The SIMSSA Project: Search as access to digital music libraries

Emily Hopkins
SIMSSA Project Manager, McGill University

Workshop on SIMSSA XV, IAML, Leipzig, 28 July 2018
• SSHRC Partnership Grant (2014-2021)

• PI: Ichiro Fujinaga (McGill University)

• Partners include the British Library, Bodleian Libraries at Oxford, Bibliothèque Nationale de France, Bavarian State Library, New York Philharmonic Archives, Alexander Street Press, RILM, and RISM Switzerland among others
How it works:

1. Library digitizes scores
2. Optical Music Recognition
3. Symbolic Encoding with MEI
4. Search and Analysis
• How do we access the scores?

• How can we teach computers to read musical scores?

• How will music search and analysis work?
How do we access the scores?
International Image Interoperability Framework
Selected items

Français 20000
Bibliothèque nationale de France

N-13752
Bibliothèque nationale de France

2 Songs, Op.91
The Internet Archive
How can we teach computers to read musical scores?
Optical Character Recognition

• Makes images of text machine-readable
• XML
Optical Music Recognition

• Makes images of sheet music machine-readable

• MIDI, MusicXML, MEI
Music Encoding Initiative (MEI)

Example borrowed from the MEI tutorial at music-encoding.org; music is Robert Schumann’s Der Abendstern.
Sonate
(in B dur)
für das Pianoforte componirt
von
FRANZ SCHUBERT.

Molto moderato.

(Componirt im September 1828.)
Pixel.js: Making ground truth data
Pixelwise Classification

Interactive Classifier: Identifying glyphs & training our OMR
Neon.js: Correcting OMR output
Scoring-up Tool

Superius

Tenor

Contratenor

Scoring-up Tool

File 1
_______
_______
_______

Superius part

File 2
_______
_______
_______

Contratenor part

File 3
_______
_______
_______

Tenor part

Scoring-up Tool

Output File
_______
_______
_______

SCORE

RODAN: Web-based distributed OMR workflow system (Hankinson 2014)
Crowdsourced OMR Correction

- Making tools more user-friendly
- Collaboration with Partner organizations and user communities
How will music search and analysis work?
Melodic Search

https://patternfinder.elvisproject.ca/
Corpus Studies


A musical ‘word’ or 2-gram: 2 vertical intervals linked by a melodic motion of the lower voice (upper voice implied).
Percentage of perfect sonorities for all pieces
MEASURING POLYPHONY

DIGITAL ENCODINGS OF LATE MEDIEVAL MUSIC

http://measuringpolyphony.org/

<table>
<thead>
<tr>
<th>Regression model</th>
<th>Mode</th>
<th>Mode family</th>
</tr>
</thead>
<tbody>
<tr>
<td>test data: leaps and outlines</td>
<td>36%</td>
<td>67%</td>
</tr>
<tr>
<td>comparison set 1: remainder notes</td>
<td>39%</td>
<td>68%</td>
</tr>
<tr>
<td>comparison set 2: pc distributions</td>
<td>45%</td>
<td>71%</td>
</tr>
</tbody>
</table>

| Experiment w/ experts                         |      |             |
| experiment 1: pc tallies                     | 35%  | 65%         |
| experiment 2: pitch, interval size & direction| 39%  | 61%         |
| Full score experiment                         | 67.5%| 100%        |
Machine learning and composer identification


<table>
<thead>
<tr>
<th>Composer Group</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>All 7</td>
<td>92.7%</td>
</tr>
<tr>
<td>Ockeghem / Busnoys / Martini</td>
<td>87.2%</td>
</tr>
<tr>
<td>Ockeghem / Busnoys</td>
<td>84.4%</td>
</tr>
<tr>
<td>Ockeghem / Martini</td>
<td>94.6%</td>
</tr>
<tr>
<td>Busnoys / Martini</td>
<td>93.8%</td>
</tr>
<tr>
<td>Josquin / Ockeghem</td>
<td>93.9%</td>
</tr>
<tr>
<td>Josquin / Busnoys</td>
<td>96.0%</td>
</tr>
<tr>
<td>Josquin / Martini</td>
<td>88.2%</td>
</tr>
<tr>
<td>Josquin / La Rue</td>
<td>85.4%</td>
</tr>
<tr>
<td>Victoria / Palestrina</td>
<td>99.9%</td>
</tr>
</tbody>
</table>
Thank you!

https://simssa.ca/
emily.hopkins@mcgill.ca

@simssaproxject
@e_a_hopkins