

Building Chopin Early Editions

Tod A. Olson

University of Chicago Library
Digital Library Development Center

Graduate School of Library and Information Science
University of Illinois at Urbana Champaign

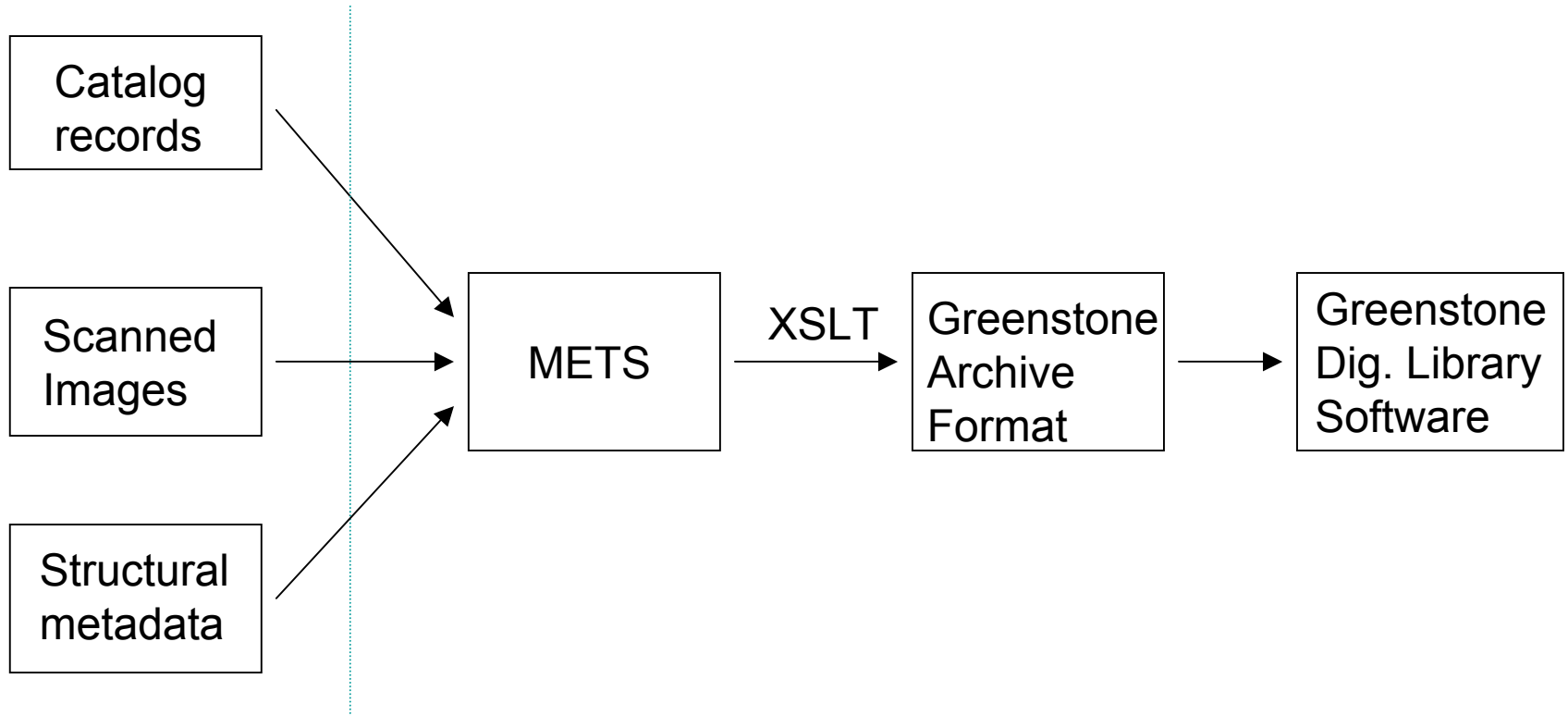
ISMIR 2003, Baltimore, MD

October 28, 2003

Introduction

- 420+ physical scores, published 1830-1880
- 370+ scanned and online
- Site live in March, 2003
- Nearly 100 hits/day avg.
- 30% traffic is international, all continents
- Highest international use:
Argentina and Brazil

Production stream



Human processing

XML-based automated processing

Catalog records

Bib #: 1561329

LDR 01253ccm 2200337 a 4500

008 981117q18481856enkncz n c

100 1 †a Chopin, Frédéric, †d 1810-1849.

240 10 †a Nocturnes, †m piano, †n op. 55

245 10 †a 15me. & 16me. nocturno / †c [composé par Frederic Chopin].

246 3 †a Quinzième et seizième nocturno

260 †a London (No. 229, Regent Street, corner of Hanover Street) :

†b Wessel & Co., importers and publishers of foreign music,

†c [between 1848 and 1856]

300 †a 10 p. of music ; †c 33 cm.

490 1 †a Wessel & Co.'s complete collection of the compositions of
Frederic Chopin for the piano forte ; †v no. 59

500 †a "Dédié à Mademoiselle J.W. Stirling"--Caption.

650 0 †a Piano music.

800 1 †a Chopin, Frédéric, †d 1810-1849. †t Piano
music (London, England) ; †v no. 59.

Catalog records

1. Descriptive metadata

- Describe scores
- Distinguish between similar scores
- Provide access to scores
- Shows how one score relates to rest of collection

2. Information taken primarily from physical score: title, composer, publisher, place published, dedication, etc.

3. Some information taken from other sources

Example: few scores have publication dates printed, take these from outside research

Inconsistent use of descriptive terms

E.g., same work published under different titles:

15me. & 16me. nocturno

2 nocturnes pour le piano, op. 55

Deux nocturnes pour le pianoforte, op. 55

Gather all versions together by *uniform title* (rules for uniform titles codified by AACR2):

Nocturnes, piano, op. 55

Related work: Functional Requirements for Bibliographic Records (FRBR) establishes a shallow hierarchical grouping of “sameness” for organizing multiply published works, see <http://www.ifla.org/>.

Scanned images

- Created according to National Archives and Records Administration guidelines.
- 400dpi, 24-bit color, uncompressed TIFF
- No touchups, rescan rather than retouch
- Produce two JPEG files from each TIFF, 2000- and 700-pixel wide
- Testing underway for DjVu versions.
- Files stored by naming convention based on score and image sequence.

Significant details in scores are preserved:



1 in.



1/4 in.

Structural metadata

Document = score, object = image within score

chopin_structural : Table								
	document	object	page	project	milestone	filesize	imagewidth	imageheight
	107	014		chopin		83.7 MB	4464	6549
	108	001		chopin		85.5 MB	4552	6559
	108	002		chopin		85.2 MB	4536	6559
	108	003	1	chopin	Nocturne, no.15	83.1 MB	4440	6539
	108	004	2	chopin		83.8 MB	4496	6508
	108	005	3	chopin		83.8 MB	4448	6570
	108	006	4	chopin		84.0 MB	4464	6580
	108	007	5	chopin		83.4 MB	4448	6549
	108	008	6	chopin	Nocturne, no.16	84.0 MB	4464	6570
	108	009	7	chopin		84.0 MB	4464	6570
	108	010	8	chopin		84.0 MB	4488	6539
	108	011	9	chopin		86.0 MB	4492	6539
	108	012	10	chopin		85.4 MB	4528	6590
	109	001		chopin		91.6 MB	4728	6765

Structural metadata

- Proper sequence of images for each score
- Features from score image
 - Page number as printed
 - Milestones: cover, title page, piece within score, etc.
- Technical and administrative metadata: files sizes, image dimensions, software and settings
 - Do not yet use this data

Metadata Encoding & Transmission Standard (METS)

- Digital library standard for encapsulating objects with their metadata
- OAIS lingo: use METS for SIP, AIP, DIP
- Share digital objects between institutions
- Share work of building tools to produce, store, display digital objects
- Library of Congress maintenance agency:
<http://www.loc.gov/standards/mets/>

METS structure

Seven sections:

- METS Header
- Descriptive metadata*
- Administrative metadata
- File list*
- Link structure
- Structure map*
- Behavioral section

* Chopin Early Editions currently uses only 3 sections

Metadata Object Description Schema (MODS)

- METS does not prescribe a descriptive metadata encoding, uses extension schemas
- Flexible XML encoding of library data
- Maintained by LoC: <http://www.loc.gov/standards/mods/>

```
<mdWrap MDTYPE="OTHER" OTHERMDTYPE="MODS">
  <xmlData>
    <mods:mods>
      <mods:titleInfo type="uniform">
        <mods:title>Nocturnes, piano,</mods:title>
        <mods:partNumber>op. 55</mods:partNumber>
      </mods:titleInfo>
      <mods:name type="personal">
        <mods:namePart>Chopin, Frédéric,</mods:namePart>
        <mods:namePart type="date">1810-1849.</mods:namePart>
      </mods:name>
    </mods:mods>
  </xmlData>
</mdWrap>
```

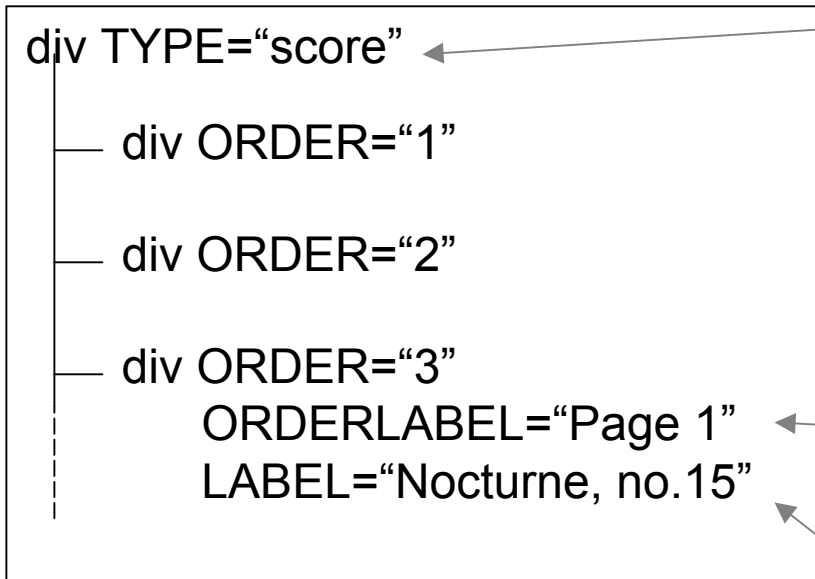
METS file list

Files can be carried internally, or linked to externally.

```
<fileSec>
  <fileGrp> <!-- 2000 pixel wide JPEGs -->
    <file ID="JPGH108002" MIMETYPE="image/jpeg">
      <FLocat LOCTYPE="URL" xlink:href="http://.../chopin108-002r.jpg"/>
    </file>
    <file ID="JPGH108003" MIMETYPE="image/jpeg">
      <FLocat LOCTYPE="URL" xlink:href="http://.../chopin108-003r.jpg"/>
    </file>
  </fileGrp>
  <fileGrp> <!-- 700 pixel wide JPEGs -->
    <file ID="JPGL108002" MIMETYPE="image/jpeg">
      <FLocat LOCTYPE="URL" xlink:href="http://.../chopin108-002q.jpg"/>
    </file>
    <file ID="JPGL108003" MIMETYPE="image/jpeg">
      <FLocat LOCTYPE="URL" xlink:href="http://.../chopin108-003q.jpg"/>
    </file>
  </fileGrp>
</fileSec>
```

METS Chopin structure

Structure map

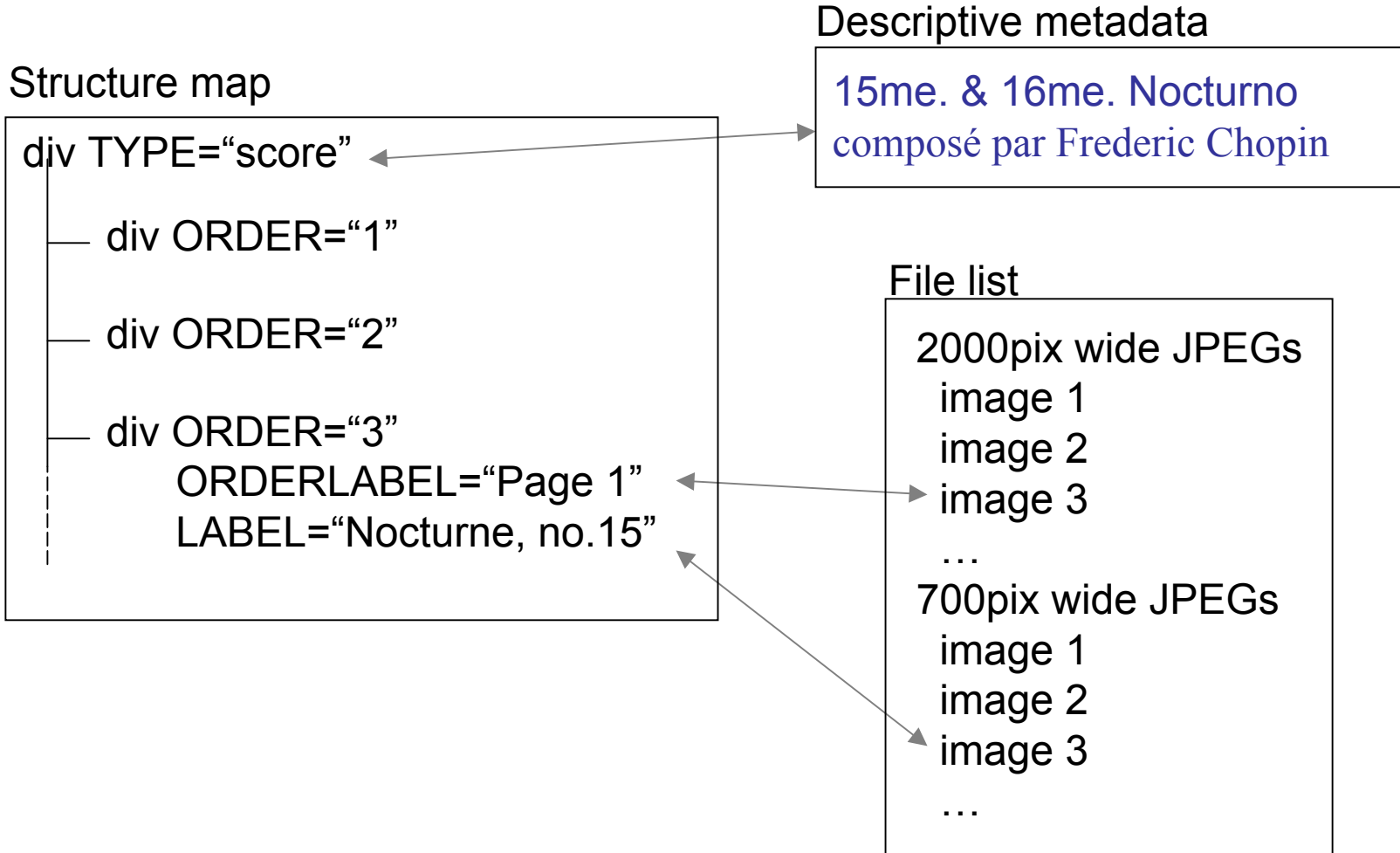


Descriptive metadata

15me. & 16me. Nocturno
composé par Frederic Chopin

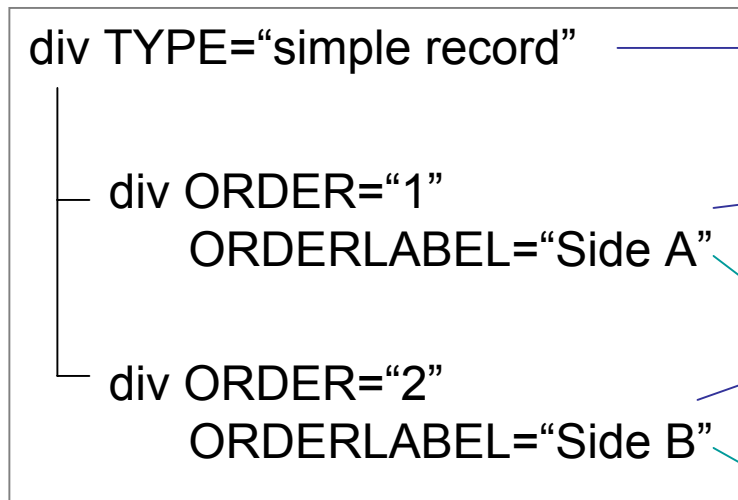
File list

2000pix wide JPEGs
image 1
image 2
image 3
...
700pix wide JPEGs
image 1
image 2
image 3
...



Example from LoC: sound records (45's)

Structure map



Descriptive metadata

Columbia Records, no. C1234
Jelly Roll Morton

King Porter Stomp

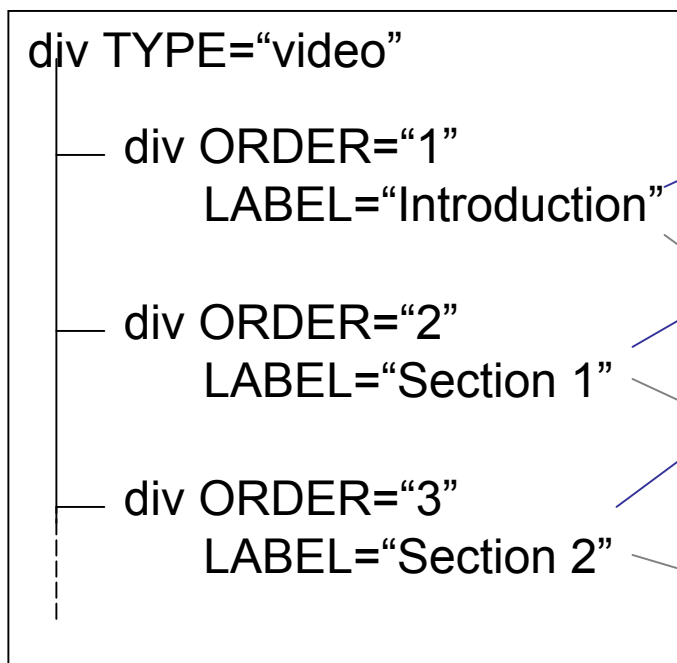
Wolverine Blues

File list

king.mpg
song2.mpg

Example from NYU: video w/ transcript

Structure map



File list

QuickTime video, 5 min.

0:00 - 1:28

1:29 - 2:33

2:34 - 4:29

...

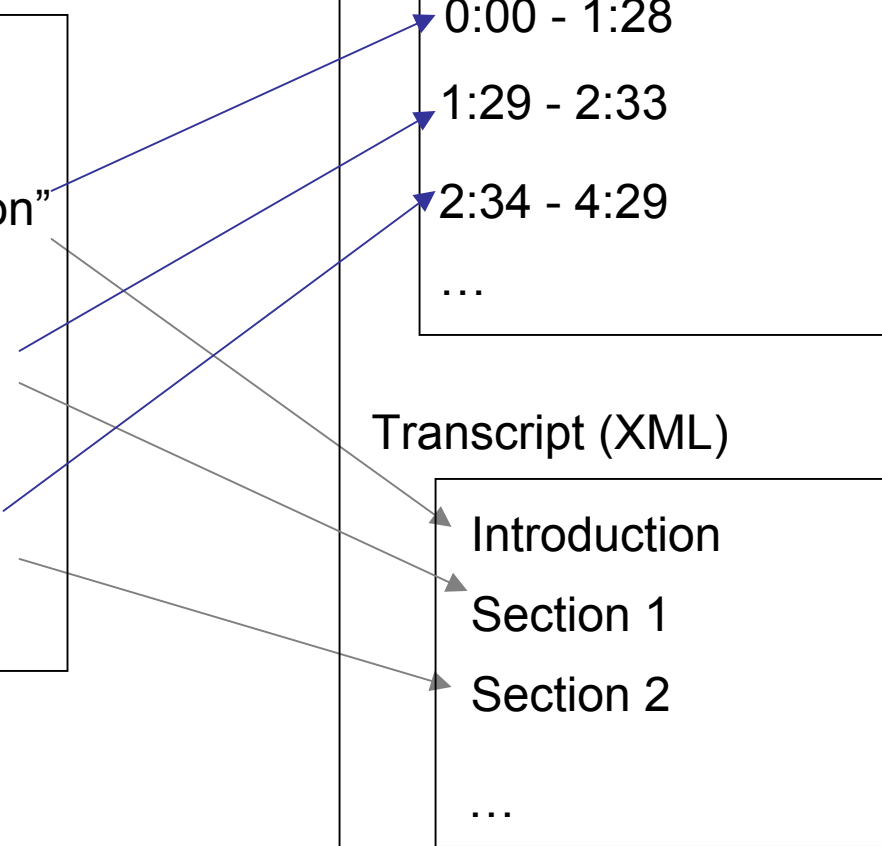
Transcript (XML)

Introduction

Section 1

Section 2

...




Greenstone

- Handles arbitrary descriptive metadata
- Supports hierarchical document structure
- Configurable user interface
- <http://www.greenstone.org/>

Greenstone Archive Format

- Matches METS hierarchical object structure
- METS transformed to GSAF via XSLT
- Metadata normalized for US keyboards:
 - Title: Quinzième et seizième nocturno
 - Titleidx: Quinzieme et seizieme nocturno
- Place names modified for improved retrieval:
 - Place: London
 - Placeidx: London London Londres


Descriptive metadata for navigating collection

Chopin | Early Editions 

search titles uniform titles genres

-  [Ballades](#), (16 scores)
-  [Barcarolle](#), (2 scores)
-  [Berceuse](#), (6 scores)
-  [Bolero](#), (5 scores)
-  [Concertos](#), (16 scores)
-  [Etudes](#), (16 scores)
-  [Ecossaisses](#), (2 scores)
-  [Fantasia](#), (6 scores)
-  [Mazurkas](#), (65 scores)
-  [Nocturnes](#), (40 scores)

Descriptive metadata for navigating collection

Chopin | Early Editions 





search

titles


uniform titles

genres

Nocturnes, (38 scores)

-  [13e. et 14e. nocturne pour piano, op. 48 /](#),
Paris (Rue Richelieu, 97) : Chez Maurice Schlesinger ; [1841]
Nocturnes, piano,op. 48. No. 1
-  [13e. et 14e. nocturne pour piano, op. 48 /](#),
Paris : Chez Maurice Schlesinger (Rue Richelieu, 97) ; [1841]
Nocturnes, piano,op. 48. No. 2
-  [13e. et 14e. nocturne pour piano, op. 48 /](#),
Paris (Rue Richelieu, 97) : Chez Maurice Schlesinger ; [1841]
Nocturnes, piano,op. 48. No. 1
-  [15me. & 16me. nocturno /](#),

Structural metadata for navigating document

Chopin | Early Editions 

HELP PREFERENCES

search

titles

uniform titles

genres

dedicatees

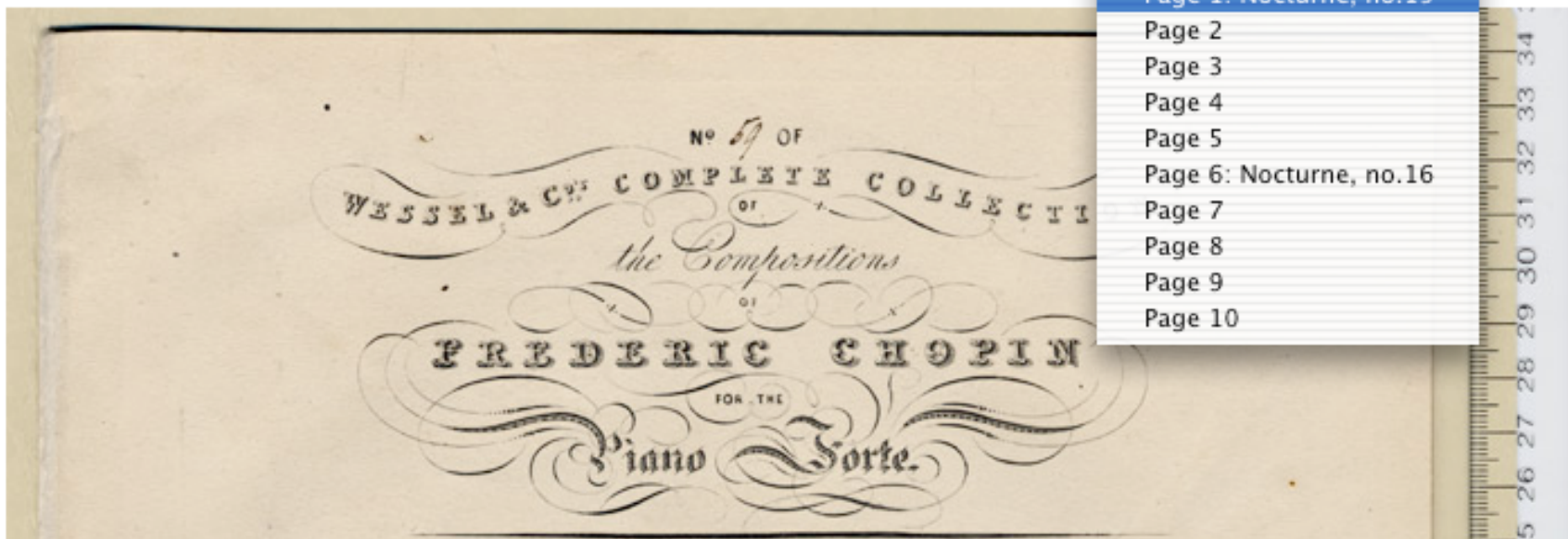
15me. & 16me. nocturno /

[next](#) ►

view score

[bibliographic description](#)

unnumbered page



Benefits

- Flexibility/extensibility
 - Accommodate different descriptive metadata sources (e.g., Dublin Core)
 - Accommodate additional types of data (e.g. sound files)
- Reuse
 - Of production stream for other projects
 - Of METS objects for different applications (e.g. OAI harvesting)

Future

- Integrate sound-based indexing (MelDEX?)
- Add representative performances
- OMR?
- User interface
 - Usability testing
 - Content-based thumbnails?
 - Sound? Piano scroll?
- ???