimento. **Ciência da informação**. Brasília, v.27, n. 2, p. 128-133, maio/ago. 1998.

VANTI, Nadia A.P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**. Brasília, v. 31, n. 2, p. 152-162, maio/ago. 2002.

WORMELL, Irene. Informetria: explorando bases de dados como instrumentos de análise. **Ciência da informação**. Brasília, v. 27, n. 2, p. 210-216, maio/ago. 1998.