

BASES DE DADOS: FONTES PARA PRODUÇÃO DE INDICADORES DO TRABALHO IMATERIAL

Nair Yumiko Kobashi

ECA-USP

NYKOBASH@USP.BR

NAIRKOBASHI@GMAIL.COM

TRABALHO IMATERIAL

Economia informacional:

as inscrições resultantes da produção da pesquisa científica e tecnológica são expressões do trabalho imaterial
(Hardt e Negri, 2002)

(M.Nélida González de Gomes)

INSCRIÇÕES/REPRESENTAÇÃO DOCUMENTÁRIAS

Para garantia de EFETIVO uso social,
devem ser ADEQUADAMENTE
identificadas e tratadas (organizadas
documentariamente)

- DISCIPLINAS relacionadas:
- Representação descritiva
- Organização da informação e do conhecimento

ORGANIZAR INFORMAÇÃO/CONHECIMENTO

- PRODUÇÃO DE REPOSITÓRIOS (EXOMEMÓRIAS)
- PRÓTESES que permitem reutilizar conhecimento produzido para produzir novos conhecimentos, tomar decisão, etc...
- REPOSITÓRIOS = CONHECIMENTOS FUNCIONALIZADOS COMO INFORMAÇÃO

REDES ELETRÔNICAS

- LOCALIZAÇÃO DE CONTEÚDOS PERTINENTES LARGAMENTE DEPENDENTES DE TRATAMENTO SEMÂNTICO (Catalogação/Indexação)
- SIST. DE REC. CONTÊM:
 - - INFORMAÇÕES REFERENCIAIS
 - - SUBSTITUTOS, REPRESENTAÇÕES
 - (com links ou não a texto integral)

REPRESENTAÇÕES DOCUMENTÁRIAS

- SÃO OBJETOS SIMBÓLICOS
- SE EXPRESSAM PELA LINGUAGEM
- NECESSÁRIO DISPOR DE PRINCÍPIOS, TEORIAS, MÉTODOS PARA FABRICÁ-LAS.

ALGUNS PRESSUPOSTOS

- -SEM LINGUAGEM NÃO HÁ COMUNICAÇÃO
FLUXO/RECEPÇÃO DE INFORMAÇÃO
- -INFORMAÇÃO PARTICIPA DE DIFERENTES
ESTRUTURAS DE SIGNIFICAÇÃO
- -ESTRUTURA E LINGUAGEM DEVEM SER
ADERENTES AOS CONTEXTOS E PÚBLICOS

SISTEMAS DE RECUPERAÇÃO D INFORMAÇÃO

- - NÃO SÃO ARTEFATOS NEUTROS (são institucionalizados)
- - DEVEM TER EQUILÍBRIO DINÂMICO (são SISTEMAS ABERTOS)
- - DEVEM SER PENSADOS NO PARADIGMA DA COMUNICABILIDADE
- - SIGNIFICADO DAS INFORMAÇÕES É ALCANÇADO INTERSUBJETIVAMENTE.

(recuperação)

É ESCAVAÇÃO DE REPOSITÓRIOS

- BUSCA, ANÁLISE E INTERPRETAÇÃO DE INSCRIÇÕES
- ARQUEOLOGIA DO SABER

A escavação dos "arquivos", discursos efetivamente pronunciados, permitem conhecer o passado, mas também possibilitam o surgimento de novos discursos (Foucault, 2000).

PRODUÇÃO DE INDICADORES D C&T

- SUPÕE ESCAVAÇÃO DAS EXOMEMÓRIAS, SEGUNDO HIPÓTESES DE TRABALHO
- PRESSUPOSTO ASSUMIDO em nossas pesquisas atuais (Grupo Scientia - cadastrado no CNPq)
- Estudos de produção científica tradicionais, realizados com artigos científicos referenciados e bases de dados internacionais, são questionáveis quanto à sua validade para mapear a ciência efetivamente produzida em cada país.

Projeto global: Institucionalização da pesquisa científica no Brasil

Grupo de pesquisas **SCIENTIA**
(<http://www.scientia.iv.fapesp.br>)

Objetivo: buscar alternativas teóricas e metodológicas para mapear a produção científica brasileira e produzir indicadores

Fundamentos: Estudos sociais da ciência
Organização da informação; Bibliometria

PROBLEMAS TEÓRICOS E METODOLÓGICOS

- Que referenciais teóricos poderão garantir:
 - a formulação de hipóteses ?
 - as escolhas metodológicas ?
 - a interpretação dos dados ?
- Que tipos de dados empíricos são adequados para produzir indicadores confiáveis de C&T?

REFERENCIAIS TEÓRICO

Institucionalização da pesquisa científica e tecnológica

Institucionalização cognitiva:

- aspectos epistemológicos, teóricos e metodológicos das disciplinas científicas

Conceitos da área:

- » legitimidade e pertinência dos problemas propostos ao campo
- » aceitabilidade das soluções encontradas
- » adequação dos métodos, técnicas e instrumentos de observação fenômenos

Institucionalização social:

- estruturas formais que demarcam os membros de uma comunidade científica
 - » grau de organização de uma área do conhecimento
 - » sua integração às estruturas de legitimação
 - » condições de acesso a programas de fomento

(Whitley, 1974, 1980)

(Bibliometria tradicional; Solla Price, Merton, GIBBONS, História da Ciência, Sociologia da Ciência, Arqueologia do saber, etc. etc.)

OPÇÕES METODOLÓGICA

- Base empírica:
 - bases de dados de
 - - dissertações e teses de programas de p
graduação do país
 - - produção técnico-científica das instituiç
- Método de abordagem:
 - técnicas bibliométricas
- Tratamento dos dados:
 - Seleção, padronização de forma e de

DESAFIOS

Na abordagem bibliométrica:

- Operacionalização das relações dicotômicas entre quantidade e qualidade.
- Fundamento da operacionalização realizada: relações entre conhecimento quantitativo e conhecimento qualitativo (Granger, 1989)
- - **CATEGORIZAÇÃO** (Subsunção de dados específicos a

DESAFIOS

Na organização e representação da informação:

- estabelecer conjuntos rigorosos
- "semelhanças de família" (Wittgenstein),
- conhecimento quantitativo e qualitativo (Granger, 1989)

DESAFIOS

Precariedade da infra-estrutura brasileira de informação

Inconsistência dos dados bibliográficos e temáticos

- Elevada dispersão de dados (bibliográficos e temáticos)
- Problemas para obtenção de clusters:
 - temas emergentes
 - cooperação

Bases bibliográficas

- USO CONSAGRADO EM ANÁLISE DE PRODUÇÃO CIENTÍFICA
- CONFIABILIDADE DOS RESULTADOS DEPENDENTE DE
 - ESTRUTURA ADEQUADA DE METADADOS
 - EXAUSTIVIDADE DE COBERTURA
 - CONSISTÊNCIA DOS REGISTROS

Problemas para seleção de fontes

- bases de dados internacionais de artigos, representam parcialmente a produção científica dos países porque voltados para estudos de impacto.

Dissertações e teses

- produção com rigor metodológico e conceitual
- avaliados em instâncias reconhecidas



Estudos sociais da *ciência*: *Bibliometria*, cientometria, informetria, etc.

permitem identificar quantitativa e qualitativamente a atividade científica e tecnológica:

- de países;
- de instituições;
- de áreas do conhecimento;

EXEMPLO DE ESTUDO

PRODUÇÃO DO PPG-IPEN

- **Base de dados da Biblioteca do Ipen**
- **Repositório:** toda a produção científica do Instituto.
- Estruturada de acordo com padrões internacionais

Corpus

Informações bibliográficas consolidadas e duas bases de dados:

- Base de dissertações e teses, com 1190 registros (1976-2005);
- Base de projetos, com 405 registros (iniciados entre 2001 e 2006).

Análise dos dados

Campos analisados:

- orientador, tipo de documento, local de defesa, ano de defesa, descritores de assunto, categorias de assunto e área de concentração.

Softwares utilizados:

- *Infotrans e Dataview*

Etapas da análise

- Análise de consistência e correção dos dados;
- Reformatação para análise bibliométrica;
- Análise dos resultados;
- Validação por especialista da área de Energia Nuclear.

Inconsistências detectadas

diferentes formas de entrada de

- **nome do orientador**
- **tipo de documento**
- **nome da instituição**

INCONSISTÊNCIAS E INCOERÊNCIAS EXCLUSIVAS DA BASE DE DADOS DO IPEN???

Inconsistências encontradas
causaram surpresa porque

- catalogadores e indexadores recebem treinamento
- a inserção de dados é cuidadosa e supervisionada

Reformatação da estrutura de dados (metadados)

Criação de novos campos

- Categorias de assuntos
- Linhas de pesquisa

REFORMATAÇÃO DE DADOS

Perspectiva teórico-pragmática:

- Neutralização das diferenças específicas dos elementos do corpus, promotoras da dispersão
- Integração de dados: categorias adequadas às finalidades do estudo.
- Adoção de metacategorias de distribuição dos dados.

Correção de erros de form

- SOFTWARE DE EMISSÃO DE LISTAGENS (DATAVIEW)
- SOFTWARE DE PADRONIZAÇÃO DE DADOS, COMPLEMENTAÇÃO DE CAMPOS, ELIMINAÇÃO DE DUPLICAÇÕES (Infotrans)
- Entradas encontradas
 - Tese, Tese de doutoramento, Doutora

Emissão de listas de frequência

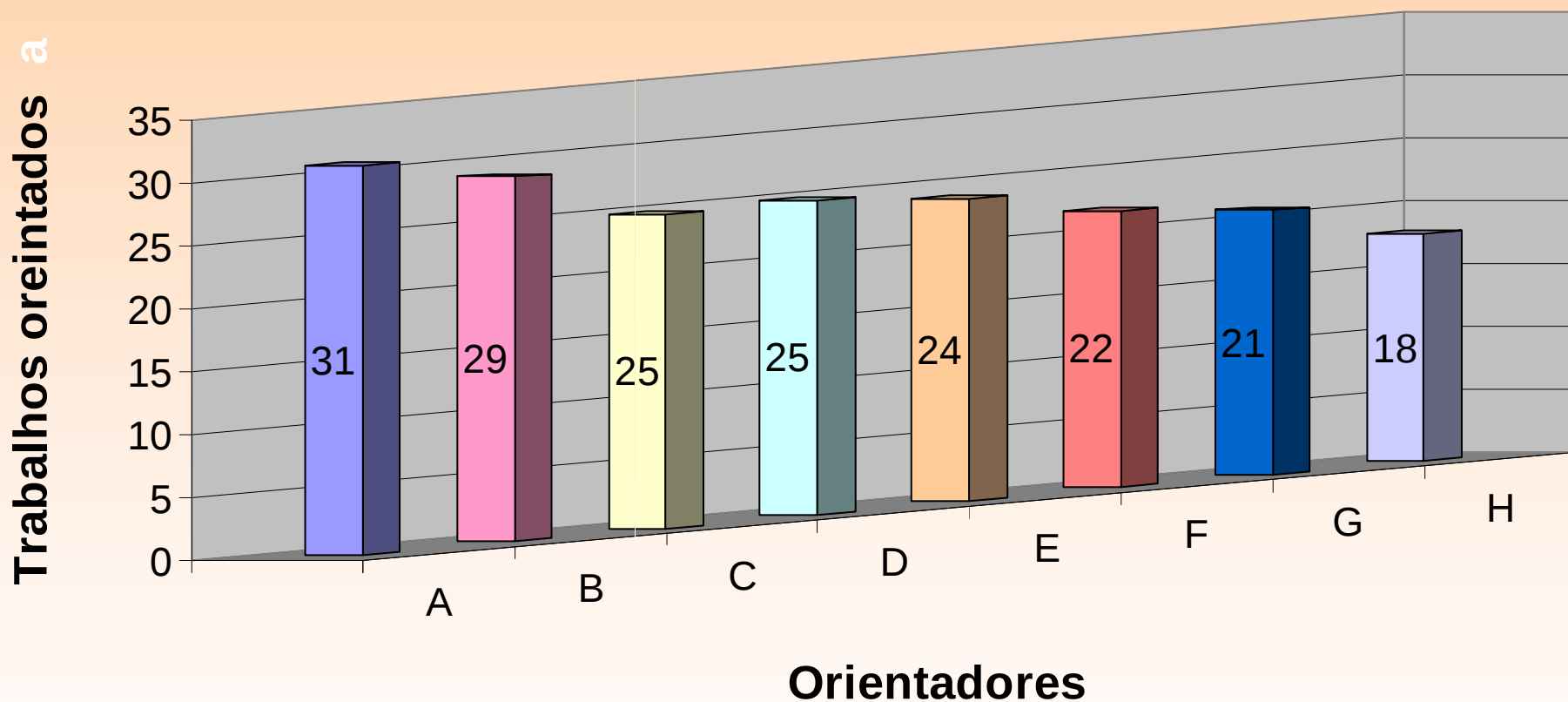
Orientadores

Linhas de pesquisa

Assuntos

Distribuição dos assuntos por quinquênios

ORIENTADORES

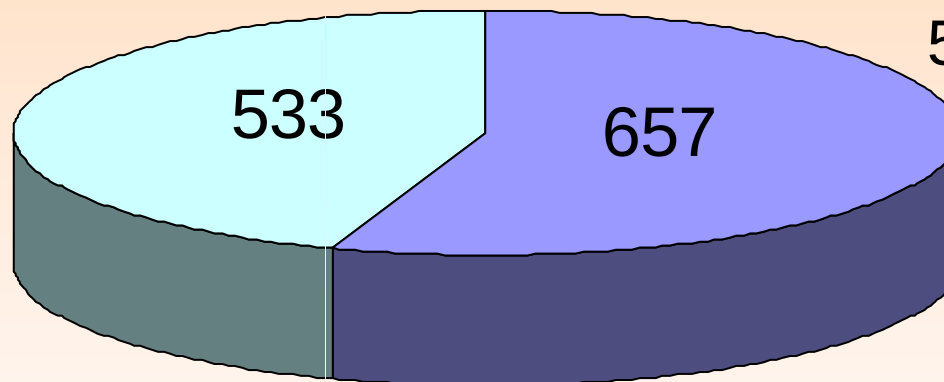


Graf. 1 - Docentes com maior número de orientações

Trabalhos por área

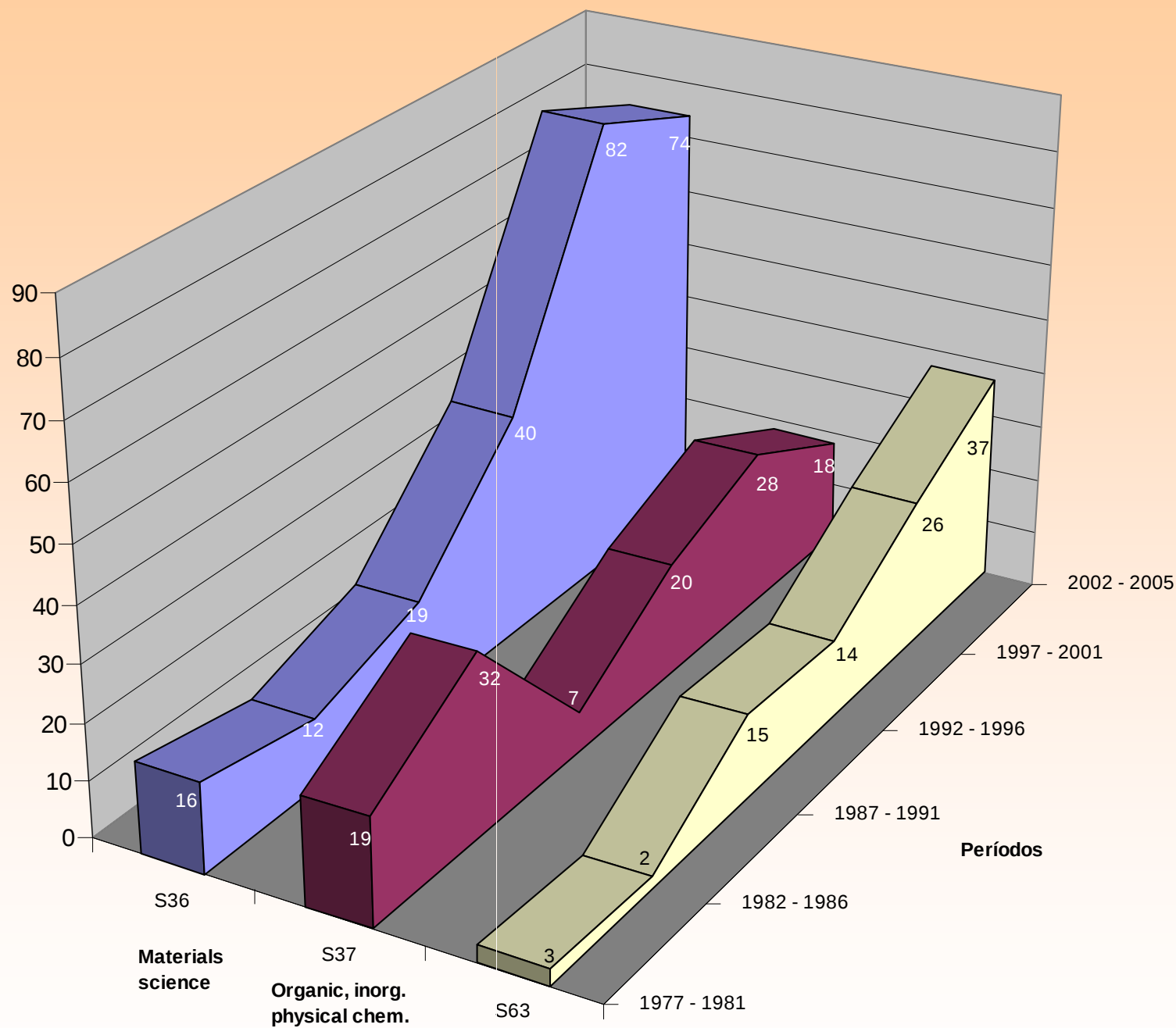
Reat. nucl. e
tecnol. do
combustível
nucl.
45%

Tecnol. nucl.
básica
55%

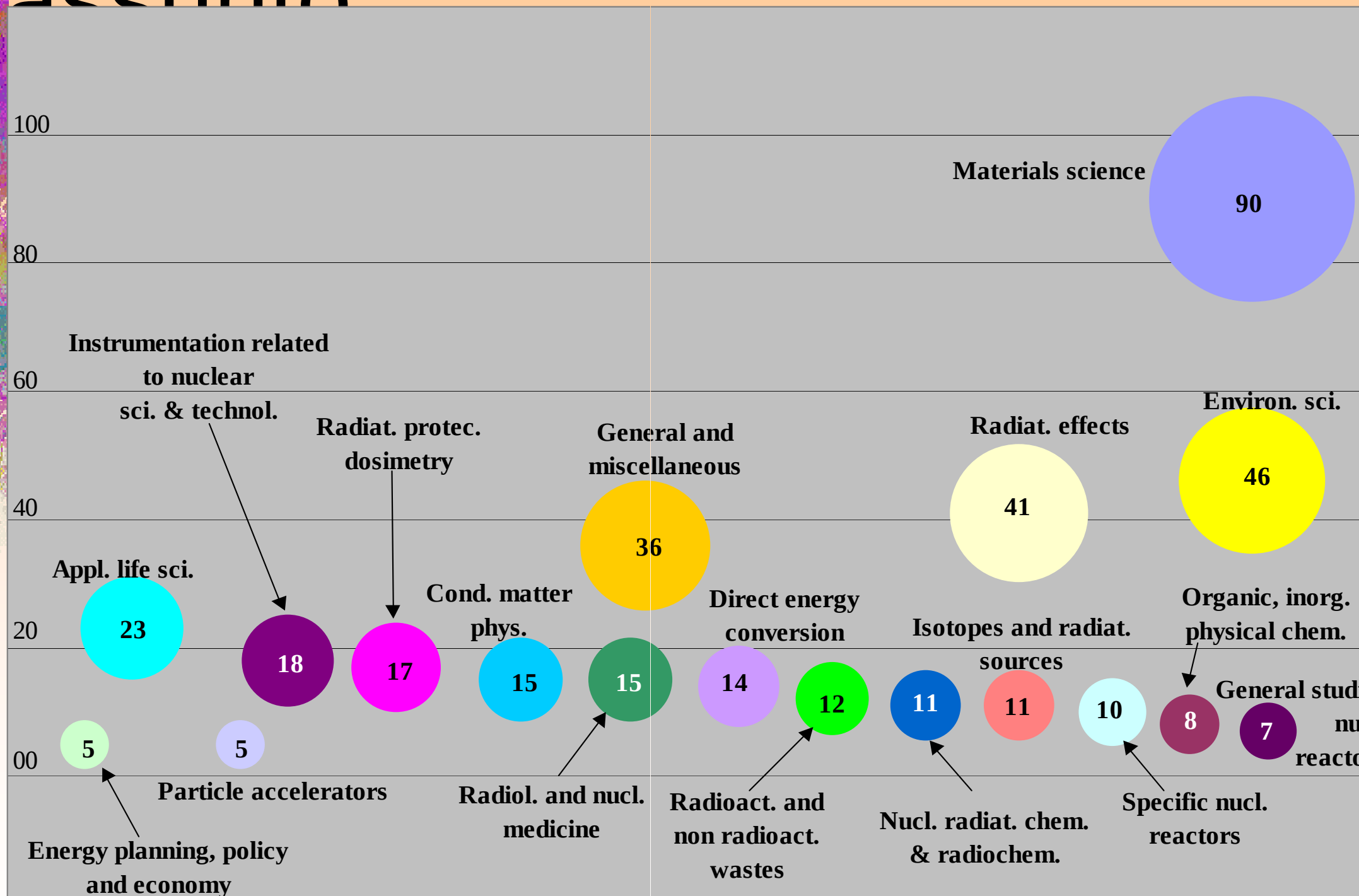


Graf. 2 - Distribuição dos trabalhos por área de concentração

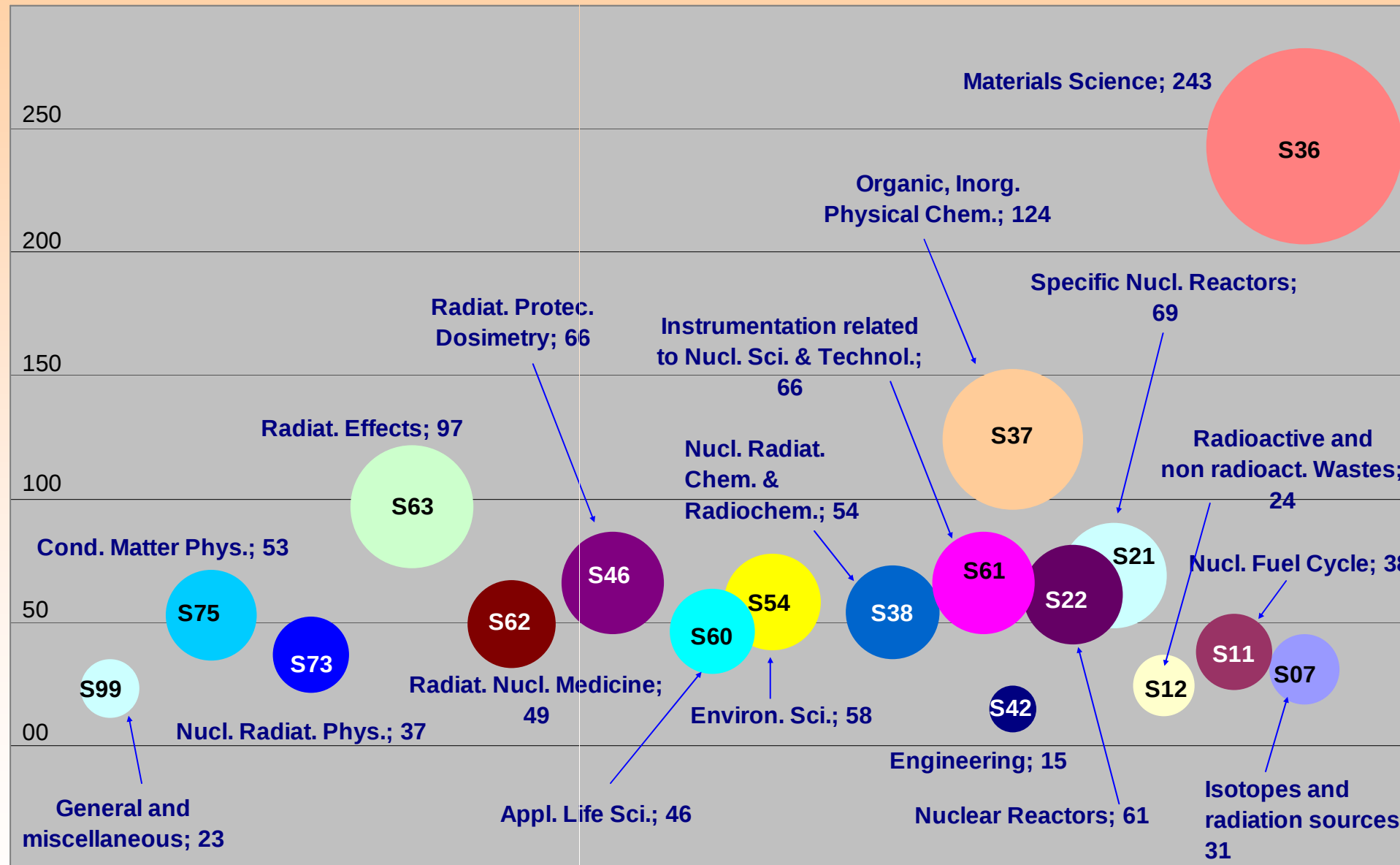
Análise do campo categoria do assunto”



Analise do campo categoria do assunto



ASSUNTOS MAIS FREQUENTES

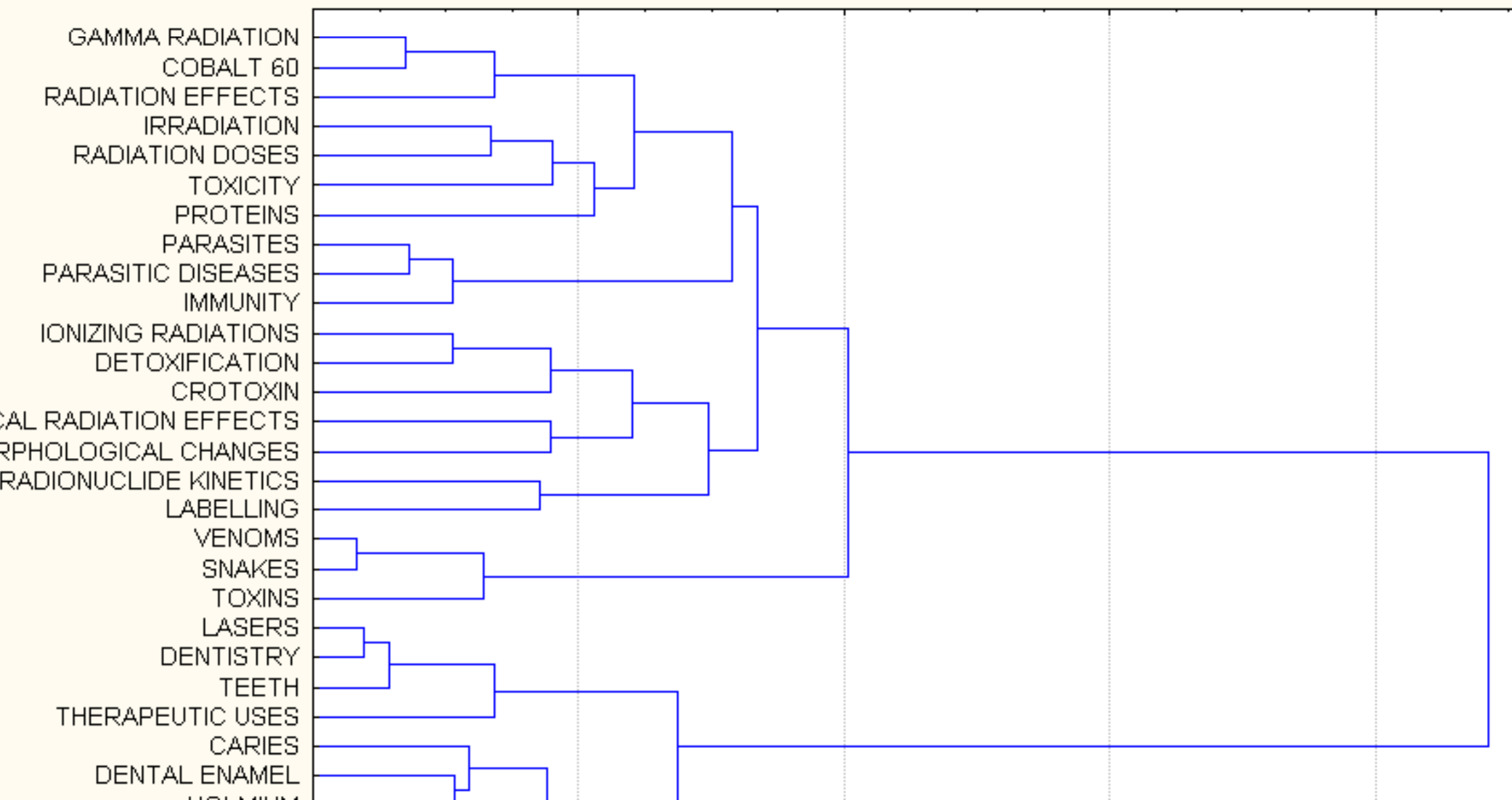


Categoria S63 - "Radiation - thermal effects - biological, material and Living organisms"

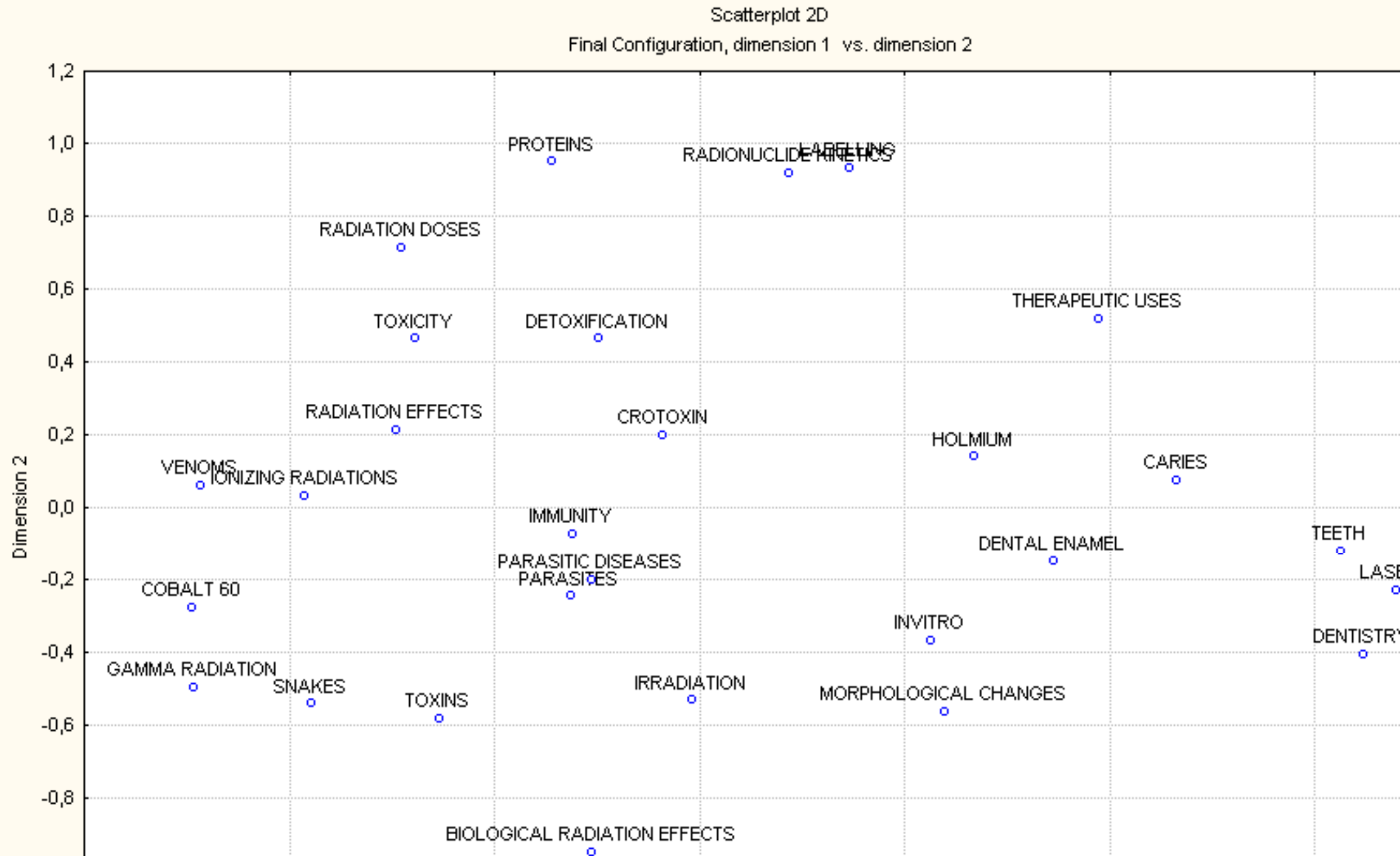
Tree Diagram for Variables

Ward's method

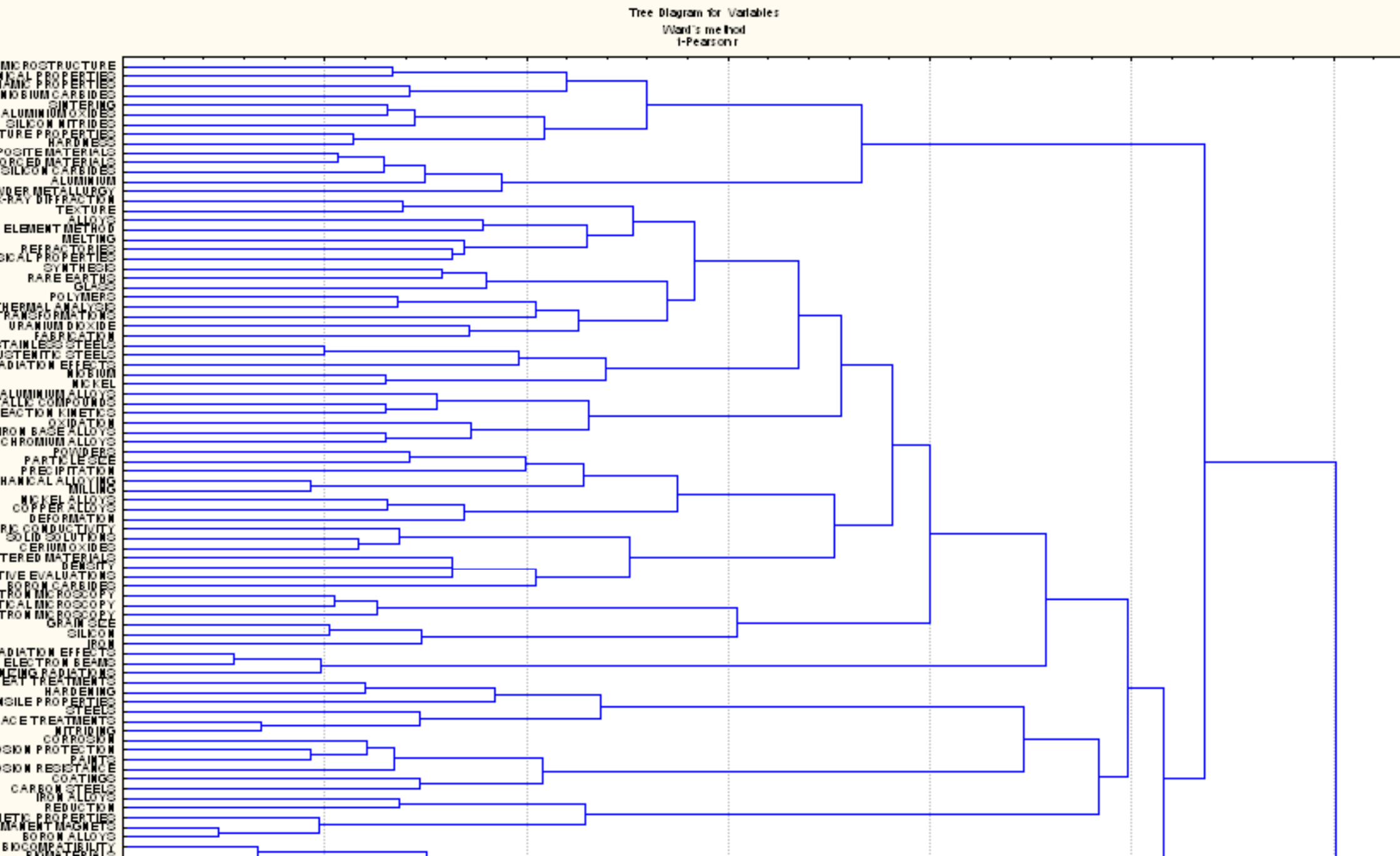
1-Pearson r

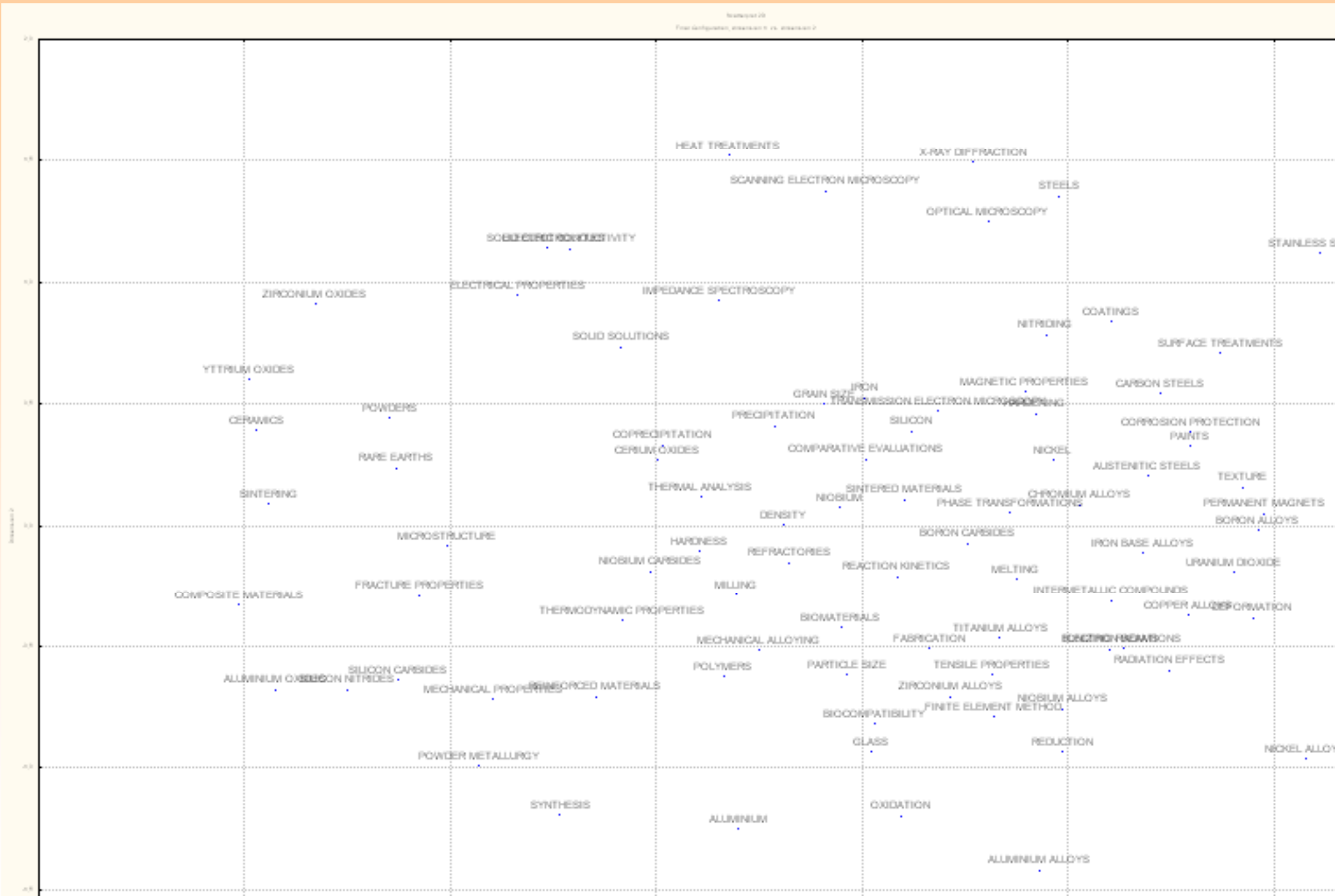


Categoria S63 - "Radiation - thermal effects - biological effects - dental uses -"



Categoria S36 - "Materials Science"





Os gráficos

- mostram as características da produção de dissertações e teses do Instituto, nos últimos anos (são retratos do passado)
- permitem fazer projeções sobre as tendências da pesquisa par

Discussão...

- Foi possível identificar:
 - os temas que estiveram em foco
 - a dinâmica das frentes de pesquisa.
- (continua)

Discussão

Resultados correlacionados aos cenários sociais, políticos e econômicos, transformam-se em informações estratégicas para a tomada de decisões relacionadas às políticas de pesquisa.

ALGUMAS CONSIDERAÇÕES FINAIS

- a estrutura de metadados e a qualidade dos registros interferem na confiabilidade
 - da recuperação de informação
 - da gestão de acervos
 - dos estudos de produção científica
 - dos indicadores de C&T
- Os *softwares* bibliométricos são ferramentas úteis para analisar a consistência das bases.

(CONTINUA)

CONSIDERAÇÕES...

- Experiências têm sido importantes para
 - aprimorar referencial teórico e metodológico
 - aprimorar as competências dos membros do grupo de pesquisa (pesquisadores seniores, alunos de pós-graduação, graduação)
- fortalecer as condições para produzir resultados que sirva de bases efetivas para avaliar a ciência e tecnologia nacionais.
 - (CONTINUA)

CONSIDERAÇÕES

- BASES DE DADOS BIBLIOGRÁFICOS
PODEM CUMPRIR MÚLTIPLAS FUNÇÕES
- Funções mais tradicionais: recuperação da
informação, gestão de acervos
- Uso crescente: estudos de produção
científica, produção de indicadores de
atividade científica, estudos sociais da
ciência

Considerações...

- Bases de Dados Referenciais
 - Úteis para gerar representações da cultura científica do país.
- importantes para realizar estudos de institucionalização cognitiva e social da ciência e expressá-las por indicadores

(CONTINUA)

Gestão de qualidade de bases de dados

- Necessita de instrumentos adequados
- *softwares* bibliométricos são ferramentas úteis para
 - analisar a consistência das bases
 - promover sua higienização

SUGESTÕES (AÇÕES)

- CRIAR CULTURA DE AVALIAÇÃO permanente da qualidade das bases de dados;
- adequar as bases de dados para a produção de indicadores de C&T.

(CONTINUA)

Das ações podem se beneficiar:

- os usuários de bibliotecas, porque a recuperação de informações será mais precisa;
- os gestores de políticas de pesquisa (coordenadores de pós-graduação, pró-reitorias de pesquisa e pós-graduação), porque poderão dispor de dados confiáveis para planejar.

Para terminar (mesmo)

- A manutenção de repositórios inconsistentes depende de inversão significativa de recursos (humanos, financeiros, materiais)
 - Investimento necessário
 - - bases bibliográficas são dispositivos participam diretamente do ciclo da produção de conhecimento
- (colaboração Delatéria Weinberg (10/2))



Muito obrigada!

(pela atenção e paciência)

Participaram desta pesquisa:

- Priscila Nozaki (aluna de IC. Atualmente, bacharel em Biblioteconomia - ECA-USP)
- Mery Igami (Mestre em C. da Comunicação-ECA-USP. Bibliotecária do IPEN-SP)