J-DISC

Sustainability Planning

June 2013

Table of Contents

Introductory Note

Ithaka S+R has worked with the J-DISC team since July 2012 on various stages of research and analysis aimed at developing a clear path for the sustainability of this project, an online discography of jazz and jazz musicians, fully cross-searchable by name, instrument, recording session, published record ("issue"), and many more fields not available or searchable elsewhere. J-DISC stands apart from other competitors in the field for its devotion to including carefully vetted metadata that permits the user to research by fields that most commercial services ignore, but that scholars and fans deeply value.

As work began in mid-2012, it became clear that before considering the long-range funding model for the project – ideas including creating a subscription service, or licensing metadata for use by other online services – J-DISC's primary challenge would be to identify and incorporate a sufficiently large and coherent set of discographic data into the database, to create a product that would be eagerly adopted by an enthusiastic audience of users.

As a result, after discussion with the J-DISC team, we decided to focus on strategies developing a content strategy and user needs, rather than pursue revenue generating options at this time. Early phases of research included a landscape assessment and review of competitive products and platforms, as well as an audience assessment. In addition, we offered a usability review, to provide feedback on user interface and navigation, both of which will be critical issues regardless of the funding model pursued. More recently, we have focused on helping the team to consider alternative content development strategies that could support their effort to first build the collection to reach critical mass. Based on that work, we have sketched out various alternative structures that could offer the potential for ongoing support.

Finally, as a means to develop a deep base of support for the project within the academic jazz community, we have recommended that J-DISC seek to strengthen its collaboration with different types of partners: *content holders* (at universities, archives, libraries, and private collections) who might find shared interest in having J-DISC ingest metadata for the content they hold; with *MIR*, the music analysis research effort now underway, that might offer some innovative avenues to enhance the reesarcher's experience working with J-DISC's discographic data; and finally and most important, with *institutional partners*, with a similarly strong devotion to jazz studies, who might prove to be strong allies in building the database, by offering their editorial guidance, as well as some labor from scholars and students.

This document includes the main deliverables for this work, including:

- **The Final Report**, a documented created to guide a group discussion with the J-DISC team and Ithaka S+R

    - **Content Strategy** addresses the challenges and some opportunities for shaping and building the database
    - **Sustainability Directions** outlines some possible avenues for long-term support

- **Research phases**, which includes documentation from earlier phases of our work, provided here as reference

Final Report

Content Strategy

Defining a content strategy has been a key step in allowing J-DISC to strategically develop the database and has involved three main areas: (1) defining the scope of content (2) helping to assess various types of data sources for their availability and usefulness and (3) ask what are the best alternatives to quickly and accurately ingest this material into the database. Over the past several months, Ithaka S+R has worked with J-DISC to clarify all three elements of this strategy. This memo serves to summarize our findings up to this point, and to indicate potential next steps for J-DISC to take in order to obtain the quality and quantity of content it needs, both for the current grant and into the future.

(1) Content Scope

While the metadata fields that are the heart of the J-DISC enterprise had already been determined by J-DISC leaders and their expert advisory committee before Ithaka S+R began its work, initial discussions with Project Manager Tad Shull and his team surfaced the need to clearly define the scope of the number and character of the individual records ("sessions") that would be included. Discussions with the project team suggested developing a content strategy that might be more clearly defined and serve as an editorial roadmap both for the current grant phase and other future stages, as well.

The new content strategy we have worked out with Tad and his team is to focus on incorporating the complete recorded output of 75 key artists whose activity ranges from 1923 to the present, rather than to attempt a random cross-section of every jazz recording regardless of its influence or quality. We and the J-DISC team believe this more focused approach will be more useful and valuable to discographers, jazz historians and theorists, librariansand educators. It is also more feasible, putting a sufficiently broad and viable corpus of data within reach in the near term. Assuming an average of 100 sessions for the 75 artists, the target for the project is then to complete 7500 sessions.   J-DISC believes that by reaching this content target, the database will be able to represent a respectable and substantial discographic source.   At the current data entry rate, J-DISC is hoping to complete 2500 entries by June 2014, about one third of the targeted amount, which should be reached in the years after that date. The artists chosen include 25 considered 'canonical' and another 50 considered 'representative', allowing the database to provide some diversity of culture, background, gender and era/subgenre. (See artist list in Appendix A).

The logic of this approach is to start off with a core of major artists within the original time frame of 1923 to the present, by including those likely to be in greatest demand in terms of record sales, scholarly and discographic treatment, or educational value.

Next steps
There are several possible ways that J-DISC could continue to develop its content base. Below are some first thoughts on this that will be further informed by ongoing discussions about potential content partnerships.

1. Consider using the same "iconic/representative" approach for adjacent time periods.
2. Continue building out sessions by expanding the network from the center out. One way to do this would be to start with a session entered because the lead player was considered "iconic" and build out the discographies of the sidemen included there.
3. Start layering in special collections (eg Record company archives; recording studio archives, etc..) as opportunity permits.
4. Expand the remit to include living artists, who could enter their own discographic data. This could be an interesting avenue to pursue, and could offer a significant boost in usership, and possibly in revenue generation. However, it would also require some additional planning (and development) to insure that it does not interfere with the vetted content on the site.
5. Whatever editorial framing is chosen, sharing at least the broad outline of the "content roadmap" with users should be part of the communications strategy. Communicating information about progress and future direction of content acquisition can be as simple as periodic announcements like: "In June 2013 J-DISC will add 50 new sessions of John Coltrane" or can be a broader statement like: "J-DISC will launch with discographic information about ## canonical jazz musicians, and will add ## musicians each year."Sharing the plan can serve both to drive people's interest, and to help manage expectations.

(2) Content Quality, Data Sources and Vetting Process

Content Quality
J-DISC seeks to differentiate itself by providing better-quality content than that of other existing jazz discographies, and our interviews with difference audience segments support this. As such, J-DISC cannot indiscriminately add data from other sources without careful review of the content, nor can it rely on a single resource to provide all of the information desired on an artist or a session. For any given recording artist, the project team has to turn to several available sources and through a multifaceted process essentially decide what information it would like to include.

At present, J-DISC's data sources include print discographies, hard copies of record albums, online discographies, artist websites, and other recording-related websites such as WorldCat or All Music Guide. (For a sample list of sources currently used and trusted by J –DISC see Appendix B). When completing a session in J-DISC, the individual entering the data has to look into several of the sources and decide what information to trust. For example, if an artist is part of a respectable print discography (i.e. The Blue Note Label" by Michel Ruppli), approximately

¾ of the information that goes into a single "session" can be found and entered into the database, but the remaining information needs to be dug out from the other sources such as the online discographies or websites dedicated to artists. For the most part, available online discographies like the Lord Jazz Discography, a commercially driven omnibus site,cannot be fully trusted on their own and thus multiple sources to better cross reference the information are still necessary.

Data Sources
Given the range of data sources, and methods of data entry, Ithaka S+R has suggested that there might be ways to identify primary source content collections that would offer valuable and accurate data for J-DISC. These might include documentation about sessions from record labels or recording studios; they might be album covers or liner notes. If such data were available, it might offer a strong selling point for the accuracy of the data in J-DISC, and an additional "hook" for promoting the content of the collection.

Additionally, if there were special collections or archives that included content rich in discographic data, there is the possibility that J-DISC could partner with those institutions to apply for grant funding to digitzed and catalog those collections.

An important first step here would be to identify any relevant collections that both fit the time/subject remit of J-DISC, and to begin discussions with the relevant archivists, directors of the collections. Some of the collections may be known to J-DISC directors and program manager; others could be identified by desk research and interviews with scholars and archivists in the field. Obtaining data directly from existing record labels appears to be a long-shot, due to their tendency to hold information closely and their evident decline; this could, however, be another approach.

Vetting system/quality control
The exact process and categories are yet to be defined by Tad and his team, nonetheless, some type of editing system will be necessary in order to maintain the accuracy of the information as the database continues to grow.

Tad has suggested setting up a way for users to know where in the review pipeline a session is. First, records will be identified in the system as either published, or unpublished, for those awaiting editorial review. When a record has been entered, but before it has been vetted, the default will be "unpublished" and it will not appear on the site. Once a session is vetted, an editor officially "published" to the site, it becomes visible to users. Then the record can be identified with these markers, to signal its stage in the process to users of the website:

- **Draft**: Sessions/issues populated with the best knowledge of the Jazz Studies Center and up for review by other individuals and by the experts. It should be understood that the recording session or commercial issue may not be complete or perfectly accurate and could thus use further information from third parties

- **Peer-reviewed**: Sessions/issues reviewed and approved by a subject authority on the artist or an editing committee. Such data is therefore confirmed to be valid and accurate.

Next steps:
1. The J-DISC team is researching potential content partners, based on collections that can help to more quickly build the database**.**

(3) **Data input strategy**

J-DISC's current content strategy consists of having staff manually entering the artists' sessions and their respective fields. According to Tad's estimate, it takes approximately one hour for a trained assistant to enter one complete single session. The sources used to complete the information are multiple and depend on the artist. Additionally, the process requires an expert to review the information and make critical decisions when for example, the required information cannot be found or the information found is contradictory. Tad has taken this editorial role and has been reviewing the data entered. He has also been training the assistants on how to enter the data onto J-DISC and where to get the information from. Going forward, we see a benefit for using graduate students with knowledge on the field performing the data entry task, as their academic training may make them better suited to more accurate/rapid creation of the records.

Currently there are two part-time assistants with strong data management experience populating the required fields with in the sessions, and the team has managed to complete 1450 sessions to date. Their current data entry rate depends on having assistants work close to 14 hours a week for an approximate of 40 weeks per year. At the pace of one hour per session and with two assistants, it would take J-DISC 5+ years to complete the project's goal. It is obvious that J-DISC must look at alternative ways in which it can faster and strategically expand its database.

For most of the sessions, not all fields or data point are able to be completed, based on J-DISC's commitment to provide extremely robust metadata, including biographical, repertory and production of recorded materials. Most commonly the person entering the data is able to track down close to 90% of the information required. The fields and data entries for each session are made to be very extensive and cover every aspect that went into a single session including and not limited to, musicians, engineers, locations, times, instruments. As the project has progressed, the team has found that this information is not made available or was not documented.

Next steps: In order to more quickly build the database, there are some possible tactics to consider:
1. **Identify potential institutional partners who can add some time to this effort.** In its early stages, J-DISC benefitted from a group of expert advisors, who have provided an ongoing source of advice and support throughout the project. That said, many of their contributions hav been as individuals, offering guidance on specific aspects, as needed by the J-DISC team. Going forward, given the opportunities that J-DISC could offer to music faculty and students, J-DISC could consider identifying a more permanent group

of institutions able to play a more significant advisory role. In addition to having access to additional support and buy-in from the academic jazz community this strategy could potentially provide expertise for editorial vetting from faculty as well as time for data entry from graduate students. (See Appendix C for a list of possible institutional partners.)

2. **Explore a crowdsourcing effort** that could help ingest the most basic of the data. At this time, reaching out to the "general public" of jazz enthusiasts is not advisable, until the specific tasks for crowdsourcing and editing are further defined. Still, once that is done, there might be a way to encourage people to help enter some basic discographic data. (For more detail on the sort of crowdsourcing efforts that might work best, see Appendix D)

Please see Appendix F for a Summary of the Content Strategy section

Sustainability Directions

At this stage of development, there are three main directions for J-DISC to pursue: (1) **partnerships** of different sorts can be valuable in (more quickly) advancing the project to reach critical mass, and also to solidify support for the project among those who will be its strongest supporters; (2) **additional grant funding** may be available, for projects that J-DISC can lead or otherwise participate in and (3) **revenue generation, which at this halfway mark in the current grant,** appears to be too early to pursue, whether subscription, sponsorship or licensing. Until the database has reached its target on content of 7500 entries, it would be difficult to consider implementing any of these strategies Nevertheless, we have included some of the ideas discussed in the course of this engagement, suggesting possible conditions that would need to be met and possible next steps to pursue, in order to get there.

1. **Seeking partnerships and other long term support**

In earlier phases of work, Ithaka S+R conducted research with several types of potential users of J-DISC. This confirmed that the most loyal source of support for the project is likely to come from those who are also deeply committed to the study of jazz and its practitioners: music scholars and students and music librarians. That said, it was also noted that the field is far from unified: even among those who have supported J-DISC in the past, the depth of their sense of commitment to the project was not always clear and we know that some discographers have been less than willing to actively share their IP with the project. And yet, securing the good will of the community will be important piece for the longterm success of this project.

At this point, we are not making a recommendation for the specific form this partnership would take; we are just recommending that active steps be made to explore how J-DISC can work more closely with the "community" of jazz scholars. Some potential values of this approach include:

- Support from other institutions with significant interest in jazz will lend further academic prestige, peer review, and projection of quality to the project
- Having a valuable sounding board and source of expert advice on editorial direction of J-DISC. This could mean offering suggestions for which artists and eras to cover in subsequent rounds, based on the courses being taught and research interests at those schools
- Offering opportunities to develop partnerships around funded projects, which might result in more metadata for J-DISC (see below)
- Offering a means for sharing information about the project with the wider jazz community, through the shared contacts of the partners
- Developing a basis of vital, direct support for the aims of the project, including potentially offering in-kind support in the form of labor for data entry and editing activities

Next steps would include identifying potential institutions to partner with, based on the reputation , areas of expertise, and focus on jazz education and research; planning and convening a meeting with these potential partners; using the meeting as a means to seek their input, and to sound them out on the possibility of other forms of support.

2. Pursuing further grant opportunities to build content and audience engagement

While we hope that ongoing grant support will not be the only element in the sustainability plan, J-DISC certainly offers several possible avenues to pursue. Below are some ideas that would be worth further exploration:

- **Innovation**. Several funders focus on the innovative advances that digital projects in the humanities can offer. Certainly this is the case for Mellon. Another possibility could be NEH's Office of Digital Humanities, which has recently launched a new grant track called "Digital Implementation Grants." In addition the close association with the MIR project might be useful here as well and there may be ways for MIR to develop a proposal that supports their more experimental work, while also resulting in added content and/or functionality for J-DISC. (See Appendix F for an update on the MIR project and for guidance as to the ways we believe MIR can enhance J-DISC)
- **Engagement.** There have in the past been some grant programs to offer funding to digital projects in the humanities and heritage fields, to encourage user/audience engagement.
- **Content partners**. Some grants (NEH, Preservation and Access, IMLS) could offer a means of partnering with a key content holder, who seeks to digitize a special collection or archive. Ithaka can imagine J-DISC being written into such a grant, as part of the sustainability plan, with detailed metadata aligned with that of J-DISC as part of the grant. This could also work for just a straight cataloging grant, like the Hidden Collections program funded by Mellon and administered by CLIR: http://www.clir.org/hiddencollections/index.html. As discussed earlier, identifying some content partners like this could both serve as a funded means to more quickly upload content into J-DISC.
- **Institutional partnerships.** While partnerships with institutions and jazz research centers could offer a variety of benefits (see above), it might not be so easy to secure the needed number of graduate students devoted to entering the metadata that J-DISC will require now and into the future. If institutions have access to these positions and choose to devote them to this task, that would be great. If not, it may be possible to seek grant funding in the form of post-doc positions, to build digital capacity in libraries. Some places to examine concerning programs like this would include IMLS, NEH and Mellon.

> **Next steps** would include further research into specific grant programs, to assess suitability and begin to frame the logic for a proposal.

3. Earned income

As mentioned above, J-DISC in its current form is not yet ready to move ahead with any revenue generating strategies. That said, there are certain types of revenue generation that appear to be a good fit for the content and service that J-DISC will eventually offer. Below are short outlines of these ideas, along with notes on what potential next steps would be worthwhile in the future to explore them further:

a) **Subscription model**: This model would involve packaging J-DISC as an online scholarly resource that could be licensed to institutions of higher learning for a monthly or annual subscription fee. The benefits of this model are substantial: developing a reliable customer base

can alleviate the need to search for new sources of support each year. But getting there is not easy: fee services undergo a great deal of scrutiny from customers (in this case, librarians), and need to have very strong offer on the table, both in terms of the value of the content, the need of the students/faculty for that content, and the ease of the user interface, not to mention a price point that is seen as fair and reasonable. Even in a best-case scenario, however, "music" tends to be a niche category, and "jazz discography" is a niche of a niche…meaning that the direct audience for this as a scholarly resource may be fairly small.

That said, there are examples of niche products that find a loyal base of supporters. Before J-DISC can consider this approach it will need to solidify its offer by:

1. Defining its content roadmap now and for the future, at least in clear conceptual terms.
2. Reaching the content target of 7500 so that users are extremely likely to find what they are looking for.
3. Investing in user interface upgrade, based on our recommendations and other user feedback the J-DISC team may choose to solicit.

Then, next steps would include conducting a pricing review of comparable resources, interviews or further review with music librarians/future customers, and revenue modeling, based on market research to size the market.

b) **Metadata licensing**: There is a great deal of activity in services that sell and stream music (iTunes/Apple, as well as the internet radio services Pandora, Spotify, Rhapsody and surely others) most do a terrible job with metadata for jazz, offering just the shallowest set of options for searching and browsing the sizable collections they have. As these services have integrated business models, they have a strong interest, in theory, to encourage people to explore and find the music they like in the most effective way possible. It could be that with an excellent and gold-standard of jazz metadata, that some music service could benefit from licensing the right to use that metadata to fill out its records for the audio files it has. Rhapsody has recently announced that it intends to include liner notes in order to better credit contributing artists, and others may well follow.

Before J-DISC can consider this approach, it will need to:

1. Further develop its database, so that it contains the coverage a commercial vendor is going to want to have. The approach based on 75 iconic artists is a good one for this purpose.
2. Confirm that the fields J-DISC has coverage of the artists, time periods and the main fields these vendors would need.
3. Work with CU IP office to confirm that licensing the metadata would not be in conflict with any IP regulations, given the sources of the original data.

> **Then, next steps would include** further research to understand how far along Rhapsody and others are in identifying and inputting this metadata already. They may find a much quicker and cheaper way to do this; it might not be as high-quality, but still be considered an improvement over what they have now. Meetings with people at the major services would serve to introduce the value that having a constantly updated, accurate database of jazz recordings would offer.

c) **Premium service for musicians:** During our phase of user research, we were interested to hear that often professional musicians do not have a single place to offer their OWN discographies, and this suggested another means of growing the database from the present day forward. While this user-generated content would pose some challenges to the J-DISC mission, it could increase overall usage of the site, and generally drive traffic. It also suggests the possibility of making this a service for musicians to manage their own recording history.

Before J-DISC can consider this approach, it will need to:
1. Determine if this is a comfortable fit, alongside the rest of the vetted content in J-DISC.
2. Develop ways to signal to users which records are self-created, and which are peer reviewed.

> **Then, next steps would include** further competitive scan of the other services offering this; interviews with musicians who might use the service; determining what new interface and display would be needed to support this new function; cost and revenue projections for this new service.

Appendix A: Priority List of Artists

The current content strategy involves focusing on a handful of artists considered either "canonical" or "representative" within the agreed upon time frame of ninety years. The following list was developed by Tad and the J-DISC team.

Canonical

Albert Ayler
Art Ensemble of Chicago
Art Tatum
Benny Goodman
Bill Evans
Billie Holiday
Bix Beiderbecke
Cecil Taylor
Charles Mingus
Charlie Christian
Charlie Parker
Coleman Hawkins
Count Basie
Dizzy Gillespie
Duke Ellington
Fletcher Henderson
Horace Silver
Jelly Roll Morton
John Coltrane
Keith Jarrett
Lester Young
Louis Armstrong
Miles Davis
Ornette Coleman
Sidney Bechet
Sonny Rollins
Sun Ra
Thelonious Monk
Wes Montgomery

Representative

Abbey Lincoln
Albert Mangelsdorff
Albert Nicholas
Arnett Cobb
Art Blakey
Benny Carter
Bill Dixon

Billy Higgins
Blue Mitchell
Cannonball Adderley
Chet Baker
Chris McGregor
Clifford Brown
Clifford Jordan
Dave Brubeck
Dexter Gordon
Don Byas
Don Patterson
Dorothy Donegan
Edmond Hall
Elvin Jones
Eric Dolphy
Gerald Wilson
Grant Green
Harold Land
Harold Mabern
Jackie McLean
Jimmy Heath
Jimmy Lyons
Joe Harriott
Joe Henderson
John Carter
John McLaughlin
Kenny Dorham
King Oliver
Krzysof Komeda
Lennie Tristano
Lucky Thompson
Mary Lou Williams
Melba Liston
Nat King Cole (piano)
Oscar Peterson
Paolo Moura
Pat Metheny
Pharaoh Sanders
Roland Kirk
Sarah Vaughan
Sidney Bechet
Tommy Flanagan
Tony Williams
Wardell Gray
Wayne Shorter

Appendix B: List of data sources to be used by J-DISC

The following data sources are among those that J-DISC is currently using to create the session entries in J-DISC.

All Music Guide. Allmusic: the Online Destination for Music Fans. Rovi Corporation. Web resource. <http://www.allmusic.com> June 22, 2012.

Allen, Walter C. Hendersonia: The Music of Fletcher Henderson and his Musicians. Highland Park, New Jersey: Walter C. Allen, 1973.

Bruyninckx, W & L. 93 years of recorded jazz & blues (1917-2010) [CD-ROM]: (prejazz 1897-1917), A to Z complete. Mechelen, Belgium: W. & L. Bruyninckx, 2010

Cecil Taylor Research Group, Cecil Taylor Discography<http://www.webmutations.com/ceciltaylor/index.html>

Cohen, Noal. Noal Cohen's Jazz History Website. Web resource. <http://www.attictoys.com/> June 22, 2012.

Cohen, Noel & Michael Fitzgerald. Rat Race Blues: The Musical Life of Gigi Gryce. Berkeley, California: Berkeley Hills Books, 2002.

Coltrane, John. The Heavyweight Champion: The Complete Atlantic Recordings. Atlantic Recording Corporation, 1995. CD Box Set.

Connor, D. Russell & Warren W. Hicks. BG On the Record: A Biodiscography of Benny Goodman. New Rochelle, New York: Arlington House, 1969.

Cuscuna, Michael & Michel Ruppli. The Blue Note label: a discography. (New York: Greenwood Press, c1988).

Dapogny, James et. al. Jelly Roll Morton <http://www.doctorjazz.co.uk/page15.html>

Lord, Tom. The jazz discography. (West Vancouver, B.C. Lord Music Reference). Licensed web-based resource at <http://www.lordisco.com/> [Available in print and CD-ROM formats.]

Fitzgerald, Michael & Steve Albin. Jazz Discography. Web resource <www.jazzdiscography.com>. June 22, 2012.

Jazz.com's Jazz Encyclopedia of Jazz Musicians. Web resource. <http://www.jazz.com/encyclopedia> June 22, 2012.

Hazevoet, Cornelis J. Don Byas, Part 2: European Recordings 1946-1972. Hazevoet, 2011. Web resource. <http://www.paristransatlantic.com/magazine/interviews/hazevoet.html> July 29, 2012.

Mingus, Charles. Passions of a Man: The Complete Atlantic Recordings 1956-1961. Atlantic Recording Corporation, 1997. CD Box Set.

Porter, Lewis ed. The John Coltrane Reference. (New York: Routledge, 2008).

Johnson, J. Wilfred. Ella Fitzgerald: an annotated discography; including a complete discography of Chick Webb. (North Carolina: McFarland & Company, Inc., 2001).

Raben, Erik, et al. Jazz Records 1942-1980: a discography. (Copenhagen, Denmark: Stainless/Wintermoon, c1980-c1995) v.1-8.

Ruppli, Michel & Bob Porter. The Clef/Verve labels: a discography (New York: Greenwood Press, 1986).

Ruppli, Michel & Bob Porter. The Prestige label: a discography (New York: Greenwood Press, 1980).

Ruppli, Michel & Bob Porter. The Savoy label: a discography (New York: Greenwood Press, 1980).

Sheridan, Chris. Dis Here: a bio-discography of Julian "Cannonball" Adderley. (Connecticut: Greenwood Press, 2000).

Szwed, John. So What: The Life of Miles Davis. London: Arrow Books, 2002.

Vanwambeke, Joel. Bix Beiderbeck <http://bixography.com/recordingsjoel/index.html>

Togashi, Nobuaki, Kohji 'Shaolin' Matsubayashi & Masayuki Hatta. The Jazz Discography Project. jazzdisco.org. Web resource <http://jazzdisco.org>

WorldCat. Web resource. <http://www.worldcat.org/>

Appendix C: Potential Institutional Partnerships

The following institutions were suggested by the J-DISC team, and could potentially be contacted to discuss possibilities for working together.

Name	Description	Potential Interest in J-DISC
Jazz Institute Darmstadt	-Well-known resource for jazz research materials, particularly text-based -Resource for discography -Reaches European, German-speaking audience	-Interested in new media metadata standards for jazz -Interested in metadata for sessionography -Interested in classroom new media
Centro Nazionale Studi Sul Jazz—Arrigo Polillo inSiena, Italy	- Archive on jazz, with emphasis on Italian jazz or Italian works on jazz -Francesco Martinelli, Director, is a J-DISC panelist	-Interested in new media metadata standards for jazz -Interested in metadata for sessionography—Martinelli has been very prolific in this field
Felix Grant Archive (FGA)	-Extensive personal collection of which Michael Fitzgerald is archivist	-Interested in idea of comprehensive discography with collaboration by experts like Michael Fitzgerald, who may be working on using his methods to catalog FGA holdings
Chicago Jazz Archive at University of Chicago Library	-Focus on jazz research -Deborah Gillaspie is J-DISC panelist -Emphasis on Chicago musicians, many of whom are world renowned	[as above]
Other Jazz Studies or related programs: -Institute of Jazz Studies. Rutgers -Leeds College of Music -University of Guelph -Jazz Institute of Chicago -Cal Institute of the Arts -University of California at San Diego -William Ransom Hogan Jazz Archive at Tulane University -UCLA Jazz Archival Collection -Manhattan School of Music -New England Conservatory -North Texas Jazz Studies -William Paterson University Jazz Program	-Focus on jazz research	-Interested in idea of comprehensive discography -Potentially interested in MIR tools for jazz
Library of Congress	-Played an advisory role in J-DISC (MARC Standards and interoperability; Jazz Sound Archives and Naming Standards)	-Interested in metadata standards for music, recordings, name authorities -Potentially interested in sessionography, which we argue documents their holdings more accurately and thoroughly -Potentially interested in MIR tools for jazz

Appendix D: Crowdsourcing for J-DISC

Crowdsourcing encompasses a wide variety of tactics to leverage the volunteer efforts of people to – in this case – enter data into an online database. The specific tasks the "crowd" performs differ widely by project, some requiring people to make judgments or evaluate information, others to enter observations, and still others to simply transcribe characters or words from a specific visual artifact. Some projects appeal to the broad public; others to devoted amateurs of a specific topic.

Given the high importance placed on having accurate data in J-DISC, it is unlikely that just anyone could or should be trusted to choose what enter discographic data to enter. Rather, should J-DISC want to pursue a crowdsourcing strategy, the first and critical step will be to clearly define what the task will be that people are being asked to perform in such a way that you benefit from their work without sacrificing the very value that J-DISC seeks to create.

Two types of crowdsourcing seem to be most likely to work in this case:

1. **Transcription from viable online discographies.** J-DISC could, for example, identify a reference work containing the needed discographic metadata fields, and have people copy from the online document into the appropriate data field.

2. **Data entry from fixed object in personal collection.** If J-DISC feels sufficient data exists in a readily available format, like CDs or LPs, it could be possible to direct people to which issues to enter, so that people can enter it from objects they own in their own personal collections. This might appeal to record collectors, for example. One risk is that published recordings may only include a small portion of the recording session metadata that is central to the project, leaving the J-DISC team to fill in many gaps. The solution would be to ask users to confirm that the disc has the correct session data first and only those would go ahead with data entry.

In short, there are several general conditions that need to be in place for the plan to work:

- The "crowdsourcing" activity must be extremely carefully defined: Where will the data come from? What fields will be entered, and how?
- Very clear instructions and other guidance must be provided to participants. A difficult interface will discourage people, as will unclear instructions
- Invest in creating a system to register those who will be contributing and to track their progress
- Build a feedback channel where the volunteers can easily communicate with the J-DISC team in case they have any additional questions
- Establish a clear workflow for vetting the crowdsourced content (QA)

Crowdsourcing: Possible Users and Tasks

Type of user	Task	Type of Source	Motivation	Disadvantage
General enthusiast	Enter data located on the physical CDs, Album covers, liner notes Basic session information such as name of artist, instruments, dates, record names, etc	Personal collections	Jazz fans who find pleasure in being part of a project they care about	Accuracy questions
Jazz Scholar/Researcher	Provide original discographic works with complete session information Review other discographic information already in J-DISC Suggest new sources	Their original work or sources they locate and trust	Scholars and researchers would like to see their works within larger collections	More senior scholars may not have the desire to manually input the data, though they could help with peer review process.
Musician	Provide their original session information, including that of the other musicians who also played in the session	Their original work	Self promotion, professional credit, securing their legacy	It is unclear whether individual musicians would have the time to manually input the data, but they have a strong incentive to do so

Note that the following case studies' popularity and success are based on the simplicity of the task required. In both cases, the public is being asked to copy, enter or review information from an image the team behind the project has previously uploaded.

What's on the menu?

> What's on the Menu (WOTM) is an initiative from the New York Public Library (NYPL) to help digitize and transcribe their anthology of New York's food menus. The collection consists of 45,000 menus from the 1840's to the present. (http://menus.nypl.org/)

> > *Helpful tips for J-DISC*
> > > o **A section for instructions and frequently asked questions by transcribers.** Recurrent concerns by individuals participating in WOTM include questions around punctuation, misspelling, multilingual items, lists and multiple prices and the website offers actual samples of these ambiguities to better explain what to do in each of these situations.

- o **Transcribers enter and review data**. Volunteers can choose whether to transcribe a new menu from scratch or verify the transcription made by another individual.

Transcribe Bentham

Transcribe Bentham is a project based from the University College London (UCL), whose aim is to digitize their entire collection of Jeremy Bentham's works. There are close to 60,000 manuscripts written by Bentham that UCL would like to make available to a larger audience. The project is collaboration between UCL's greater Bentham Project, UCL Library Services, UCL's Centre for Digital Humanities, and the University of London Computer Centre. (http://www.transcribe-bentham.da.ulcc.ac.uk/td/Transcribe_Bentham)

Helpful tips for J-DISC

- o **Transcribers create a user account which will allows them to access the manuscripts, along with access to a set of tools and guidelines on how to transcribe.** Having individual accounts allows the team behind the project to keep track on who is helping with the collaborative project, especially to separate those individuals whose work is reliable and of good quality, to those that may need some extra guidance or tools. There are 3,361 registered users in the Bentham's project, however only a minority is an active.
- o **Allowing transcribers to interact with each other by creating social profiles** where they are able to message each other, privately through messages or publicly though discussion boards. The team behind the project is essentially bringing together like-minded individuals, providing them with a platform to communicate which each other, to share their experience with transcribing and to form a community.

Appendix E: Summary Content Grid

Data Source	Ingest method	By whom?	Benefits/Considerations	How to cover cost?
Reliable discographies (print or online)	-Data entered manually by an individual	-Graduate or undergraduate could be trained to do this -Some fields may require expert guidance	✔ Using trusted sources ✔ Information from discographies is compatible with J-DISC fields ➡Information usually covers only 75% of the fields and must look elsewhere for remaining 25%	-Graduate assistantships -Undergraduate research stipends -Post-docs
Released CDs, Album covers, liner notes	-Data entered manually by an individual	-Graduate or undergraduate could be trained to do this -Some fields may require expert guidance	✔ Using trusted sources ➡Information does not cover more than 50-75% of the fields ➡ Iindividual must understand what is relevant data in the physical item	-Graduate assistantships -Undergraduate research stipends -Post-docs
Archival material from record labels or recording studios (print records)	- Data entered manually by an individual -Would need to be digitized first, possibly with a grant, and then uploaded manually or as a batch	-Graduate or undergraduate could be trained to do this -Some fields may require expert guidance	✔Could offer unique and reliable data source and appealing selling point to J-DISC ➡ Need to identify such archives ➡ Digitization process can take up many resources ➡Information would not cover 100% of the fields	-If content is not yet digitized, this could be a grant in partnership with content holder, to include into J-DISC
Digitized Archival material	- Data entered manually by an individual -Could be basis for an online transcription project -Batch uploading	-Graduate or undergraduate could be trained to do this -Some fields may require expert guidance -Via crowdsourcing	✔Could offer unique and reliable data source and appealing selling point to J-DISC ➡Information would not cover 100% of the fields ➡ Crowdsourcing requires much better interface and content vetting system	-If already digitized, cost is to prepare materials and instructions for data entry
Original data from artist	- Artists enter their own discographic records manually	-Musicians	✔Could potentially be great way to drive traffic, usage and build content ➡ Will require review of interface and content some small degree of editing	-No direct cost, but would require an upgrade of the interface -Could potentially be a revenue generator

Meeting with Dan Ellis of MIR Team
June 5, 2013

On June 5, we met with Tad Shull and John Swed to discuss the final stage of the Ithaka S+R consulting engagement. At the meeting, we were joined by Dan Ellis, MIR Director, who is leading the Music Information Retrieval (MIR) research that is part of the current Mellon grant to J-DISC.

Dan reported on the progress of the research, and these notes summarize our understanding of that work, as well as earlier conversations with Douglas Repetto, and specifically how we think that the MIR team's work could be most valuable to the long-term strategy of J-DISC.

Overview
Since the grant began in July 2102, the team (which also includes Douglas Repetto) hired a full time post-doc, Brian, to specifically focus on the application of MIR to Jazz. Brian and Dan have dedicated their time to trying to apply research tools, techniques and advancements done in the MIR field, to jazz. MIR, traditionally focused on Pop music, has proven to be more complicated than anticipated in Jazz, given that music recognition algorithms don't seem to "do as well" with jazz tracks given the improvisational nature of the genre. It is understood however, that the impact MIR can potentially have in Jazz and discography is very important. With MIR algorithms working successfully, a computer could generate "objective " information about an audio track including and not limited to, identification of instruments an chords, classification of beat and rhythm, and recognition of artist and solos - all analyzable information that could encourage new methods of intellectual inquiry into the music.

Areas of focus
Dan and his team have been developing MIR tools by working with an audio collection that contains 20,000 jazz tunes, named "Porkpie". The J-DISC grant project funded the purchase of the server and the creation of this sound database. To date, the team's research has been mainly focused on four areas: instrumental identification, rhythm tracking and chord sequence recognition, visualization. Below is a summary of progress, as provided by Dan:

1) Instrument identification –the MIR team has been developing techniques to identify which instruments are playing on a given track. Dan reports that this is working to some extent, but the level of accuracy and confidence is low, compared to the results with Pop music.
   a. The audio collection that the team has been working on is too small, but they've managed to track a database in Portugal that they can use in order to further polish this tool. Essentially, we will be able to get from a single track a list of instruments performed in that particular track along with their respective levels of confidence.

2) Rhythm tracking – This has been focused on exploring the different sequence of drums and beats intrinsic to Jazz (e.g. drummer behind or in front of beat). Unfortunately, this tool has not proven as successful and reliable in Jazz as it does in Pop. Given jazz' nature and way it is played, the computer is constantly mislead and picks up the beat of the soloist, which in turn is not

necessarily following the drummer. Brian and Dan have been trying to solve this problem, but are still unable to do so. The specific usage of this tool in jazz is still unknown – "Quite ambitious to think that the computer can turn out with a list of all drummers that are behind the beat within the database".

3) Chord analysis – Again, this exercise is an easy analysis to run in Pop music, which only has 24 chords, but very complicated for Jazz. Main hurdle presented by Dan is the lack of a correct data set from which they can work with. Dan stated that they possibly need to a redefine their strategy in this area of the analysis.

4) Visualization - Brian has also been working on a web based visualization of the audio files in Porkpie. Unfortunately this audio base is not linked to J-DISC at the moment, but could be an application in the future.

Implications for J-DISC

There are many potential ways in which the work being done by MIR could potentially help to fuel the growth and development of J-DISC. At this time, however, few of those have been possible, due to the early state of exploration of the MIR team, and the "experimental" nature of the questions they have identified and are pursuing. These constraints were understood at the time of the grant; but nevertheless, we would like to simply highlight the areas where the two projects could benefit from deeper integration, even in the short term.

- **The audio archive.** The audio archive – Porkpie – that has been created in order to test MIR tools is not aligned with the sessions Tad and his team have entered into J-DISC. To some extent this is driven by practical considerations; the audio database needed to be developed quickly, and J-DISC entries are often derived from paper sources, or web-based documents that have no audio attached. However, down the road, it will be important for there to be greater alignment between the metadata J-DISC holds and the audio files in Porkpie. This could be done at the point of data entry (in those cases when J-DISC transcribes data from an audio recording, such as a CD) or afterwards, if MIR can develop ways to retrospectively connect recordings to their metadata. Either way, there are significant benefits to having greater overlap here, including:
  - Using MIR to derive new, useful ways of searching audio files (by instrument, genre, drum beat, riff, etc…).
- **Using MIR innovation to enhance J-DISC**. As MIR tools develop, some will be of interest to scholars and students, who will want to use them in answering their own research questions. Having the ability to integrate a search of the sound archive, and showing results in terms of artists, or sessions would be very valuable for the core audience of J-DISC. This could be done behind the scenes (as a means to introduce new data fields), or offered as a major feature to users (eg. "use this MIR tool to search the archive for sessions that include this jazz riff" or similar)
- **Using J-DISC metadata to provide the ground-truth for the tests MIR runs on its algorithms**. As MIR tools permit new ways to understand the audio data, it is valuable to know which recordings are being referenced. The accurate session-based recording metadata in J-DISC

could provide/store the "ground truth" for the audio files in Porkpie, so having the audio linked to the discographic data will be important.

- **Developing visualizations for data contained in J-DISC**. Among the "quick wins" for J-DISC users –and this point was brought up by Douglas in an earlier meeting—would be to develop some visualizations to understand the data already in J-DISC. This, in fact does NOT need to rely on MIR algorithms at all, but it is hoped that the team there could be instrumental in helping develop a priority list of visualizations and possibly in creating them as well or at least helping to define them, should an outside developer be brought in.

- **Pursuing new innovation grants together**. There may be ways for J-DISC and MIR to partner, specifically with the goal of increasing value to researchers. We can imagine some general directions for this:
  - **Innovation grant for J-DISC** to integrate advanced MIR tools into its platform to offer researchers new ways of searching music content.
  - **MIR grant to expand the results of early research for greater public impact.** This would be a similar idea, from the other direction: MIR could position itself as being in a position to more widely share useful tools among scholars of music, by developing a user interface that encourages users to test the tools, share their research needs and interests, and share the results of their experimentation with the tools the MIR team makes available to them. This "space" could be a part of the larger J-DISC website, and this new innovation space would lend considerable attention to the basic discographic search, which would still be the form in which results were displayed.

Research Phases

Additional materials

This section serves to summarize the work that Ithaka S+R has previously delivered to J-DISC, from August 2012 to April 2013. This section highlights the extensive analysis completed in order to better understand J-DISC's competitive landscape, the potential end users of the database, and finally an assessment of the navigation and ease of use of the user interface.

Landscape review and competitive analysis
November 2012

There have been several attempts at creating a comprehensive jazz discography. However, due to the vast amount of information and the difficulties associated with locating and organizing the necessary data, the existing jazz discographies leave something to be desired. J-DISC is positioned to create a database that overcomes many of the shortcomings of the previous attempts. The purpose of this document is to highlight current direct and indirect competitors to provide a sense of how J-DISC will fit into the current landscape.

In this document we divide the existing products and services into five categories:
- direct competition provided for a fee
- direct competition provided for free
- related services
- potential collaborators: free services
- potential collaborators: paid services

For each category we offer a brief description of the main products or services, as a first step to thinking strategically about how J-DISC can position itself within this environment.

Direct Competition Provided for a Fee:
The two most prominent discographies in this category are: Tom Lord's *Jazz Discography* and Brian Rust's *Jazz and Ragtime Records (1896-1942)*:

- **Tom Lord's *Jazz Discography*** is an extensive jazz discography that is available in three formats: online, CD and a 34-volume reference work. The database is compiled and edited by Lord. However, he notes that there is a global community of contributors and he accepts feedback in the form of notification of errors and omissions. The data presented for each song includes: musician name, supporting artists, instrumentation, date and place of recording, and matrix/issue number. In total Lord has presented information on 1,137,000 musicians and 1,232,000 songs making it one of the most comprehensive jazz discographies available. The digital formats both have very advanced search capabilities allowing the user to search multiple musicians and tunes at once— this offers a substantial advantage over the printed volumes. The various formats are for sale for different prices. The CD is available for $350, the online membership is $9.99 a month and the printed format runs about $100 per volume.
- **Jazz and Ragtime Records** is a jazz discography covering songs recorded between 1896 and 1942. Though originally printed, the discography is now only available in CD format. The discography was compiled by Brian Rust and contains information on over 33,000 songs. It provides full metadata including full personnel, instrumentation, vocalists, arrangers, conductors, accompanists, recording dates and locations, catalog, matrix and take numbers, and title and artist indexes. Compared to Lord's CD, Rust's version has somewhat limited search capabilities allowing you to only search one song or artists at a time. The CD is sold for $99.

Though the two discographies are quite different they share a few important strengths and weaknesses. Both contain the desired depth of metadata such as matrix number and supporting musicians. In addition, Lord's discography in particular contains advanced search capabilities providing a useful research tool for scholars. Though the data is accurate overall, the more obscure entries contain errors, and this has been pointed out by many critics of the source. In addition, neither discography provides an exhaustive set of all jazz sessions ever recorded.

J-DISC stands as a direct competitor to these two discographies, and has the potential to improve upon them. The collaborative nature of the project will allow for greater accuracy and a more exhaustive list of sessions. It is worth noting, however, that despite all its errors, Lord's discography is still used by many and considered to be a useful research tool. The fact that J-DISC is offered for free may not be enough to attract users from Lord's discography, until J-DISC can begin to amass a collection closer to the scope of Lord's.

Direct Competition Provided for Free
The discographies below all appear online and are sustained by both advertisements and the efforts and funding of individuals. The main discographies that we chose to present here are: jazzdiscography.com and jazzdisco.org:

- **JazzDiscographies.com** is an online jazz discography created by Michael Fitzgerald and Steve Albin. The site provides access to BRIAN, a relational database application, which allows users to create discography entries. JazzDiscographies.com hosts over 400 band leader entries; many of them written by Fitzgerald and Albin but some written by other enthusiasts. Each entry has full discographic data containing: full personnel, instrumentation, date, location, CD released on, matrix number and issue number. However, most if not all of the pages do not provide an exhaustive list of all the sessions each musician ever played in. The search capabilities of the website are basic allowing the user to scroll through a list of artist names as well as search for individual artists by text. With no advertisements or outside funding, the website is sustained by the efforts of Michael Fitzgerald and Steve Albin.
- The *Jazz Discography Project* (jazzdisco.org) is created and maintained by Nobuaki Togashi, Kohji 'Shaolin' Matsubayashi, and Masayuki Hatta. It is a website that contains extensive discographic information on recorded sessions from over 50 band leaders and many more supporting musicians. The data included on each session includes full personnel, instrumentation, data, location, record label and issue number. The search capabilities are limited allowing users to scroll through an alphabetical list or simply search a single artist by name. It contains several advertisements which fund the site.

The free discographies share several key strengths and weaknesses. Their main strength is that they are provided for free and typically foster a collaborative community. The major weakness of these sites is that they are mostly maintained by jazz enthusiasts and are not able to compile a sufficiently exhaustive dataset. In addition, they do not contain the desired search capabilities—something that can severely limit the usefulness to scholars. J-DISC will compete directly for users of these sites. Because the J-DISC team consists of jazz scholars they will likely be able to create a more complete and accurate database. Along

with the advanced search capabilities and larger team of contributors J-DISC seems like it will become an obvious improvement to these earlier attempts at free jazz discographies.

Related Services

These products are related to jazz discography but not necessarily discographies themselves. The prime example of this is rainerjazz.com.

- Rainerjazz.com is a website designed to create an online environment where jazz scholars/enthusiasts can come together to compile metadata, photos, and biographical information. The site is intended to foster an online community of jazz scholars. Despite the sites large ambitions it does not host very much actual data. There are a few photographs of artists and old records but most of the data—for example almost all of the discographies—is provided through links to other websites. The website is funded by advertisements.

Rainerjazz.com and other sites like it offer a community for jazz enthusiasts to share information. The people who frequent this type of site may be interested in using J-DISC and connecting with these types of users may help to promote usage of J-DISC.

Potential Collaborators: Free Services

In this group we include organizations that provide metadata and music discovery services for free. These sites and companies do not necessarily contain jazz discographies but rather they tend to provide less rigorous information on music from all genres. In addition, they provide a broader range of data from biographical information to audio files. The top sites here include: musicbrainz.org, last.fm, and allmusic.com.

- **MusicBrainz** is a website that hosts discographies of a large number of musicians who belong to a wide spectrum of styles including but not limited to jazz. The site has an extensive listing of musicians but more limited information on each recording session. In fact, because the site is intended for the general public it does not contain full discographic data that scholars would require. The information is uploaded and edited Wikipedia style—drawing upon the many users of the website to build to and maintain the existing content. All information on the website is freely accessible. The website does not host any advertisements and relies on outside funding from both individuals and organizations. Groups such as Google, grooveshark and last.fm are all recent contributors to MusicBrainz.
- **Last.fm** is a website that hosts a vast amount of information on musicians and their music. The website contains a multitude of aspects to it including: a Pandora-like radio station, concert information, lists of popular and new songs, a social media tool and biographical information on musicians. Last.fm covers a wide spectrum of musical genres including but not limited to jazz. Though the site contains information on many jazz artists, the metadata provided is not complete enough to satisfy scholars. For each artist there is a short biography, list of available CD's with links to purchase, and comprehensive list of songs. To create their growing database Last.fm draws upon its larger user base letting them add pictures, videos and biographical information to the page. There is no fee for accessing any of the features on the website (although this is not true for all countries outside the US). To afford the team of editors and programmers the company generates revenue through online advertisements.

- **AllMusic.com** is an online music guide that contains metadata on over 6 million songs. It has an extensive collection of artists, albums and songs from an extremely wide range of genres. It provides biographical information on artists as well as information about individual albums. This information includes album reviews and links to purchase the albums/songs. The website is geared towards people who are curious about learning more about a particular artist and possibly shopping for a new album. It also aims to help people discover new music. From the vast number of visitors to the website AllMusic is able to generate some profit from advertising. However All Music also offers an information service. Working with such music sales companies like Windows Media Player, AllMusic provides album artwork and information.

Though these sites do provide metadata on musicians and their songs, they are not direct competitors to J-DISC. These sites do not hold up to the academic rigor that scholars and enthusiast demand from a true jazz discography. It is feasible that there could be a partnership between J-DISC and some or many of these organizations. J-DISC could potentially contribute data. Furthermore, it may be useful to look further into these websites studying both what makes them so popular and how they have managed to generate profits. In particular, many of these sites use algorithms to suggest relevant music or musicians to users. This feature, if added to J-DISC, may help it appeal to a larger audience.

Potential Collaborators: Paid Services
These companies tend to provide data to larger music companies such as iTunes as opposed to providing data to individual users. The prime example here is Gracenote—although allmusic.com also provides some services for purchase.
- **Gracenote** is a company that possesses a massive database of music and video metadata. They use this information to provide a variety of services to large companies such as iTunes and HTC smart phones. Their services range from identifying song names, artists and album artwork from CDs to providing an audio recognition program that allows users to identify songs by recording a ten second clip on their phone. The company has an extensive database that contains information on over 130 million songs. Gracenote does not offer full discographic information like the names of supporting musicians and date and location of the original recording.

Though GraceNote possesses a very large database of metadata it still lacks information on some artists. Their database grew from crowd sourcing the immense task in the early 90's. Thus it's possible that they lack information on more obscure jazz CD's. They may be potentially interested in buying metadata from J-DISC.

Implications for J-DISC and Directions for Ithaka S+R Research
Because of the shortcomings associated with the existing jazz discographies J-DISC is poised to make a much needed contribution: providing a definitive record of all recorded jazz. Its free access and advanced search capabilities will be a useful tool for many jazz scholars and enthusiasts alike. Below we list a few of the key findings to take away from this report. We hope that this information will help form a discussion around future steps for J-DISC as well as further research to be conducted by ITHAKA.
- **There is direct competition out there**. Though not without flaws, Lord's discography is probably the most comprehensive jazz discography and J-DISC must keep in mind that though there are many complaints in the field about the inaccuracies of the database, it remains an important competitor, due to the near-comprehensive scope of its entries. J-DISC has the

potential to appeal to users for a more community-base development, more accurate data, and free access.

- **The free options out there are not at the level of accuracy scholars need**. The current free discographies that rely on user contribution have not been able to compile exhaustive databases. J-DISC must make efforts to ensure that they foster a community of users and that they have enough active contributors. Even if all users are not jazz scholars J-DISC could employ an editing system like Music Brainz to increase contributor size.

- **But lots of commercial and fan sites have lots of users**. Commercial sites like iTunes, Spotify and Rhapsody draw many users, though their music metadata is not particularly deep. Still, examining other music discovery sites could offer valuable insight on sustainability strategies, particularly around driving usage, if advertising (for example) were to be an option. Having direct links to listen to or purchase audio is the draw for these sites. Thinking about ways to possibly link into them is a direction that J-DISC may want to pursue.

(See Appendix A for competitor landscape grid)

User needs assessment

As part of the user needs phase of Ithaka S+R's engagement with J-DISC, we conducted a series of interviews with potential users in order to understand the types of information they sought, the resources they used, their expectations from discographies, and their opinions on a crowdsourcing approach. The goal of this phase was to develop an understanding of users' needs, in order to inform J-DISC's content priorities moving forwards and to evaluate the possibility of a crowdsourcing approach to content development.

We interviewed 12 people in this phase, representing the perspectives of academics, musicians, and other music professionals, including a publisher of music textbooks, and a radio programmer. Though each person is counted just once, many were able to address more than one perspective. For example, a professional jazz trumpet player was also commercial composer of music. The segments chosen correspond to those considered most likely to be future users and potential contributors to J-DISC: academics; professional musicians; and others deeply involved in various aspects of the music business.

Category	Number of interviewees
Academics (scholars + students)	4
Musicians	4
Music professionals: Collectors	2
Music professionals: Publishers	1
Music professionals: Industry (radio)	1

Interviewee segmentation
Interviews took place via telephone, and lasted approximately one hour. The interviews were semi-structured, which is to say that we both followed the questions in our interview guide, and allowed ourselves to pursue useful digressions as needed. The Interview Guide is included here as Appendix B. Anonymized notes from each interview are included in Appendix C. This memo summarizes the main findings from these interviews and suggests some implications for J-DISC.

Why do people look for information about jazz musicians and their recordings?
Interviewees cited a range of reasons for seeking out data about jazz artists and their recordings.

- *Conducting broad research to identify jazz trends or seek out recordings.* Academics (professors and graduate students) and collectors stressed the importance of formal, structured discographic information to guide their research or collecting interests. Scholars valued discographic data to understand who played together or where artists were playing and recording over time, while collectors mentioned valuing many smaller details – some beyond the scope of standard discographies – such as reissue information, label information, or original issue format.
  - *Understanding outputs of specific performers*. On occasion, this research was directed in a more specific way, targeting specific performers. Interviewees both inside and outside of academia indicated that they often sought information about specific performers. Scholars valued this information to comprehensively understand subjects of research, while collectors and casual fans (i.e., those whose interests in jazz were not professional

or did not have a structured collector's interest) indicated that they wanted to learn more about their favorite artists or promising new performers for reasons of personal interest.

- *Finding interesting stories or anecdotes to enhance the understanding of a performer, recording, or album.* A couple interviewees – a collector and a casual fan– pointed to Ashley Kahn's *Kind of Blue* as a model for this type of information. While they were interested most of all in jazz music, they appreciated interesting bits of information in order to provide their understanding of their favorite artists and sessions with more color.

- *Finding music to listen to.* Musicians in particular mentioned seeking out music – whether, in the case of musicians, it was for new music to perform or adapt for their own performances. This came as no surprise; after all, all interviewees first began their study or performance or industry careers because of a love of jazz. Perhaps more interesting, rather, was the fact that this was not conspicuously mentioned by scholars or industry professionals, who may have had more realistic expectations of where and how they would find audio files.

What resources do users currently use to do this?

People we spoke with rely on a variety of resources to serve these various needs. Many said they found information and resources on music by searching Google and consulting colleagues, as a starting point. The results of these searches are inconsistent, however; some interviewees struggled to name the different sites they found through these mechanisms. One enthusiast/musician justified this practice, saying that she had "no bias toward any specific sources."

There are some exceptions to this, however. When interviewees were asked about the specific resources that they used, seven different resources were cite by name:

- *The Jazz Discography by Tom Lord* was the most frequently mentioned resource. Interviewees praised its comprehensiveness above all, but indicated that its user interface was difficult to understand. It seemed to be the first choice for a jazz discography – particularly in the academic sector, where multiple interviewees commented on how using the Lord discography was a necessity for serious academic research.

- *JazzDiscography.com* and *Brian Rust's Jazz and Ragtime Records* were each mentioned a couple of times as alternatives to the Lord discography; those who brought them up praised their depth and accuracy. However, they did not seem to be go-to sources on the level of the Lord discography, but rather were reliable sources that complemented other discographic resources.

- A few users expressed a preference for *AllMusic.com*, calling out in particular its user-friendly interface and its inclusion of information beyond basic discographic data. Interviewees admitted, however, that they did not see it as a formal scholarly resource, and that it might not be appropriate for deep research.

- *iTunes, Youtube, and Spotify* were all mentioned as online destinations for accessing and listening to music. Interviewees noted distinct weaknesses around metadata (particularly around full personnel) and session information, but the value of being able to listen to the audio itself seemed to be valuable enough that many users referenced these sources. In general, these

What do users value in a discography?

Metadata

Accurate personnel information was by far the most important information for all interviewees, regardless of role, but metadata fields for location, recording date, recording label, and instrumentation were critical as well. Deeper levels of metadata, however – such as matrix numbers or playback channels– were largely unimportant to most audiences. Record collectors were extremely interested in detailed metadata about the recordings themselves, including catalog numbers, format type, and other factors that might be able to help distinguish original pressings from subsequent issues.

Accuracy

All interviewees cited reliability and accuracy as a primary value for a discography. While the current landscape offers a multiplicity of resources, none has fully won the trust of any of the users we interviewed.

As a result, people described all sorts of strategies for determining if the data they found was accurate, often relying on their own sense of a website's value. Cross-referencing information against other resources to ensure accuracy was cited by many as a frequent practice to "safeguard" against unreliable information and is an important part of users' research habits. Even information from notable resources was subject to scrutiny. One interviewee felt that even information in Brian Rust's discography was fallible, perhaps just not fully current: "he was very trustworthy, but there's information that has come to light since then." The fact that Rust issued corrections was encouraging, but at the same time, highlighted the need to cross-reference data. Others were quick to point out that there are just some issues that may be debated for a long time. According to one interviewee, "some [information about jazz] will never be resolved."

Easy navigation

While important to any website, *navigability and user interface* was mentioned as a particularly important quality for J-DISC by interviewees across all categories. In addition, a common critique of Lord as opaque or "hard to navigate." Suggestions here included an effective search mechanism or clear links between pages.

Other possible values of J-DISC

Academics indicated that a high quality discography would be useful and valuable, though some additional features could help to further differentiate J-DISC. Industry professionals we interviewed seemed to suggest that some additional content would need to be included in order for them to find some unique value in a database of jazz recording data. While these may be beyond the remit of this grant period, perhaps this offers food for thought for future directions:

- *Biographical/historical information about jazz recordings and artists*. Some interviewees, in particular those whose responsibilities involved catering to a large population of jazz enthusiasts (e.g., radio DJs) or who were considering using the database for their personal learning, told us about an interest in more qualitative, subjective information about jazz. One radio station general manager, for instance, indicated that facts or stories behind artists and recordings could be useful for on-air DJs looking to provide interesting content on their programs. Interviewees from all sectors indicated this was an appealing possibility. As one musician described it, anything that

connects "the actual story to the recordings" would be helpful, as there as "so many incredible stories in jazz."

- *Additional types of content.* While nearly all interviewees indicated that they might have some use for pure discographic data, the idea of having additional content, as a way to "enhance the [information they drew from] scholarly resources" drew enthusiastic reactions from all. Interviewees indicated strongest interest access to audio files of the recorded music; nearly all interviewees indicated that being able to listen to sessions or tracks alongside discographic information would be the most valuable possibility. In some cases, this stemmed from an interest in listening to the music itself for the sake of listening to tracks that they already knew. In other cases, however, users expressed interest in finding information about recordings (or the recordings themselves), in order to identify similar styles of music or other potential pieces to perform. The aim, in these cases, was to identify certain qualities of the music itself that could not be captured by discographic data alone. Other content types cited include liner notes, album art, or session logs.

What content should J-DISC cover; how should it define its breadth?
Interviewees had widely differing opinions about what the "best" subset of jazz recordings would be. Most agreed that there was a set of influential artists that were most popular and well-known, but indicated also that these artists have already been covered in other discographies and resources. Furthermore, while the average user may most often search for these most popular artists, J-DISC seeks to reach a higher-level audience of scholars, musicians, industry professionals, and collectors (among others), many of whom have more specific, research-driven interests beyond the most well-known artists. One scholar indicated, for instance, that she was interested in an area of jazz that overlapped with popular music, and found the "relatively narrow definition of jazz" frustrating; because she was interested in what she described as a "fringe area," she often found it difficult to find information that she could use. Along similar lines, one interviewee indicated that collectors, for the most part, had well-developed ideas of the material they wanted to collect. A resource that focused solely on the most popular artists and their most popular recordings might, then, prove unhelpful to scholars focused on very specific sub-categories of jazz or collectors interested in.

Interviewees, particularly scholars and collectors, indicated that adding content by label would likely be the most effective strategy, particularly because of the consistency and overlap in styles and performers. One particularly well-known collector noted, in fact, that many collectors organize or seek artists by labels—so, for instance, he said, collectors knew that Sonny Rollins recorded on Prestige until 1957, and that defined a particular era for him.

This approach furthermore presented an intriguing possibility for sourcing data. A couple of interviewees indicated that approaching these labels directly might allow J-DISC to collect and integrate valuable primary source data, such as session logs, liner notes, or cover art. One industry professional indicated that mid-size labels, such as Concord or Independent might be the best places to approach for this type of endeavor.

How do users feel about crowdsourcing?
Interviewees expressed a range of views about the viability of this model. Many felt that enthusiasm and goodwill in the jazz community could potentially be leveraged toward a crowdsourcing strategy for J-DISC. Among the most enthusiastic were scholars and graduate students, the core audience for J-DISC. Some noted that hiring graduate students would be a viable means to develop content. Scholars and students, we heard, would mostly likely be motivated by a desire to help the community and share information. Particularly for scholars, some sort of professional acknowledgment could be a potential motivator, with some mentioning that, in a best-case scenario, contributions might develop into a meaningful activity for tenure-track faculty.

Professional musicians and those in related professions seemed less likely to contribute, but we did hear one possibility worth further exploration: some practicing musicians mentioned that they would be willing to enter their own discographies into the database, and could imagine others would, too. One musician who has led some bands, but played as a sideman in many more, pointed out that sites like iTunes or the radio only report the leadperson. On the major music sites, he says, "You have [information about] extremely famous people, and … everything else is a disaster."

For crowdsourcing of any type to work, some conditions were cited as being necessary:
- **Develop a community around the resource**. J-DISC has the advantage of being an open-access resource run by and associated with respected, well-known scholars. This is critical; while we heard that interviewees might be excited and willing to support an open, collaborative resource, we expect that for-profit businesses or resources would have much more trouble forming a community of users.
- **Design suitable tasks for the crowd**: successful crowdsourcing projects have been able to design tasks or contributions that not only ensure that they can contribute in effective and accurate ways, but also capture users' interests in a direct way. Asking users to conduct independent research or simply transcribe existing text records may not fit these criteria, so developing a more effective way to frame the crowdsourced task will be essential.
- **Maintain the high standards of accuracy of the data**. While interviewees indicated that crowdsourcing would not necessarily compromise the quality of the data in their view, the methodology would need to be very sound and presented in a clear manner in order for J-DISC to remain trustworthy.

That said, the tone of most interviewees concerning crowdsourcing suggests cautious optimism. There may be people who would contribute, but the process cannot compromise the accuracy of the data and must involve some benefit for contributors. Before we can assess the likelihood of participation, J-DISC must first be able to clearly define the task expected of them.

Conclusion
Interviewees all expressed enthusiasm for J-DISC's mission and indicated that, well executed, J-DISC could provide real value for its users. There is real interest from the academic community, as well as further potential interest from a wider range of users, including musicians and industry professionals. More general music enthusiasts are unlikely to engage with J-DISC as a discography.

While there are a handful of widely-used discographies already, J-DISC has the potential to set itself apart with its openness, its relational navigation, and the possibility of creating a community resources. Interviewees all agreed that accuracy was very important to them, but that no single resource had proven to be infallible; this suggests a way for J-DISC to distinguish itself, as utterly reliable, last-word on jazz recording data. It may take time to establish this reputation and at the start we imagine users will gradually start using it as just one source among many others.

While creating a corpus of accurate discographic data is the priority, there may be value in considering how the database could offer access to other content types, whether session logs, liner notes, album art, or audio links – paired with discographic data. If the right content source were identified, this might help offer a source for the metadata to be entered into J-DISC as well as complementary aspect to the metadata. In our next phase of work, we hope to further discuss content strategy and potential data sources with you.

(See Appendix D for grid on J-DISC potential audiences)

Usability assessment
May 2013

Observations of the current J-DISC website

Ithaka S+R was asked to develop a short list of issues concerning the J-DISC interface that could prove difficult for users (audience and contributors). This list is intended just to "raise flags"; it is hoped that J-DISC will be able to work with tech staff at Columbia to examine the site for user experience and to prioritize elements to resolve.

The notes below are observations made by ITHAKA S+R team members. These individuals are not experts in jazz or discography, but they have extensive experience in reviewing academic websites and database projects. Their observations are strictly based on their experience in viewing and navigating the site as first-time users.

Interface and presentation
HOME PAGE
- The opening page has a great deal of text already; difficult to scan the large box of copy, which is the first thing one sees on the site. It could be more powerful if the text focused directly on the "what." The "how" part could be included in the "about us" section
- Is it necessary to include two links to the glossary section on the home page?

"Discussions"
- It is unclear what the "Discussions" menu and page contains. Discussions often means on-going dialogues among several individuals, but that does not seem to be the case here.
    o What is the main objective of this section?
    o Is there a way to better organize the entries? Right now, they appear to be single, random annotations, rather than threads of conversation/discussion

"About us"
- Too much text
    o Suggest inserting images and/or headings to break up the text

Navigation/language
- Some of the language choices are not intuitive and could cause misunderstanding and search delays
    o "Search issues" - could be interpreted as how to troubleshoot a search
    o "Search session" – could be interpreted as a log of past searches
- It is unclear which texts are hyperlinks and which are regular texts
    o Suggest changing color or underlying hyperlinks

Other aspects to consider that could enhance navigation and encourage user feedback:

- *Contact Us:* A direct contact to Project Director and team could be useful for users that have questions about the site or would like to know more information about the program and the database.
- *Feedback* Section: An excellent communication method to better understand how the users are interacting with J-DISC. A constructive tool for J-DISC where users can express what aspects of the site they find useful and what aspects need improvements. Additionally, this link could potentially encourage users to share information that might be absent from some of the sessions.
- *Blog* Section – A platform where the users can easily communicate with each other and share information about the sessions and jazz in general. This could increase the popularity of the site by creating a community around it.

Search interface
- There is an overwhelming amount of search fields
  - This, we understand, is a strength of the J-DISC concept, but there are probably ways to simplify the choices for users.
    - For example, it will probably turn out that most people mainly use a handful of fields, and perhaps those could be for grounded
    - Other fields will require some explanation, and that is currently lacking
- Are search fields mutually exclusive?
  - Possible solution could be considering doing a "basic search" with fewer fields and a separate "advanced search"
  - Considering drop-down menus rather than including an open-text field for every search option
  - Including some guidance in the form of "search tips" could help the user find the information more effectively and overall enhance their experience with the database
- The lists that appear to the right of the search fields are not clearly labeled. Are they an illustration? A sample record? Something else entirely?

Results display
- The results can be difficult to interpret given that the output pages are not uniform and clearly labeled

Data Entry Screen
- We do not have comments on the data entry interface at this time, but this will be a very important aspect to consider and design, particularly if J-DISC chooses to pursue some form of crowd-sourced data entry

Appendix A: Competitive landscape grid

	Relation to J-DISC	Format	Funding model	Data Entry Method	Data Provided	Size of Database (number of musicians)	Audience Size (how many users)	Intended Audience	styles covered/ time periods	Date created	Staff	Strength	Weaknesses
J-DISC	-	online	free	team of experts/grant funded assistants/exploring electronic import	full discographic date	exhaustive	-	jazz discographers, academics and researchers, jazz enthusiasts	all jazz	2012	Columbia staff in collaboration with other experts	-	-
Lord's Discography	direct competition	Print, CD, online	one-time purchase for CD/print. Monthly subscription to online.	Tom Lord and various contributors, all approved by Lord	full discographic data	extensive: contains 1,137,000 musicians and 1,232,000 songs	-	jazz discographers, academics and researchers, jazz enthusiasts	fairly comprehensive (not blues or gospel) but covers 1917-current	1992	Tom Lord and his personal paid staff	Comprehensive range, but with primitive advanced search capabilities, allowing you to search only 2 artists at once with few other descriptor options	more obscure information is not entirely accurate, rank, demonstrable, and even avowed plagiarism, limited independent fact-checking, pay for access, search functions and metadata can be very confusing
Rust's Discography	direct competition	Print/CD	one-time purchase	data compiled by Brian Rust	full discographic data	contains information on over 33,000 songs	-	jazz discographers, academics and researchers, jazz enthusiasts	1897-1942	1954 (but continuously updated)	Brian Rust	large list of recording sessions, fairly accurate data for most entries	has limited search capabilities; can only search one artist at a time, unreliable data for obscure entries, pay for access, does not cover most of the jazz
JazzDiscography.com	direct competition	online	free/self-financed	Wikipedia style	ranges upon entry, some of full discographic some limited to just artist name	much more limited-- mostly just big name musicians	-	jazz discographers, jazz enthusiasts	focused on modern jazz (particularly hard bop), to the exclusion of early, experimental, fusion, and non-US jazz	-	Michael Fitzgerald and Steve Albin	Uses BRIAN, a relational database application that allows users to easily and freely upload discographic information. Free access	has inconsistent information on different artists, very incomplete, limited search capabilities. Not designed to serve as a comprehensive database with search capabilities or collaborative features. Expressly built on print model.
JazzDisc.org	direct competition	online	advertisement and free access	entered and edited by three creators	full discographic information on each session	50 band leaders--but information on many more musicians-- though not complete list of their works	-	jazz discographers, jazz enthusiasts	varies widely not a focused selection of styles and time periods	-	Nobuaki Togashi, Kohji 'Shaolin' Matsubayashi, and Masayuki Hatta	contains a lot of information on the individual entries.	only has complete chronological entries on 50 musicians.
MusicBrainz.org	Weaker competition: not just jazz but less complete data	online	sources include Google, Grooveshark; and last.fm has free access	Wikipedia style	limited information: artist name, song length and album, release date	artists: 679,878 tracks: 11,411,792	618,665 editors + additional users	general public: anyone who wants to look up information about any song/artist.	contains all types of music, not just jazz	2002	team of 10 programmers	huge user base that continuously edits and contribute to the material. nice visual layout, very easy to navigate	does not have full discographic information, no supporting artists
Rainerjazz.com	Weaker competition: less data but targeted at wider audience of jazz enthusiasts	online	advertisement and free access	posted by website creator--links to other sites	very little discographic data provided. There are also a few links that go to other discographies mentioned above	-	-	jazz enthusiasts: not enough information to be truly useful to academics	no focused style	-	-	has a variety of interesting information. Really geared towards creating an online community of jazz enthusiast	very limited discographic information (mostly consists of links to other discographies)
Last.fm	Potentially Collaboration	online	advertisement and free access	users are able to add content which is edited by last.fm	short biography, list of CD's with links to purchase, and list of songs. Not full discographic data	very large--unclear on actual size. Has mostly more popular musicians not the obscure ones	over 30 million users	general public-- music discovery/entertainment	all styles and all time periods	2002	staff of 63	very user friendly has many features including: Pandora-like radio station, concert information, lists of popular and new songs, social media tools and biographical information. Free to all users	does not have full discographic information--not of much use to academics who want to learn about recording sessions or check historical facts
AllMusic.com	Potential Collaboration	online	some advertisement but also sells information services.	created by employees	not full discographic information but contains: bio, reviews, links to purchase songs, list of albums	over 6 million songs		general public-- music discovery/entertainment	all styles and all time periods	1991	-	free provides lots of information about artists and their work--review and biographies. Most info is accurate	Does not provide the discographic information that academics desire. No way to cross-search artists, or combine other search terms.
Gracenote	Potential Collaboration	online	sell information services to large companies	initially used crowd sourcing now the information is privately owned and managed	unclear--not full discographic info though. Most likely just song, artist, album and artwork	over 130 million tracks. Also contains info on TV and movies	they supply to a couple large companies (iTunes, Ford sync, etc.) so audience is very large	large corporations	all styles and all time periods	1998	very large staff	massive database with accurate information, variety of provided services	Serious problems with metadata and accuracy. Does not even include main side personnel. Full version or available freely to the public
Online platforms: Rhapsody Music and Spotify	Potential Competitor or Collaboration (too early to determine)	online	Pay and freemium models	-	-	-	very large	general public	all music styles and all time periods	-	-	As of spring 2013, both Rhapsody and Spotify publicly announced that they are looking into the creation of a more comprehensive metadata for their music platforms. Whether or not their effort will be comprehensive and accurate is yet to be determined.	-

Appendix B: J-DISC user interview guide

- What is your background in music, specifically jazz?
- We'd like to ask you a few questions about the sorts of resources you turn to when you want to find information about music.
    - How often do you search for information about jazz? (Musicians, records, sessions, etc.)
    - What types of information, specifically, are you seeking?
    - What resources do you turn to for this information?
    - What purpose is this for? (Potential answers could include general interest, research, collection-building, etc.)
    - What do you find most valuable about these resources? What unique value do they have?
    - How do you find the resources that you use?
    - Are there any sorts of information that you *cannot* seem to find, that is generally very difficult to find, or that has any other problems around it? Why?
    - How do you keep track of the records you own?
- Do you have any experience consulting discographies?
    - What, specifically, do you use these discographies for?
    - What are the most important types of information (metadata) that you'd want to see in the data?
    - What are the most important categories of performances that you'd want to see in the data?
    - What are the biggest positives or strengths associated with discographies you've seen?
    - What are the biggest frustrations or negatives associated with the discographies you've seen?
    - Is there any additional functionality, features, or paired content that would make the resources significantly more valuable?
- The resource we're working with is considering drawing data through a crowdsourcing mechanism.
    - If this were to be the case, would you consider contributing?
    - What factors would lead you to contribute, or lead you to not contribute?
    - What incentives would be most compelling?
    - How much time would you anticipate contributing per week?
    - What marketing channels would be most effective to reach you for this sort of effort?

Appendix C: Anonymized interview notes

JAZZ COLLECTOR
Background
This interviewee is a jazz record collector who's been collecting records since 1970. He's also a journalist (though not a music journalist) and a "really amateur" musician; he's given a few guest lectures at NYU on bebop, but isn't a formal academic in any sense.

He's a prominent figure because he owns and writes (all by himself) a fairly well-visited website, which is an outlet for him to "share information and write about something he loves." This isn't a business for him – just a hobby. Nevertheless, it's picked up a big following – about 1,000 page views every day – especially considering that he's had no advertising or formal training on how to run the site. He started the site about 10 years ago, first as an email newsletter, which was converted to a blog about 5 years ago. The original goal was for this to become a business, but he realized that it was just a hobby.

On collecting
Collecting has traditionally been a very secretive hobby: if you're a collector, you don't want to share your information. If you walked in to a store and saw an original Blue Note selling for $5, you didn't want to tell anybody – there's limited quantity!

Ebay has "totally changed the dynamic of the market," however, by "making the market really global." Before the advent of eBay, collecting was a bit of "a secret world" – some people (such as collectors) knew that the stuff was valued, but many people in the record stores didn't necessarily know. Collectors will still look for bargains through garage sales, Craigslist, etc., but because of eBay, these sorts of deals are harder to come by, because people can see what records are worth.

The jazz vinyl market right now is relatively hot, and in fact, the value of collectible records seems to be going up on a regular basis. The collectibles market first developed in Japan, starting with Japanese collectors placing importance on original pressings. People really value "the tactile and other experiences that go with the original pressing of a jazz record." This is the key thing about this market: identifying what is or isn't an original pressing.

When this interviewee was buying records (he doesn't anymore, as he said his collection was relatively complete), he would primarily go through listings on eBay. (Similarly, now, when he's writing his posts, if he's in a rush, he'll just run a search for the highest priced records.) In general, people start with a passion for the music then identify artists and eras of music they like. From there, "a combination of serendipitous discovery while knowing what you want."

The information he searches for
He began by saying that a database for collectors would be very different from a standard discography. There are "many incremental pieces of information you'd want to include from a collector's standpoint that you wouldn't necessarily want in a standard discography." For instance, Charlie Parker's 1948 recording of Scrapple from the Apple released on Savoy was originally recorded a 78. Collectors would want to know what the original 78 was, what the first issuance on vinyl was, whether it was 10" or 12", what the record label was, and any other subsequent re-issues. If there were original factors that could distinguish original pressings from second and subsequent pressings, that would be important too.

That said, the resources he uses are very collector specific. Fred Cohen, who is "probably the #1 dealer of jazz records in the world," wrote a collector's guide to Blue Note that contains everything you'd want to know about collecting Blue Note records. He'll also use eBay as a source of information because – even if it's not always accurate – when someone sells a record, they'll put information in about it that will try

to entice collectors. He takes some of this with a grain of salt (since they are trying to sell something), but if a reputable person is selling, then he'll believe it. The pictures provided are often helpful though – "all you have to do is look at the picture and you'll know that it's not an original pressing." Websites gathering communities of collectors are also good resources, because there's a wide readership that is often willing to answer questions posted on the site. The community of jazz collectors is really the best place to go; "there is more expertise collectively in the community than from a guy like Lord."

Some people use the Lord discography, but he has found it to be somewhat cumbersome. It also didn't have the information he was looking for, and not in an easy manner. If there are side performers that aren't listed, Lord is good for finding those personnel.

That said, collectors often don't find discographic data helpful just by itself. "Collectors believe it's all connected – it's hard to separate the things they're interested in from the music itself."

Content to add

From a collector's perspective, content should be added by label. Collectors compartmentalize the artist – so he knows, for example, that Sonny Rollins recorded on Prestige from 1951-1957. It covers a period of Sonny Rollins's career, but also defines a particular era for him.

Other obvious one way to start adding content is by artist.

Recording format might also be interesting. When 12" LP became a standard format, everything changed. There are eras in music, but these are harder to define. You can say "the era of bebop" or "post-bop," but these are much more perspective.

Crowdsourcing

Crowdsourcing, in this interviewee's opinion, is a possibility – but not for sure. He brought up a question about credit, recognition, acknowledgement, or money. Serious collectors have a vested interest in what they produce. Some level of acknowledgement would be of value. All that said, he thought it was "highly likely that you could get the community to participate" but you'd have to explore the incentives and some other things you'd need to investigate.

Most importantly, "people are willing to contribute to a community, they have confidence in the resource." People feel confident putting information on this interviewee's site because he has a reputation for being straightforward, for not making money, etc. There are hardly any inhibiting factors for them sharing information on his site.

Of course, there's a big question of accuracy if users are contributing. Who's the arbiter of who's correct if users disagree? No clear answer, in his opinion.

STUDENT

Background

This interviewee is a graduate student in musicology who earned his master's in jazz studies. His interests are in South American jazz, improvisation, and jazz history; he has also performed across the country.

Existing resources

He firmly believes that the only way a public database could be useful is if it were complete. The problem, though, is that no one can agree on what "complete" is – there's an idea that there's a "canon" in jazz, but this defines jazz too narrowly. Because of this, there isn't enough information about some areas – like early jazz – for him to find discographies useful. (He noted that Columbia is already guilty of this in particular.)

The Tom Lord discography does a pretty good job right now, because it's really broad: "the Wikipedia of jazz discographies." He uses it to find a lot of information, but it doesn't solve all his questions.

Value of a new resource

A new discography would need to be built in a way that can be used to think "more holistically and horizontally" with "mutually intersecting sections." Being able to open up that data to reveal more connections underneath would be tremendous. At the moment though, if you're not in those circles of jazz scholars, and you're not socially connected, you don't have access to that data! A space for a wider swath of interested parties would be valuable/important – could be an important tool for a lot of different people, including musicians, record collectors, and "jazz nerds."

An aggregator would be the most useful thing. It'd be best if it operated under the assumption that there were no "most important musicians." As long as it included good metadata on date, location, and name, that would be sufficient for him. Artists and geographies are the most interesting pieces of data. And by opening up the data, and using algorithms to look at all that data, you could turn a lot of assumptions about early jazz around. That said, it'd be great to include things like age, gender, size of the band, musical information like tempo and key.

You'd want to be able to go back and see all the sources for these individual discographies being aggregated.

That could all go toward being a site for discourse as well – like the talk page of a Wikipedia page.

How does it get in there?

Lord is currently the only one who makes all this data available, and he charges a lot. This interviewee isn't sure how Lord did it, however… he expects it would take a lot of manual labor. Deputizing grad students through work-study might be effective. But because he thinks a directory that aggregates discographies is the best idea, they wouldn't need to add individual data points.

Misc

This interviewee is *very* interested in hearing how this project develops, and is enthusiastic to help however possible.

SCHOLAR

Background

This interviewee is a professor at an institution that, as a distance-learning institution, is very reliant on digital resources. Before that, she taught at a practice-based institution where students were fundamentally focused on performing or composing – which, as you can imagine, used a very different group of resources! Her research now focuses on jazz in Britain, focusing on early pre-WWII periods. (She's also recently written on Benny Goodman.) Most of the time, however, she needs to find recordings that aren't as easy to come by as most American mainstream music.

Resources she's used

Because of the context of her institution, they have access to amazing digital resources – better than any other university she knows of. They have access to articles, e-books, online encyclopedias like Grove (with audio clips and such on the OUP site as well), and the Lord discography.

She primarily uses the Lord discography, "which can be useful to a certain extent." (It's certainly better in a digital version than the paper equivalent, however, since it's so much more searchable.) She finds it a bit frustrating in how it cuts off though, since it has a "relatively narrow definition of jazz." Her work overlaps into popular music though, so she thinks it might not quite be normal.

She also relies on Russell Connor's discographies for Benny Goodman. Cross-referencing was quite a headache though, especially since there were different sorts of information in both. Older print discographies, for instance, might cite places where recordings could be found, but they might not be available in digital form yet. It's important to her, though, to listen to available versions – if any sort of discography could include that, it would be incredibly helpful. For instance, a discography where you could click through to Spotify would be fantastic.

"Information in liner notes is also helpful, but how do you know about that information if you're buying it digitally?" She used one label in particular – the Chronological Classics label – specifically because it had that information.

Accuracy

Accuracy is important though: you need to be sure they're the right things; it's "a bit of a minefield" in that way. After using resources for a while, "you get a sense of what looks like it could be reasonably well trusted." She also looks at where the resource is based, who's working on it, how it's being developed, and what sort of review there is. Still, she tries to be quite skeptical, even of resources like Lord, which "is not error-free."

To solve this problem, she does a lot of cross-referencing. According to this interviewee, cross-referencing is a typical practice – she uses a lot of different sources to get to what she needs to know. The Benny Goodman print biodiscography is acknowledged as *the* resource, including a vast variety of different recordings, including things privately recorded, etc. and things that aren't commercially available. Lord is also good because it's easy to search, but sometimes doesn't have precise details, especially about recordings released on different CDs that couldn't be found.

Discovery

She finds resources through standard internet searching – just runs searches to see what comes up, including fan pages and such. She doesn't rule out the possibility of small sites; sometimes they hold vital information. Besides this, she relies on consulting her peers, contacting people who she thinks will know, or reaching out to the jazz research list. She generally tries to keep aware of what resources other people are using.

Value of a discography

Personnel information is the most important thing. If you're trying to trace through the movements of a particular musician, being able to look at recording sessions is also important: what they were doing, who they were working with, where they were recording. Instrumentation is also quite important. This all goes directly into research – primarily for historical purposes. It'd also be helpful to be able to see how recordings were put out – were they on a particular album? What were they coupled with? What countries were they released in?

Being able to search for musicians by leader vs. support might also be valuable. Being able to cross-reference (in this case, it sounded like she meant easily navigable) arrangers and composers to search by tunes by composer or arranger would be great – "people expect things to be linked in that way."

Adding on to that basic functionality, she mentioned that the ability to save searches might be helpful. She also has trouble, at the moment, identifying who's playing solos, so that could be a unique value. If there were images that could be linked to performers or videos or audio – "things to build that picture" – they would all help to enhance the scholarly resource.

Initial sample

From a personal perspective, she's very interested in the earlier periods of jazz, pre-1935. There are things like Brian Rust and Red Hot Jazz that cover this to some extent, but not always with the level of detail that she'd like. That said, she acknowledged that if you want to attract users, it would be best to include mainstream stuff, like Miles Davis albums, which everyone knows about. "Louis Armstrong, combined with some relatively rare things."

Crowdsourcing

She thinks contributions would definitely be a possibility, and would personally contribute herself – though it might be a minimal amount – and would encourage others to contribute as well. The incentive would be helping the scholarly community, on which she relies for much of her research herself. Because of that, an open model would make more sense to her – and in fact, would potentially make the resource stand out from those that are closed down by creating a community.

The risk, of course, is that of accuracy, but she thinks that if contributors had strong direction given by someone reputable, she'd find that acceptable. In projects she's done, grad students have had a lot of responsibility, but with a structure to assure quality control – this might be a viable model.

JAZZ COLLECTOR
Background
This interviewee is a long-running jazz collector and has been heavily involved in an organization that has allowed him to interact with many other collectors.

Many serious record collectors have very large collections. In fact, if it weren't for record collectors, many universities/research archives would not have the records they do; collectors often donate or sell records to institutions.

His cohorts tend to be interested in pre-war jazz, but there are of course some members interested in post-war jazz and contemporary jazz. This interviewee's personal collection starts with 1917 but goes all the way up to the 1980s, primarily in the form of 78s and LPs

Once collectors become interested in music, they start wondering about when it was made, additional takes, and other similar or collaborator musicians. Discographies are essential for this, as tools to better understand. Historical discographies that document recording sessions, theater appearances, etc. are very helpful for this. Individual artist discographies are perhaps most accurate and detailed here.

Resources he uses
He tends to use print discographies more than online ones, however, though being able to search digital versions is very helpful. He turns to Lord occasionally, but the issue of inaccuracy there is compounded by having drawn on other discographies, thus perpetuating errors in other ones. More often, he finds other discographies through Google, finding resources serendipitously. No systematic approach – "Google and see where it takes you."

Value
At a minimum, a discography should have the correct and full title of the leader of the recording and the date. Matrix numbers are very helpful. It should include information about alternate takes, whether they were issues, if there were rejections or alternate titles. Recording time is also helpful, though resources often disagree on playback times. Recording date and location (the city at the very least) are also important to include. Instrumentation, correct names and spelling of all personnel, and any special notes (e.g., gap in matrix numbers, note indicating issued under a different name).

Accuracy is very important: any research effort is only as good as the source material that the researcher is able to access. As a seasoned writer and discographer himself, however, he naturally gets a good sense of where the misdirection and inaccuracies are.

He currently has trouble dealing with bootleg labels – unauthorized issues of recordings particularly with CDs, nightclubs, etc. These shouldn't be denoted as a separate category, but rather included under general artist information.

Things they can't find: dealing with bootleg labels—these are unauthorized issues of recordings particularly with CDs, nightclubs, etc. hall of fame was a bootleg label where a lot of tunes appeared on the jazztunes label. Would like them included as part of the artist information, not as a separate category.

Crowdsourcing
A small segment of the population might be responsive to crowdsourcing. It's worth noting that in the history of doing discography, most of the work has been done by individuals who are plugged in to small, like-minded groups of collectors. They do this out of a sense of civic responsibility, but credit might also be appreciated, particularly for career enhancement.

RADIO STATION MANAGER

Background

This interviewee is the general manager of a jazz radio station based in the Northeast. This turned out to be a fairly productive call, and he sounded enthusiastic about the prospect of a discography and its potential uses on the air. Notably, the most appealing thing about J-DISC to him seemed not to be the detailed metadata, but rather the annotated part with interesting factoids. Also potentially interesting: collaborations for content—they have huge a lot of physical music content that might serve as a place to get data sources for crowdsourcing.

General thoughts on a jazz discography

"An online accurate jazz discography would be fabulous," as radio stations are always talking about artists, dates, record dates. They try not to overtalk, but they're nevertheless very keen on getting good information about the artists they do play. They already look to see if computers have something online about the sessions they're playing. They're looking for interesting facts about the recording—e.g., if they were sick, if they were in a bad mood. Of course, they're also interested in more traditional factual information around alternate takes and such. Ideally, this would include as much about the musicians as you can fit in without making it a novel. "When you're doing a show, you're looking for succinct information, not reams and reams of information." They preach "keep it short, but keep it interesting." There aren't any particular features or pieces of metadata that would be significantly more important than others. "Interesting tidbits" – any interesting facts, especially ones tied to the recording. If J-DISC lived up to its potential – would be using it a lot.

The most important things: If they want radio to jump in and be part of this – make it easy to negotiate and navigate. If they're going from Coltrane to Buddy Rich – make it easy for him to find. In radio, it's all about keeping on top of the clock. If it's complicated, convoluted, then it won't work for on-air people.

Current usage

This interviewee isn't on the air anymore, however, so he can't speak to what's going on right now. When he was DJing, he found discographies really useful—he could plot whole shows out, make notes about significant tunes, etc. He suspects, now, that there are some online sources that people use. He wasn't able to name any specific ones, but he expects that they have a few go-to resources, and find the rest through Google. They also still use CDs and LPs at his station, so they consult the records themselves.

Monetization and other value

They'd consider subscribing to J-DISC, but this would depend on the cost and service. But budgets are tough, so it'd be a tough sell. "Is it going to be a $200 service or a $1200 service?" They don't currently subscribe to any research tools like J-DISC—all their subscriptions are programming and other things directly related to the service. (e.g., fee to NPR, other programming like Jazz @ Lincoln Center.) And if you offer something that's too expensive – much of that information can be found elsewhere. Penny-pinching radio stations will look elsewhere if it's too expensive.

One option would be to consider bartering for marketing—trading the service for radio stations mentioning it on air. This could be an effective marketing channel: about 40,000 people who tune in every week to his station. This costs $30-35 per one-time spot. Clients usually spend from $500 to $10,000, depending on how long/how many spots they purchase. In general, you need to be on 10-14 spots per week to be effective, rotating throughout the broadcast.

Q re: content scope

A big question: how will you constrain/restrict the scope of the resource? Everybody is recording content these days – you can't believe how much content we're receiving these days. They're going to have to decide at some point – who to include on this discography, who to leave off.

Content sourcing

They'd be willing to provide information from their files, but would want to know exactly would be involved. If it takes hours and hours to copy this stuff, then they just don't have to time to do that. They just have 3 full-time people, 4-5 part-timers, and the rest volunteers. They'd consider finding volunteers/interns to do that sort of work. This sounds like a good project for a college intern – if they could do that, they wouldn't find any money. Other managers might say that it's too much work for them to get done. You've got other sources already if you're only going up to the 1970s.

MUSICIAN

Background

This interviewee is an editor for an online jazz publication. She's a vocalist who performs primarily in Philadelphia, New Jersey, New York, and Maryland. She performs primarily as part of a quartet, quintet, or big band.

Resources she uses

As a vocalist, she's constantly searching for new materials and older recordings to perform. It's often difficult to track down music that has vocalists on it; databases are extremely important for this purpose. She uses the word "database" broadly to include books that include information about compositions and actual compositions, discographies, resources through Google, and things like iTunes and Spotify.

For example, when she hears an arrangement done in a movie, she'll look for information about it – particularly recordings. She started with Google for artist or song names to see what hits she gets. She'll click on whatever's out there – "no bias toward any specific sources." She'll then see if any other artists have performed it before, and if so, in what sorts of styles. From there, she'll look for sheet music.

It didn't sound, however, like she was particularly interested in discographic data – instead, the information she sought was for the purpose of finding new performance material.

Value in jazz resources

As such, she placed a premium on the navigability of resources. Simple search fields that returned appropriate results were important, as well as a well-developed corpus of content. (She "just doesn't want to see a red X that says 'search not found'.") Accuracy is also important, but again for supporting her musical career – she wants to accurately know who wrote the work. It's worth noting, however, that no resources currently live up to this standard: she said that "if she found a reliable, comprehensive, simple, easily navigable resource, she would go to that all the time."

Beyond that, she thought it might be helpful to include some history and information about rights. Because she'd be using these resources as more than just references, the rights to the information in the resource would be important for her.

Authority

Her perspective on authority was murkily defined—"if it were developed by someone who had knowledge in the field, she'd be inclined to trust it." Graduate students in jazz would be okay, but she'd be suspicious of random contributors.

Value: choices and knowing what those choices were without having to search for more.

Content to include

Like many other interviewees, her thoughts on the most important content to include came from a very personal perspective. She was personally interested in swing, then older-modern jazz (pre-1945-1950s), then modern jazz least of all. In her view, the most important period is from 1935-1945, when ragtime and blues crossed over into large band ensembles. With that said, she thinks it's important to include seminal artists like Miles Davis, John Coltrane, et al.

Crowdsourcing

She's cautiously optimistic about the possibility of crowdsourcing. The online jazz publication she works for has a user profile section, but that's entirely for yourself – not sure that's a great analogy. It really depends on how broad and large a task it would be. People are willing to talk about it and dno't mind putting up information they like, but she's just "not sure you can have an *expectation*." On a weekly basis, she might expect that people would be willing to contribute an hour and a half or so of time.

SCHOLAR

Background

This interviewee is a professor in jazz studies and ethnomusicology; he chairs the jazz studies program at his university. He teaches jazz history, jazz research methods, and jazz styles and analysis.

He had some familiarity with J-DISC already, as he'd heard about it on the jazz research list. His initial impressions were fairly positive—he was excited by the idea of a resource that could be edited and updated by the community.

How he uses discographies

He occasionally makes assignments for his graduate classes that ask students to make thorough discographic searches for artists or sets of artists. He could imagine people spending a lot of time on this, tracking performances around the same time. They might also use discographies for very quick checks of discographic information.

There's a real value to discographic data. Students sometimes say they find information on Spotify, but he realizes that Spotify doesn't provide a lot of metadata. Discographies can be important in filling in more metadata.

Furthermore, he expects students to become experts on the history of their instrument in jazz. Part of this involves knowing the tracks as they're released and being able to relate to repertoire in meaningful groupings rather than isolated tracks. This metadata is used to build a context for listening.

How they find resources

If students are conscientious, they go to the library and find standard discographies. His institution has the Lord bound volumes, but not the online version, which can be a pain. Sometimes students also go to Wikipedia, sometimes Spotify (but, as said earlier, this is often not helpful in terms of metadata). The library catalog carries a certain amount of metadata with the library record, particularly if they have the album themselves and have the liner notes. Still, the risk of separation of metadata from audio data is a real one, especially as issues become digital only. Jazzdiscography.com and BRIAN are also fantastic. Accuracy is an important quality – they're very conscious of this. Students are assigned to read Kornfeld and Rye, which discusses problems of accuracy and plagiarism. Michael Fitzgerald's review of the Lord discography is also influential. But the issue of duplication isn't necessarily a problem: if J-DISC can be made available in an authoritative fashion and editable by knowledgeable people, it will stand out as a significant resource.

Value from discography

Other valuable features could include the ability to "save" a session and download it to a file for further work with it – sort of like downloading a JSTOR article. Scholars often want to do more processing with data, and you can imagine something like a research project where they want to do an analysis of the social network of musicians. This isn't possible with the built-in tools of discographies now. If there were some sort of coherent export format – such as Bibtex, which is a standard bibliographic format for citations – that gave scholars an exportable structured format for scholars who want to work quantitatively, that would be really valuable. This could allow for a quantitative approach to jazz discography that would allow people to work across the entire corpus. If the tools are there for new questions, then this might open a door for additional research.

Including liner notes would also be valuable – no resources currently do this!

Outreach: how might J-DISC reach people?

The first thought is to reach out to scholarly societies that jazz scholars are currently involved with: AMS, SCM, SMT, Society for American music, Jazz Education Network. Ads in jazz education publications

might also be a possibility. Other possibilities include the Music Library Association, the Association for Recorded Sound Collections (who also have a journal), or other European contacts (via Wolfram Nauer). If the quantitative work with the data cropped up, a showcase of research uses of the tool – to let people know what's possible – would be fantastic. Music library association. Association for recorded sound collections (and they have a journal).

Crowdsourcing

This interviewee generally feels this is a realistic possibility. They already do comparable things – prepublication reviews for scholarly services, book reviews for no money, contributions to Grove. "There's a precedent for this sort of sourcing from a large group of people." The important thing is the "vetting" process – a mechanism for people to be recognized as contributors. If it were reviewed that way, some status might also accrue to being a vetted contributor, such as for tenure-track faculty. They do have to start making an argument, however, that "discography is more than simply typing in data." This might become a very real possibility if there were peer review for this!

MUSICIAN
Background
This interviewee has been a professional musician for 24 years, dealing with jazz and other genres. At this point, he spends more time recording than in live concerts—something like 60-70% of his time is recording now.

Nuances of sessions/performances
Performances everywhere (even random live ones) are increasingly recorded and archived in places. More and more venues do this as a matter of course; some make these outputs available to anyone, while others make the recordings available to the artists for a fee. Still, some of these aren't made available. And there are a lot of other types of performances that aren't available either—outtakes and tracks may not be put out at all. A lot of records are just in development, and there's a lot of interested, unreleased music that's recorded. This unreleased music probably just lives unmixed in hard drives all over the world. He "can't imagine trying to chase down every performance and every recording and outtake." He doesn't keep track of that at all—even with his own groups.

As a listener, he occasionally gets access to outtakes and alternate tracks, but only when they're released commercially. And occasionally people will have access to multitracks and they'll appear online on Youtube or torrent sites.

Resources he uses right now
He seeks out information on jazz pretty often. Most of it is *not* historical, but he occasionally is curious to see who played on what record. Or, if he's playing with someone he doesn't know, he'll search out information on who they are, what they've done.

For this, he usually winds up on Allmusic or Discogs; Wikipedia is also a frequently consulted resource for more historical information. He's usually looking to see who played on what record, who wrote a particular song, or when two different musicians worked together. This is really for background, to learn more about whatever the obsession of the week is—there's no particular goal. These aren't always in-depth resources though, and might not have all the data he'll need (particularly for Discogs, which he suspected wouldn't be as good for older music).

There's no consistent source for other sorts of information. He'll turn to individual discographies on occasion, but this is a very ad hoc search process—he'll just go to Google and find whatever looks right.

Needs from resources
He likes when things are full cross-referenced. IMDB is great in this regard—you can click on a person and see his full information. In terms of general characteristics, it's important that a resource be accurate, well laid out, and chronological. He has his discography available on Discogs; when asked what he'd value between breadth vs. depth/accuracy, he said he would prefer something smaller and more accurate, *so long as it includes all commercial releases*—those commercial releases are a priority.

Inaccuracy is a real problem with existing resources. Records that have been released in different territories or under other names, for example, is a problem. Tracks that show up on compilations can also present complications. As a result, it might look like you have 200 records to go through, when the actual number of records is only like 80. These types of problems can be immediately evident—you can see that misspellings are listed as different artists. Checking his own discography, for example, he can see how certain things are repeated.

There isn't really any information that's *impossible* to find: "everything is probably somewhere at this point."

Other things that might be nice: audio clips might be helpful. It might also be good to have some jazz writing. For example, he doesn't have the patience to read all Miles Davis biographies; having access to anecdotes, different musical relationships, etc. would be helpful. Anything that connects the actual story to the recordings—there are "so many incredible stories in jazz." This could also be helpful for listeners when going back to the recordings.

Collections

He has a fairly significant collection, but no great way to keep track of records. At the moment, he thinks most people are just using their iTunes library and keeping records organized in alphabetical order, but these aren't great solutions.

Contributions

Very plausible that jazz fans would contribute. "Jazz fans are probably pretty obsessive—this would be a perfect outlet." And there are a lot of stories that need to be told, that people are excited to contribute details to.

The concern would be over policing accuracy—the site's value would be enhanced if these things were reviewed. (And he'd want to know *how* it was confirmed.) Furthermore, these types of user contributions might need to be set apart from "official" records. This is mostly just an issue for informal performances where you can't rely on a formal check in the same way as a commercially released recording (where there's a hard copy you can check against).

Musicians would also want to be included, especially if it established itself as *the* discography for jazz. Might be a useful marketing mechanism if so, since at the moment all his personal marketing efforts are basically word of mouth. (He doesn't even have a website right now!)

STUDENT

Background

This interviewee is an MLIS student who also studied jazz history previously. As part of his master's program in jazz history, he did a lot of archival work, which led him to library school. He maintains a few connections back in the jazz community, particularly because he now works on an online jazz data project.

Resources he used

He didn't actually use discographies extensively. In most cases, he found it easier to use AllMusic vs. Lord's; the searchability and use was much easier than Lord's. He did have to use the Lord discography for his master's thesis on Sonny Sharrock, however. It's a fairly comprehensive resource, and "it's just what they'd use at [his institution]." It's difficult to use, however, because you have to be really precise with what you search. The results also aren't great, and they're ordered in an unintuitive way. AllMusic was much easier in terms of searching for parameters and also had a much more modern interface. That said, "if you want to do deep serious research, definitely Lord."

For Lord, he's normally search for a certain musician or group, looking for all the stuff they'd produce, particularly looking for personnel information. This would be a discovery tool, in order to find out what else those performers have done. (He did, however, say that he hadn't done a huge amount of scholarly discography work.)

Other than Lord and AllMusic, he hasn't used any other extremely comprehensive discographies. All other independent resources are found through Googling and seeing what comes up, or asking other musicians or scholars about websites that have discographies. In order to verify those results, however, he finds other resources and sees how the information matches up. Resources that are obviously crowdsourced, that have a lot of people involved in their creation, are seen as especially reliable.

In addition to these, he follows Twitter and other blogs to find information on what's going on in jazz.

Discography value

Potential strengths for a discography could include ease of use (modern UI, easily navigable), sample tracks, extra metadata or enhanced content (such as album cover) that "adds to the experience of the record," and hyperlinks to other links to content.

Negatives or frustrations include when searches don't function properly, ambiguities in track titles, missing information (e.g., liner notes, genre)

Content to include

There's a "who's who" of jazz that the PI for this resource should know. As a jazz enthusiast though, the main focus, in his mind, should be bebop and pre-bebop.

That said, there are advantages to focusing on more obscure music. "Everyone knows all these recordings of Miles and Coltrane – they might want more obscure recordings." So having earlier and more obscure recordings might have some advantages, benefiting researchers who are doing things outside of what's already been done, or musicians delving deeper into jazz music.

Crowdsourcing

The interviewee thinks it would be plausible to rely on a contribution system, and would consider contributing himself. Jazz researchers pride themselves in knowing things that other people don't, so would be motivated to contribute. Some people just enjoy putting information there to be put out there, because this passion for the music is so strong in these researchers – that might be enough. But that said, the specifics of what they'd do would depend on the individual researcher; some are strict discographers, while others are more relaxed.

Individual recognition might also help, but there's a potential danger in terms of people butting heads and disagreeing over information. And the problem of conflicting information would be difficult to resolve. The important thing would be to make the interface easy enough.

He'd actually look favorably on a crowdsourcing mechanism – in some ways, even better than a big partnership because it gets the community involved. His opinion of this wouldn't change if they were uncredentialed; it's important to keep in mind that people can be experts without having a piece of paper denoting them as such.

MUSICIAN
Background
This interviewee is a jazz trumpet player and composer. For a while, he played gigs very regularly – four nights a week or so – before he took a job at a music publisher. Around that period, he worked a variety of jobs: gigs at night, working at a college, composing jingles, performing at weddings, pop gigs, etc. He became a composer of commercial jingles in 2000, at a firm where he stayed for about 11 years, before starting his own company. He now no longer considers himself a full time player or teacher, but rather focuses on writing.

Information searching
As a jazz musician, he hears music on the radio (oftentimes WKCR) that's not in the mainstream – not just jazz, but other music as well. He'll search on the internet using the artist as a keyword, then listen to recordings via iTunes. iTunes, in his view, is pretty extensive, though he "can't say that iTunes has everything." Particularly as a composer, the ability to listen to the music is very valuable, so iTunes and Youtube are very helpful. He's also starting to use Spotify a little bit. It's a great tool in his view, but his main issue with it is how they treat artist royalties – fractions of pennies for the performers. All of these resources are also not fantastic in terms of dealing with variations.

The only other information he's ever really concerned about is personnel, which he says iTunes and Spotify are terrible with. They (and radio stations as well) will usually just say the lead and won't include the whole band – the whole personnel is what he draws from jazz resources. There are cases, however, where he'll be really inspired by someone or something and want to delve in much further. This is mostly up to other audiences though, such as educators and collectors, who are "just obsessed." (At this point, he jokingly laughed at those who would care about metadata fields like "record number.")

Still, anything that helps delineate the history of our music properly is a huge help. A lot of this stuff could get washed away – it's not popular anymore – and he thinks it would be a shame if it wasn't archived in its form.

Discographies
Perhaps most interestingly, he revealed that he didn't actually use discographies! He described how in college, he did a paper on jazz in the civil rights movement and about half the information he got was from liner notes. He also mentioned that at one point, for a few of his very favorite players, he could recite the order in which they recorded records and where they were recorded – but again, that was from liner notes. (He then mentioned that he wasn't sure how kids listen to music anymore, if they even look at liner notes, or if they're fully available as parts of electronic packages.)

Other than that, he has never really used products like Lord, BRIAN, Rust, etc.

Scope
He personally doesn't agree with the 1955-1970 scope. In his view, the biggest changes in the music occurred from mid-20s to mid-50s; music moved most quickly in this early era.

Initial impressions of J-DISC
"Amazing resource" – he liked the way it was laid out. Having all this sort of information would be nice, he said, and educationally it seemed pretty awesome. The question of sources was unclear though, and he recommended that we speak with Phil Schaap about how he finds sources for his information.

He also noted that if this could be connected to people's recordings and videos, that'd be great, especially since there are communities based around Youtube and such.

Crowdsourcing

He doesn't really think crowdsourcing is a great possibility. He expects that the younger population will need a quicker solution and won't be willing to commit time to contribute – but they're also the primary users. And even if you do, he's skeptical that J-DISC would be able to get enough people.

That said, the idea of an interactive site sounded "amazing."

Other future possibilities

He thinks the idea of having his own recordings in there would be very helpful for him, and he expects other people would agree. He only has a couple records as a leader, but he's played on a lot of sideman projects. There aren't any great sources to find that sort of information though, because places like iTunes or the radio only report the leadperson. On the major music sites, including iTunes, he says: "You have extremely famous people, and everything else, and that everything else is a disaster."

In that sense, J-DISC could have a future as a music network. It could be a testbed for more contemporary musicians and partner with places like Smalls. This could be especially helpful since many musicians don't market themselves at all – they rely on reputation or word of mouth.

PUBLISHER

Background

This interviewee is a publisher of music textbooks for the higher education market, and in previous positions published many trade and reference works on jazz. He is also a published author on country and roots music and an accomplished amateur musician.

The market for J-DISC

In his view, a discography would be most appropriate for record collectors and people who are very interested in hearing about who recorded when. For the casual listener, though, it likely wouldn't be too helpful, unless you had a big audio database that they could listen to. Institutional users could be schools with really deep jazz programs – but this is a subset of music schools, which is already a small group. And most of these are performance programs rather than scholarship programs, so it would be hard to find a huge audience.

Biographical information would be helpful. But by far the most important thing would be sound and video to make it attractive, which is the most difficult to grab, especially because a lot of audio services don't include jazz. Only a handful of places, like Rutgers-Newark, would be really excited about this.

Other resources

Grove Jazz might be interesting to take into consideration, since it doesn't have deep discographical information. AllMusic Guide is also a competitor, though it's aimed at consumers and not very scholarly. Various label discographies are also relevant.

If all these were consolidated in one place in an authoritative way, that would be of some value to people in the industry. Without the information about the performers themselves though, it really assumes that know who the performers are and assumes you just want to find out who played with Charlie Parker; this assumption heavily limits the audience though. If something like J-DISC had the definitive biography of Charlie Parker linked to deep discographic information, however, it might be more relevant to more people. Deep discographic information will always be the most high-end type of information, appealing to a very small base.

Considering partnerships with these sorts of resources might be very important going forward: places like Alexander St. Press, Smithsonian Folkways, etc.

Value

Accuracy is important, especially in the field of jazz. He tries to independently verify everything because there is so much disagreement. That said, you can often judge by the quality of the stuff that's up there. He noted also that if someone cares enough to put up a site about this sort of thing, there's a good chance they're probably a nut about it and usually that information is reliable. And names like Brian Rust are really reliable: if you say something like "we have Brian Rust…" or someone who did the authoritative discography of Blue Note—then why recreate that?

Crowdsourcing

The notion of crowdsourcing information is very thin right now. The resource needs to be indisputably accurate, but he thinks a crowdsourced resource would be reliable enough – this is the sort of thing people argue about all the time already.

The 10-15 people who get into it (Wikipedia style) would argue endlessly, bringing up a fundamental question: is this a site facilitating discussion, or will it publish an ultimate decision? He would prefer the latter – "I'm not interested in reading people arguing". One compromise might be to open the record, but leave the commenting field open to the public.

That changes, however, if you could link to real recording logs – for example, by making deal with Sony (Victor and Columbia) to get and digitize logs. (Ancestry.com links to real census data, for instance.) This would legitimize the data as well as add to the "fetish value." IJS has filing cards from old recordings, and big labels and studios might have that data as well. Downbeat, Google Books, Grove, might also be potential partners here.

If you look at something like the book on the making of Kind of Blue and see its popularity – that's really "the ultimate fetish book." J-DISC could tap into that type of enthusiasm this way.

Content scope

Releasing by label seems to make the most sense. You can bill it as "the definitive database of everything recorded by Prestige" and release in sections that way. One way to start might be to choose one of the medium-sized labels (such as Concord, Independent), which might have paperwork from the 1920s on. There's a decent chance J-DISC could get that paperwork and start testing this idea.

This leads to a really cool possibility of getting information that has never been available or digitized: session logs, photos of the session (a lot of these studios had house photographers), notes, information gathered during the session, among other things.

This would come with a lot of information – Music Library Association would have guidelines here – and would also set it apart as more than just a discography. Some people might bristle at this, since there are two schools of thinking: some people who think it's just music, and other who think it's sociocultural.

Not into collection-tracking like GoodReads. He's not a collector, so he wouldn't care. People are pretty secretive about what they own, so not sure they'd want it to be a public thing. And some of them are really traditionalist – one person uses a typewriter!! So many database programs out there that he doesn't see the value. Maybe some sort of social media value there?

MUSICIAN
Background
This interviewee is a fairly renowned professional jazz musician who has also written a couple books and lectured at various universities. He has a particularly keen interest in jazz history and the sociocultural side of jazz music.

Resources he uses
He's worked mostly with Rust; don't have a copy of Lord. The Rust book is quite good – amazed that he got so much information in there. He's very trustworthy, but there's information that has come to light since then. He printed corrections every now and then as well, and he sees it as probably about 90% reliable. Some of it will never be resolved, some of it can be by scholarship, listening to records closer, consulting other records.

Whenever possible, he likes to consult primary sources as well. The IJS is a great resource for this. He also likes first person accounts and oral histories, and wishes they were digitized – he trusts these more than other secondary sources. One problem with sources from the 1950s and 1960s is that they often didn't have footnotes or sources, whereas newer outputs from academic presses do.

Needs from a discography
He's mostly happy with the amount of detail on existing sources; he's mainly interested in personnel, date, studio, and label, which are in most discographies. There isn't really anything that he's been unable to find, or that's been overly difficult to find. Completeness of information and authority are the two most important qualities.

It would be nice, however, if something had "everything that was ever said by X who were X's contemporaries all in one source, so you knew everything that was there." It's also really nice if there's a lot of cross-referencing, so you can click from link to link. There should also be an easy way to correct data or suggest changes; the idea of a Wikipedia-style debate page sounds interesting. "Ease of use" also figured heavily into how he'd judge a discography.

Content to include
No strong position on the content to include. It really "depends on what people's interests are." He would start by "dumping everything in" and seeing where there are discrepancies and disagreements. The issue there is with copyright – he wonders if there are any copyright issues here. In any case, before you release this to the public, consolidate it and designate where the information is coming from.

Organizing by label and such would be helpful, though he "thinks there would be Blue Note discographies already." You want to start out with a narrow slice of something.

Crowdsourcing
He thinks that people would be willing to contribute/correct/etc. There are scholars who just want to share their work, and there are people who just want to hoard their work.

Appendix D: Potential audiences

Guide used by Ithaka S+R to better understand the end users. This grid could serve J-DISC for future promotional efforts.

Category	Name	Size	About the organization	How they would use J-DISC	What they value in J-DISC
Enthusiasts	Down Beat (www.downbeat.com)	70,000 subscribers	Magazine that contains news and articles and jazz that is both old an new.		
	Jazz Times	100,000 subscribers 56,738 members of the online community	Primarily a magazine, but also has an associated website that contains many articles. Has an online community for discussion		
	www.JazzReview.com	10,692 members in community	Website that has lots of reviews on jazz CD's. Also has an online community for discussions and information about performances. Reviews contain some discographic information but this does not seem to be the main attraction	• Look up specific details about recording sessions (like personnel) • research historical relationship between musicians • find a community of other jazz enthusiasts	
	www.jazzjournal.co.uk	Has a forum that has some activity. Last comment was posted in may. Every couple of months they get a comment	Journal that has reviews, news and articles about jazz. They claim to have a large jazz discography as well		
	http://www.allaboutjazz.com/	Average monthly Visits: 2,300,000 Unique: 870,000	Free online journal. Provides reviews, articles and links to songs and videos. Lots of data on recent jazz. Discussion board where people discus, promote CDs, review jazz and sell things		• Interesting interface • Useful organization of data • Collaborative community • exhaustive list of musicians and sessions • Accuracy would be important but not the most important
	Jazz Society of Pensacola		provide an online community around jazz performance, education and enjoyment for listeners and musicians. Website that has many features. They are also creating a library of jazz CD's		
	Society for American Music		This organization is focused on American music in general and not just jazz. The mission of the organization is to stimulate: appreciation, performances, listening and study of American music.	• Look up information on both artists and the origin of particular songs • Study how particular styles/musicians evolved and influenced each other • find other fans of America music	
	America's Finest City Dixieland Jazz Society	1,100 members in the San Diego area	dedicated to the preservation and perpetuation Dixieland music. They have monthly concerts, annual festivals, newsletters and work shops	• research historical information about Dixieland musicians	
	New York Jazz Report				
	Jazz Societies				
	Blogs: DoTheMath and AboutLastNight			• Devoted community who have not seen their interests addressed effectively in discographies, hopefully J-DISC can meet their needs	
	Listserv: Jazz listserv, Miles Davis listserv, Latin jazz listserv				
Educators	Jazz Educators Journal ('93-'03)	Estimated 28,000 readers	Publication for jazz educators	• Use as teaching resource • Reference for research • Tool to get students excited and interested in jazz history	• Accuracy of information • Useful organization of information
Industry Professionals	Jazz Record Labels: Columbia Records, Blue Note, Prestige, Impulse, Verve	Very large? Hard to quantify—demand for jazz metadata is unknown	Organizations sell metadata to companies like iTunes, providing services like audio recognition and matching album data with CD's	• J-DISC could sell metadata to these companies	• Accuracy of information • useful organization of metadata

Appendix D: Potential audiences (continued)

Category	Name	Size	About the organization	How they would use J-DISC	What they value in J-DISC
Scholars & Academics	America Musicological Society (AMS)	3,300 individuals and 1,200 institutions	Purpose is to create a community of music scholars. They focus on all American music, not just jazz. They also publish a journal and have annual conventions.	•authors could use as reference and research • readers could use J-DISC to check facts • examine relationship between musicians and over time • find academic community around jazz discography	• Accuracy of information • useful organization of metadata. Both is easy to search specific facts but also organized in an intuitive way that fosters research--allowing scholars to find new connections between musicians over time • foster collaborative work
	Current Research in Jazz http://www.crj-online.org/	Universities and libraries around the world	This is an open source academic journal. Mostly aimed at presenting historical information.		
	http://jjs.libraries.rutgers.edu	18 members on the editing and publishing team	Open source academic journal. It is aimed at academics and enthusiasts alike. The journal contains articles on analysis as well as historical articles.		
	http://www.equinoxpub.com/JAZZ/index	24 on the editorial team	This is an academic journal. They publish articles on a wide variety of subjects related to jazz.		
	College or University Fine Arts Faculty (United States only)	165,178, but this figure includes all fine arts faculty. Music is a subset within this and jazz studies is an even smaller subset within this	Fine arts faculty in institutions of higher education in the US. Note that jazz music faculty would be a very small subset of this list.	•faculty could use a reference material • use as a research tool--finding new relationships between musicians over time • use as a teaching tool for students (more likely to be uses in teaching here as opposed to K-12 • create scholarly community around jazz discography	
Discographers	The IAJRC Magazine http://www.iajrc.org	415 likes on Facebook 4 people in conversation on their page—page created in 2009	Quarterly magazine aimed at collectors and discographers. It contains articles about jazz and discographic information	•learn about each recording session--find additional data • discover new songs/musicians •find new ties between musicians •	• Accurate data above all else • Useful organization of data • Collaborative community • exhaustive list of musicians and sessions
	International Association of Sounds and Audiovisual Archives (Discography committee)	400 members in over 70 countries	Their stated mission is to preserve audio and audiovisual recordings. However, within the organization they have a discography branch--concerned with documenting metadata. Vincent Pelote, of the Rutgers Institute of Jazz Studies was the president.	•learn about each recording session--find additional data • discover new songs/musicians • add artists from their country--nationalist motivations • assists record companies to publish full discographic information	
Jazz Musicians	Jazz musicians in New Orleans	1,700		• input their own discographic information: both serve as an organizational tool for themselves as well as advertising--help them to spread their music •research older musicians--both for historical research and research of musical ideas • look for other musicians they would be interesting in collaborating with	• useful organization of the data • accuracy of data • exhaustive list of musicians and sessions • community--ability to connect with other musicians
	Jazz Musicians in New York	33,000	active fulltime jazz musicians in each city		
	Musicians Union				
	Jazz musicians in San Francisco	18,700			