Visualizing Linked Jazz: A web-based tool for social network analysis and exploration

Matthew Miller
Pratt Institute,
144 West 14th St New York, NY 10011
mmille18@pratt.edu

Jeff Walloch
Pratt Institute,
144 West 14th St New York, NY 10011
jwalloch@pratt.edu

M. Cristina Pattuelli
Pratt Institute,
144 West 14th St New York, NY 10011
mpattuel@pratt.edu

ABSTRACT
This paper intends to demonstrate that recent improvements in web-based technology have made the creation of web-based network visualizations a viable and appealing possibility. We showcase a network visualization tool made for the Linked Jazz Project which utilizes a number of interactive and multimedia features to deliver an engaging experience to a wide audience. We briefly explore the history, benefits, and possible limitations of this web-based approach.

Keywords
social network visualization, linked data, jazz, cultural heritage, web technology

INTRODUCTION
Linked Jazz is an ongoing project that investigates the potential application of Linked Open Data (LOD) technology to enhance discovery and visibility of digital cultural heritage materials (http://www.linkedjazz.org/). The main purpose of the project is to develop methods that explore the applicability of Linked Open Data technology to digital archives of jazz history to expose social and professional relationships between musicians and reveal their community's network.

Supporting research through the visualization and analysis of this network is one goal of the project; however, we felt the richness of the data presented a unique opportunity to develop a tool that could be useful and engaging to a wider audience of jazz enthusiasts, including the general public. To achieve this goal, we developed a web-based network visualization written in the JavaScript programing language. Our visualization (http://www.linkedjazz.org/network) demonstrates what is currently possible for web-based network graphs, and shows new novel possibilities to enhance discovery and visibility of digital cultural heritage materials.

ASIST 2012, October 28-31, 2012, Baltimore, MD, USA.

FROM THE DESKTOP TO THE WEB
Social network analysis is one of the largest and fastest growing fields of research today (Marin & Wellman, 2009). The field's growing popularity has led to the creation of many tools to facilitate data analysis. The majority of these tools are desktop-based software packages. Tools such as Pajek[1] or Gephi[2] are examples of desktop software used for analyzing and visualizing network graphs. These programs excel at network analysis, but are constrained to the individual's computer. Visualizations can be shared through the exporting of static images, yet provide no possibility for interaction with the end viewer.

Early web-based network visualizations introduced the ability for interaction through the development of Java Applet and Adobe Flash-based network graphs. A tool such as IBM's Many Eyes[3] network graph is an example of a Java Applet-based visualization. While the visualization is web-based, it relies on user installed plugins which have recently fallen out of favor (Roberts, 2007) and often provide a disjointed user experience.

Continued advancements in web browsers such as Mozilla's Firefox and Google's Chrome, and specifically their JavaScript processing engines, have enabled the creation of JavaScript-based network graphs. The emergence of JavaScript libraries has facilitated the development of complex visualizations. For our network tool, we made use of d3.js[4], a JavaScript library built by Michael Bostock which implements a force-directed graph utilizing simple constraints (Dwyer, 2009).

DESCRIPTION
The Linked Jazz network graph is a hybrid combination of a visualization and a simple analysis tool (Figure 1). The

[1] http://pajek.imfm.si/doku.php

[2] https://gephi.org/

[3] http://www-958.ibm.com

[4] http://d3js.org/

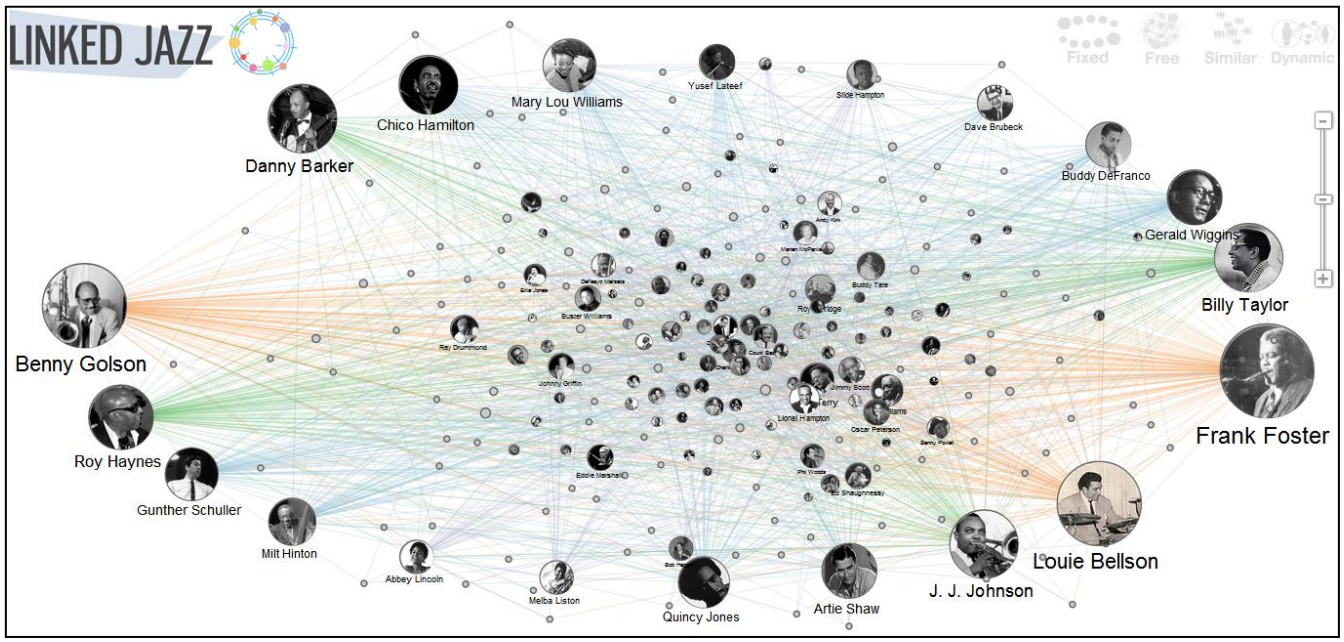

Figure 1. Initial view of the network graph.

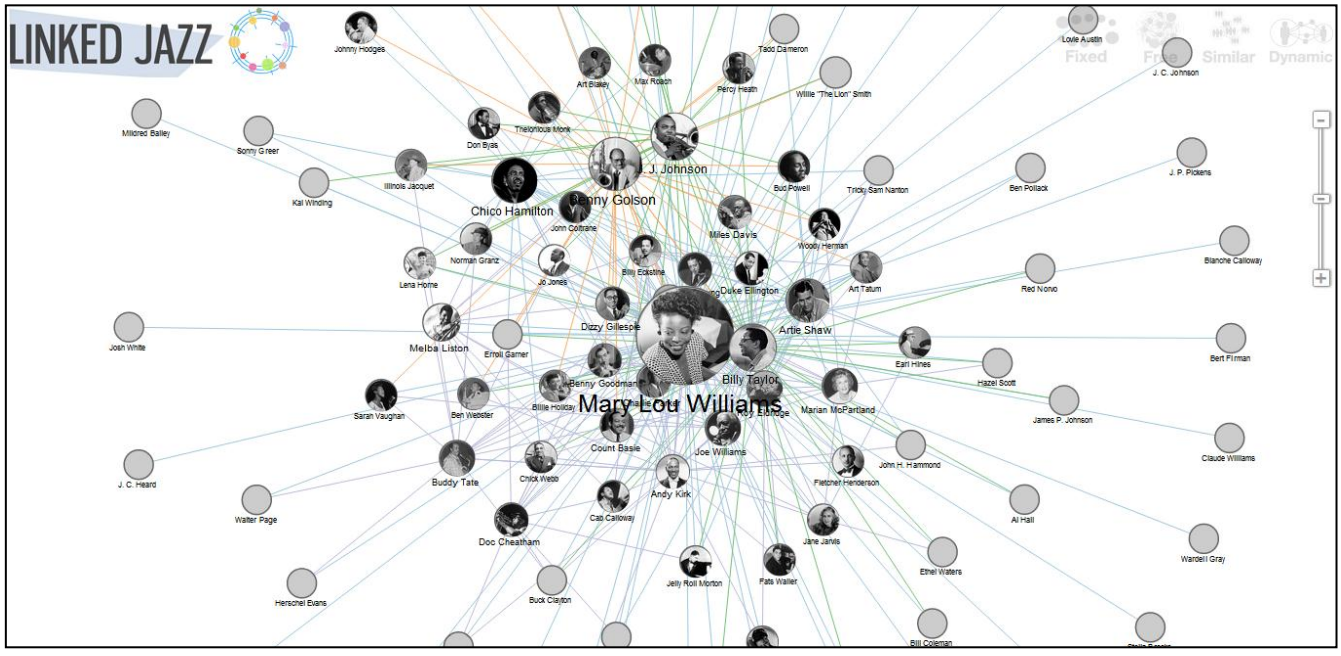

Figure 2. Radial view of Mary Lou Williams.

tool ingests our current dataset in the form of an RDF triple store consisting of about 3,000 triples. All data processing and analysis is done on the client side in JavaScript. Each jazz musician is represented by their photograph within a circular node. Edges are drawn between the musicians' nodes based on their relationships to other musicians. From our current triple dataset, we produce and map about 1,000 nodes and 3,000 edges. The visualization is highly interactive, allowing the user to pan and zoom, isolate individuals and their connections, retrieve short biographical descriptions, and even play a video clip of the musician. The analytical aspect of the tool allows for the user to switch the algorithm used to organize the network. For example, there is a view resulting from the application of a cosine similarity-like-filter which organizes individuals into cliques based on their shared connections. Another analytical function of the tool enables the user to drill down to the individual level and see their cohort and the shared connections within their group (Figure 2). One additional analytical feature allows the user to create their own ad hoc network by dynamically adding individuals to the graph, which they can then use to compare the selected artists' shared or dissimilar relationships.

WEB-BASED BENEFITS

A web-based network graph offers a number of benefits over the static network visualization alternative. The principal benefit is interactivity. Instead of presenting a curated network graph, the web-based version provides a faceted tool to explore the dataset. Another advantage of being web-based is the support for multimedia. The inclusion of elements such as photographs, text, audio, and video transforms the visualization into an engaging cultural experience. This engagement in turn has the potential to attract a wider audience who can glean value from the visualization. An additional benefit is accessibility. Anyone with internet access and a web browser can easily interact with the visualization.

LIMITATIONS

While JavaScript-based visualizations have improved dramatically in recent years, they still present some performance issues. For example, in a modern browser such as Google's Chrome, our tool has no issue rendering a network of 1,000 individuals in a synchronous simulation, when the rendering takes place behind the scenes and then the results are displayed. However, asynchronously, or in real-time, any more than a couple hundred nodes creates a noticeable slowdown. Another limitation is that while libraries do exist to aid in creating these visualizations, significant development time is still required to realize a custom-made web-based network graph.

FUTURE WORK

As the Linked Jazz Project continues to develop, we plan to extend enhancements to the network visualization. A major improvement will be providing a more granular representation of the type of relationships between artists based on an extended version of the Friend of a Friend (FOAF) ontology. For example, the degrees of two individuals' personal and professional relationships will be reflected in our visualization. Other possible features include adding network organizing algorithms focused on musical instruments and geographical facets, as well as a time-based filter.

CONCLUSION

This paper showcases a web-based visualization tool developed to visualize and support the analysis of the network of the jazz artist community, the focus of the Linked Jazz Project. Although web-based network visualizations are still in the very early stages of development, our tool demonstrates the potential of this new generation of applications by introducing elements such as interactivity and multimedia to the traditional network graph.

REFERENCES

Dwyer, T. (2009). Scalable, Versatile and Simple Constrained Graph Layout. *Computer Graphics Forum*, *28*(3), 991–998.

Marin, A., & Wellman, B. (2010). Social Network Analysis: An Introduction. In P. Carrington & J. Scott (Eds.), *The Sage Handbook of Social Network Analysis*, London: Sage.

Roberts, E. (2007). Resurrecting the applet paradigm. *ACM SIGCSE Bulletin*, *39*(1), 521.