Using Linked Open Data to Map Relationships Among Musicians

Bill Levay, Pratt Institute
48th Annual ARSC Conference, May 16, 2014
Overview

• Project goals
• Why Linked (Open) Data?
• What has Linked Jazz done so far?
• Where is Linked Jazz going in the future?
The Project

- Experiment with applying Linked Open Data technology to archival materials in order to enhance visibility and access
- Develop tools to facilitate discovery and analysis of the archives of jazz history
- Make Linked Jazz dataset openly available on the web
Linked Data

From a web of documents...

Current landscape:

- Lots of linked documents
- Databases are silos of information
- User interprets how documents are related
Linked Data

...to a web of data

- Databases are interconnected
- Data is structured in a standardized way
- Relationships between data are made explicit
- URIs point to data  
  <http://dbpedia.org/resource/John_Coltrane>
John Coltrane

From Wikipedia, the free encyclopedia

"Coltrane" redirects here. For other uses, see Coltrane (disambiguation).

John William Coltrane, also known as "Trane" (September 23, 1926 – July 17, 1967),[1] was an American jazz saxophonist and composer. Working in the bebop and hard bop idioms early in his career, Coltrane helped pioneer the use of modes in jazz and was later at the forefront of free jazz. He organized at least fifty recording sessions as a leader during his career, and appeared as a sideman on many other albums, notably with trumpeter Miles Davis and pianist Thelonious Monk.

As his career progressed, Coltrane and his music took on an increasingly spiritual dimension. His second wife was pianist Alice Coltrane and their son Ravi Coltrane is also a saxophonist. Coltrane influenced innumerable musicians, and remains one of the most significant saxophonists in music history. He received many posthumous awards and recognitions, including canonization by the African Orthodox Church as Saint John William Coltrane and a special Pulitzer Prize in 2007.[2]
Linked Data

RDF Triples

subject

<http://dbpedia.org/resource/John_Coltrane>

<http://purl.org/ontology/mo/composer>

object

<http://dbpedia.org/resource/A_Love_Supreme>

composed

<http://dbpedia.org/resource/John_Coltrane>

<http://purl.org/ontology/mo/composer>

<http://dbpedia.org/resource/A_Love_Supreme>
Linked Open Data

Best Practices

1. Available on the web with an open license
2. Available as structured data (formatted, editable)
3. Use non-proprietary formats (CSV instead of Excel)
4. Use URLs to denote things so people can ‘point’ to it
5. Link your data to other data and provide context
Linked Jazz, so far

• Get the names of musicians
• Find the names in oral history transcripts
• Describe the relationships and map them

First phase was funded through an OCLC/ALISE Library and Information Science Research grant.
Linked Jazz Name Directory

- **Data processing:**
  Extract and ingest names from DBpedia and name authority files (LC & VIAF) via Python scripts

- **Human curation:**
  Manually refine results with the Curator Tool

- **Resulting Linked Jazz Name Directory:**
  ~9,000 individuals represented by triples, e.g.:

  `<http://dbpedia.org/resource/Mary_Lou_Williams>`
  `<http://xmlns.com/foaf/0.1/name>  "Mary Lou Williams"`
Curator Tool

<table>
<thead>
<tr>
<th>Name</th>
<th>Years</th>
<th>Verified</th>
<th>Approved</th>
<th>Flagged</th>
<th>Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big Joe Williams</td>
<td>1903-1982</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Claude Williams</td>
<td>1908-2004</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Fess Williams</td>
<td>1894-1975</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Harold Ivory Jr.      Williams</td>
<td>1949-2010</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Joe Williams</td>
<td>1918-1999</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Johnny Williams</td>
<td>1905-1984</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Mary Lou Williams</td>
<td>1910-1981</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Paul Williams</td>
<td>1915-2002</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Richard Williams</td>
<td>1931-1985</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Robert Pete Williams</td>
<td>1914-1980</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Skippy Williams</td>
<td>1916-1994</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Spencer Williams</td>
<td>1889-1965</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Tony Williams</td>
<td>1945-1997</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>Bob Wills</td>
<td>1905-1975</td>
<td></td>
<td></td>
<td>III</td>
<td></td>
</tr>
</tbody>
</table>

Mary Lou Williams (May 8, 1910 – May 28, 1981) was an American jazz pianist, composer, and arranger. Williams wrote hundreds of compositions and arrangements, and recorded more than one hundred records (in 78, 45, and LP versions). Williams wrote and arranged for such bandleaders as Duke Ellington and Benny Goodman, and she was friend, mentor, and teacher to Thelonious Monk, Charlie Parker, Miles Davis, Tadd Dameron, Dizzy Gillespie, and many others.

http://id.loc.gov/authorities/names/n82025133
altLabel | Scruggs, Mary Elfrieda, 1910-1981
prefLabel | Williams, Mary Lou, 1910-1981
Transcript Analyzer

- Oral history transcripts from the Smithsonian, Rutgers, Hamilton College, University of Michigan
- Automated named-entity recognition enabled through the use of natural language processing
- Splits interview content into question-and-answer segments, later used in the Linked Jazz 52nd Street crowdsourcing tool
Transcript Analyzer

Transcript:

Going with Bubbles. He'd come to Pittsburgh and he'd say, "Now I'm going to teach you some of my things." Buck was a great teacher for me. We'd be in the Subway jamming and he'd say, "Hey, look, girl, you're like me. Come over here. I'm going to teach you something." He showed me how. Later on after I met Count Basie and Art Tatum, Buck showed me a run that Art Tatum - it was his famous run. He made it from top to bottom and Buck had taught me that run. He said, "Now when you get in front of Count Basie and Art Tatum, you make it like this," and he turned it around. One night I was jamming with Art Tatum. Art played and I was playing so I played this run that Buck had taught me and Art Tatum said, "Hey, what's that? Play it again." For about 3 or 4 years he and Count Basie were making this run the wrong way. That's the type of musician Buck was. He'd show me something but he didn't want me to play it when they were around, 'cause he didn't want them to learn it. They came around and would hear me turn it around and make it some other way.
52nd Street Crowdsourcing Tool

Let's get started!

- choose a musician
- explore the conversation
- uncover relationships
- grow our network

We need your feedback, tell us what you think
It was beautiful to see the expression on you face and Osie’s and Hank, when the new music was put in front of you. I mean like “ayyy, yeah, hey look what we got,” you know. Wow, like it was like somebody bought you a new gift. Cause Jimmy Jones was writing then, and...

Oh, yeah, well he was contractor, you know, Jimmy Jones. We could really read music and that was the thing that I thank my mother this very day for that, because I wanted to play everything, and she says "you got to learn to read that music." And that’s what we could do, we could read. And young Cannonball Adderley’s on the boat here, and yesterday on the Meet the Stars he was telling some people about a picture I have, I have a picture I took of Cannonball, sitting down with a stack of music and it’s falling down on the music stand, all over the floor. And he says "I got that picture of yours." I say well “that was on a date with Barry Galbraith on guitar, Art Farmer on trumpet, Cannonball Adderley on saxophone, and yours truly on bass. No piano, no drums. I said "it’s the hardest record date I think I ever made in my life." John Benson Brooks wrote a thing called "Alabama Suite" and [inaudible] off everybody. And Cannonball Adderley was sitting there looking at that music, and he played every note on it, it was so beautiful.
Network Visualization Tool
Network Visualization Tool
Network Visualization Tool
What’s Next?

• Women in Jazz
  • Rutgers Women in Jazz project
  • Wikipedia Edit-a-Thons
What’s Next?

- Women in Jazz
- Tulane University Hogan Jazz Archive Photo Collection
Mahalia Jackson

View Description
What’s Next?

- Women in Jazz
- Tulane University Hogan Jazz Archive Photo Collection
- Freely available jazz discographies: Columbia University’s J-DISC
God Bless The Child

Song:
God Bless the Child

Playtime: 2:33

Session Information

Session: Count Basie August 31, 1950
Order: 2
Track Production Status: Master

Session Personnel (complete session personnel)

Clark Terry (trumpet)
Buddy DeFranco (clarinet)
Wardell Gray (tenor saxophone)
Count Basie (piano)
Freddie Green (guitar)
Jimmy Lewis (bass)
Gus Johnson (drums)

Track Personnel Changes (variances from session personnel)

Billie Holiday (vocal)
What’s Next?

• Women in Jazz

• Tulane University Hogan Jazz Archive Photo Collection

• Freely available jazz discographies: Columbia University’s J-DISC

• Mashups with other LOD sets — Musicbrainz, WhoSampled.com, etc.
What’s Next?

• Women in Jazz

• Tulane University Hogan Jazz Archive Photo Collection

• Freely available jazz discographies: Columbia University’s J-DISC

• Mashups with other LOD sets — Musicbrainz, WhoSampled.com, etc.

• Archival metadata from EAD finding aids
What’s Next?

• Women in Jazz

• Tulane University Hogan Jazz Archive Photo Collection

• Freely available jazz discographies: Columbia University’s J-DISC

• Mashups with other LOD sets — Musicbrainz, WhoSampled.com, etc.

• Archival metadata from EAD finding aids

• Redesign tools based on user feedback
You Can Help

• Explore our network maps and our tools at linkedjazz.org

• Use our 52nd Street Crowdsourcing Tool and give us feedback

• Use our data: linkedjazz.org/api

• Suggest other openly available data we could mash up with our dataset
Thanks for listening!

Visit linkedjazz.org for more information.

Many thanks to Dr. Pattuelli and the entire Linked Jazz team.

Bill Levay | wjlevay@gmail.com | wjlevay.net
References


Image Credits

• John Coltrane

• “A Love Supreme” label

• Linked Jazz screenshots
  http://linkedjazz.org