

# **UCLA**

## **UCLA Historical Journal**

### **Title**

Revealing the Relationships of the Jazz Community: Review of Linked Jazz

### **Permalink**

<https://escholarship.org/uc/item/6pp8x12n>

### **Journal**

UCLA Historical Journal, 29(1)

### **Authors**

Crandall, Alison  
Accettola, Anna

### **Publication Date**

2018

### **Copyright Information**

Copyright 2018 by the author(s). All rights reserved unless otherwise indicated. Contact the author(s) for any necessary permissions. Learn more at <https://escholarship.org/terms>

Peer reviewed

# Revealing the Relationships of the Jazz Community: Review of *Linked Jazz*

Alison Crandall and Anna Accettola

Linked Jazz is a project of many parts. Part network visualization, part “application of Linked Open Data technologies,”<sup>1</sup> part socio-cultural interrogation, and part open data experimentation—this project covers a variety of areas of interest to Digital Humanities. The overarching goal, however, is to integrate different types of tools with data about the jazz community, not to answer specific questions, but to allow for the possibility of new types of research. Bringing together interviews of famous jazz musicians, connections are found through the artists’ own words. But these connections are limited. Because there is no overarching research question to guide the construction of the project, Linked Jazz is only a framework. Each artist may have their information categorized and be linked to another, but so far no larger pattern or project has come to the fore to benefit from this work. The main detractor from this interesting collection of data is that it only demonstrates the tools available, but does not add to a historical question or conversation.

The attention to jazz music and musicians is an important one for American cultural history. Beginning with the Blues in the 1800s and, then, the blend of 1900s New Orleans sounds which gave birth to jazz, this music has inspired generations and worked its way into the fabric of American history. Artists were cultural icons, political activists, and heroes to people, black and white, local and immigrant, across the nation. How they connect to one another and how they connected to their communities left a deep impact on our culture and it continues into the modern day. This dataset gives a glimpse of that complex world and helps to set it in its historical context—giving voice to the artists themselves and

letting them share the connections that most impacted the way they think about their own history.

The source material for the project is largely based on oral histories and DBPedia, which is a Wikipedia-type of Linked Open Data (LOD). Official archives at the Smithsonian, UCLA, University of Michigan, Rutgers, and Hamilton College provided the original 50 interviews. Then an additional 9300 names were taken from DBPedia.<sup>2</sup> Text analysis tools were used in order to find all possible names through these transcripts and documents. We found this system interesting as it depending on a very small number of archives and is heavily dependent on open access information. As such, the dataset is restricted to the information that was produced during interviews, but is likely not a complete set. The inclusion of other archives, for example the David Brubeck collection at the University of the Pacific in Stockton, California, may be able to illuminate other connections which oral interviews lacked. The Brubeck Collection, in particular, was established specifically “to preserve, promote, and pass on the rich legacy of jazz to future generations. The Brubeck Collection is a unique accumulation of materials representing the creative life of one of jazz’s most renowned practitioners.”<sup>3</sup> Brubeck was a widely traveled and immensely popular musician until his death in 2012. But, in the Linked Jazz Curator tool, he is only ranked a “medium” in terms of information and connections known. Likely, this resource would provide significantly more information if tapped. Moreover, in the end, interviews on their own are restricted to the questions asked and answers received, but cannot possibly contain all possible interpersonal experiences over a lifetime.

On the technical side of the project, linked data plays a crucial role in the construction of this network of information. Once uploaded into databases, the data of the linked networks is outsourced to volunteers. Then, a second analytical tool is put into effect as they use a network visualization tool with “fixed,” “similar,” “gender,” and “dynamic” categories. The fixed and similar modes groups the jazz musicians from the data in the visualization by connections—that is, associations from interviews and documents. “Fixed” shows the musicians with the most connections in a ring around the rest to demonstrate artists with the most activity or popularity. The “similar” view utilizes the same information as the “fixed,” but arranged in groups of shared connections, showing relationship clusters—which may be indicative of artists related by social group or style. The “dynamic” visual allows for any visitor to choose musicians from a list in order to examine specific connections in more detail, a useful concept for anyone who might have a new line of inquiry to examine against the amassed data. Even though these connections are visible, they are just raw data. No theme is presented by the project to shape that information into a larger statement or purpose.

“Gender,” the newest category, was organized once the developers realized that the list incorporated almost entirely men. There they asked a research question: “Could we preliminarily assume that jazz women mention other women

in the context of their lives and careers more often than men in jazz mention women?” Utilizing network visualization, they added new data of female jazz musicians to examine the results. Though no clear answer lies in the still male-dominated visual, the inquiry is ongoing, and will update as new databases and research lines are incorporated. The database, however, generated the question after the visualization was produced. This is in line with the project’s goal of providing opportunities for new research paths, but not creating their own. However, many of the tools of this project do not speak for themselves, but visualize data without giving enough context for a new user. It creates a difficulty in accessing the complex system of networks going on behind the scenes (or shown through the visualization maps).

Several other tools were developed by the project to form a suite of accessible technology for looking at the oral history of jazz musicians. In particular, the “Name Mapping and Curator Tool” provides access to spreadsheets of all the names in the dataset. Though still a prototype, this tool is an interesting, if clunky, means of accessing more specific information. It also is not yet complete. Names are organized alphabetically with links to the appropriate Library of Congress page. They are also divided up by “completeness”—defined as “perfect,” “high,” “medium,” “low,” “many,” or “none.” Again, however, there is little explanation about the purpose of these labels, except where they seem obvious. The difference between “low” and “many,” for example, remains obscure. Furthermore, of the several thousand names, only 21 have been tagged as “Verified.” While this is clearly an ongoing section of the project, a more thorough introduction may be useful to a new user.

One of the overarching themes that we found in this site is the number of incomplete datasets. Much of this seems to be due to the project’s reliance on volunteer contributions. But, we do not see this as a stumbling block for the project, but a boon. While it may take longer to achieve a final project, the use of volunteers opens the project to people of various backgrounds, which makes contributions more dynamic. Quality, also, is not sacrificed, as the project has overarching control over the process of knowledge production. “Edit-A-Thons” are one way that forward progress continues under the project’s control. Reported on the Linked Jazz website on July 21, 2015, “Our recent Women of Jazz Wikipedia Edit-A-Thon at the New York Public Library for the Performing Arts focused on filling the gaps of Wikipedia’s coverage of female musicians. Participants contributed to existing Wikipedia pages, as well as adding several new pages. . .”<sup>4</sup> Such events involve various communities and advance the project, showing the directions which are still available to the dataset and tools.

The tools are continually refined and the project intends for them to be adopted by other initiatives and projects of contexts, even beyond jazz. Individual tools and their various versions are available online through the developer (Matthew Miller)’s GitHub to promote accessibility. The primary concept of the network analysis tool is based on the W3C’s Linked Open Data (LOD) community project,

so similar projects exist, though these rely primarily on existing metadata. Linked Jazz's custom tools have yet to be incorporated by any freestanding project, though they are often used for projects and presentations of a more educational or pedagogical nature. They are also open to different types of educational spheres, outside of academia and inclusive of the larger jazz community.<sup>5</sup>

Linked Jazz also has had an impressive number of publications associated with it from its beginning in 2011. These publications cover topics on people-mapping, LODLAM, and "Linking Networks of Information with Networks of People," but seem largely metaphysical overall. Linked Jazz is a case study for the theoretical application of new analytical tools. But again we come back to the lack of a research questions. Once the tools are shown to have a practical application, they need to go to work. As stated before, the project has become something of a microcosm of how to put data together and apply different tools. They have even made their methods of data production and source material open to all through an interactive and comprehensive website. But, at this point, it comes to something of an end.

As a Humanities project, Linked Jazz has attracted the attention of grants and collaborations for the curation of its content, which will lead to future research in the subject matter. As a project of the Digital Humanities, its development is primarily concerned with the visualization tool itself and the accumulation of data, which is useful, but lacks some of the academic direction desired in mature projects. Since its first release in 2011, the extensive outreach initiatives at professional and educational venues represents the primary success of the system. As the project continues to expand its dataset through archives and collaboration, the potential of the material grows too. As with the gender-based visualization, more questions will arise as new information becomes available to scholars and interest groups alike. In its new initiative with Tulane University, the project seeks to add connections as seen in photographs to the data, allowing for the interpretation of more abstract information and data, in addition to existing metadata.

Overall, this project has built an impressive framework to highlight the many ways in which Digital Humanities can apply itself to datasets. Furthermore, it has done so in a way that opens the information to both the academic and lay communities. The intensive interrogation of a dataset in order to form linked data, visualization graphs, and a workable theoretical basis for research projects provides the foundation for many possible avenues of research. Thus, the stated goals of the project seem to have been met. Unfortunately, it does not appear that the secondary goal of putting this dataset into practical application has yet come to pass. Perhaps the project, once the tools are finished and the connections are verified, will be able to pose research questions driven by internal queries formed through the processing of this data. This seems to be the most practical next step for the project--not only to host the data, but to find a way to integrate into the current conversation.

## NOTES

<sup>1</sup> “Home,” <http://www.linkedjazz.org>.

<sup>2</sup> “Data Sources,” <http://linkedjazz.org/data-sources/>.

<sup>3</sup> “The David Brubeck Collection,” Last modified February 2017, <http://www.pacific.edu/Library/Find/Holt-Atherton-Special-Collections/Brubeck-Collection.html>.

<sup>4</sup> “News,” <http://linkedjazz.org/category/news/page/2/>.

<sup>5</sup> This is evidenced by the project director (Cristina Pattuelli)’s recognition by the Jazz Education Network in early January 2017. “News,” <http://linkedjazz.org/category/news/>.

