

8.2 Network Analysis and Other Tools

Table 8.1 provides an overview of existing tools used in scientometrics research, see also (Fekete and Börner-chairs 2004). The tools are sorted by the date of their creation. Domain refers to the field in which they were originally developed such as social science (SocSci), scientometrics (Scientom), biology (Bio), geography (Geo), and computer science (CS). Coverage aims to capture the general functionality and types of algorithms available, e.g., Analysis and Visualization (A+V), see also description column.

Table 8.1 Network analysis and visualization tools commonly used in scientometrics research.

| Tool | Year | Domain | Coverage | Description | UI | Open Source | Operating System | References |
|----------------------|------|------------|---------------------------|---|--------------|-------------|------------------|----------------------------------|
| S&T Dynamics Toolbox | 1985 | Scientom | A + V | Tools from Loet Leydesdorff for organization analysis, and visualization of scholarly data. | Command-line | No | Windows | (Leydesdorff 2008) |
| In Flow | 1987 | SocSci | A + V | Social network analysis software for organizations with support for what-if analysis. | Graphical | No | Windows | (Krebs 2008) |
| Pajek | 1996 | SocSci* | A + V | A network analysis and visualization program with many analysis algorithms, particularly for social network analysis. | Graphical | No | Windows | (Batagelj and Mrvar 1998) |
| MANET | 1996 | Statistics | A + V | MANET is a tool for exploring data, providing a range of graphical tools for studying multivariate features. | Graphical | Yes | Mac | (Unwin 1996) |
| ExplorN | 1997 | Statistics | A + V | ExplorN is a data analysis and visualization tool that supports scatterplot matrices, parallel coordinate plots, incon-enhanced three-dimensional setoscopic plots, and more. | Graphical | Yes | All Major | (Carr et al. 1997) |
| XGobi | 1998 | Statistics | A + V | XGobi is a data visualization system for viewing high-dimensional data. | Graphical | Yes | Windows, Linux | (Swayne et al. 1998) |
| UCInet | 2000 | SocSci* | A + V | Social network analysis software particularly useful for exploratory analysis. | Graphical | No | Windows | (Borgatti, Everett et al. 2002) |
| XploRe | 2000 | Statistics | A + V | XploRe allows for multidimensional analysis, non- and semiparametric modelling, and analysis of financial markets. | Graphical | Yes | Windows, Linux | (Härdle et al. 2000) |
| Boost Graph Library | 2000 | CS | Analysis and Manipulation | Extremely efficient and flexible C++ library for extremely large networks. | Library | Yes | All Major | (Siek, Lee et al. 2002) |
| nViZn | 2000 | Statistics | A + V | nViZn is a Java foundation for analytical graphics, best understood as a geometric analytical engine that allows for the visualization of statistical data. | Command-line | No | All Major | (Wilkinson et al. 2000) |
| Common GIS | 2001 | GeoVis | A + V | CommonGIS is a tool for visualizing spatial data and allows for exploratory data analysis. | Graphical | Yes | Web based | (Andrienko et al. 2003) |
| Visone | 2001 | SocSci | A + V | Social network analysis tool for research and teaching, with a focus on innovative and advanced visual methods. | Graphical | No | All Major | (Brandes and Wagner 2008) |
| GeoVISTA | 2002 | Geo | GeoVis | GIS software that can be used to lay out networks on geospatial substrates. | Graphical | Yes | All Major | (Takatsuka and Gahegan 2002) |
| Cytoscape | 2002 | Bio* | Visualization | Network visualization and analysis tool focusing on biological networks, with particularly nice visualizations. | Graphical | Yes | All Major | (Cytoscape-Consortium 2008) |
| Mondrian | 2002 | Statistics | A + V | Mondrian is a general purpose data visualization program particularly useful when working with categorical data, geographical data and large data sets. | Graphical | Yes | All Major | (Theus 2002) |
| NetworkX | 2002 | Networks | A + V | NetworkX is a Python language software package that allows for the analysis and visualization of complex networks. | Command-line | Yes | All Major | (Hagberg, Swart, & S Chult 2008) |
| Tulip | 2003 | CS | Visualization | Graph visualization software for networks over 1,000,000 elements. | Graphical | Yes | All Major | (Auber 2003) |
| iGraph | 2003 | CS | Analysis and Manipulation | A library for classic and cutting edge network analysis usable with many programming languages. | Library | Yes | All Major | (Csárdi and Nepusz 2006) |
| CrystalVision | 2003 | Statistics | A + V | ExploRn is a data visualization program that focuses on parallel coordinate plots, scatterplots, and grand tour animations. | Graphical | Yes | All Major | (Wegman and Dorfman 2003) |
| CiteSpace | 2004 | Scientom | A + V | A tool to analyze and visualize scientific literature, particularly co-citation structures. | Graphical | Yes | All Major | (Chen 2006) |
| HistCite | 2004 | Scientom | A + V | Analysis and visualization tool for data from the Web of Science. | Graphical | No | Windows | (Garfield 2008) |
| R | 2004 | Statistics | A + V | A statistical computing language with many libraries for sophisticated network analyses. | Command-line | Yes | All Major | (Ihaka and Gentleman 1996) |

| | | | | | | | | |
|-------------------|------|-----------------------|------------------------------|---|--------------|-----|---|--|
| GraphViz | 2004 | Networks | Visualization | Flexible graph visualization software. | Graphical | Yes | All Major | (AT&T-Research-Group 2008) |
| OpenGeoDa | 2005 | Geo | A + V | OpenGeoDa is a tool for exploratory spatial data analysis (ESDA) on lattice data (points and polygons). | Graphical | Yes | All major | (Anselin 2005) |
| Prefuse | 2005 | Visualization | Visualization | A general visualization framework with many capabilities to support network visualization and analysis. | Library | Yes | All Major | (Heer, Card et al. 2005) |
| VisuaLyzer | 2005 | Networks | Visualization | VisuaLyzer is a tool for importing data from Edgelist/Edgearray, Excel or GraphML formats and generating network visualizations. | Graphical | No | All Major | (Reshef 2009) |
| NWB Tool | 2006 | Bio, SocSci, Scientom | A + V | Network analysis & visualization tool conducive to new algorithms supportive of many data formats. | Graphical | Yes | All Major | (Huang 2007) |
| BibExcel | 2006 | Scientom | A + V | Transforms bibliographic data into forms usable in Excel, Pajek, NetDraw, and other programs. | Graphical | No | Windows | (Persson 2008) |
| tnet | 2006 | Networks | A + V | Tnet is a tool for social network analysis of weighted, two-mode, and longitudinal networks. | Graphical | Yes | All Major | (Barrat et al. 2004) |
| KrackPlot | 2006 | Networks | A + V | KrackPlot is a network analysis program designed specifically for the analysis of social networks. | Graphical | Yes | All Major | (McGrath and Blythe 2004) |
| libSNA | 2006 | Networks | A + V | This open-source social network analysis tool is under development by Abe Usher. | Graphical | Yes | All Major | (Terna 2008) |
| GUESS | 2007 | Networks | Visualization | A tool for visual graph exploration that integrates a scripting environment. | Graphical | Yes | All Major | (Adar 2007) |
| Publish or Perish | 2007 | Scientom | Data Collection and Analysis | Harvests and analyzes data from Google Scholar, focusing on measures of research impact. | Web-based | No | Windows, Linux | (Harzing 2008) |
| Commetrix | 2007 | Networks | A + V | Commetrix software allows users to dynamically analyze and create rich network maps. | Graphical | Yes | Windows (any system that supports Java) | (Trier and Bobrik 2007) |
| PEGASUS | 2008 | Networks | A + V | Overview PEGASUS is a Peta-scale graph mining system, fully written in Java that provides large scale algorithms for important graph mining tasks. | Command-line | Yes | All Major | (Kang, Tsourakakis, and Faloutsos, 2009) |
| GraphStream | 2008 | Networks | Visualization | GraphStream is a dynamic graph library written in Java that allows for the presentation of dynamic graphs. | Command-line | No | All Major | (Pigné et al. 2008) |
| Gephi | 2009 | Networks | A + V | Gephi is an interactive visualization and exploration platform for all kinds of networks and complex systems, dynamic and hierarchical graphs. | Graphical | Yes | All Major | (Bastian et al., 2009) |
| VOSviewer | 2009 | Networks | Visualization | VOSviewer is a visualization tool useful for analyzing bibliometric networks. | Graphical | Yes | All Major | (Eck and Waltman 2011) |
| GraphInsight | 2009 | Networks | Visualization | GraphInsight is a freely available program that allows for visualization of complex data structures into graphs to provide business intelligence for organizations. | Graphical | No | All Major | (Dallachiesa and Nicolini 2009) |
| NodeXL | 2010 | SocSci | A + V | NodeXL is a free, open-source template for Excel 2007 and 2010 that lets you enter a network edge list, click a button, and see the network graph, all in the Excel window. | Graphical | No | Windows | (Hansen et al. 2010) |
| PlotViz | 2010 | Networks | Visualization | PlotViz is a tool for the visualization of large-scale high dimensional data | Graphical | | Windows | (Choi et al. 2010) |
| TINA | 2011 | Scientom, Networks | A + V | Tool for interactive assessment of projects portfolio and visualization of scientific landscapes. | Graphical | Yes | All Major | (Cointet 2008) |
| Spotfire 4.0 | 2011 | Statistics | A + V | Spotfire allows users to analyze and visualize their data using complex and predictive statistics. | Graphical | No | All Major | (Shneiderman 2007) |
| CitNetExplorer | 2014 | Scientom, Networks | A + V | CitNetExplorer ('Citation Network Explorer') is a software tool for analyzing and visualizing citation networks of scientific publications. | Graphical | No | All Major | (van Eck and Waltman 2014) |
| MuxVi | 2014 | Multiplex networks | A + V | MuxVi ('multilayer analysis and visualization platform') is a framework for the multilayer analysis and visualization of networks. | Library | Yes | All Major | (Domenico 2014) |
| VennMaker | 2013 | Networks | A + V | VennMaker is a software tool which allows users to collect network relationships from an actor's point of view and render for comparative and quantitative analysis. | Graphical | No | All Major | (Schönhuth, Kronenwett, Gamper, Stark, 2013) |

Many of these tools are very specialized and capable. For instance, BibExcel and Publish or Perish are great tools for bibliometric data acquisition and analysis. HistCite and CiteSpace each support very specific insight needs – from studying the history of science to the identification of scientific research frontiers. The S&T Dynamics Toolbox provides many algorithms commonly used in scientometrics research and it provided bridges to more general tools. Pajek and UCINET are very versatile, powerful network analysis tools that are widely used in social network analysis. Cytoscape is excellent for working with biological data and visualizing networks.

The Network Workbench Tool has fewer analysis algorithms than Pajek and UCINET, and less flexible visualizations than Cytoscape. Network Workbench, however, makes it much easier for researchers and algorithm authors to integrate new and existing algorithms and tools that take in diverse data formats. The OSGi (<http://www.osgi.org>) component architecture and CIShell algorithm architecture (<http://cishell.org>) built on top of OSGi make this possible. Cytoscape is also adopting an architecture based on OSGi, though it will still have a specified internal data model and will not use CIShell in the core. Moving to OSGi will make it possible for the tools to share many algorithms, including adding Cytoscape's visualization capabilities to Network Workbench.

Several of the tools listed in the table above are also libraries. Unfortunately, it is often difficult to use multiple libraries, or sometimes any outside library, even in tools that allow the integration of outside code. Network Workbench, however, was built to integrate code from multiple libraries (including multiple versions of the same library). For instance, two different versions of Prefuse are currently in use, and many algorithms use JUNG (the Java Universal Network/Graph Framework). We feel that the ability to adopt new and cutting edge libraries from diverse sources will help create a vibrant ecology of algorithms.

Although it is hard to discern trends for tools which come from such diverse backgrounds, it is clear that over time the visualization capabilities of scientometrics tools have become more and more sophisticated. Scientometrics tools have also in many cases become more user friendly, reducing the difficulty of common scientometrics tasks as well as allowing scientometrics functionality to be exposed to non-experts. Network Workbench embodies both of these trends, providing an environment for algorithms from a variety of sources to seamlessly interact in a user-friendly interface, as well as providing significant visualization functionality through the integrated GUESS tool.

Many other tools are available outside the scope of network analysis that are still useful for studying the data of science. One such tool is the web-based [Data Science Toolkit](#), a web-based collection of open-source data sets and tools which allows the user to query for geographical data, parse text, and run named entity recognition.