SDB Release Notes v0.8

This page is under construction as part of the coming release

September 28, 2012

The Scholarly Database team at Indiana University is pleased to announce the release of version 0.8 of the Scholarly Database. This release adds two new data sources as well as provides updates to existing data.

New Data

Clinical Trials

The National Institutes of Health runs ClinicalTrials.gov, a central registry of clinical trials. The database includes both publicly and privately funded studies around the world. ClinicalTrials.gov was established by law in 1997 and made available to the public in 2000. The database includes data on trials, related diseases or conditions, interventions, eligibility criteria, locations and contacts. The Clinical Trials wiki page gives more detailed information about the dataset, including the table schema and data coverage.

National Endowment for the Humanities Awards

The National Endowment for the Humanities was created in 1965 to award grants to promote excellence in the humanities and awareness of the lessons of history. It is one of the largest funders of humanities programs in the United States. The NEH wiki page gives more detailed information about the dataset, including the table schema and data coverage.

Data Update

National Institutes of Health

The National Institutes of Health data has been updated to include data current through the Summer of 2012. In addition, new information linking grants to patents and publications is available. Information on the financial value of grants has also been added, where available.

Data Summary

Dataset	# Records	Years Covered	Regular Update
MEDLINE Papers	19,039,860	1865-2010	Yes
USPTO Patents	4,178,196	1976-2010	Yes
NIH Awards	2,490,837*	1972-2012	Yes
Clinical Trials	119,144	1900-2012	Yes
NSF Awards	453,687	1952-2010	No
NEH Awards	47,197	1970-2012	No
Total	**26,328,921**		

The number of NIH awards was not aggregated by base project, it includes subprojects. Some projects have up to 3,000 subprojects.

The Scholarly Database (SDB) at Indiana University aims to serve researchers and practitioners interested in the analysis, modeling, and visualization of large-scale scholarly datasets. The online interface at http://sdb.cns.iu.edu provides access to six datasets: MEDLINE papers, Clinical Trials, U.S. Patent and Trademark Office patents (USPTO), National Science Foundation (NSF) funding, National Endowment for the Humanities funding, and National Institutes of Health (NIH) funding – over 26 million records in total. Users can register for free to cross-search these databases and to download result sets as dumps for scientometrics research and science policy practice.

If you would like to learn more about the datasets of Scholarly Database, please visit the SDB wiki at http://sdb.wiki.cns.iu.edu.