### 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>
<thead>
<tr>
<th>Tool Name</th>
<th>Year</th>
<th>Domain</th>
<th>Coverage</th>
<th>Description</th>
<th>UI</th>
<th>Open Source</th>
<th>Operating System</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MANET</td>
<td>1996</td>
<td>Statistics</td>
<td>A+V</td>
<td>MANET is a tool for exploring data, providing a range of graphical tools for studying multivariate features.</td>
<td>Graphical</td>
<td>Yes</td>
<td>Mac</td>
<td>(Unwin 1996)</td>
</tr>
<tr>
<td>ExplorN</td>
<td>1996</td>
<td>Statistics</td>
<td>A+V</td>
<td>ExplorN is a data analysis and visualization tool that supports scatterplot matrices, parallel coordinate plots, and more.</td>
<td>Graphical</td>
<td>Yes</td>
<td>All Major</td>
<td>(Carr et al. 1997)</td>
</tr>
<tr>
<td>XGobi</td>
<td>1998</td>
<td>Statistics</td>
<td>A+V</td>
<td>XGobi is a data visualization system for viewing high-dimensional data.</td>
<td>Graphical</td>
<td>No</td>
<td>Windows, Linux</td>
<td>(Swayne et al. 1998)</td>
</tr>
<tr>
<td>Boost Graph Library</td>
<td>2000</td>
<td>CS</td>
<td>Analysis and Manipulation</td>
<td>Extremely efficient and flexible C++ library for extremely large networks.</td>
<td>Library</td>
<td>Yes</td>
<td>All Major</td>
<td>(Siek, Lee et al. 2002)</td>
</tr>
<tr>
<td>nViZn</td>
<td>2000</td>
<td>Statistics</td>
<td>A+V</td>
<td>nViZn is a Java foundation for analytical graphics, best understood as a geometric analytical engine that allows for the visualization of statistical data.</td>
<td>Command-line</td>
<td>No</td>
<td>All Major</td>
<td>(Wilkinson et al. 2000)</td>
</tr>
<tr>
<td>CommonGIS</td>
<td>2001</td>
<td>GeoVis</td>
<td>A+V</td>
<td>CommonGIS is a tool for visualizing spatial data and allows for exploratory data analysis.</td>
<td>Graphical</td>
<td>Yes</td>
<td>Web based</td>
<td>(Andrienko et al. 2003)</td>
</tr>
<tr>
<td>Visone</td>
<td>2001</td>
<td>SocSci</td>
<td>A+V</td>
<td>Social network analysis tool for research and teaching, with a focus on innovative and advanced visual methods.</td>
<td>Graphical</td>
<td>No</td>
<td>All Major</td>
<td>(Brandes and Wagner 2008)</td>
</tr>
<tr>
<td>GeoVISTA</td>
<td>2002</td>
<td>Geo</td>
<td>GeoVis</td>
<td>GIS software that can be used to lay out networks on geospatial substrates.</td>
<td>Graphical</td>
<td>Yes</td>
<td>All Major</td>
<td>(Takatsuka and Gahegan 2002)</td>
</tr>
<tr>
<td>Cytoscape</td>
<td>2002</td>
<td>Bio</td>
<td>Visualization</td>
<td>Network visualization and analysis tool focusing on biological networks, with particularly nice visualizations.</td>
<td>Graphical</td>
<td>Yes</td>
<td>All Major</td>
<td>(Cytoscape-Consortium 2008)</td>
</tr>
<tr>
<td>Mondrian</td>
<td>2002</td>
<td>Statistics</td>
<td>A+V</td>
<td>Mondrian is a general purpose data visualization program particularly useful when working with categorical data, geographical data, and large data sets.</td>
<td>Graphical</td>
<td>Yes</td>
<td>All Major</td>
<td>(Theus 2002)</td>
</tr>
<tr>
<td>NetworkX</td>
<td>2002</td>
<td>Networks</td>
<td>A+V</td>
<td>NetworkX is a Python language software package that allows for the analysis and visualization of complex networks.</td>
<td>Command-line</td>
<td>Yes</td>
<td>All Major</td>
<td>(Hagberg, Swart, &amp; S Chui 2008)</td>
</tr>
<tr>
<td>Tulip</td>
<td>2003</td>
<td>CS</td>
<td>Visualization</td>
<td>Graph visualization software for networks over 1,000, 000 elements.</td>
<td>Graphical</td>
<td>Yes</td>
<td>All Major</td>
<td>(Auber 2003)</td>
</tr>
<tr>
<td>iGraph</td>
<td>2003</td>
<td>CS</td>
<td>Analysis and Manipulation</td>
<td>A library for classic and cutting edge network analysis usable with many programming languages.</td>
<td>Library</td>
<td>Yes</td>
<td>All Major</td>
<td>(Csárdi and Nepusz 2006)</td>
</tr>
<tr>
<td>CrystalView</td>
<td>2003</td>
<td>Statistics</td>
<td>A+V</td>
<td>ExiPower is a data visualization program that focuses on parallel coordinate plots, scatterplots, and grand tour animations.</td>
<td>Graphical</td>
<td>Yes</td>
<td>All Major</td>
<td>(Wegman and Dorfman 2003)</td>
</tr>
<tr>
<td>CiteSpace</td>
<td>2004</td>
<td>Scientom</td>
<td>A+V</td>
<td>A tool to analyze and visualize scientific literature, particularly co-citation structures.</td>
<td>Graphical</td>
<td>Yes</td>
<td>All Major</td>
<td>(Chen 2006)</td>
</tr>
<tr>
<td>HistCite</td>
<td>2004</td>
<td>Scientom</td>
<td>A+V</td>
<td>Analysis and visualization tool for data from the Web of Science.</td>
<td>Graphical</td>
<td>No</td>
<td>Windows</td>
<td>(Garfield 2008)</td>
</tr>
<tr>
<td>Tool</td>
<td>Author/Year</td>
<td>Type</td>
<td>Description</td>
<td>Interface</td>
<td>Compatibility</td>
<td>Visualization Support</td>
<td>Graphical Support</td>
<td>Requires Programming</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>------</td>
<td>-------------</td>
<td>-----------</td>
<td>---------------</td>
<td>----------------------</td>
<td>-----------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Gephi</td>
<td>Barrat et al., 2004</td>
<td>A + V</td>
<td>An interactive visualization and exploration platform for all kinds of networks</td>
<td>Yes</td>
<td>Windows, Linux</td>
<td>Yes</td>
<td>Windows (any system that supports Java)</td>
<td>(Trier and Bobrik, 2007)</td>
</tr>
<tr>
<td>TINA</td>
<td>Commetrix, 2008</td>
<td>A + V</td>
<td>Commetrix software allows users to dynamically analyze and create rich network maps.</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>(Eck and Waltman, 2011)</td>
</tr>
<tr>
<td>NodeXL</td>
<td>Kronenwett, Gamper, et al., 2010</td>
<td>SocSci</td>
<td>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.</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>(Hansen et al., 2010)</td>
</tr>
<tr>
<td>PlotViz</td>
<td>Choi et al., 2010</td>
<td>Networks</td>
<td>PlotViz is a tool for the visualization of large-scale high dimensional data.</td>
<td>Yes</td>
<td>Windows</td>
<td>Yes</td>
<td>Windows</td>
<td>(Choi et al., 2010)</td>
</tr>
<tr>
<td>TINA</td>
<td>Commetrix, 2008</td>
<td>A + V</td>
<td>Tool for interactive assessment of projects portfolio and visualization of scientific landscapes.</td>
<td>Yes</td>
<td>Windows</td>
<td>Yes</td>
<td>Windows</td>
<td>(Chair et al., 2010)</td>
</tr>
<tr>
<td>VisuLyzer</td>
<td>VOSviewer, 2005</td>
<td>Networks</td>
<td>VOSviewer is a visualization tool useful for analyzing bibliometric networks.</td>
<td>Yes</td>
<td>Windows, Linux</td>
<td>Yes</td>
<td>Windows</td>
<td>(Choi et al., 2010)</td>
</tr>
<tr>
<td>VOSviewer</td>
<td>VOSviewer, 2005</td>
<td>Networks</td>
<td>VOSviewer is a visualization tool useful for analyzing bibliometric networks.</td>
<td>Yes</td>
<td>Windows, Linux</td>
<td>Yes</td>
<td>Windows</td>
<td>(Choi et al., 2010)</td>
</tr>
<tr>
<td>VOSviewer</td>
<td>VOSviewer, 2005</td>
<td>Networks</td>
<td>VOSviewer is a visualization tool useful for analyzing bibliometric networks.</td>
<td>Yes</td>
<td>Windows, Linux</td>
<td>Yes</td>
<td>Windows</td>
<td>(Choi et al., 2010)</td>
</tr>
<tr>
<td>VOSviewer</td>
<td>VOSviewer, 2005</td>
<td>Networks</td>
<td>VOSviewer is a visualization tool useful for analyzing bibliometric networks.</td>
<td>Yes</td>
<td>Windows, Linux</td>
<td>Yes</td>
<td>Windows</td>
<td>(Choi et al., 2010)</td>
</tr>
<tr>
<td>VOSviewer</td>
<td>VOSviewer, 2005</td>
<td>Networks</td>
<td>VOSviewer is a visualization tool useful for analyzing bibliometric networks.</td>
<td>Yes</td>
<td>Windows, Linux</td>
<td>Yes</td>
<td>Windows</td>
<td>(Choi et al., 2010)</td>
</tr>
<tr>
<td>VOSviewer</td>
<td>VOSviewer, 2005</td>
<td>Networks</td>
<td>VOSviewer is a visualization tool useful for analyzing bibliometric networks.</td>
<td>Yes</td>
<td>Windows, Linux</td>
<td>Yes</td>
<td>Windows</td>
<td>(Choi et al., 2010)</td>
</tr>
<tr>
<td>VOSviewer</td>
<td>VOSviewer, 2005</td>
<td>Networks</td>
<td>VOSviewer is a visualization tool useful for analyzing bibliometric networks.</td>
<td>Yes</td>
<td>Windows, Linux</td>
<td>Yes</td>
<td>Windows</td>
<td>(Choi et al., 2010)</td>
</tr>
<tr>
<td>VOSviewer</td>
<td>VOSviewer, 2005</td>
<td>Networks</td>
<td>VOSviewer is a visualization tool useful for analyzing bibliometric networks.</td>
<td>Yes</td>
<td>Windows, Linux</td>
<td>Yes</td>
<td>Windows</td>
<td>(Choi et al., 2010)</td>
</tr>
<tr>
<td>VOSviewer</td>
<td>VOSviewer, 2005</td>
<td>Networks</td>
<td>VOSviewer is a visualization tool useful for analyzing bibliometric networks.</td>
<td>Yes</td>
<td>Windows, Linux</td>
<td>Yes</td>
<td>Windows</td>
<td>(Choi et al., 2010)</td>
</tr>
</tbody>
</table>

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.