Historical Recorded Sound Database Gets a Reboot

A team of researchers at the UCSB Library recently finished an exciting enhancement project to the Discography of American Historical Recordings (DAHR), an online database of over 314,000 master recordings made by American record companies during the 78rpm era. DAHR is the preeminent searchable database documenting historical commercial sound recordings in the US, with audio streaming of more than 30,000 recordings, available free-of-charge. It boasts an impressive 200,000 users each year and continues to grow in popularity, particularly as more digitized recordings are added.

With funding from the Library of Congress National Recording Preservation Board, a team of editors at the UCSB Library Special Research Collections spent a year integrating "Linked Data" into DAHR's features. At its core, Linked Data is a mechanism for identifying relationships between different online entities, making the data more meaningful and useful. It's a way of saying the Duke Ellington in our database is the same as the Duke Ellington any number of other databases and is becoming a standard across the web.

"Users of DAHR will immediately notice a richer experience," says Performing Arts Curator and project lead David Seubert. "Individual records now offer a much more complete picture of the performer and provide a valuable context for the historic audio recordings."

DAHR uses an identifying number assigned by the Library of Congress to establish these identities and linkages, which in turn can be linked to any number of other web databases. This technology is the core of many open source projects on the web, including Wikipedia.

By harvesting data from around the web using the Library of Congress identity, records in DAHR now have enhanced features like Wikipedia biographies, photographs, and links to many other databases. For example, Duke Ellington's DAHR listing has links to information in other online sources including Discogs, where you might research LP reissues, Allmusic, where you could read record reviews, or Spotify, where you could access streaming audio of Ellington.

Library catalogers can also traverse across the datasphere to almost any library authority database in the world through the Virtual International Authority File (VIAF) link. If DAHR has an artist's VIAF number (and it does for over 18,000 names), they can

determine who this person is in the authority files used by nearly all other libraries around the world, a tremendous benefit for the library community.

And for music lovers and collectors, this enhancement to DAHR will make it easier to find and understand rare early recordings, and enable a deeper understanding of the artist and their biographical history.

Used by researchers, musicologists and general music fans, DAHR had its genesis in 1964 as the Encyclopedic Discography of Victor Recordings, a project founded by private collectors Ted Fagan and William R. Moran. It initially documented recordings made by the Victor Talking Machine Company from original documents in the Victor (now Sony) Archives. It became an online database at UCSB in 2008 and streaming audio was added through a partnership with the Library of Congress in 2011. In 2013, DAHR became the primary project of the UCSB-based American Discography Project (ADP)— with the expansion of the project to document seven other early record companies with support from the Packard Humanities Institute.

Today, DAHR plays a critical role in researching early sound recordings and provides scholars around the world with accurate data compiled from primary sources, detailed biographical information, and audio content on a scale that didn't exist 10 years ago. By enhancing the data represented in DAHR, the UCSB Library has amplified our commitment to the access, understanding, and appreciation of this incredible treasure trove of our collective history.