

NOTICIA
DE INVESTIGACIÓN

UNA PERSPECTIVA DE BIG DATA PARA LA MÚSICA DE SALÓN CHILENA

*Dr. José Manuel Izquierdo König
Pontificia Universidad Católica de Chile*

*Dra. © Fernanda Vera Malhue
Universidad Metropolitana de Ciencias de la Educación
Chile*

La noción de “Big Data” se ha instalado con fuerza en las humanidades en los últimos años. Big Data, un concepto originado en la década de 1990, refiere a la capacidad de procesar enormes conjuntos de datos que, sin la ayuda de un *software* computacional, serían imposibles de analizar por parte de seres humanos. Naturalmente, debido al aumento explosivo de datos digitales en el mundo esto abre enormes nuevas posibilidades para la investigación en todo aspecto. Si bien en principio refiere al trabajo con enormes volúmenes de datos (por sobre los varios terabytes de información), progresivamente se ha ido aplicando a una forma de análisis por la cual podemos contemplar, mediante recursos estadísticos, volúmenes de data relativamente menores, pero que, de igual forma, no podrían ser contemplados si no es con la ayuda de un software dedicado, específico para alguna tarea. ¿Cómo podría una persona analizar millones de entradas, o incluso miles o cientos, estableciendo una comparación juiciosa?

Más allá de las perspectivas de uso de los datos, existen ciertos conceptos claves cuando pensamos en *Big Data*. En primer lugar, debe entenderse que esto no implica solamente un análisis estadístico. El análisis de enormes volúmenes de datos permite también encontrar patrones, así como establecer consistencias e inconsistencias en una serie de objetos que se consideran relativamente homogéneos. Por lo mismo, si bien uno podría pensar que el procesamiento de datos es una tarea más o menos automática (ejercida por un computador), en realidad dista mucho de serlo. Normalmente debe haber cierta consciencia por parte del investigador sobre qué es lo que se pretende encontrar, sobre

el volumen y variedad de los datos, las posibilidades de procesamiento, o la veracidad de la información. Existe una necesidad por ser crítico frente al volumen de información, permitiendo un análisis generado desde la perspectiva de la investigación.

Desde la musicología, la noción de “Big Data” ha tomado un vuelo importante en los últimos años. Si bien es difícil trazar exactamente cómo se ha instalado, pueden señalarse una serie de hitos importantes. En 2000 la ISMIR (*International Society for Music Information Retrieval*) se instaló como un foro de trabajo, estableciendo varios congresos durante la misma década y oficializándose como organización el año 2008.¹ El 2005 Nicholas Cook ya presagiaba que la creación de “data sets” de música, iba a “abrir áreas completamente nuevas para la musicología” en el futuro cercano.² El interés de esta asociación está en la pregunta sobre los modos en que la música puede transformarse en “datos”, en información que puede ser accedida, normalizada, sistematizada. Por otra parte, RISM (*Répertoire International des Sources Musicales*), comenzó a desarrollar en 1997 un proyecto de catálogo y base de datos online, pero solo el año 2008 ofreció por primera vez sus entradas (tradicionalmente impresas en papel) como base de datos, a través de EBSCO. El año 2013 la base de datos fue liberada como *open data*.³

Fue ese gesto, la liberación de los datos de RISM (ya rastreables a través de un buscador desde 2010 sin cargo) como información utilizable para proyectos de Big Data, el que realmente parece haber hecho despegar la relación entre musicología y esta nueva aproximación a la investigación. Desde entonces el problema se ha instalado con fuerza, a través de la serie de *International Digital Libraries for Musicology Workshop* (iniciadas en Londres el 2014 y ya en su cuarta versión), y publicaciones asociadas. “Big Data for Musicology” fue un paper colectivo publicado el 2014, que instaló varios de los problemas centrales:⁴ ¿Qué definimos como información o dato musical? ¿Cómo pasamos de la transcripción de música a los datos? ¿Qué hacemos con las enormes bases de datos musicales privadas, como Spotify o iTunes?

El año 2015 ocurren dos eventos de importancia para la relación entre musicología histórica y *Big Data*. El primero es el vínculo entre la Royal Holloway de Londres y la British Library para la generación del proyecto “A Big Data History of Music”, y el seminario “Digital Strategies for Historical Musicologists” del 11 de marzo de 2015, al cual uno de nosotros pudo asistir.

¹ <http://www.ismir.net/>. Consultado el 5 de abril de 2018.

² Nicholas Cook, “Towards the compleat musicologist?”, Charla realizada en ISMIR 2005. <http://ismir2005.ismir.net/documents/Cook-CompleatMusicologist.pdf> Consultado el 5 de abril de 2018.

³ <http://www.rism.info/organisation/project-history.html> Consultado el 10 de abril de 2018.

⁴ <http://openaccess.city.ac.uk/4077/6/Big%20Data%20for%20Musicology.pdf> Consultado el 8 de abril de 2018.

Como parte de este proyecto, el catálogo de música impresa y manuscrita de la British Library fue transformado en Open Data para que cualquier musicólogo pueda trabajarlo.⁵ El segundo evento fue la publicación, por la revista *Early Music* de un segundo volumen de “Early music and modern technology”. Allí, dentro de una serie de otros artículos relevantes (varios de los cuales planteaban un centro en “datos”, aparece el trabajo -derivado del proyecto anterior- de Stephen Rose, Sandra Tuppen y Loukia Drosopoulou “Writing a Big Data history of music” (p.649-660), que realmente instala y visibiliza las posibilidades que se generan con este nuevo tipo de análisis, replanteado importantes aspectos sobre la historia de la música impresa entre 1500 y 1700 en Europa.

Por otra parte, y en otro lugar del mundo, en Chile en el mismo periodo se ha generado un importante interés por catalogar colecciones musicales a niveles nunca vistos antes. Alejandro Vera comienza con su proyecto sobre la Catedral de Santiago (pronto a publicarse) hacia el año 2004, y en 2009, mediante un Proyecto Bicentenario, comienza también la catalogación de las colecciones del Teatro Municipal de Santiago, por parte de uno de los autores. Un año más tarde este interés nuestro se prolonga a una serie de proyectos, principalmente desarrollados por Fernanda Vera como la catalogación y puesta en valor del Fondo Musical del Seminario Pontificio Mayor de Santiago y la Sección Partituras del Archivo central Andrés Bello. El mismo equipo, liderado por Víctor Rondón, realiza la misma labor con la Colección Musical de la Biblioteca Patrimonial de la Recoleta Dominica de Santiago.

El volumen de datos generados solamente por estas tres últimas colecciones, supera con creces las tres mil partituras, principalmente del siglo XIX y comienzos del XX. Debido al enorme volumen de material a interpretar es que nació la idea de realizar un trabajo de análisis de “Big Data” sobre aquellas partituras. Pero, ¿Cómo iniciar el trabajo, hacia dónde enfocarlo? La propuesta, trabajada durante el año 2017 mediante un Fondo de la Música, tuvo como objetivo principal generar una base de datos única, la cual pudiese ser limpiada (homologada y homogeneizada lo más posible sin alterar los datos) para cruzar la información sobre un corpus musical específico conformado por la música secular impresa en Chile durante el largo siglo XIX (esto es, aproximadamente desde 1840 hasta 1910).

⁵ <http://www.bl.uk/bibliographic/download.html> Consultado el 10 de abril de 2018.

En primer lugar, se creó una base original a partir de una tabla de *Microsoft Excel* con los siguientes campos, homologando la información desde las bases de datos ya existentes:

ARCHIVO	Lugar donde se encuentra la partitura
AUTOR	Autor señalado de la "obra"
ARR.	Si es arreglo y no obra original
EXT.	Si la persona es extranjera (no chilena)
AUTOR 2	En caso de arreglo o colaboración
NACIONALIDAD	Del autor/a
GÉNERO	Del autor/a (hombre, mujer)
AÑO	Año conocido de impresión
TÍTULO	Título simplificado de la pieza
INSTRUMENTO	Instrumento para el que está escrita
GÉNERO	Género musical señalado
OPUS	Si es que hay indicación de opus
TEMÁTICA	Temática asociada a la obra (ver más abajo)
COLECCIÓN	Si pertenece a una colección más amplia
LETRA	Autor de la letra
DEDICATORIA	Si es dedicada a una persona
EDITOR	Quién es el editor/impresor
LUGAR DE EDICIÓN	Dónde fue editada la pieza
PLANCHA	Si hay número de plancha

Esta información fue completada tanto desde las bases de datos que ya teníamos (mencionadas anteriormente) así como también de la base de datos de acceso público de la Biblioteca Nacional, la cual se complementó con trabajo de archivo en el Archivo de Música de la misma Biblioteca. Además, se sumó información de otros archivos regionales, como el Museo Histórico de Valdivia, el Museo Regional de Atacama, la Biblioteca del Departamento de Música de la Universidad de la Serena, y también del centro DAE del Teatro Municipal de Santiago, más otras colecciones privadas. Se utilizó una entrada por cada impresión original de obra realizada en Chile, y se cotejó información con dos bases de datos antiguas en papel: la *Bibliografía musical* de Ramón Laval⁶ y *Bibliografía* de Eugenio Pereira Salas,⁷ que cubren partituras impresas en Chile en el mismo periodo, en un rango más reducido. A la fecha de esta publicación,

⁶ Ramón Laval, *Bibliografía Musical. Composiciones impresas en Chile y Composiciones de autores chilenos publicadas en el extranjero*. Segunda parte 1886- 1896. Santiago: Establecimiento Poligráfico Roma, 1898.

⁷ Eugenio Pereira Salas, *Bibliografía musical de Chile desde los orígenes a 1886*. Santiago: U. de Chile, 1978.

esto nos ha dado un volumen aproximado de 1500 obras/ediciones únicas realizadas en Chile de música que podemos considerar “chilena”.

Como podrá verse en la tabla superior, nuestra decisión sobre la base de datos estaba ajustada a nuestros propios intereses como investigadores, por lo que se desarrollaron campos que no son entregados explícitamente por las partituras señaladas. Por ejemplo, puede referirse el campo de “Género del Autor”, que nos permitiría establecer vínculos entre ciertos repertorios y una prevalencia en ellos de compositores mujeres u hombres. Del mismo modo podríamos hacerlo con “géneros musicales”, o con la nacionalidad de los autores. El espacio para “temáticas” fue decidido para tratar de entender de qué “hablan” los títulos de las partituras de salón del periodo: sentimentales, políticas, religiosas, progresistas, naturaleza, lugares, personas, comerciales, o simplemente no llevan título y solo descripción de género (vals, polka).

En este sentido, se debe recalcar que nuestra base de datos no es un catálogo ni inventario, sino que un cuerpo de datos (de información recabada y homologada) que permite un análisis complejo de un gran volumen de información puesto que estamos hablando de más de 13000 entradas únicas de información en total, un volumen que ya podemos considerar, sin mayor esfuerzo, como de Big Data. Además, decidimos incorporar nueva información que creímos ha sido frecuentemente omitida de estudios similares, en particular el número de plancha, información vital si nuestro centro está enfocado en ediciones impresas e impresores, puesto que revela una cronología de impresión, así como también si se trata de la misma o distinta edición de una obra, entre otros detalles interesantes.

Si bien no estamos aún en condiciones de entregar un resultado definitivo de esta investigación, si consideramos que esta base de datos, que esperamos hacer pública el año 2019 como *open data*, nos ha permitido releer la música del siglo XIX bajo nuevos lentes. Utilizando el software Tableau (de libre uso para trabajos sin fines de lucro) y como resultado de esta primera aproximación a partir de un proyecto del Fondo de la Música Nacional, ya podemos destacar ciertas tendencias que nos parecen relevantes y que difícilmente cambien a futuro:

1. Es evidente que la colección más grande de música de este tipo es la de la Biblioteca Nacional de Chile, pero la misma es seguida de cerca por el Archivo Andrés Bello y el Museo Regional de Atacama. Este último, en particular, fue una sorpresa para nosotros.
2. Un importante número de autores que aparecen entre los primeros lugares en número de ediciones y sus obras son totalmente desconocidos para la musicología nacional y aparecen estadísticamente como más

relevantes que otros más conocidos. Ejemplos clave de esto son Rodolfo Lucero Villegas, E.J. Hermosillas y Guillermo Wetzer.

3. Pareciera que existe una cierta asociación entre géneros musicales e instrumentación que no había sido destacada antes y que podría necesitar un análisis mayor. En particular, creemos que debe destacarse el vínculo entre la guitarra y la mazurca, el que además se expande a la relación de la mazurca con compositoras mujeres.
4. Valparaíso duplica, al menos, las ediciones musicales realizadas en Santiago de Chile, y es probable que, si reducimos las partituras solo a aquellas antes de 1890, el número sea aún mayor; esta preponderancia de Valparaíso como centro editorial también necesita un análisis mayor.
5. Dentro de las temáticas asociadas a las partituras, es la “sentimental” y personal la más frecuente, seguida luego por la temática política, siempre frecuente, pero de marcada presencia en años de conflictos, como la Guerra Civil de 1891.

Lo interesante de estas conclusiones es que no habríamos podido llegar a ellas, o afirmarlas de modo contundente sin el trabajo con Big Data, sin un volumen de datos interrogado a través de un modelo de software. Creemos que esta perspectiva podría con facilidad aplicarse a otros trabajos locales y regionales de importancia que ya han acumulado enormes cantidades de datos: la base de datos de programas y funciones del Teatro Municipal de Santiago, o las de discos generadas por proyectos como Fonoteca Nacional de Chile, la que incluye el Cancionero Discográfico de Cuecas Chilenas, el Archivo de Música Tropical Chilena y el Archivo de Jazz Chileno. Este trabajo también nos lleva a preguntarnos por las posibilidades de libre acceso a estas bases de datos, en un formato *open* que permita su trabajo estadístico y el análisis de información por investigadores externos a tales proyectos. ¿Son exportables las bases de datos chilenas? ¿Qué pasa con las bases de datos de instituciones como el Pequeño Derecho de Autor, o la misma SCD? ¿Qué tipos de investigaciones podrían generarse sobre estas bases, y como consolidar esta nueva veta de investigación en Chile, también pionera para América Latina? Esperamos que nuestro trabajo temprano sobre Big Data y procesamiento de información, aquí reseñado en sus problemas fundamentales (a la espera de un artículo futuro con mayores resultados específicos de tipo historiográfico), sea un aliciente a otros trabajos y perspectivas en estas mismas líneas.