• Established in 2008 to provide scholars with sets of public data that are easy to find, acquire, understand and use

• Four major data sets are available
  – US Patents
  – MEDLINE publications
  – NIH Grants
  – NSF Grants

• Over 25 million total records
• Allows for federated searching of these datasets by creator, title, abstract and year
<table>
<thead>
<tr>
<th>Source</th>
<th>Authors/Creators</th>
<th>Year</th>
<th>Title</th>
<th>Score (out of 3.07)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDLINE</td>
<td>Grishok</td>
<td>2005</td>
<td>RNAi mechanisms in Caenorhabditis elegans.</td>
<td>3.07</td>
</tr>
<tr>
<td>USPTO</td>
<td>Mello et al.</td>
<td>2007</td>
<td>RNA interference pathway genes as tools for targeted genetic interference</td>
<td>3.01</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>Berkhout and Haasnoot</td>
<td>2006</td>
<td>The interplay between virus infection and the cellular RNA interference machinery</td>
<td>2.89</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>Thompson and Viney</td>
<td>2008</td>
<td>Two hypotheses to explain why RNA interference does not work in animal parasitic nematodes.</td>
<td>2.88</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>Gallili and Tang</td>
<td>2004</td>
<td>Using RNAi to improve plant nutritional value: from mechanism to application.</td>
<td>2.85</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>Berkhout and de Vries</td>
<td>2008</td>
<td>RNAi suppressors encoded by pathogenic human viruses.</td>
<td>2.85</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>Li et al.</td>
<td>2005</td>
<td>[Experimental strategies of the application of RNAi technique in mammalian cells]</td>
<td>2.70</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>Chiang and Lee</td>
<td>2008</td>
<td>RNA interference: new therapeutics in allergic diseases.</td>
<td>2.70</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>Kim and Rossi</td>
<td>2009</td>
<td>Overview of gene silencing by RNA interference.</td>
<td>2.70</td>
</tr>
</tbody>
</table>
• Our goal is to provide data that is easy to understand and easy to use
  – Data is available as dumps and as network files
    • Co-author, co-investigator, citations
  – Data dictionaries are available for all sets via the SDB site and via the wiki
  – Flat files are provided in the standard csv, while network files use the NWB format.
  – All files are easily imported into our network science tools
    • Network Workbench
    • Science of Science (Sci2) Tool
Download Results

- Download all
- Data Dictionary
- Sample File

Download 20000 records starting at record 1 from the following databases:

**MEDLINE Database:**
- MEDLINE author table
- MEDLINE co-author table (nwb format)
- MEDLINE master table
- MEDLINE MeSH heading table
- MEDLINE MeSH qualifier table

**NIH Database:**
- NIH master table

**NSF Database:**
- NSF co-investigator table (nwb format)
- NSF master table

**USPTO Database:**
- USPTO agent table
- USPTO assignee table
- USPTO citation table (nwb format)
- USPTO claims table
- USPTO co-inventor table (nwb format)
- USPTO inventor table
- USPTO master (burst format)
- USPTO master table
- USPTO Patent Cooperation Treaty table

Download
• Still growing and adding data
  – Ongoing updates of existing datasets
  – Provide new linkage data between NIH grants and publications
  – National Endowment for the Humanities grant data
• Sign up at http://sdb.cns.iu.edu
• Questions/comments? lightr@indiana.edu