

## 2.5 Sample Datasets

Meaningful analysis requires carefully collected data. The Sci2 Tool can import many varieties of scientometric data, outlined in section 4.2 [Data Acquisition and Preparation](#), but the tool also comes bundled with several sample datasets. This sample data will be used throughout sections 4 [Workflow Design](#) and 5 [Sample Workflows](#), and includes:

- Geo
- Network
- Scientometrics
  - Bibtex
  - CSV
  - Endnote
  - ISI
- Models
  - TARL
- NIH
- NSF
- Properties
- Scopus
- SDB
  - RNAi
- IVMOOC Sample Data

### Files available:

Full Sample Dataset

### Geo

[usptoInfluenza.csv](#)[usptoInfluenza.csv](#) – Heavily pre-processed and geocoded data covering USPTO patents awarded containing the keyword 'Influenza'.

### Network

[convertGraph\\_v0.8.0.graphml](#)

[convertGraph\\_v0.8.0.graphml](#)

[convertGraph\\_v0.8.0.graphmlkidscontest.net](#)

[kidscontest.netnetsci06-conference.net](#)

[netsci06-conference.netseiyu.graphml.xml](#)

[seiyu.graphml](#)

### Scientometrics

#### *Bibtex*

[Bibsonomy.bib](#)

[LaszloBarabasi.bib](#)

#### *CSV*

[LaszloBarabasi.csv](#)

[20publications.csv](#)

[Medline\\_master\\_table.csv](#) (most recent version is always available from the [Scholarly Database](#) site)

#### *Endnote*

[KatyBorner.enw](#)– 146 publications authored or co-authored by Katy Börner from 1992-2013.

[LaszloBarabasi.enw](#)

## ***ISI***

AlessandroVespignani – 101 publications authored or co-authored by Alessandro Vespignani from 1990-2006.

EugeneGarfield

FourNetSciResearchers – 361 publications spanning 52 years and four network scientists: Albert-László Barabási, Eugene Garfield, Alessandro Vespignani, & Stanley Wasserman. This data was collected in 2007 from Web of Science.

LaszloBarabasi

StanleyWasserman

Test5Papers

Scientometrics.isi – All articles (2,126) published in the journal *Scientometrics* from 1978-2008.

ScientometricsNewFormat.isi – Papers published in *Scientometrics* from 2010.

## **Models**

### ***TARL***

Agingfunction

iniscrypt.tarl

### **NIH**

CTSA2005-2009 – 2,546 papers and 534 grants dealing with Clinical and Translational Science from 2005-2009.

### **NSF**

- Refer to the CShell [Datasets](#) wiki page

### **Properties**

- Refer to the CShell [Datasets](#) wiki page

### **Scopus**

**BrainCancer.scopus**

### **SDB**

### ***RNAi***

Medline\_co-author\_table\_(nwb\_format)

USPTO\_citation\_table\_(nwb\_format)

### **IVMOOC Sample Data**

AlessandroVespignani – 101 publications authored or co-authored by Alessandro Vespignani from 1990-2006.

Authors.csv

dynamic\_data\_example1.gexf

dynamic\_data\_example2.gexf

Florentine.pdf

PNAS\_top50\_words.net

SciSIPFunding.csv