

## **The J-DISC Project: Background, Mission, and Recent Work**

J-DISC is built on the idea that deep, rich knowledge about jazz can be discovered through its recorded legacy. The recording studio is an essential locus of artistic production for improvisers: it is where they create, recreate, and strive to perfect their work, mixing composed repertoire with new melodic material, and constantly shifting instruments and personnel. Above all, the recording process preserves creative sparks that arise from their interaction. The knowledge of the people, repertoire, and production and consumption processes involved in the domain of jazz recording can help in turn study or explore the professional and social networks, diverse subcultures, and artistic choices of jazz musicians across the history of the music. By incorporating dispersed sources into a single repository now available at [jdisc.columbia.edu](http://jdisc.columbia.edu), we submit that knowledge about jazz recording to be readily shared, updated, and edited.

J-DISC features advanced searches that allow users to demonstrate important relationships between diverse recordings. J-DISC is flexible and extensible: built from open-source code such as Drupal and MySQL, it allows our team, technical administrators, and users to continuously update and modify and add new data and metadata as they become available. J-DISC is open-access: all of the content will be available to the public in searchable form on JSO and the underlying methodology will be fully documented there.

Finally, [jdisc.columbia.edu](http://jdisc.columbia.edu) is collaborative, to an extent never possible in previous jazz discography. It allows discographic experts not just to enhance the data, but also to share information about data that may be used and improved. The database thus offers scholars the tools to compare improvisations on a given song, trace the development and artistic collaborations of a given artist, observe trends in social and historical factors in jazz recordings, and explore data on the dissemination of jazz by the music industry. Scholars may also annotate and enhance the data, evaluate sources for their accuracy, and share information about their research within the application.

The project to create an online jazz discography began in 2010, and a significant new phase was added to our work in 2012. The J-DISC team began to apply Music Information Retrieval techniques to jazz discography by developing tools for providing access to and analyzing digital jazz audio files. It is now within our reach to link and use rich textual information on jazz recordings, the corresponding audio recording files they describe, and an array of tools to generate insight from the musical content of those recordings in the not-so-distant future.

## **Music Information Retrieval and Jazz: Current Work and Capabilities**

In August 2012, the J-DISC MIR team hired a Postdoctoral Researcher, Brian McFee, to perform research and development work for the MIR portion of the project. Brian addressed several challenges for any contemporary jazz history researcher accessing digital jazz recordings; and

also some related challenges for the field of MIR, which has hardly ever addressed the distinctive musical challenges of jazz. We believe this work will serve as foundation for the more advanced set of tools mentioned above that can manage and analyze jazz recordings.

There is now a huge corpus of online resources with poor metadata on the soloists, accompanists, repertoire, composers, and other essential artistic elements of jazz. To begin to address that challenge, Brian is developing means to enhance and manage the text annotation of these recordings, using computational tools and metadata best practices. We are creating a sample dataset of about 10,000 songs chosen from Columbia's digital collections to test these methods and to serve as a model for managing audio resources that will complement and can eventually be linked to the text database at [jdisc.columbia.edu](http://jdisc.columbia.edu) or to online jazz audio repositories.

The J-DISC MIR team also explored signal processing methods to probe the digital audio files of jazz performances themselves and identify, describe, and analyze their contents. For our purposes, audio signal processing refers to the alteration of [auditory signals](#), or [sound](#), through mathematical operations on the digital representation of that signal. The avoidance in jazz performance of obviously repeated material, its constant variation on the same basic repertoire, and its subtle rhythmic nuances and layering all make it difficult to train computers to recognize rhythmic and melodic patterns that human listeners identify and reconstruct when they hear the same audio signal. Nevertheless, the sheer volume of available audio files and the data they contain opens up the prospect of mining large quantities of files to uncover musically significant trends or tropes.

Brian has proceeded by refining beat-tracking algorithms currently used in MIR, which simply identify the loudest notes and assume that significant structural musical events coincide with them. The existing techniques do not apply well to jazz, where significant beats or rhythmic events are almost never the loudest ones, due to syncopation and rhythmic flux. Human listeners can detect the significant beats by filtering out louder events that occur on off beats and other "noise" or chaotic elements, but machines cannot. In response, Brian developed an algorithm that is more sensitive to the presence of loud, off-beat acoustic events by pinpointing and emphasizing events that occur together in time, rather than simply the loudest ones, and inferring that the former represent markers of a common underlying pulse. Based on the resulting data, we can more reliably construct a basic rhythmic infrastructure, or "virtual beat," for the whole audio sample, then track minute deviations from it at any given point. As our work advances, we can in turn use the beat-tracking tool to create a frame of reference to search for musically significant events such as chords or melodic fragments that may fall on those beats.

Brian is also developing and enhancing techniques for segmentation, which detect structural elements of a musical performance by aggregating small-scale temporal samples of an audio file. Brian developed an algorithm that produces a hierarchical decomposition of the song into progressively finer segments. That process facilitates simultaneous structural annotation at many layers of musical texture or points along the audio time continuum, and allows comparison of the machine-generated segments against the judgments of musical structure by human annotators (which we also want to make more use of in testing). With finely tuned criteria for performance segments, we will have the tools to reliably identify larger musical structures such as chord sequences or melodies. Another benefit of developing these criteria is to provide an interface for users to quickly and visually digest the structure of the piece.

Finally, Brian implemented some basic algorithms to predict the predominant instrument in a given recorded performance, by developing audio descriptors which characterize the timbre of individual instruments. This model will be used to identify the active instruments at each point in time throughout a recording, which will in turn be linked to performer metadata to provide a highly specific and detailed index of each performer's recordings.

We expect to use our model dataset to analyze audio data using signal processing tests during the next year of the project. Capabilities we believe we can achieve include detection of heads vs. solos, chords, and basic rhythmic figures (to indicate an artist's style or a genre). Brian created a demonstration that suggests what our rhythmic, structural, and feature extraction capability will be when it becomes fully functional, which can be found at <http://porkpie.ee.columbia.edu:5000>.

In the future, to link the program to a given dataset at any location, we can make use of the Echo Nest firm's Project Rosetta Stone to cross-link our data with other publicly available music databases (e.g., MusicBrainz) and streaming radio services (e.g., Spotify or Rdio). That would allow researchers or musicians to plug them into any web program that plays the song they are interested in, and helps us refer playback options for people who want to listen to examples we discuss in or around J-DISC. It should be noted they would still not be able to analyze the data within the file itself.

### **Future Plans for J-DISC and J-DISC MIR**

We believe our current work in the J-DISC MIR project can make a substantial contribution to the understanding of music as well as to information management for music resources. Building on work in the present project, we aim to develop a set of powerful, integrated computational tools to represent, manage, and analyze jazz recordings in digital audio form. The resulting tools will generate musical insights, aggregate and organize evidence, and provide

access to resources for researchers, performers, and educators engaged with jazz, improvisation, musicology, or acoustic science. Our audio-derived computational tools will create a representation of a musical performance, analogous to a traditional musical score. They will highlight structural features by noting musical elements such as keys, bars, and repeated motifs. But they will also capture nuances of pitch, timbre, accent, and timing, which are not easily shown in conventional scores and yet are vital to jazz performance (and to most other forms of music).

Our ultimate aim is to give music researchers an objective document that they can use to analyze and discuss structural features and idiomatic expressive elements, without having to rely solely on partial and subjective human transcriptions of improvised performances. The tools will help generate and aggregate this evidence across a very large sample of jazz recordings and provide access to relevant examples. Based on our current work, we expect to be able to identify the song forms and chord structures typically guiding jazz improvisations; the soloists' phrases, such as the sequences horn players or singers construct between breaths; or definite, repetitive melodic fragments improvisers use to construct their large-scale improvisations. Finally, we want to be able to capture subtle musical practices that are vital to jazz expression but are difficult to capture in traditional notation, such as speech-like bends and glissandos, accents, and subtle discrepancies in timing that provide rhythmic energy to jazz.