

204

**BUSCA E PADRONIZAÇÃO DE DADOS PARA ANÁLISES CIENTOMÉTRICAS.** *Marina Plentz, Ana Maria Mielniczuk de Moura, Sonia Elisa Caregnato (orient.) (UFRGS).*

O estudo é parte do projeto “Comunicação e Produção Científica em Grupos de Pesquisa de IES do Rio Grande do Sul”, cujo objetivo é analisar a organização dos grupos e a produção científica de docentes oriundos de diversas áreas do conhecimento. Relata os procedimentos utilizados na busca e na padronização de dados para subsidiar estudo cientométrico que visa analisar a interação entre produção científica e tecnológica na área de Biotecnologia. Nesse tipo de abordagem, a padronização é fundamental para a precisão dos resultados, principalmente quando envolve a comparação entre tipos de documentos (artigos e patentes) e o uso de diferentes bases de dados, como é o caso deste estudo. As buscas foram realizadas em duas fontes: primeiro na base de patentes do INPI, utilizando os códigos do Scoreboard 2005 da classificação de OECD para a área da Biotecnologia; e posteriormente na Science Citation Index, da Web of Science, pelo nome dos autores identificados como inventores de patentes brasileiras. O período das publicações foi de 2001 a 2005, com delimitação geográfica para Brazil na Web of Science e sem esta delimitação na base do INPI. As ferramentas utilizadas para a padronização dos dados foram Microsoft Excel, Bibexcel, de Olle Persson, e Textocompara, da MV Infomática. Os totais de registros recuperados foram: 3571 patentes, sendo 244 brasileiras e 3327 estrangeiras, e 4299 artigos. Após a retirada dos registros duplicados e com informações incompletas, os valores totalizaram em: 2525 patentes, sendo 194 brasileiras e 2331 estrangeiras, e 2584 artigos. Observou-se uma diferença significativa nas quantidades e frequências de autores e instituições, após a padronização dos artigos, o que reforça a necessidade de padronização de dados para estudos cientométricos. (CNPq).