



# UTILIZAÇÃO DO SOFTWARE VANTAGEPOINT NA EXTRAÇÃO DE DADOS PARA O ANUÁRIO ESTATÍSTICO DA USP

Eixo temático: MÉTODOS, TÉCNICAS E FERRAMENTAS PARA ESTUDOS BIBLIOMÉTRICOS E CIENTOMÉTRICOS

Adriana Nascimento Flaminio ([flaminio@usp.br](mailto:flaminio@usp.br)) Laucivaldo Cardoso de Oliveira, Roseli Koizimi Matsuda, Sibebe Fausto  
Departamento Técnico do Sistema Integrado de Bibliotecas da USP (DT/SIBiUSP)

## INTRODUÇÃO

O Anuário Estatístico da Universidade de São Paulo (USP) é um veículo de divulgação das mais diversas atividades das unidades da Universidade, somando um total de 42 indicadores de desempenho divulgados à sociedade que a mantém e que lhe dá condições para atingir seus objetivos.



Desde 1987 o Sistema Integrado de Bibliotecas da USP (SIBiUSP) realiza a coleta de alguns dados estatísticos das bibliotecas que constam de seu Sistema, de forma padronizada e sistêmica para fornecer informações para a elaboração de tabelas que constam na publicação do Anuário Estatístico da Universidade de São Paulo. A Divisão de Gestão de Tratamento da Informação (DGTI) do Departamento Técnico do SIBiUSP (DT-SIBiUSP) gerencia o Banco de Dados Bibliográficos da USP – DEDALUS, que representa os diversos acervos das bibliotecas distribuídas entre as 46 Unidades da USP. É a partir da extração de dados do DEDALUS que são disponibilizados os números da Produção Científica da Universidade, representados em tabelas e gráficos no Anuário Estatístico.

Atualmente o novo contexto de cooperação e internacionalização das Universidades (itens fundamentais para melhores posicionamentos nos diversos Rankings Universitários Mundiais), impõe que as instituições de ensino busquem indicadores que demonstrem o seu nível de participação neste novo contexto. Esses indicadores somente poderão ser apresentados a partir da extração de dados consistentes. Soma-se a esta necessidade, no âmbito das instituições públicas, garantir a autenticidade, integridade e o acesso às informações, prevista na Lei Federal nº 12.527, de 18 de novembro de 2011 (BRASIL, 2011), a partir de uma gestão transparente da informação.

## JUSTIFICATIVA E OBJETIVO

O Programa USP Internacional tem por objetivo fortalecer a presença da Universidade no exterior por meio de várias ações, sendo uma delas a implementação e consolidação das parcerias com instituições de ensino superior estrangeiras (A USP ultrapassa..., 2013). Nesse contexto, as informações divulgadas no Anuário Estatístico da USP por meio da antiga Tabela 9.03, publicada até 2012, são extremamente importantes e estratégicas, pois refletem de forma muito positiva a Produção Científica da USP com a participação de coautores estrangeiros, formando uma rede de colaboração científica muito rica, diversificada e interdisciplinar.

Tabela 9.03 - Participação de Instituições e de Coautores Estrangeiros em Artigos Científicos de Periódicos Internacionais Publicados por Docentes/Provisores da USP, em 2011

Unidade	Nº de Trabalhos	Nº de Participações Estrangeiras*	Países Envolvidos (Fusão dos Anos)	
EE	18	29	15	Canadá, Chile, Colômbia, Espanha, Inglaterra, EUA, Portugal, Peru
EEHE	9	1	1	Alemanha
BERP	32	21	31	Argentina, Bolívia, Cabo Verde, Canadá, Chile, Colômbia, Costa Rica, El Salvador, Espanha, Espanha, Honduras, Holanda, México, Moçambique, Nicarágua, Venezuela
HEC	18	211	12	Alemanha, Austrália, Bélgica, China, Espanha, EUA, França, Portugal, Rússia

Deste modo, a retirada dessa informação estratégica que eleva e reflete a internacionalização da USP do seu veículo de divulgação à sociedade em geral, de forma aberta e gratuita na Internet, torna-se muito preocupante, demandando soluções tecnológicas imediatas para retornarmos com esses dados ao Anuário Estatístico.

Sendo assim, procuramos avaliar a utilização da ferramenta tecnológica VP para a análise dos dados textuais estruturados retirados do DEDALUS e, a partir desta análise, a extração dos dados estatísticos para alimentar a supracitada Tabela 9.03 do Anuário Estatístico da USP, e posteriormente demais tabelas e gráficos. Futuramente, os dados que constarão no Anuário Estatístico da USP deverão ser extraídos do DEDALUS e da BDPI.

Este trabalho objetiva fornecer dados estatísticos consistentes sobre a rede de colaboração dos autores da USP com autores de instituições estrangeiras, como forma de contribuir para a internacionalização da USP, analisando se as informações contidas nos registros bibliográficos do DEDALUS estão devidamente estruturadas e se são passíveis de serem contabilizadas pelo VP, automatizando a extração de dados para o Anuário Estatístico da USP.

## PROCEDIMENTOS METODOLÓGICOS

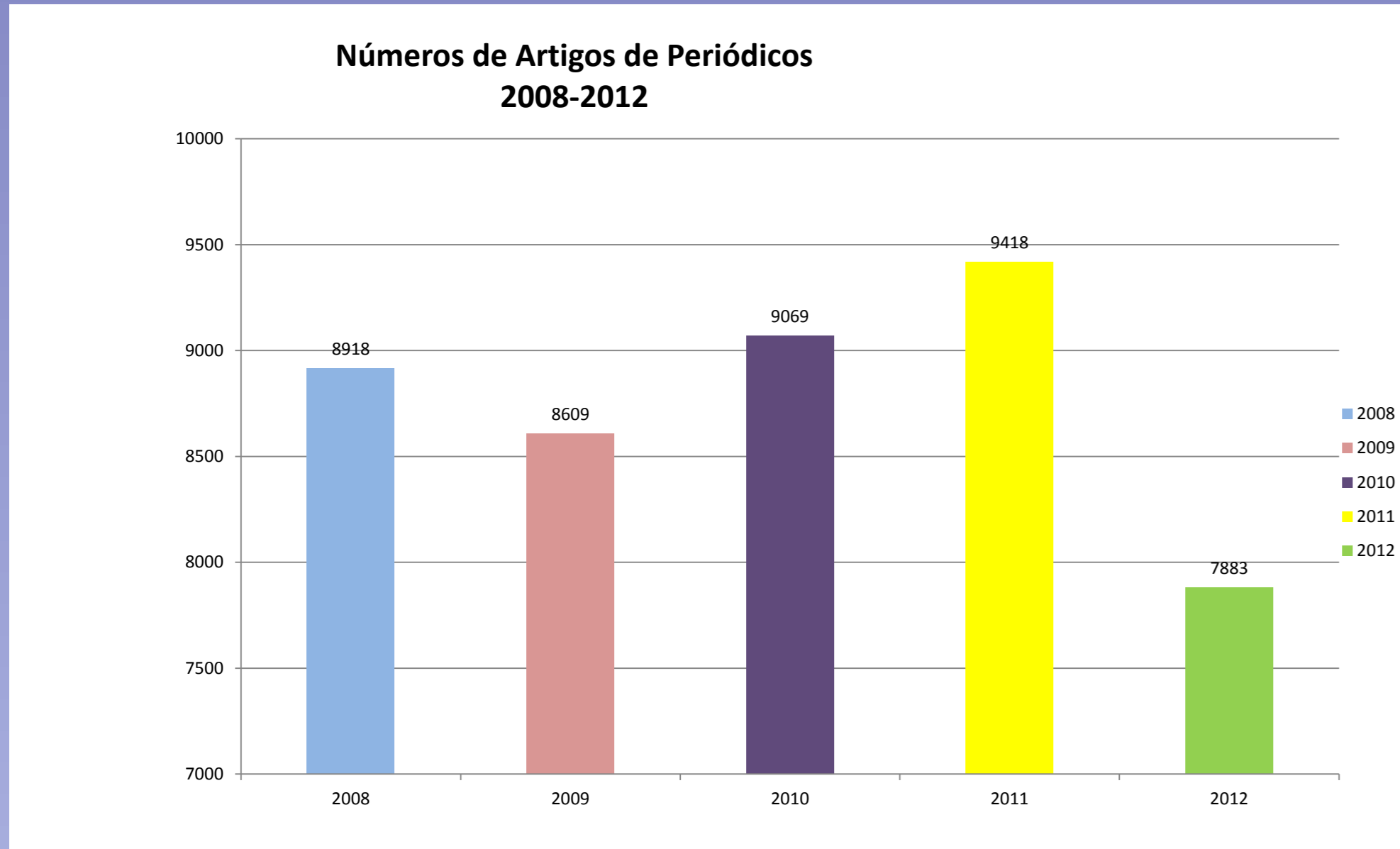
O corpus da análise utilizado foram os dados dos registros bibliográficos do Banco DEDALUS. Primeiramente foram realizados alguns levantamentos necessários para uma maior compreensão de como são extraídos atualmente os dados para as tabelas do Anuário, com os seguintes levantamentos:

- Levantamento das tabelas do Anuário Estatístico da USP referentes à Produção Científica da USP, títulos e campos utilizados para a extração de dados;
- Levantamento de tipologia documental definidos para cadastramento na Base de Produção Científica no Banco DEDALUS;
- Levantamento de unidades da USP previstas para o Banco DEDALUS;
- Levantamento de campos e subcampos utilizados nos registros bibliográficos da Base de Produção Científica, no Banco DEDALUS;

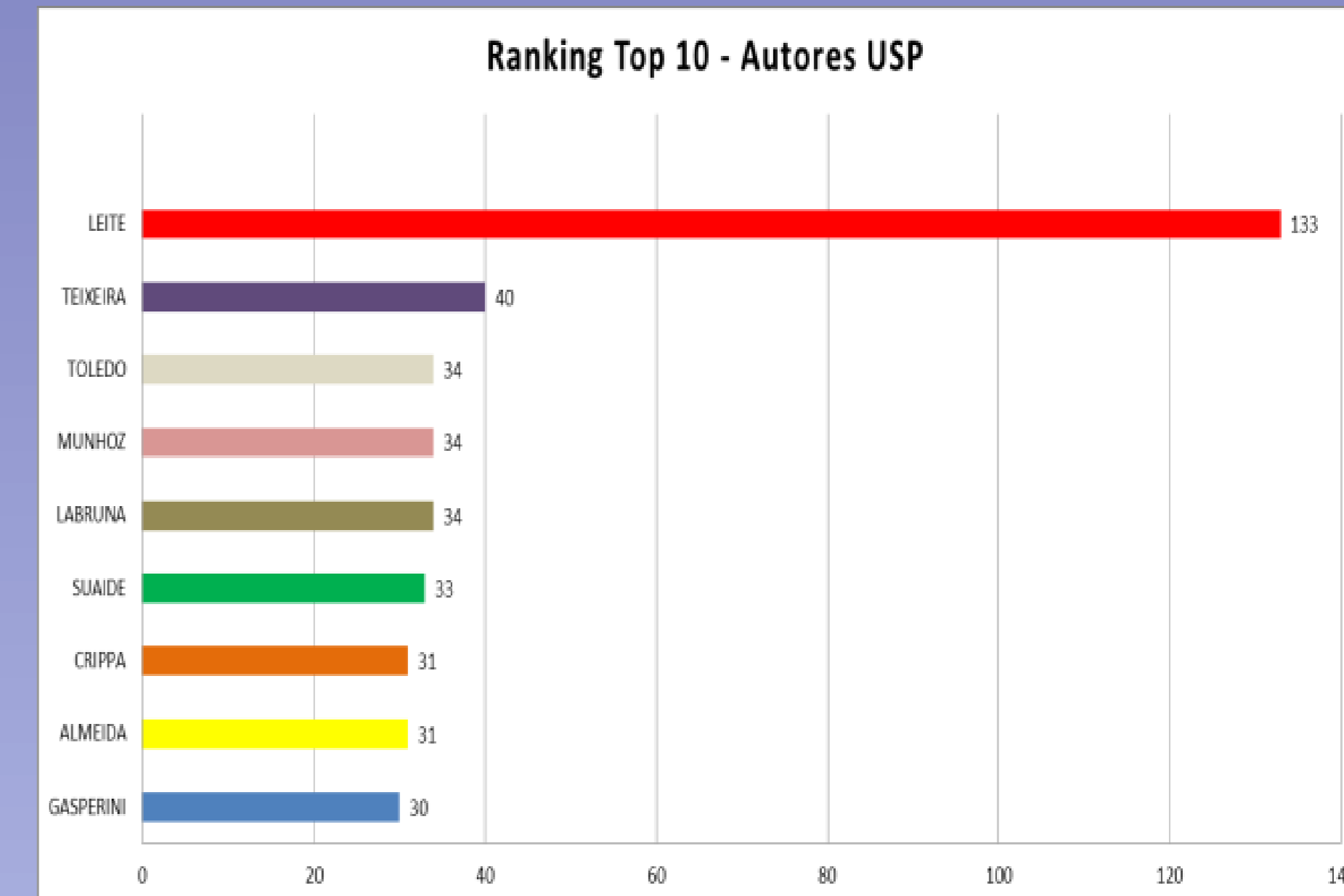
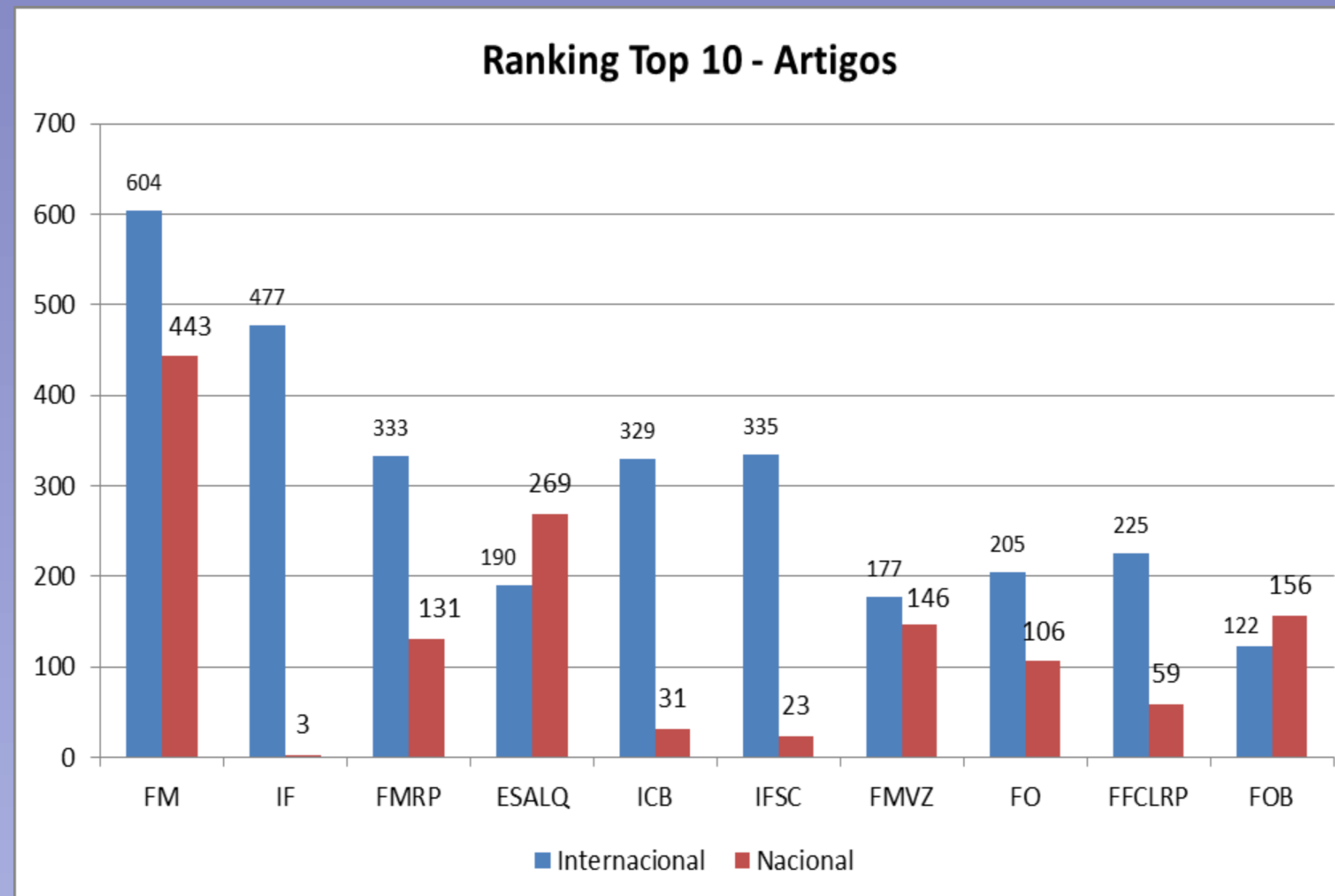
Após esses levantamentos, procedeu-se a extração de dados através do VantagePoint utilizando os filtros elaborados para análise dos dados. Os registros foram salvos em formato MARC, no bloco de notas com extensão .txt. No VP, importamos os dados. Durante esse processo, surgiram questionamentos em relação ao tratamento dos dados utilizando o VP, que são discutidos junto aos resultados obtidos.

## RESULTADOS

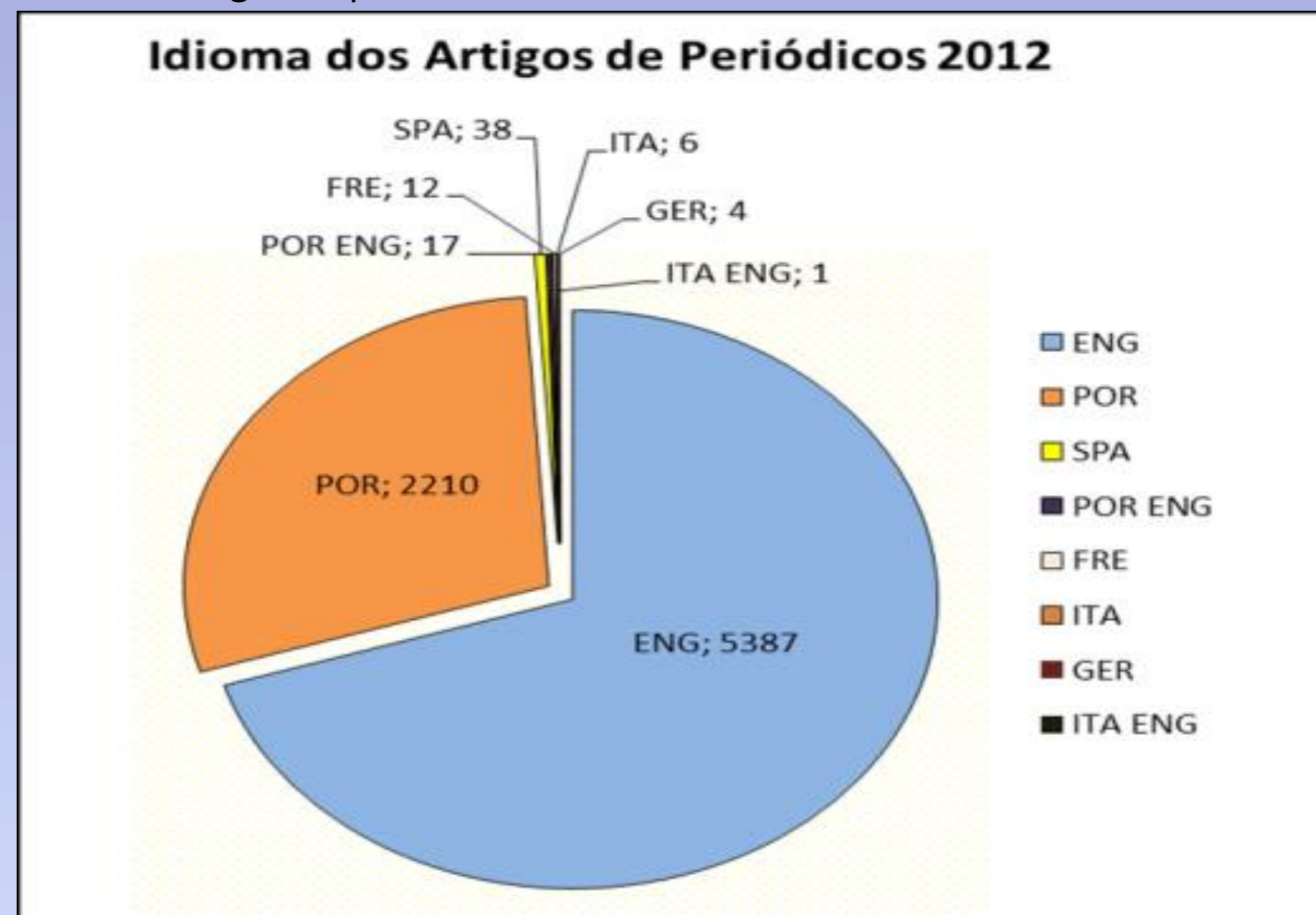
Gráfico 1: Números de artigos de periódicos 2008 a 2012



Gráficos 2 e 3: Ranking top 10 - total de artigos de periódicos por unidade e autores em 2012



Gráficos 4 e 5: Coautores nacionais e internacionais e Idiomas dos artigos de periódicos - 2012.



## CONSIDERAÇÕES FINAIS

O software VP é uma ferramenta para análise de dados extraídos de uma determinada fonte, e não tem por objetivo a normalização ou padronização dos dados. Portanto, se a fonte dos dados não estiver padronizada, a extração dos dados estatísticos pelo VP fica inviável e/ou inconsistente. Sendo assim, a extração da informação "país" dos registros bibliográficos da produção científica docente do DEDALUS para alimentar a Tabela 9.03 do Anuário Estatístico da USP, no momento, não é totalmente viável, uma vez que a informação não está padronizada na fonte. Neste contexto, é preciso avaliar o custo/benefício para a padronização do subcampo \$8 dos campos MARC 100 e 700 para a efetiva extração dos dados. Assim, a definição de novas políticas e diretrizes por parte do DT-SIBiUSP é fundamental para a padronização dos dados no DEDALUS e BDPI para a geração de dados consistentes para a apresentação em relatórios gerenciais internos e externos, relatórios demandados pela Reitoria da USP, INEP, Anuário Estatístico da USP, entre outros. Aumentando, desta forma, a credibilidade do DT-SIBiUSP na qualidade e consistência das informações divulgadas, no gerenciamento eficaz e efetivo do acervo das bibliotecas do Sistema, da produção científica da Universidade, no uso das fontes de informações adquiridas pelo DT-SIBiUSP, na prestação de serviços a comunidade USP e à sociedade em geral, contribuindo assim, de forma concreta para a internacionalização da USP.

Destaca-se que não foi objetivo deste trabalho levantar críticas ou problemas relacionados ao desempenho do software VantagePoint, mas sim, foi um exercício inicial, fruto do Trabalho de Conclusão de Curso de Capacitação de Bibliotecários em Análise Bibliométrica para Apoio à Gestão da Pesquisa em Universidade Pública, no âmbito da parceria SIBiUSP e Escola Técnica e de Gestão da USP, visando obter parâmetros e subsídios para futuros desenvolvimentos e customizações do aplicativo para atender a diversas demandas do SIBiUSP com a aquisição do referido software.

## REFERÊNCIAS

- A USP ultrapassava fronteiras. *Boletim USP Destaques*, n. 71, 09 Abr. 2013. Disponível em: [http://www.usp.br/imprensa/wp-content/uploads/USP-Destaques\\_71.pdf](http://www.usp.br/imprensa/wp-content/uploads/USP-Destaques_71.pdf). Acesso em 28 abr. 2013.
- BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso xxxiii do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da constituição federal; altera a lei nº 8.112, de 11 de dezembro de 1990; revoga a lei nº 11.111, de 5 de maio de 2005, e dispositivos da lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. *Diário Oficial da União*, 18/11/2011, p. 1 (Edição Extra). Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/l12527.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm). Acesso em 17 mai. 2013.
- MENEGHINI, R.; PACKER, A. L. *Is there science beyond English? Initiatives to increase the quality and visibility of non-English publications might help to break down language barriers in scientific communication*. *EMBO Rep.* v. 8, n. 2, p. 112-116, fev. 2007.
- PORTER, A.; PALOP, F. *Mineração de textos para decisões de gestão de pesquisa e tecnologia - tech mining com a ajuda de software Vantage Point*. In: Encontro Brasileiro de Bibliometria e Cientometria, 3., Gramado, RS. Anais... Gramado: UFRGS, 2012. Disponível em: <http://www.ufrgs.br/ebbc2012/arquivos/workshop-2>. Acesso em 01 Mai. 2013
- UNIVERSIDADE DE SÃO PAULO (USP). *Anuário Estatístico*. Apresentação. Disponível em: <https://uspdigital.usp.br/anuario/apresentacao.jsp?codmenu=2781>. Acesso em: 28 Abr. 2013.
- Reitoria. Resolução nº 6444, de 22 de outubro de 2012. Dispõe sobre diretrizes e procedimentos para promover e assegurar a coleta, tratamento e preservação da produção intelectual gerada nas Unidades USP e pelos Programas Conjuntos de Pós-graduação, bem como sua disseminação e acessibilidade para a comunidade. *Diário Oficial do Estado*, 23 de out. 2012a. Disponível em: <http://www.usp.br/dtrh/novo/legislacao/doi2012/res-usp6444.html>. Acesso em 10 Abr. 2013.
- Reitoria. Portaria GR-5917, de 22 de outubro de 2012. Altera dispositivo da Portaria GR-2.922/1994, de 16 de novembro de 1994. *Diário Oficial do Estado*, 23 out. 2012b. <http://www.leginf.usp.br/?portaria=portaria-gr-no-5917-de-22-de-outubro-de-2012>. Acesso em 04 mai. 2013.
- Reitoria. Portaria GR-2922, de 16 de novembro de 1994. Regulamenta o funcionamento do Banco de Dados Bibliográficos da Universidade de São Paulo e dá outras providências correlatas. *Diário Oficial do Estado*, 18 nov.1994. Seção I, p. 58. Disponível em: [http://cltrus.uspnet.usp.br/sibi/Portaria-Resolucao/port\\_gr\\_2922.htm](http://cltrus.uspnet.usp.br/sibi/Portaria-Resolucao/port_gr_2922.htm). Acesso em: 22 dez. 2012.