
Learning Analytics: Tool Matrix

(Adapted from David Dornan)

One of the challenges with getting started with learning analytics is getting a sense of tools and resources. **For this assignment**, review the structure provided below and:

1. Research learning analytics tools, both open source and proprietary
2. Add tools that you encounter in your search to the table below (the categories provided are sample categories – extend/add on as is warranted). The categories below have been seeded with a few examples to get you started. Add additional categories or
3. Share your updated document on your blog, in ProSolo, in edX forums, or in Google Docs (if you're on Twitter, please share using #dalmooc tag)

The *goals of this assignment* are to introduce you to the range of tools and to begin engaging with tool sets based on intended functionality, analytics activity to be undertaken, and internal or organizational capacity.

Tool (URL)	Description	Opportunities in Learning Analytic Solutions	Weaknesses/Concerns/Comments
Data Cleansing/Integration Prior to conducting data analysis and presenting it through visualizations, data must be acquired (extracted), integrated, cleansed and stored in an appropriate data structure. We will look at this in more detail in Week 2 of #DALMOOC. Given the need for both structured and unstructured data, the ideal tools will be able to access and load data to and from data sources including RRS feeds, API calls, RDMS and unstructured data stores such as Hadoop.			
Pentaho Integration	Pentaho Data Integration (PDI) is a powerful easy to learn open source ETL tool that supports acquiring data from a variety of data sources including flat files, relational	PDI provides a versatile ETL tool that can grow with the evolution of an institutions learning analytics program. For example, initially a LA program may start with institutional	Two concerns with with PDI: 1. Pentaho does not have built in integration with R statistics. Instead Pentaho data mining integration focuses on a WEKA module.

	databases, Hadoop databases, RSS Feeds, and RESTful API calls. It can also be used to cleanse and output data to the same list of data sources.	data that is easily accessible via institutional relational databases. As the program grows to include text mining and recommendation systems that require extracting unstructured data outside the institution, the skills developed with PDI will accommodate the new sources of data collection and cleansing.	2. Pentaho is moving away from the open source model. Originally PDI was an open source ETL tool called Kettle developed by Matt Casters. Since Pentaho acquired Kettle (and Matt Caster), it has become a central piece to their subscription based BI Suite and the support costs are growing at a rapid pace.
MS Excel	The standard spreadsheet for managing data.	Simple to use, huge amount of support, NodeXL template for exploring network graphs, opens and saves to open csv format. I used this to clean up MOOC comment data for my summer dissertation project. I also used it to carry out Euclidean Closeness analysis.	

Statistical Modeling

There are three major statistical software vendors: SAS, SPSS (IBM) and R. All three of these tools are excellent for developing analytic/predictive models that are useful in developing learning analytics models. This section focuses on R. The open source project R has numerous packages and commercial add-ons available that position it well to grow with any LA program. R is commonly used in many data/analytics MOOCs to help learners work with data. We opted for Tableau during week 1 & 2 due to ease of use and relatively short learning curve.

<p>R</p>	<p>R is an active open source project that has numerous packages available to perform any type of statistical modeling.</p>	<p>R statistics strength is the fact that it is a widely used by the research community. Code for analysis is widely available and there are many packages available to help with any type of analysis and presentation that might be of interest. Some of these include:</p> <ol style="list-style-type: none"> 1) Visualization: <ol style="list-style-type: none"> a) ggplot provides good charting functionality. b) googlevis provides an interface between R and the Google Visualization API 2) Text Mining: <ol style="list-style-type: none"> a) tm provides functions for manipulating text including stripping whitespace and stop words and removing suffixes (stemming). b) openNLP identifies words as nouns, verbs, adjectives or adverbs 	<p>Two issues that may be of concern to some universities:</p> <ol style="list-style-type: none"> 1) Lack of Support - only Revolution R provides support for the R product 2) High Level of Expertise Required to Develop and Maintain R. How does a university retain people that have the skill required to develop and maintain R/RevoDeployR.
<p>IBM SPSS</p>	<p>As mentioned above, this is a standard package for statistical analysis.</p>	<p>Fats, massive support, useful visualisations. I was taught how to use this on my masters in web science course and used it on my summer project. It's expensive for</p>	

		non-institutional users, but is easy to get the hang of and doesn't break.	
PSPP	An open source version of SPSS.	I briefly check this out as an alternative to SPSS. It appears to work well enough, but as it currently lacks visualization functionality, I didn't take it very far.	
<p>Network Analysis</p> <p>Network Analysis focuses on the relationship between entities. Whether the entities are students, researchers, learning objects or ideas, network analysis attempts to understand how the entities are connected rather than understand the attributes of the entities. Measures include density, centrality, connectivity, betweenness and degrees. We will spend time in #DALMOOC on these topics in week 3 & 4.</p>			
SNAPP	Social Networks Adapting Pedagogical Practice (SNAPP) is a network visualization tool that is delivered as a 'bookmarklet'. Users can easily create network visualizations from LMS (Blackboard, Moodle, and D2L) forums in real time.	<p><i>Self Assessment Tool for Students</i></p> <p>SNAPP provide students with easy access to network visualizations of forum posting. These diagram can help students understand their contribution to class discussions.</p> <p><i>Identify at Risk Students/ Monitor Impact of Learning Activity</i></p> <p>Network Analysis visualizations can help faculty identify students that may be isolated. They can also be</p>	

		used to see if specific activities have impacted the class network.	
Other Tools for Analysis			
NVivo	The top proprietary app for analyzing unstructured data.	Excellent tool for supporting content analysis, again with a lot of online support available. I used it for keyword in context analysis on the MOOC comments I analysed for my project.	
Networkit 3.3	An open-source Python module for network analysis.	It provide tools for the analysis of large networks in the size range from thousands to billions of edges. Requires some experience of coding in Python.	
Linked Data			
<p>If Tim Berners-Lee vision of linked data (http://www.ted.com/talks/tim_berniers_lee_on_the_next_web.html) is successful in transforming the internet into a huge database, the value of delivering content via courses and programs will diminish and universities will need to find new ways of adding value to learning. Developing tools that can facilitate access to relevant content using linked data could be one way that universities remain relevant in the higher learning sector.</p>			
Ontologies e.g. DBPedia	Ontologies are essentially an agreed upon concept map for a particular domain of knowledge.	<i>Dynamically Deliver Relevant Content</i>	

Protégé	An open source app for building ontologies.	Still in Beta – and, in my experience, it definitely shows it. Despite its limitations this app is probably the only worthwhile tool to use for building variations of OWL and carrying out SPAQL queries. My experience of using it for coursework has put me off this subject for life.	
WikiData	An online, community tools for connecting Wikipedia data and making it machine readable.	Not strictly speaking an ontology builder (as far as I’m aware it doesn’t use standard W3C-approved formats), but it connects in triples, is relatively simple to use, and (in my experience) doesn’t break. There are also a few querying tools out there. A great project that is accessible and aims to engage people in large numbers – essential for building the semantic web.	
<p>Visualization</p> <p>The presentation of the data after it has been extracted, cleansed and analyzed is critical to successfully engage students in learning and acting on the information that is presented.</p>			
Google Visualization API’s (http://code.google.com/apis/chart/)	Google Visualization provides an API to their chart library allowing for the creation of charts and other visualizations. They have recently	<p><i>Interactive Learning Dashboards</i></p> All of these tools are useful for creating visualizations for learning	Learning how to use these tools/libraries requires a fair amount of effort. Developer retention is a

	released an API to add interactive controls to their charts.	feedback systems such as dashboards.	risk for system maintenance and enhancement.
Protovis D3	Protovis and D3 are JavaScript frameworks for creating web-based visualizations. Protovis is no longer an active open source project. It has been replaced by D3.	All of these tools can present data as a heat maps, network analysis diagrams and tree maps. Here's a link to an example dashboard created in D3, presenting university admission data. http://keminglabs.com/ukuni/	
NodeXL	An open source template for MS Excel.	"...displays and analyzes network graphs created from edge and vertex lists stored in an Excel workbook. It is compatible with Excel 2007, Excel 2010 and Excel 2013."	
Immersion	An online network visualization tool.	Developed at MIT, this tool provides network graphs of your email accounts, and some surprisingly accurate clustering.	
36 best tools for data visualisation	Apparently people are interested in visualizing data and the nice people at Creative Bloq have provided a great big list of apps that do this.	Here's that list in full: 01. Dygraphs, 02. ZingChart, 03. InstantAtlas, 04. Timeline, 05. Exhibit, 06. Modest Maps, 07. Leaflet, 08. WolframAlpha, 09. Visual.ly, 10. Visualize Free, 11. Better World Flux, 12. jQuery Visualize, 13. jqPlot, 14. Dipity, 15. Many Eyes, 16. D3.js, 17. JavaScript	

		InfoVis Toolkit, 18. jpGraph, 19. Highcharts, 20. Google Charts, 21. Excel, 22. CSV/JSON, 23. Crossfilter, 24. Tangle, 25. Polymaps 26. OpenLayers, 27. Kartograph, 28. CartoDB, 29. Processing, 30. NodeBox, 31. R, 32. Weka 33. Gephi, 34. iCharts, 35. Flot 36. Raphaël	
--	--	---	--